MP/MP*

Exercices d'algèbre et de probabilités

David Delaunay

- RÉSUMÉS DE COURS
- MÉTHODES
- 3 NIVEAUX D'EXERCICES :
  - apprentissage
  - entraînement
  - approfondissement
- CORRIGÉS DÉTAILLÉS PAS À PAS

deboeck SUPÉRIEUR

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : **www.deboecksuperieur.com**

Rue du Bosquet, 7 B-1348 Louvain-la-Neuve

1ère édition, 2017
1er tirage, 2017

Imprimé aux Pays-Bas.

Dépôt légal:
Dépôt légal France : juin 2017
Dépôt légal Belgique : 2017/13647/088

ISBN : 978-2-8073-0627-1

La pratique d'exercices est essentielle à l'apprentissage du cours de mathématiques : il n'est pas de meilleure façon de mémoriser et de comprendre un théorème que d'en faire usage !

Cet ouvrage regroupe sur 9 chapitres 317 exercices portant sur le programme d'algèbre et de probabilités en classe de MP. Il respecte strictement le programme en cours et vient compléter l'ouvrage d'*analyse* que l'on retrouvera dans la même collection.

Chaque chapitre commence par un rappel des principales définitions et des résultats essentiels du cours. Il se poursuit avec des exercices aux corrigés détaillés regroupés sur trois niveaux :

- *Les exercices d'apprentissage* servent à l'acquisition des concepts fondamentaux du cours. Ce sont souvent des sujets faciles où j'ai choisi volontairement de ne faire figurer que peu de technicité.
- *Les exercices d'entraînement* permettent de poursuivre l'acquisition du cours, trois niveaux d'étoiles servent à anticiper leur difficulté. Ces sujets ont été choisis pour leur intérêt, leur classicisme ou ont été inspirés par des questions rencontrées aux écrits et aux oraux des différents concours.
- *Les exercices d'approfondissement* sont les plus ambitieux, ils nécessitent souvent de passer par une phase de recherche ou entrent en résonance avec d'autres chapitres du programme. Ces sujets sont inspirés de questions rencontrées aux concours les plus ambitieux.

Les corrections des exercices sont accompagnées de *méthodes*. Celles-ci servent à souligner les idées récurrentes ou bien à mettre en exergue la démarche qui va être suivie pour résoudre la question posée. Le lecteur pourra prendre appui sur celles-ci pour amorcer une résolution ou pour reprendre la main lors de sa lecture d'une correction. Afin d'aider le lecteur dans son étude, il est fait référence aux théorèmes utilisés lors de leurs premiers usages. Les notes de bas de pages complètent les résolutions en présentant des démarches alternatives ou font le lien avec d'autres sujets présents dans l'ouvrage.

Je remercie vivement Olivier RODOT d'avoir initié ce projet, François PANTIGNY pour son expertise TeXnique et Sébastien MARCOTTE pour sa relecture attentive ainsi que les compléments apportés.

Je dédicace cet ouvrage à mon fils Nathan.

David DELAUNAY

CHAPITRE 1

Groupes

1.1 Structure de groupe

1.1.1 Groupe

Définition

On appelle *groupe* tout couple $(G, \star)$ formé d'un ensemble G et d'une loi de composition interne $\star$ sur G associative, possédant un neutre et pour laquelle tout élément de G est symétrisable.

$(\mathbb{C}, +)$ et $(\mathbb{C}^*, \times)$ sont des groupes commutatifs de neutres respectifs 0 et 1.
L'ensemble $\mathcal{S}_E$ des permutations d'un ensemble E est un groupe pour la composition des applications $\circ$. Son neutre est l'application identité Id_E.
L'ensemble $\mathrm{GL}_n(\mathbb{K})$ des matrices inversibles de taille n est un groupe multiplicatif de neutre I_n.
L'ensemble $A^\times$ des inversibles d'un anneau est un groupe multiplicatif.

1.1.2 Structure produit

Définition

Si $\star_1, \ldots, \star_n$ sont des lois de composition internes sur des ensembles $E_1, \ldots, E_n$, on appelle *loi produit* sur $E = E_1 \times \cdots \times E_n$ la loi $\star$ définie par

$$(x_1, \ldots, x_n) \star (y_1, \ldots, y_n) \overset{\text{déf}}{=} (x_1 \star_1 y_1, \ldots, x_n \star_n y_n).$$

Théorème 1
Si $(G_1, \star_1), \ldots, (G_n, \star_n)$ sont des groupes de neutres respectifs $e_1, \ldots, e_n$ alors $G_1 \times \cdots \times G_n$ muni de la loi produit $\star$ est un groupe de neutre $e = (e_1, \ldots, e_n)$.

Le symétrique d'un élément $(x_1, \ldots, x_n)$ de $G_1 \times \cdots \times G_n$ est alors $\left(x_1^{-1}, \ldots, x_n^{-1}\right)$.

1.1.3 Sous-groupes

Définition
On appelle *sous-groupe* d'un groupe $(G, \star)$ toute partie H de G non vide et stable par composition avec le symétrique[1] : $x \star y^{-1} \in H$ pour tous x et y dans H.

Rappelons que si H est un sous-groupe de $(G, \star)$ alors $(H, \star)$ est un groupe.

Théorème 2
Si $(H_i)_{i \in I}$ est une famille de sous-groupes d'un groupe $(G, \star)$ alors l'intersection

$$H = \bigcap_{i \in I} H_i$$

est un sous-groupe de $(G, \star)$.

En revanche, la réunion de deux sous-groupes peut ne pas être un sous-groupe[2].

1.1.4 Sous-groupe engendré par une partie

Soit $(G, \star)$ un groupe.

Définition
On appelle *sous-groupe engendré par une partie* A de G l'intersection de tous les sous-groupes de $(G, \star)$ contenant A. On le note $\langle A \rangle$.

Théorème 3
$\langle A \rangle$ est un sous-groupe de $(G, \star)$ contenant A.
De plus, tout sous-groupe de $(G, \star)$ contenant A contient aussi $\langle A \rangle$.

Au sens de l'inclusion, $\langle A \rangle$ est le plus petit sous-groupe de $(G, \star)$ contenant A.

Lorsque a désigne un élément de G, le groupe engendré par $\{a\}$, simplement noté $\langle a \rangle$, est l'ensemble des itérés[3] de a

$$\langle a \rangle = \left\{a^k \mid k \in \mathbb{Z}\right\}.$$

Lorsque $\langle A \rangle = G$, on dit que A est une *partie génératrice* de G.

1. Si la loi du groupe est notée additivement, on lit $x - y \in H$ pour tous x et y dans H.
2. Voir sujet 1 p. 8.
3. Si la loi du groupe est notée additivement, on lit $\langle a \rangle = \left\{ka \mid k \in \mathbb{Z}\right\}$.

1.1.5 Sous-groupes de $(\mathbb{Z}, +)$

Théorème 4
Les sous-groupes de $(\mathbb{Z}, +)$ sont les $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$ avec $n \in \mathbb{N}$.

1.2 Morphismes de groupes

Soit $(G, \star)$ et $(G', \top)$ deux groupes de neutres e et e'.

1.2.1 Définition

Définition
On appelle *morphisme* du groupe $(G, \star)$ vers le groupe $(G', \top)$ toute application $\varphi : G \to G'$ vérifiant $\varphi(x \star y) = \varphi(x) \top \varphi(y)$ pour tous x et y de G.

L'exponentielle est un morphisme du groupe $(\mathbb{C}, +)$ vers $(\mathbb{C}^*, \times)$.
Le logarithme est un morphisme du groupe $(\mathbb{R}_+^*, \times)$ vers $(\mathbb{R}, +)$.
Le déterminant définit un morphisme du groupe $(\mathrm{GL}_n(\mathbb{K}), \times)$ vers $(\mathbb{K}^*, \times)$.
La signature d'une permutation définit un morphisme de $(\mathcal{S}_n, \circ)$ vers $(\{1, -1\}, \times)$.
Les applications linéaires entre espaces vectoriels sont des morphismes de groupes additifs.

Théorème 5
Si $\varphi : G \to G'$ est un morphisme de groupes alors $\varphi(e) = e'$, $\varphi(x^{-1}) = \varphi(x)^{-1}$ et, plus généralement, $\varphi(x^n) = \varphi(x)^n$ pour tout $x \in G$ et tout $n \in \mathbb{Z}$.

La composée de deux morphismes de groupes est un morphisme de groupes.

1.2.2 Image et noyau

Théorème 6
L'image directe (resp. réciproque) d'un sous-groupe par un morphisme de groupes est un sous-groupe.

Définition
Si φ est un morphisme du groupe $(G, \star)$ vers le groupe $(G', \top)$, on introduit :
— son *noyau* $\mathrm{Ker}(\varphi) = \varphi^{-1}(\{e'\})$ qui est un sous-groupe de $(G, \star)$;
— son *image* $\mathrm{Im}(\varphi) = \varphi(G)$ qui est un sous-groupe de $(G', \top)$.

Théorème 7
Soit φ un morphisme du groupe $(G, \star)$ vers le groupe $(G', \top)$.
a) φ est injectif si, et seulement si, $\mathrm{Ker}(\varphi) = \{e\}$.
b) φ est surjectif si, et seulement si, $\mathrm{Im}(\varphi) = G'$.

1.2.3 Isomorphisme de groupes

Définition

On appelle *isomorphisme* de groupes tout morphisme de groupes bijectif.

Lorsqu'il existe un isomorphisme entre deux groupes, ceux-ci sont dits *isomorphes* : ils sont alors parfaitement identiques d'un point de vue calculatoire.

Théorème 8

La composée de deux isomorphismes de groupes est un isomorphisme de groupes et la bijection réciproque d'un isomorphisme de groupes est un isomorphisme de groupes.

Définition

On appelle *automorphisme* du groupe $(G, \star)$ tout isomorphisme du groupe $(G, \star)$ dans lui-même.

L'ensemble des automorphismes d'un groupe $(G, \star)$ est un sous-groupe de $(\mathcal{S}_G, \circ)$.

1.3 Groupes monogènes

n désigne un entier naturel non nul.

1.3.1 L'ensemble $\mathbb{Z}/n\mathbb{Z}$

La congruence modulo n définit une relation d'équivalence sur $\mathbb{Z}$:

$$a \equiv b \ [n] \iff n \mid (b - a).$$

On note $\bar{a}$ la classe d'équivalence de $a \in \mathbb{Z}$ pour cette relation.

Définition

L'ensemble des classes d'équivalence pour la relation de congruence modulo n est noté $\mathbb{Z}/n\mathbb{Z}$.

Théorème 9

$\mathbb{Z}/n\mathbb{Z}$ est un ensemble fini à n éléments : $\overline{0}, \overline{1}, \ldots \overline{n-1}$.

Définition

On définit une opération d'addition [1] et de multiplication sur $\mathbb{Z}/n\mathbb{Z}$ en posant pour tous a et b dans $\mathbb{Z}$

$$\overline{a} + \overline{b} \overset{\text{déf}}{=} \overline{a+b} \quad \text{et} \quad \overline{a} \times \overline{b} \overset{\text{déf}}{=} \overline{ab}.$$

1. Dans l'égalité $\overline{a} + \overline{b} = \overline{a+b}$ le symbole $+$ désigne deux opérations différentes. Dans le premier membre, il s'agit de l'addition dans $\mathbb{Z}/n\mathbb{Z}$ que l'on définit. Dans le second membre, il s'agit de l'addition sur $\mathbb{Z}$. On a évidemment la même remarque concernant le produit.

Les résultats de ces opérations ne dépendent pas des représentants choisis des classes d'équivalence car la congruence modulo n est compatible avec les opérations d'addition et de multiplication :

$$\begin{cases} a \equiv a' \ [n] \\ b \equiv b' \ [n] \end{cases} \Longrightarrow \begin{cases} a+b \equiv a'+b' \ [n] \\ ab \equiv a'b' \ [n]. \end{cases}$$

1.3.2 Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$

Théorème 10
$(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe abélien de neutre $\overline{0}$.

L'opposé de $\overline{a}$ est $\overline{(-a)}$. Plus généralement, on vérifie $k\overline{a} = \overline{ka}$ pour tout $k \in \mathbb{Z}$.

1.3.3 Groupes monogènes, groupes cycliques

Définition
Un groupe $(G, \star)$ est dit *monogène* lorsqu'il existe un élément a qui l'engendre :

$$G = \langle a \rangle = \{a^k \mid k \in \mathbb{Z}\}.$$

Un tel élément a s'appelle alors un *générateur* du groupe $(G, \star)$.

$(\mathbb{Z}, +)$ est un groupe monogène engendré [1] par 1 (ou par -1).

Un groupe monogène est assurément commutatif car $a^k \star a^\ell = a^{k+\ell} = a^\ell \star a^k$ pour tous k et ℓ de $\mathbb{Z}$.

Définition
Un groupe $(G, \star)$ est dit cyclique lorsqu'il est monogène et fini.

Théorème 11
$(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe cyclique de cardinal n.
Ses générateurs sont les $\overline{m}$ pour tout entier m premier avec n.

On peut alors décrire *à isomorphisme près* les groupes monogènes :

Théorème 12
Tout groupe monogène infini est isomorphe à $(\mathbb{Z}, +)$.
Tout groupe monogène fini de cardinal n est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

En particulier, on peut considérer le groupe $(\mathbb{U}_n, \times)$ des racines n-ièmes de l'unité :

$$\mathbb{U}_n = \{\omega^k \mid k \in [\![0\,;n-1]\!]\} \quad \text{avec} \quad \omega = \mathrm{e}^{2\mathrm{i}\pi/n}.$$

Ce groupe est cyclique engendré par ω. Il est par conséquent isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

1. La loi étant additive, on comprend $\langle 1 \rangle = \{k1 \mid k \in \mathbb{Z}\} = \mathbb{Z}$.

1.3.4 Ordre d'un élément dans un groupe

Définition

On dit qu'un élément a d'un groupe $(G,\star)$ est d'*ordre fini* lorsqu'il existe[1] $n \in \mathbb{N}^*$ vérifiant $a^n = e$. On appelle alors *ordre* de a le plus petit n de $\mathbb{N}^*$ vérifiant $a^n = e$.

Théorème 13

Si un élément a est d'ordre fini égal à n alors, pour tous k et $\ell \in \mathbb{Z}$,

$$a^k = a^\ell \iff k \equiv \ell \ [n].$$

En particulier, $a^k = e$ si, et seulement si, n divise k.

L'ordre n de l'élément a apparaît alors comme le cardinal du groupe $\langle a \rangle$ et ce dernier est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

Théorème 14

Dans un groupe fini de cardinal n tous les éléments sont d'ordre fini et leur ordre divise le cardinal du groupe.

Si $(G,\star)$ est un groupe fini à n éléments, on a assurément $a^n = e$.

1.4 Exercices d'apprentissage

Exercice 1

Soit H et K deux sous-groupes d'un groupe $(G,\star)$. Montrer que $H \cup K$ est un sous-groupe de $(G,\star)$ si, et seulement si, $H \subset K$ ou $K \subset H$.

Solution

On raisonne par double implication.

($\Longrightarrow$) Si H est inclus dans K, la réunion de H et K est égale à K et c'est donc un sous-groupe de $(G,\star)$. Si K est inclus H, c'est analogue.

($\Longleftarrow$) Supposons[2] $H \cup K$ sous-groupe de $(G,\star)$.

méthode

Pour montrer une disjonction « $\mathcal{P}$ ou $\mathcal{Q}$ », on suppose non($\mathcal{P}$) et l'on établit $\mathcal{Q}$.

Supposons $H \not\subset K$. Il existe un élément h dans H qui n'appartient pas à K. Montrons l'inclusion de K dans H. Soit k un élément arbitraire choisi dans K.

1. On dit que l'élément est d'*ordre infini* sinon.

2. On peut aussi raisonner par contraposée : si H n'est pas inclus dans K, ni K dans H, on observe que $H \cup K$ n'est pas stable en étudiant la composition par $\star$ d'un élément de H qui n'est pas dans K avec un élément de K qui n'est pas dans H.

méthode

Si $x \star y$ appartient à un sous-groupe H, on ne peut pas affirmer que x et y appartiennent à H. Cependant, si l'on sait de plus que x est élément de H, on peut affirmer que $y = x^{-1} \star (x \star y)$ appartient à H.

En tant que composé de deux éléments du sous-groupe $H \cup K$, l'élément $h \star k$ appartient à $H \cup K$. Cependant, il n'appartient pas à K. En effet, si par l'absurde $h \star k \in K$, on peut affirmer que h est élément du sous-groupe K en écrivant $h = (h \star k) \star k^{-1}$: ceci est contraire au choix de h. On a donc $h \star k \in H$ puis $k = h^{-1} \star (h \star k) \in H$ par opérations dans le sous-groupe H. On peut donc conclure $K \subset H$.

Exercice 2 (Groupe engendré par deux éléments)

Soit a, b deux éléments d'un groupe G noté multiplicativement et

$$H = \left\{ a^{k_1} b^{\ell_1} a^{k_2} b^{\ell_2} \dots a^{k_n} b^{\ell_n} \mid n \in \mathbb{N}, k_1, \ell_1, k_2, \ell_2, \dots, k_n, \ell_n \in \mathbb{Z} \right\}$$

l'ensemble des produits finis d'itérés de a et de b.

(a) Montrer que H est le sous-groupe engendré par $\{a, b\}$.

(b) Simplifier la description de H lorsque a et b commutent.

Solution

(a) **méthode**

On raisonne par double inclusion : une première est acquise en observant que H est un sous-groupe qui contient $\{a, b\}$, l'inclusion réciproque s'obtient en constatant que tout élément de H appartient à $\langle a, b \rangle$.

H est une partie non vide de G. H est stable par composition car le composée de deux produits finis d'itérés de a et de b définit un produit fini d'itérés de a et b. H est aussi stable par passage à l'inverse car

$$\left(a^{k_1} b^{\ell_1} \dots a^{k_n} b^{\ell_n}\right)^{-1} = a^0 b^{-\ell_n} a^{-k_n} \dots b^{-\ell_1} a^{-k_1} b^0 \in H$$

H est donc un sous-groupe de G. De plus, H contient les éléments a et b car il est possible d'écrire $a = a^1 b^0$ et $b = a^0 b^1$. On en déduit $\langle a, b \rangle \subset H$ (Th. 3 p. 4).

Inversement, les itérés de a sont éléments du sous-groupe $\langle a, b \rangle$ et il en est de même des itérés de b. Les produits d'itérés de a et b sont aussi éléments de $\langle a, b \rangle$ et donc $H \subset \langle a, b \rangle$ puis l'égalité.

(b) Si a et b commutent, on peut regrouper entre eux les itérés de a et b et l'on obtient

$$\langle a, b \rangle = \left\{ a^k b^\ell \mid k, \ell \in \mathbb{Z} \right\}.$$

Exercice 3

Soit a un élément d'un groupe $(G, \star)$.

(a) Montrer que l'application $\varphi\colon k \mapsto a^k$ définit un morphisme du groupe $(\mathbb{Z}, +)$ vers $(G, \star)$.

(b) Déterminer l'image et le noyau de φ.

Solution

(a) **méthode**

> On vérifie qu'une application est un morphisme en observant que « l'image d'un composé est la composée des images ».

Soit k et ℓ dans $\mathbb{Z}$. Par opérations sur les itérés d'un élément

$$\varphi(k+\ell) = a^{k+\ell} = a^k \star a^\ell = \varphi(k) \star \varphi(\ell).$$

(b) **méthode**

> L'image d'un morphisme s'obtient en déterminant l'ensemble des valeurs prises et le noyau en résolvant l'équation $\varphi(x) = e$.

Les valeurs prises par φ sont les a^k avec k parcourant $\mathbb{Z}$: selon que a est d'ordre fini ou non, les itérés de a peuvent comporter des répétitions ou non. Le noyau de φ s'obtient en résolvant l'équation $\varphi(k) = e$, c'est-à-dire $a^k = e$, d'inconnue $k \in \mathbb{Z}$. La résolution de cette équation nécessite aussi de discuter selon l'ordre de l'élément a.

Cas : l'élément a est d'ordre infini. L'équation $a^k = e$ a pour seule solution $k = 0$ et les a^k ne comportent aucune répétition :

$$\mathrm{Ker}(\varphi) = \{0\} \quad \text{et} \quad \mathrm{Im}(\varphi) = \left\{a^k \mid k \in \mathbb{Z}\right\} = \langle a \rangle.$$

Cas : l'élément a est d'ordre fini égal à n. On sait que $a^k = a^\ell$ si, et seulement si, n divise $k - \ell$ (Th. 13 p. 8). On en déduit

$$\mathrm{Ker}(\varphi) = n\mathbb{Z} \quad \text{et} \quad \mathrm{Im}(\varphi) = \left\{e, a, \ldots, a^{n-1}\right\} = \langle a \rangle.$$

Exercice 4

Soit a et b deux éléments d'un groupe[1] $(G, .)$.

(a) Montrer que a et bab^{-1} ont le même ordre.

(b) On suppose ab d'ordre fini égal à n. Que dire de ba ?

(c) On suppose a d'ordre fini égal à n. Pour $k \in \mathbb{Z}$, quel est l'ordre de a^k ?

1. Le groupe est ici noté multiplicativement.

Solution

(a) **méthode**

‖ On calcule l'ordre d'un élément en déterminant ses itérés égaux au neutre.

Soit $p \in \mathbb{N}$. On remarque

$$(bab^{-1})^p = \underbrace{(bab^{-1})(bab^{-1})\dots(bab^{-1})}_{p \text{ facteurs}}$$
$$= ba(b^{-1}b)a\dots(b^{-1}b)ab^{-1} = ba^pb^{-1}$$

et donc

$$\begin{aligned}(bab^{-1})^p = 1 &\iff ba^pb^{-1} = 1 \\ &\iff a^pb^{-1} = b^{-1} && \text{en multipliant par } b^{-1} \text{ à gauche} \\ &\iff a^p = 1 && \text{en multipliant par } b \text{ à droite.}\end{aligned}$$

On en déduit que les éléments a et bab^{-1} ont le même ordre (fini ou infini).

(b) En multipliant l'égalité $(ab)^n = 1$ à gauche par b, on obtient $b(ab)^n = b$ soit encore $(ba)^nb = b$. En multipliant à droite par b^{-1}, on obtient $(ba)^n = 1$. Ainsi, ba est d'ordre fini au plus égal à n. Un raisonnement symétrique établit que ab est d'ordre inférieur à celui de ba et donc ab et ba ont le même ordre.

(c) Soit $p \in \mathbb{N}$. Puisque l'élément a est d'ordre n, $(a^k)^p = a^{kp} = 1$ si, et seulement si, l'entier n divise kp.

méthode

‖ On introduit le PGCD de k et de n.

On pose $d = k \wedge n$ et l'on factorise ce PGCD des entiers k et n pour écrire $k = dk'$ et $n = dn'$ avec k' et n' premiers entre eux. L'entier n divise kp si, et seulement si, n' divise $k'p$. Par le théorème de Gauss [1], cette condition s'exprime encore n' divise p. On en déduit que a^k est d'ordre $n' = n/(n \wedge k)$.

Exercice 5

Soit $n \in \mathbb{N}^*$ et $\omega = \mathrm{e}^{2i\pi/n}$.

(a) Montrer que l'application $\varphi\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{U}_n$ déterminée par $\varphi(\bar{k}) = \omega^k$ est un isomorphisme du groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ vers $(\mathbb{U}_n, \times)$.

(b) Quelles sont les générateurs du groupe $\mathbb{U}_n$?

1. Pour a, b et c entiers, si a divise bc et si a et premier avec b alors a divise c.

Solution

(a) **méthode**

> On commence par vérifier que l'application φ est bien définie en observant que la valeur $\varphi(\bar{k})$ ne dépend pas du représentant k choisi pour la classe $\bar{k}$.

Si k' est un autre représentant de la classe $\bar{k}$, on a $k' \equiv k \ [n]$ ce qui permet d'écrire l'égalité $k' = k + pn$ avec $p \in \mathbb{Z}$. On vérifie alors

$$\omega^{k'} = \omega^{k+pn} = \omega^k \left(\omega_{=1}^n\right)^p = \omega^k \quad \text{car} \quad \omega^n = 1.$$

Ainsi, la valeur $\varphi(\bar{k}) = \omega^k$ ne dépend pas du représentant choisi pour figurer $\bar{k}$.

On vérifie ensuite que φ est un morphisme de groupes car, pour k et ℓ dans $\mathbb{Z}$,

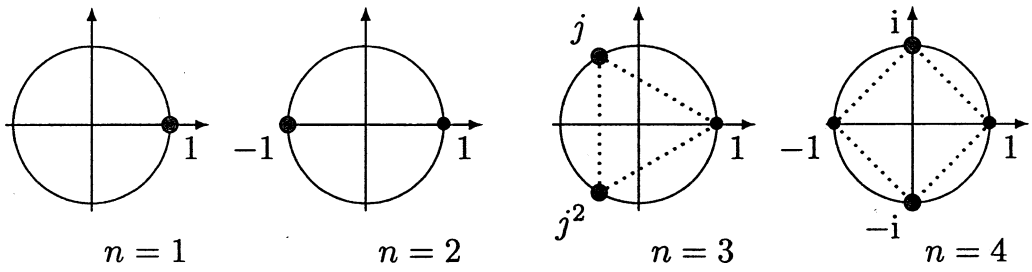
$$\varphi\left(\bar{k} + \bar{\ell}\right) = \varphi\left(\overline{k+\ell}\right) = \omega^{k+\ell} = \omega^k \omega^\ell = \varphi(k)\varphi(\ell).$$

L'application φ est aussi surjective car les racines n-ièmes de l'unité sont les ω^k avec k parcourant $[\![0\,;n-1]\!]$ et ce sont donc des valeurs prises par φ. Enfin, l'application φ est bijective[1] car surjection entre deux ensembles finis de mêmes cardinaux : c'est un isomorphisme.

(b) **méthode**

> Un isomorphisme échange les générateurs des groupes entre lesquels il opère.

Les générateurs de $\mathbb{Z}/n\mathbb{Z}$ sont les $\overline{m}$ avec m entier premier avec n (Th. 11 p. 7), les générateurs de $\mathbb{U}_n$ sont donc les $\omega^m = \mathrm{e}^{2\mathrm{i}m\pi/n}$ avec m entier premier avec n : on les appelle les *racines primitive n-ièmes de l'unité.*



$n = 1$ $n = 2$ $n = 3$ $n = 4$

Les racines primitives de l'unité pour les premières valeurs de n.

Exercice 6

Montrer que tout groupe fini de cardinal un nombre premier p est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

1. L'injectivité peut aussi être acquise directement : si $\omega^k = \omega^\ell$, il suffit considérer des arguments de ces deux nombres complexes pour obtenir $k \equiv \ell \ [n]$.

Solution

Soit $(G, \star)$ un groupe fini de cardinal p avec p nombre premier.

méthode

‖ On étudie l'ordre des éléments de G.

Puisqu'un nombre premier est au moins égal à 2, il existe dans G un élément a différent du neutre. Celui-ci est d'ordre fini divisant le cardinal p de G (Th. 14 p. 8). L'élément a n'étant pas le neutre, son ordre n'est pas 1 et vaut donc p. Le groupe engendré par a est un sous-groupe de G à p éléments, il est donc égal à G. On peut conclure que G est un groupe cyclique à p éléments et donc isomorphe à $(\mathbb{Z}/p\mathbb{Z}, +)$ (Th. 12 p. 7).

1.5 Exercices d'entraînement

1.5.1 Groupes finis

Exercice 7 *

Décrire les sous-groupes finis de $(\mathbb{C}^*, \times)$.

Solution

Soit H un sous-groupe fini de $\mathbb{C}^*$ et n son nombre d'éléments.

méthode

‖ On montre par cardinalité que $H = \mathbb{U}_n$ en observant une inclusion.

Soit z un élément de H. Celui-ci est d'ordre fini divisant le cardinal de H (Th. 14 p. 8). On a donc $z^n = 1$ ce qui signifie que z est une racine n-ième de l'unité. On a ainsi l'inclusion $H \subset \mathbb{U}_n$ puis l'égalité car ces deux ensembles sont de mêmes cardinaux finis.

Ainsi, les sous-groupes finis de $\mathbb{C}^*$ sont les groupes $\mathbb{U}_n$ des racines de l'unité, $n \in \mathbb{N}^*$.

Exercice 8 **

Soit A une partie d'un groupe fini $(G, .)$. On suppose $2\,\mathrm{Card}(A) > \mathrm{Card}(G)$. Montrer que tout élément de G peut s'écrire comme le produit de deux éléments de A.

Solution

méthode

‖ Pour $g \in G$, l'application $x \mapsto gx^{-1}$ est injective sur A.

Soit $g \in G$. L'application $x \mapsto gx^{-1}$ est une permutation[1] de G, sa restriction au départ de A réalise une injection de A dans G. L'image de A par cette application est une partie à $\mathrm{Card}(A)$ éléments incluse dans G : celle-ci ne peut pas être disjointe de A car $2\,\mathrm{Card}(A) > \mathrm{Card}(G)$. Il existe donc un élément a dans A tel que ga^{-1} appartient à A. Ceci permet d'écrire $g = \left(ga^{-1}\right)a$ produit de deux éléments de A.

1. C'est la composée des permutations $x \mapsto x^{-1}$ et $y \mapsto gy$.

Exercice 9 ** (Théorème de Lagrange)

Soit $(G,.)$ un groupe fini et H un sous-groupe de G.
Pour tout $a \in H$, on note $aH = \{ah \mid h \in H\}$.

(a) Montrer que les ensembles aH sont disjoints ou confondus.

(b) En déduire que le cardinal de H divise celui de G.

(c) Application : Décrire $H \cap K$ lorsque H et K sont deux sous-groupes de G de cardinaux premiers entre eux.

Solution

(a) Soit a et b dans G. Si aH et bH ne sont pas disjoints, il existe x élément commun à aH et bH : celui-ci s'écrit à la fois ah et bh' avec $h, h' \in H$. On montre alors l'égalité de aH et bH par double inclusion.

Soit $y = ak \in aH$ avec $k \in H$. On peut écrire

$$y = \underbrace{xh^{-1}}_{=a} k = \underbrace{bh'}_{=x} h^{-1}k = b\underbrace{h'h^{-1}k}_{=k'} = bk' \in bH.$$

Ainsi, $aH \subset bH$. Par un raisonnement symétrique, on obtient l'autre inclusion et donc l'égalité.

(b) **méthode**

‖ On partitionne G à l'aide des parties aH qui ont toutes le cardinal de H.

L'élément $a = a1$ appartenant à aH, on peut affirmer que G est la réunion des aH :

$$G = \bigcup_{a \in G} aH.$$

Dans cette union, les parties aH ne sont pas forcément distinctes. En regroupant entre elles celles qui sont identiques, on écrit

$$G = \bigcup_{i=1}^{p} a_i H$$

avec $a_1, \ldots, a_p$ des éléments de G choisis de sorte que les $a_i H$ soient distincts et que l'on y retrouve toutes les parties aH possibles. Par la question précédente, on sait que les parties $a_i H$ sont alors deux à deux disjointes [1] et donc

$$\mathrm{Card}(G) = \sum_{i=1}^{p} \mathrm{Card}(a_i H).$$

1. En fait, les parties aH sont les classes d'équivalence de la relation définie par $x \mathcal{R} y \iff xy^{-1} \in H$: celles-ci réalisent une partition de G.

Enfin, pour chaque $a \in G$, l'application $x \mapsto ax$ définit une bijection[1] de G vers G et celle-ci transforme H en aH. On en déduit $\text{Card}(aH) = \text{Card}(H)$ puis

$$\text{Card}(G) = \sum_{i=1}^{p} \text{Card}(H) = p\,\text{Card}(H).$$

(c) $H \cap K$ est un sous-groupe de H et un sous-groupe de K, son cardinal divise donc à la fois celui de H et celui de K. Ces derniers étant premiers entre eux, $H \cap K$ est un singleton, nécessairement constitué du neutre.

1.5.2 Groupe engendré

Exercice 10 *

Dans le groupe $(\mathcal{S}_E, \circ)$ des permutations de $E = \mathbb{R}\backslash\{0,1\}$, déterminer le sous-groupe engendré par les fonctions f et g définies par

$$f(x) = 1 - x \quad \text{et} \quad g(x) = 1/x.$$

Solution

Les fonctions f et g sont des applications définies sur E et à valeurs dans E. On vérifie $f^2 = \text{Id}_E$ et $g^2 = \text{Id}_E$, elles sont donc bijectives d'applications réciproques égales à elles-mêmes. Les fonctions f et g sont bien des permutations de E.

méthode

On étudie les composées possibles de f et g et leurs inverses jusqu'à ce qu'il n'apparaisse plus d'éléments nouveaux.

On calcule $h = f \circ g$, $j = g \circ f$ et $k = f \circ g \circ f$:

$$h(x) = \frac{x-1}{x}, \quad j(x) = \frac{1}{1-x} \quad \text{et} \quad k(x) = \frac{x}{x-1} \quad \text{pour tout } x \in E.$$

Les autres compositions possibles ne font apparaître aucune nouvelle fonction notamment car $g \circ f \circ g = k$. On peut alors dresser une table de composition (en notant i le neutre Id_E) :

$\circ$	i	f	g	h	j	k
i	i	f	g	h	j	k
f	f	i	h	g	k	j
g	g	j	i	k	f	h
h	h	k	f	j	i	g
j	j	g	k	i	h	f
k	k	h	j	f	g	i

Par cette table, on peut affirmer que $H = \{i, f, g, h, j, k\}$ est un sous-groupe de $(\mathcal{S}_E, \circ)$ et c'est le plus petit qui contient f et g : c'est le sous-groupe engendré par $\{f, g\}$.

1. On vérifie que $x = a^{-1}y$ est l'unique solution de l'équation $y = ax$ ce qui permet d'affirmer la bijectivité de l'application $x \mapsto ax$.

Exercice 11 *

Soit H et K deux sous-groupes d'un groupe abélien $(G,+)$. Montrer que le groupe engendré par $H \cup K$ est

$$H+K=\{h+k \mid h \in H \text{ et } k \in K\}.$$

Solution

méthode

On vérifie que $H+K$ est un sous-groupe de G et que c'est le plus petit contenant à la fois H et K.

Les parties H et K étant toutes deux non vides, $H+K$ est une partie non vide de G. Soit x et y deux éléments de $H+K$. On peut écrire $x=h_1+k_1$ et $y=h_2+k_2$ avec $h_1,h_2 \in H$ et $k_1,k_2 \in K$. On a alors par commutativité

$$x-y=(h_1+k_1)-(h_2+k_2)=\underbrace{(h_1-h_2)}_{\in H}+\underbrace{(k_1-k_2)}_{\in K} \in H+K.$$

Ainsi, $H+K$ est un sous-groupe de G. De plus, celui-ci contient $H \cup K$. En effet, pour tout $x \in H$, on peut écrire $x=x+0 \in H+K$ car 0 est élément du sous-groupe K. On vérifie de même que K est inclus dans $H+K$. On en déduit l'inclusion (Th. 3 p. 4)

$$\langle H \cup K \rangle \subset H+K.$$

L'inclusion réciproque est quant à elle évidente car tout élément de $H+K$ est élément de $\langle H \cup K \rangle$ par somme d'éléments de ce sous-groupe. On a donc l'égalité $\langle H \cup K \rangle = H+K$.

Exercice 12 **

Soit H un sous-groupe d'un groupe $(G,\star)$ distinct de G. Déterminer le groupe engendré par le complémentaire $\overline{H}$ de H dans G.

Solution

méthode

La composition d'un élément de H par un élément de $\overline{H}$ détermine un élément de $\overline{H}$.

Montrons $\langle \overline{H} \rangle = G$. On a évidemment $\langle \overline{H} \rangle \subset G$ et il s'agit d'établir l'inclusion réciproque. Soit g un élément de G.

Si g n'appartient pas à H, c'est un élément $\overline{H}$ donc de $\langle \overline{H} \rangle$.

Supposons maintenant que g appartient à H. Puisque H est une partie distincte de G, son complémentaire $\overline{H}$ est non vide ce qui permet d'introduire un élément $a \in \overline{H}$. Le composé $a \star g$ est alors élément de $\overline{H}$. En effet, si par l'absurde, $a \star g$ appartient à H alors $a=(a \star g) \star g^{-1}$ est aussi élément de H par opérations dans le sous-groupe H : ceci contredit la définition de a. Sachant $a \star g \in \overline{H}$, on peut alors écrire $g=a^{-1} \star (a \star g) \in \langle \overline{H} \rangle$ par opérations dans le sous-groupe $\langle \overline{H} \rangle$.

On peut conclure $\langle \overline{H} \rangle = G$.

Exercice 13 **

Soit $n \in \mathbb{N}$ tel que $n \geqslant 3$. Dans le groupe $(\mathcal{S}_n, \circ)$ des permutations de $[\![1\,;n]\!]$, on considère la transposition $t = \begin{pmatrix} 1 & 2 \end{pmatrix}$ et le cycle $c = \begin{pmatrix} 1 & 2 & \dots & n \end{pmatrix}$ de longueur n.

(a) Justifier que l'ensemble $\{c, t\}$ constitue une partie génératrice de $(\mathcal{S}_n, \circ)$.

(b) Existe-t-il une partie génératrice de $(\mathcal{S}_n, \circ)$ formée d'un seul élément ?

Solution

(a) **méthode**

> On sait que toute permutation de $[\![1\,;n]\!]$ peut s'écrire comme un produit de transpositions $\tau = \begin{pmatrix} i & j \end{pmatrix}$ avec $1 \leqslant i < j \leqslant n$. Il suffit alors de savoir écrire une telle transposition à partir de compositions de c et t pour conclure.

On remarque [1]

$$c \circ t \circ c^{-1} = \begin{pmatrix} 2 & 3 \end{pmatrix}, \quad c^2 \circ t \circ c^{-2} = \begin{pmatrix} 3 & 4 \end{pmatrix}, \text{ etc.}$$

On en déduit que les transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ (avec $1 \leqslant i \leqslant n-1$) appartiennent chacune au sous-groupe engendré par c et t. Or, pour $1 \leqslant i < j \leqslant n$, on peut écrire la décomposition :

$$\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} i & i+1 \end{pmatrix} \circ \begin{pmatrix} i+1 & i+2 \end{pmatrix} \circ \cdots \circ \begin{pmatrix} j-1 & j \end{pmatrix} \circ \cdots \circ \begin{pmatrix} i+1 & i+2 \end{pmatrix} \circ \begin{pmatrix} i & i+1 \end{pmatrix}.$$

Toutes les transpositions $\begin{pmatrix} i & j \end{pmatrix}$ appartiennent donc au sous-groupe engendré par c et t et l'on peut conclure que celui-ci se confond avec $\mathcal{S}_n$.

(b) Le groupe $(\mathcal{S}_n, \circ)$ n'est pas commutatif car $n \geqslant 3$: il n'est donc pas monogène.

Exercice 14 **

Montrer qu'un groupe fini G de cardinal $n \geqslant 2$ possède une partie génératrice constituée d'au plus $\log_2 n$ éléments.

Solution

Adoptons une notation multiplicative pour la loi de G.

méthode

> On construit une partie génératrice $\{g_1, \dots, g_p\}$ en choisissant, pour $k \geqslant 2$, l'élément g_k en dehors du groupe engendré par $g_1, \dots, g_{k-1}$.

Soit g_1 un élément de G différent du neutre, g_2 choisi en dehors de $\langle g_1 \rangle$, g_3 choisi en dehors de $\langle g_1, g_2 \rangle$, etc. Le groupe G étant fini, ce processus s'arrête avec la détermination de $g_1, \dots, g_p$ vérifiant :

$$G = \langle g_1, \dots, g_p \rangle \quad \text{et} \quad g_k \notin \langle g_1, \dots, g_{k-1} \rangle \quad \text{pour tout } k \in \{2, \dots, p\}.$$

1. Plus généralement, on a $s \circ \tau \circ s^{-1} = \begin{pmatrix} s(i) & s(j) \end{pmatrix}$ pour toute permutation s de $\mathcal{S}_n$.

Considérons ensuite l'application $\Phi\colon \{0,1\}^p \to G$ définie par

$$\Phi(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n) = g_1^{\varepsilon_1} g_2^{\varepsilon_2}\dots g_p^{\varepsilon_p} \quad \text{avec} \quad g_i^{\varepsilon_i} = \begin{cases} g_i & \text{si } \varepsilon_i = 1 \\ 1 & \text{si } \varepsilon_i = 0. \end{cases}$$

Vérifions que l'application Φ est injective. Supposons $\Phi(\varepsilon_1,\dots,\varepsilon_p) = \Phi(\varepsilon'_1,\dots,\varepsilon'_p)$ avec $(\varepsilon_1,\dots,\varepsilon_p)$ et $(\varepsilon'_1,\dots,\varepsilon'_p)$ dans $\{0,1\}^p$. On a donc l'égalité

$$g_1^{\varepsilon_1} g_2^{\varepsilon_2}\dots g_p^{\varepsilon_p} = g_1^{\varepsilon'_1} g_2^{\varepsilon'_2}\dots g_p^{\varepsilon'_p}. \qquad (*)$$

Quitte à échanger, on peut supposer $\varepsilon'_p \geqslant \varepsilon_p$ et ordonner les membres pour écrire

$$g_p^{\varepsilon'_p-\varepsilon_p} = \left(g_{p-1}^{-\varepsilon'_{p-1}}\dots g_1^{-\varepsilon'_1}\right)\left(g_1^{\varepsilon_1}\dots g_{p-1}^{\varepsilon_{p-1}}\right) \in \langle g_1,\dots,g_{p-1}\rangle.$$

Puisque g_p est choisi en dehors du groupe engendré par $g_1,\dots,g_{p-1}$, on a nécessairement $\varepsilon'_p = \varepsilon_p$. On peut alors simplifier l'identité $(*)$ pour écrire

$$g_1^{\varepsilon_1} g_2^{\varepsilon_2}\dots g_{p-1}^{\varepsilon_{p-1}} = g_1^{\varepsilon'_1} g_2^{\varepsilon'_2}\dots g_{p-1}^{\varepsilon'_{p-1}}.$$

Il suffit ensuite de reprendre le processus pour pouvoir affirmer $(\varepsilon_1,\dots,\varepsilon_p) = (\varepsilon'_1,\dots,\varepsilon'_p)$.

L'application Φ est donc injective et conséquemment

$$\mathrm{Card}\left(\{0,1\}^p\right) \leqslant \mathrm{Card}(G) \quad \text{c'est-à-dire} \quad 2^p \leqslant n.$$

On en déduit $p \leqslant \log_2 n$.

1.5.3 Morphismes de groupes

Exercice 15 * (Morphisme de conjugaison)

Soit G un groupe noté multiplicativement. Pour $a \in G$, on note τ_a l'application de G vers G définie par $\tau_a(x) = axa^{-1}$.

(a) Montrer que τ_a est un morphisme du groupe G vers lui-même.

(b) Vérifier $\tau_a \circ \tau_b = \tau_{ab}$ pour tous a et b dans G.

(c) Montrer que τ_a est bijective et exprimer son application réciproque.

(d) En déduire que $\mathcal{T} = \left\{\tau_a \middle| a \in G\right\}$ muni du produit de composition est un groupe.

Solution

(a) Soit $x, y \in G$. On a par associativité [1]

$$\tau_a(x)\tau_a(y) = (ax)(aya^{-1}) = ax(a^{-1}a)ya^{-1} = axya^{-1} = \tau_a(xy).$$

L'application τ_a est donc un morphisme du groupe $(G,.)$ vers lui-même [2].

1. Le groupe G n'est pas *a priori* commutatif : il faut être attentif à l'organisation des calculs et ne pas simplifier axa^{-1} en x.

2. Comme pour les applications linéaires, on peut parler d'*endomorphisme de groupe*.

(b) On vérifie l'égalité de deux applications en constatant celle-ci en tout point. Pour tout $x \in G$,

$$(\tau_a \circ \tau_b)(x) = \tau_a\big(bxb^{-1}\big) = a\big(bxb^{-1}\big)a^{-1} = (ab)x(ab)^{-1} = \tau_{ab}(x).$$

On a donc $\tau_a \circ \tau_b = \tau_{ab}$.

(c) **méthode**

> On peut montrer que τ_a est bijective en étudiant injectivité et surjectivité, ou en résolvant l'équation $\tau_a(x) = y$ d'inconnue x. Ici, il est cependant plus à propos de proposer un candidat pour l'application réciproque de τ_a.

On a l'intuition que $\tau_{a^{-1}}$ est la bijection réciproque de τ_a. Vérifions-le[1] :

$$(\tau_a \circ \tau_{a^{-1}}) = \tau_1 = \mathrm{Id}_G \quad \text{et} \quad (\tau_{a^{-1}} \circ \tau_a) = \tau_1 = \mathrm{Id}_G.$$

On en déduit que τ_a est bijective et $(\tau_a)^{-1} = \tau_{a^{-1}}$.

(d) **méthode**

> On vérifie que $\mathcal{T}$ est un sous-groupe du groupe des permutations $(\mathcal{S}_G, \circ)$.

On a vu que les τ_a sont des permutations de G, l'ensemble $\mathcal{T}$ constitue donc une partie non vide de $\mathcal{S}_G$. Pour f et g dans $\mathcal{T}$, on peut écrire $f = \tau_a$ et $g = \tau_b$ avec $a, b \in G$. On a alors

$$f \circ g^{-1} = \tau_a \circ \big(\tau_b\big)^{-1} = \tau_a \circ \tau_{b^{-1}} = \tau_{ab^{-1}} \in \mathcal{T}.$$

Ainsi, $\mathcal{T}$ est un sous-groupe de $\big(\mathcal{S}_G, \circ\big)$, c'est donc un groupe pour le produit de composition des applications.

Exercice 16 **

Soit φ un morphisme non constant d'un groupe fini $(G, \star)$ vers $\big(\mathbb{C}^*, \times\big)$. Calculer

$$\sum_{x \in G} \varphi(x).$$

Solution

méthode

> Pour $a \in G$, l'application $x \mapsto a \star x$ est une permutation de G qui permet de réorganiser la somme à calculer.

Puisque φ n'est pas constante, il existe un élément a dans G tel que $\varphi(a) \neq 1$. L'application $x \mapsto a \star x$ étant une permutation[2], on peut réordonner les termes de la somme et écrire

$$\sum_{x \in G} \varphi(x) = \sum_{x \in G} \varphi(a \star x).$$

1. Sauf argument de cardinalité, il est indispensable de vérifier que les deux compositions donnent l'identité.

2. En effet, l'équation $a \star x = y$ d'inconnue $x \in G$ possède une unique solution $x = a^{-1} \star y$.

Or l'application φ est un morphisme et donc $\varphi(a \star x) = \varphi(a)\varphi(x)$ puis

$$\sum_{x \in G} \varphi(x) = \sum_{x \in G} \varphi(a)\varphi(x) = \varphi(a) \sum_{x \in G} \varphi(x).$$

Sachant $\varphi(a) \neq 1$, on conclut

$$\sum_{x \in G} \varphi(x) = 0.$$

Exercice 17 **

Soit f un morphisme de groupes au départ d'un groupe G fini. Montrer la formule

$$\mathrm{Card}(G) = \mathrm{Card}\big(\mathrm{Im}(f)\big) \times \mathrm{Card}\big(\mathrm{Ker}(f)\big).$$

Solution

Notons G' le groupe d'arrivée du morphisme f et adoptons une notation multiplicative pour les deux groupes G et G'.

méthode

‖ On dénombre l'ensemble des antécédents de chaque valeur prise par f.

Énumérons les valeurs prises par f : $y_1, \ldots, y_p$ avec $p = \mathrm{Card}\big(\mathrm{Im}(f)\big)$. Soit $j \in [\![1\,;p]\!]$ et x_j dans G un antécédent de y_j par f. Pour $x \in G$

$$\begin{aligned} f(x) = y_j &\iff f(x) = f(x_j) \\ &\iff f(x_j)^{-1} f(x) = 1_{G'} \\ &\iff f\big(x_j^{-1}x\big) = 1_{G'} \qquad \text{car } f \text{ est un morphisme} \\ &\iff x_j^{-1}x \in \mathrm{Ker}(f). \end{aligned}$$

L'ensemble des antécédents de y_j est donc

$$f^{-1}\big(\{y_j\}\big) = \{x_j h \mid h \in \mathrm{Ker}(f)\}.$$

L'application $h \mapsto x_j h$ étant injective, le cardinal de $f^{-1}\big(\{y_j\}\big)$ vaut celui de $\mathrm{Ker}(f)$. Enfin, G est la réunion des $f^{-1}\big(\{y_j\}\big)$ et ces parties sont deux à deux disjointes donc

$$\mathrm{Card}(G) = \sum_{j=1}^{p} \underbrace{\mathrm{Card}\Big(f^{-1}\big(\{y_j\}\big)\Big)}_{=\mathrm{Card}(\mathrm{Ker}(f))} = p\,\mathrm{Card}\big(\mathrm{Ker}(f)\big) = \mathrm{Card}\big(\mathrm{Im}(f)\big) \times \mathrm{Card}\big(\mathrm{Ker}(g)\big).$$

Exercice 18 **

Parmi les groupes suivants, lesquels sont isomorphes ?

(a) $(\mathbb{Z}/4\mathbb{Z}, +)$	(b) $(\mathbb{U}_4, \times)$	(c) $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +)$
(d) $(\mathbb{Z}/6\mathbb{Z}, +)$	(e) $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, +)$	(f) $(\mathcal{S}_3, \circ)$
(g) $(\mathbb{Q}, +)$	(h) $(\mathbb{Q}^*, \times)$	(i) $(\mathbb{Q}_+^*, \times)$.

Solution

méthode

Deux groupes isomorphes sont en bijection et ont donc le même cardinal.

Par cet argument de cardinalité, seuls les groupes (a), (b), (c) dans un premier temps, les groupes (d), (e), (f) dans un second temps et les groupes (g), (h), (i) dans un dernier temps sont susceptibles d'être isomorphes.

Cas : Groupes à 4 éléments. On étudie (a), (b) et (c).

Le groupe $\mathbb{U}_4$ est cyclique donc isomorphe à $\mathbb{Z}/4\mathbb{Z}$. En revanche, le groupe

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} = \{(\bar{0},\bar{0}),(\bar{0},\bar{1}),(\bar{1},\bar{0}),(\bar{1},\bar{1})\}$$

n'est pas cyclique car tous ses éléments x vérifient $x + x = (\bar{0},\bar{0})$: il n'est pas isomorphe aux précédents.

Cas : Groupes à 6 éléments. On étudie (d), (e) et (f).

Le groupe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ est constitué des éléments

$$(\bar{0},\hat{0}),(\bar{0},\hat{1}),(\bar{0},\hat{2}),(\bar{1},\hat{0}),(\bar{1},\hat{1}) \text{ et } (\bar{1},\hat{2})$$

en notant $\bar{k}$ les classes d'équivalence dans $\mathbb{Z}/2\mathbb{Z}$ et $\hat{k}$ celles dans $\mathbb{Z}/3\mathbb{Z}$. Ce groupe est cyclique engendré par $(\bar{1},\hat{1})$: il est isomorphe à $\mathbb{Z}/6\mathbb{Z}$.

Le groupe S_3 n'est pas commutatif, a *fortiori* non cyclique, il n'est pas isomorphe aux précédents [1].

Cas : Groupes infinis. On étudie (g), (h) et (i).

méthode

On montre que ces trois groupes ne sont pas isomorphes en observant que des équations analogues ne proposent pas les mêmes descriptions d'ensembles solutions.

Dans $(\mathbb{Q},+)$, l'équation $x + x = 0$ possède une seule solution. L'équation analogue dans $(\mathbb{Q}^*,\times)$ s'écrit $x \times x = 1$, elle possède deux solutions. Les groupes $\mathbb{Q}$ et $\mathbb{Q}^*$ ne sont pas isomorphes. Cet argument peut être repris pour affirmer que $(\mathbb{Q}^*,\times)$ et $(\mathbb{Q}_+^*,\times)$ ne sont pas non plus isomorphes. Reste à étudier si $(\mathbb{Q},+)$ et $(\mathbb{Q}_+^*,\times)$ peuvent être isomorphes [2].

Pour y élément de $(\mathbb{Q},+)$, l'équation $x + x = y$ possède assurément une solution dans $\mathbb{Q}$. Pour y élément de $(\mathbb{Q}_+^*,\times)$, l'équation analogue $x \times x = y$ peut ne pas posséder de solutions dans $\mathbb{Q}_+^*$, c'est le cas par exemple quand $y = 2$ où l'équation n'a pas de solution car $\sqrt{2}$ est irrationnel. On en déduit que $\mathbb{Q}$ et $\mathbb{Q}_+^*$ ne sont pas isomorphes.

1. Le sujet 28 p. 29 précisera quels sont les groupes à six éléments possibles.

2. $(\mathbb{R},+)$ et $(\mathbb{R}_+^*,\times)$ sont isomorphes via la fonction exponentielle mais celle-ci n'induit pas d'isomorphisme sur les nombres rationnels.

1.5.4 Éléments d'ordres finis

Exercice 19 *

Soit $(G,\star)$ un groupe de cardinal n et k un entier premier avec n. Montrer que l'application $\varphi\colon x \mapsto x^k$ est une permutation de G.

Solution

méthode

Par le théorème de Bézout, on détermine $u \in \mathbb{Z}$ tel que $x^{ku} = x$ pour tout x de G.

Les entiers n et k étant premiers entre eux, il existe u et v entiers tels que $ku+nv = 1$. Le groupe G étant de cardinal n, on a $x^n = e$ pour tout x dans G (Th. 14 p. 8). On vérifie alors

$$x^{ku} = x \star x^{-nv} = x \star \left(x^n\right)^{-v} = x \star e^{-v} = x.$$

Considérons ensuite l'application $\psi\colon x \mapsto x^u$ définie sur G. Par le calcul qui précède, on peut affirmer $\varphi \circ \psi = \psi \circ \varphi = \mathrm{Id}_G$: l'application φ est bijective et ψ est sa bijection réciproque.

Exercice 20 **

Soit a un élément d'un groupe G noté multiplicativement. On suppose a d'ordre pq avec p et q dans $\mathbb{N}^*$ premiers entre eux. Déterminer des éléments b et c d'ordres respectifs p et q tels que $a = bc = cb$.

Solution

méthode

Par le théorème de Bézout, on introduit $u, v \in \mathbb{Z}$ tels que $pu + qv = 1$ puis on détermine par une analyse b et c comme des puissances de a.

Analyse : Si a est le produit bc avec b et c commutant et d'ordres respectifs p et q, on a $a^p = b^p c^p = c^p$ puis $c = c^{pu+qv} = \left(c^p\right)^u = a^{pu}$. De même, on détermine $b = a^{qv}$.

Synthèse : Posons $b = a^{qv}$ et $c = a^{pu}$. On a immédiatement $a = bc$ avec b et c qui commutent puisque ce sont des itérés de a. Il reste à déterminer leurs ordres [1].

Soit $n \in \mathbb{N}^*$. On a $b^n = 1$ si, et seulement si, $a^{nqv} = 1$. L'élément a étant d'ordre pq, cette égalité a lieu si, et seulement si, pq divise nqv (Th. 13 p. 8) ce qui revient à dire que p divise nv. Or p et v sont premiers entre eux en vertu de l'égalité $pu + qv = 1$. Par le lemme de Gauss, on parvient à la condition : p divise n.

Finalement, b est d'ordre p et l'on montre de la même manière que c est d'ordre q.

1. On reproduit ici en contexte la démonstration déjà menée dans le sujet 4 p. 10.

Exercice 21 ***

Soit a et b deux éléments d'ordres respectifs m et n d'un groupe abélien $(G,.)$.

(a) On suppose que m et n sont premiers entre eux. Montrer que ab est d'ordre mn.

(b) On ne suppose plus m et n premiers entre eux, l'élément ab est-il nécessairement d'ordre $m \vee n$?

(c) Soit d un diviseur de m. Montrer qu'il existe un élément d'ordre d dans G.

(d) Existe-t-il dans G un élément d'ordre $m \vee n$?

Solution

(a) Par commutativité du groupe, on peut organiser le calcul des itérés

$$(ab)^{mn} = \underbrace{(ab)(ab)\dots(ab)}_{mn \text{ facteurs}} = a^{mn}b^{mn} = \left(a^m\right)^n\left(b^n\right)^m = 1^n 1^m = 1.$$

Inversement, soit $r \in \mathbb{N}^*$ tel que $(ab)^r = 1$. On a alors

$$a^{nr} = \left(a^r\right)^n = \left(b^{-r}\right)^n = \left(b^n\right)^{-r} = 1^{-r} = 1.$$

L'élément a étant d'ordre m, celui-ci divise nr. Or m et n sont premiers entre eux et donc m divise r par application du lemme de Gauss. *Mutatis mutandis*[1], n divise r et donc mn divise r car m et n sont premiers entre eux.

Finalement, ab est un élément d'ordre mn exactement.

(b) Si $b = a^{-1}$ alors a et b ont le même ordre m et $ab = 1$ est d'ordre 1 : il est possible que le produit ab ne soit pas d'ordre $m \vee n$.

(c) On écrit $m = dm'$ et l'on introduit $x = a^{m'} \in G$. On a

$$\begin{aligned} x^k = 1 &\iff a^{km'} = 1 \\ &\iff m \mid km' && \text{car } a \text{ est d'ordre } m \\ &\iff d \mid k && \text{en simplifiant par } m' \neq 0. \end{aligned}$$

L'élément x est donc d'ordre d.

(d) **méthode**

> Pour chaque facteur premier p de m ou n, on détermine un élément de G d'ordre $\max\left(v_p(n), v_p(m)\right)$ où v_p désigne la valuation p-adique.

À partir des décompositions en facteurs premiers de m et n, on peut écrire en autorisant les exposants à être nuls

$$m = p_1^{\alpha_1} \dots p_r^{\alpha_r} \quad \text{et} \quad n = p_1^{\beta_1} \dots p_r^{\beta_r} \quad \text{avec} \quad \alpha_i = v_{p_i}(m), \beta_i = v_{q_i}(n).$$

1. Locution latine signifiant « En modifiant ce qui doit être changé ».

On sait qu'alors

$$m \vee n = p_1^{\max(\alpha_1,\beta_1)} \dots p_r^{\max(\alpha_r,\beta_r)}.$$

Par la question précédente, il est possible de déterminer un élément x_i d'ordre $p_i^{\max(\alpha_i,\beta_i)}$ car cet ordre divise m ou n. Puisque les $x_1, \dots, x_r$ sont d'ordres deux à deux premiers entre eux, l'élément $x = x_1 \dots x_r$ est d'ordre $m \vee n$ par applications répétées du résultat de la première question.

1.5.5 Groupes cycliques

Exercice 22 **

On désire établir que tout sous-groupe d'un groupe cyclique est lui-même cyclique. Soit $(G, .)$ un groupe cyclique de générateur a et H un sous-groupe de G.

(a) Justifier l'existence d'un plus petit entier naturel non nul n tel que $a^n \in H$.

(b) Établir que H est alors le groupe engendré par a^n.

Solution

(a) **méthode**

> On montre l'existence d'un plus petit entier vérifiant une propriété en observant que l'ensemble des entiers concernés est une partie non vide de $\mathbb{N}$.

Notons A l'ensemble des $n \in \mathbb{N}^*$ vérifiant $a^n \in H$. Cet ensemble est une partie non vide de $\mathbb{N}$ car $a^{\mathrm{Card}(G)} = 1 \in H$ (Th. 14 p. 8). Il existe donc un plus petit $n \in \mathbb{N}^*$ tel que a^n est élément de H.

(b) **méthode**

> On reproduit avec les notations en cours la preuve du théorème caractérisant les sous-groupes de $(\mathbb{Z}, +)$.

Posons $b = a^n$. Puisque b appartient au sous-groupe H, on sait déjà $\langle b \rangle \subset H$. Considérons ensuite $x \in H$. Il existe $p \in \mathbb{Z}$ tel que $x = a^p$ car tous les éléments de G sont des itérés de a. Réalisons alors la division euclidienne de p par n : $p = nq + r$ avec $0 \leqslant r < n$. On a

$$a^r = a^{p-nq} = a^p \left(a^n\right)^{-q} = xb^{-q} \in H.$$

Or n désigne l'exposant de la plus petite puissance strictement positive de a appartenant à H et r est strictement inférieur à n donc $r = 0$. Par suite $x = a^{nq} = b^q$ et donc $x \in \langle b \rangle$.

Finalement, $H = \langle b \rangle$ et le sous-groupe H est cyclique.

Exercice 23 **

Soit G un groupe cyclique de cardinal n. Montrer que, pour tout diviseur[1] $d \in \mathbb{N}^*$ de n, il existe un unique sous-groupe de cardinal d dans G.

1. On a vu que le cardinal d'un sous-groupe divise le cardinal du groupe (voir sujet 9 p. 14) : l'hypothèse que d divise n est nécessaire à l'existence d'un sous-groupe de cardinal d.

Solution

méthode

Par isomorphisme, on peut supposer $G = \mathbb{Z}/n\mathbb{Z}$ ce qui rend l'étude plus concrète [1].

Soit d un diviseur de n et d' l'entier tel que $dd' = n$. Le sous-groupe engendré par $\overline{d'}$ dans $\mathbb{Z}/n\mathbb{Z}$ est

$$H = \langle \overline{d'} \rangle = \left\{ \overline{0}, \overline{d'}, 2\overline{d'}, \ldots, (d-1)\overline{d'} \right\}$$

car $\overline{d'}$ est un élément d'ordre [2] d dans $\mathbb{Z}/n\mathbb{Z}$. Ainsi, il existe au moins un sous-groupe à d éléments dans $\mathbb{Z}/n\mathbb{Z}$.

Inversement, considérons un sous-groupe H à d éléments. Pour tout $\overline{x}$ de H, l'ordre de $\overline{x}$ divise le cardinal de H et l'on a donc $d\overline{x} = \overline{0}$. On en déduit que n divise dx puis que d' divise x. Ainsi,

$$\overline{x} \in \left\{ \overline{0}, \overline{d'}, 2\overline{d'}, \ldots, (d-1)\overline{d'} \right\}.$$

On a alors l'inclusion $H \subset \left\{ \overline{0}, \overline{d'}, 2\overline{d'}, \ldots, (d-1)\overline{d'} \right\}$ puis l'égalité car les deux ensembles ont le même nombre fini d'éléments.

Exercice 24 **

Soit H et K deux sous-groupes d'un groupe abélien $(G, .)$ de cardinaux p et q nombres premiers distincts. Montrer que $HK = \left\{ hk \;\middle|\; h \in H \text{ et } k \in K \right\}$ est un sous-groupe cyclique de G.

Solution

On vérifie que HK est un sous-groupe de G : HK est une partie non vide de G et, si $x = hk$ et $y = h'k'$ sont deux éléments de HK (avec $h, h' \in H$ et $k, k' \in K$), on a par commutativité du groupe G

$$xy^{-1} = hk(h'k')^{-1} = hk\left(k'^{-1}h'^{-1}\right) = \left(hh'^{-1}\right)\left(kk'^{-1}\right) \in HK.$$

Au surplus, on peut affirmer que HK possède au plus pq éléments [3].

méthode

On construit un élément d'ordre pq dans HK par produit d'un générateur de H et d'un générateur de K.

Les sous-groupes H et K sont cycliques car leurs cardinaux sont premiers [4] : on peut introduire a et b générateurs respectifs de ces groupes, a est d'ordre p et b d'ordre q. Considérons ensuite $x = ab \in HK$ et déterminons son ordre.

1. Ce n'est cependant pas une nécessité. En introduisant x un générateur de G, on vérifie par une étude analogue à celle qui suit que $H = \left\{ e, x^{d'}, \ldots, x^{(d-1)d'} \right\}$ est l'unique sous-groupe de G de cardinal d.

2. On vérifie aisément que $d\overline{d'}$ est le premier itéré additif de $\overline{d'}$ à valoir $\overline{0}$. On peut aussi utiliser le résultat du sujet 4 p. 10.

3. L'application de $H \times K$ vers HK qui envoie (h, k) sur hk est surjective : il y a donc moins d'éléments dans HK que dans $H \times K$. Cette application est même un isomorphisme de groupes notamment car son noyau est réduit au neutre (voir sujet 9 p. 14).

4. Voir sujet 6 p. 12.

Dans un premier temps, on vérifie par commutativité $x^{pq} = (a^p)^q (b^q)^p = 1$ et donc l'ordre de x est inférieur à pq. Soit $n \in \mathbb{N}^*$. Si $x^n = 1$ alors $a^n b^n = 1$. En élevant à la puissance q, il vient $a^{nq} b^{nq} = 1$ avec $b^{nq} = 1$ car b est d'ordre q. On en déduit que $a^{nq} = 1$ et p divise donc nq. Or p et q sont premiers entre eux car il s'agit de nombres premiers distincts, on a donc p diviseur de n. *Mutatis mutandis*, q divise aussi n et donc pq divise n car p et q sont premiers entre eux.

Finalement, x est d'ordre exactement [1] pq et l'on peut conclure que HK est cyclique.

Exercice 25 **

Soit G et G' deux groupes notés multiplicativement. On suppose que le groupe G est cyclique engendré par un élément a.

(a) Soit b un élément de G'. Montrer qu'il existe un morphisme $\varphi\colon G \to G'$ vérifiant $\varphi(a) = b$ si, et seulement si, b est d'ordre fini divisant celui de a.

(b) Combien existe-t-il de morphismes de $(\mathbb{Z}/n\mathbb{Z}, +)$ dans lui-même ?

(c) Combien existe-t-il de morphismes de $(\mathbb{Z}/n\mathbb{Z}, +)$ dans $(\mathbb{C}^*, \times)$?

Solution

On introduit n le cardinal de G ce qui est aussi l'ordre du générateur a.

(a) Supposons qu'il existe un morphisme $\varphi\colon G \to G'$. Ce morphisme transforme l'égalité $a^n = 1_G$ en $\varphi(a)^n = 1_{G'}$ et donc $b = \varphi(a)$ est d'ordre fini divisant n.

Inversement, supposons que b soit un élément de G' d'ordre fini divisant n.

méthode

Une petite analyse amène à définir φ de sorte que $\varphi(a^k) = b^k$ pour tout exposant k de $[\![0\,;n-1]\!]$.

Les éléments de G sont les a^k pour $k \in [\![0\,;n-1]\!]$ et ces derniers sont deux à deux distincts, on peut donc construire une application $\varphi\colon G \to G'$ en posant

$$\varphi(a^k) = b^k \quad \text{pour tout } k \in [\![0\,;n-1]\!].$$

L'application φ vérifie évidemment $\varphi(a) = b$. Montrons qu'il s'agit d'un morphisme.

Pour $x, y \in G$, on peut écrire $x = a^k$, $y = a^\ell$ avec $k, \ell \in [\![0\,;n-1]\!]$. On a alors $xy = a^m$ avec $m \equiv k + \ell \ [n]$ et $m \in [\![0\,;n-1]\!]$. Dès lors $\varphi(x)\varphi(y) = b^{k+\ell}$ et $\varphi(xy) = b^m$. Or l'ordre de b divise n et divise donc aussi $m - (k+\ell)$. On en déduit $b^{k+\ell} = b^m$ puis $\varphi(x)\varphi(y) = \varphi(xy)$.

Finalement, φ détermine un morphisme de G vers G' transformant a en b.

(b) $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe cyclique de cardinal n engendré par $a = \bar{1}$. Tous les éléments de $\mathbb{Z}/n\mathbb{Z}$ étant d'ordre divisant n, à chaque valeur b dans $\mathbb{Z}/n\mathbb{Z}$ correspond un morphisme de $\mathbb{Z}/n\mathbb{Z}$ dans lui-même transformant a en b. Il y a donc exactement n morphismes possibles.

1. Voir aussi sujet 21 p. 23.

(c) Si φ est un morphisme de $(\mathbb{Z}/n\mathbb{Z}, +)$ dans $(\mathbb{C}^*, \times)$, la valeur $z = \varphi(\bar{1})$ doit être d'ordre divisant n et donc vérifier $z^n = 1$: c'est une racine n-ième de l'unité. Inversement, chaque racine n-ième est d'ordre divisant n et il lui correspond un morphisme envoyant $\bar{1}$ sur elle-même. Il y a donc autant de morphismes possibles que de racines n-ièmes de l'unité, c'est-à-dire n.

1.6 Exercices d'approfondissement

Exercice 26 * (Groupe p-quasi-cyclique de Prüfer)
Soit p un nombre premier. On pose

$$\mathbb{U}_{p^\infty} = \left\{ z \in \mathbb{C} \,\middle|\, \exists k \in \mathbb{N},\ z^{p^k} = 1 \right\}.$$

(a) Montrer que $\mathbb{U}_{p^\infty}$ est un groupe multiplicatif dont tous les éléments sont d'ordre finis.

(b) Montrer que les sous-groupes propres[1] de $\mathbb{U}_{p^\infty}$ sont cycliques.

Solution

(a) On vérifie que $\mathbb{U}_{p^\infty}$ est un sous-groupe de $(\mathbb{C}^*, \times)$. L'ensemble $\mathbb{U}_{p^\infty}$ est la réunion des sous-groupes $\mathbb{U}_{p^k}$ des racines p^k-ièmes de l'unité. L'ensemble $\mathbb{U}_{p^\infty}$ est donc une partie non vide de $\mathbb{C}^*$ stable par passage à l'inverse. De plus, les $\mathbb{U}_{p^k}$ sont en progression croissante :

$$\mathbb{U}_{p^k} \subset \mathbb{U}_{p^{k+1}} \quad \text{car} \quad z^{p^k} = 1 \implies z^{p^{k+1}} = \left(z^{p^k}\right)^p = 1.$$

L'ensemble $\mathbb{U}_{p^\infty}$ est donc stable par le produit car la multiplication de deux éléments de $\mathbb{U}_{p^\infty}$ peut s'interpréter comme le produit de deux éléments d'un même $\mathbb{U}_{p^k}$ en choisissant k assez grand.

Finalement, $\mathbb{U}_{p^\infty}$ est un sous-groupe de $(\mathbb{C}^*, \times)$ donc un groupe multiplicatif. Par définition, ses éléments sont tous d'ordres finis et l'ordre de chacun vaut une puissance de p.

(b) Soit H un sous-groupe propre de $\mathbb{U}_{p^\infty}$.

méthode

‖ On montre que $H = \mathbb{U}_{p^k}$ en introduisant[2] le plus grand $k \in \mathbb{N}$ tel que $\mathbb{U}_{p^k} \subset H$.

Si par l'absurde, il existe une infinité de $k \in \mathbb{N}$ vérifiant $\mathbb{U}_{p^k} \subset H$ alors $H = \mathbb{U}_{p^\infty}$ car l'ensemble $\mathbb{U}_{p^\infty}$ est la réunion croissante des $\mathbb{U}_{p^k}$. Ceci étant exclu, on peut introduire le plus grand $k \in \mathbb{N}$ vérifiant $\mathbb{U}_{p^k} \subset H$.

1. Un sous-groupe *propre* d'un groupe $(G, \star)$ est un sous-groupe non trivial, c'est-à-dire distinct de $\{e\}$ et G.

2. On peut aussi introduire, lorsqu'il existe, un élément d'ordre maximal parmi les éléments de H.

Tous les éléments de $\mathbb{U}_{p^{k+1}} \setminus \mathbb{U}_{p^k}$ engendrent $\mathbb{U}_{p^{k+1}}$. Or ce groupe n'est pas inclus dans H et donc H ne contient aucun élément de $\mathbb{U}_{p^{k+1}} \setminus \mathbb{U}_{p^k}$. *A fortiori*, H ne contient aucun élément de $\mathbb{U}_{p^\ell} \setminus \mathbb{U}_{p^k}$ pour $\ell > k$ et l'on en déduit $H \subset \mathbb{U}_{p^k}$ puis $H = \mathbb{U}_{p^k}$. Le sous-groupe H est donc cyclique.

Exercice 27 ** (Sous-groupes additifs de $\mathbb{R}$)

Soit H un sous-groupe de $(\mathbb{R}, +)$ non réduit à $\{0\}$. On pose $a = \inf\{h \in H \mid h > 0\}$

(a) On suppose $a > 0$. Montrer que a appartient à H puis que $H = a\mathbb{Z}$.

(b) On suppose $a = 0$. Montrer que H est une partie dense de $\mathbb{R}$.

(c) En déduire[1] $\overline{\{\cos n \mid n \in \mathbb{N}\}} = [-1\,;1]$.

Solution

Commençons par souligner l'existence du réel a. Dans le sous-groupe H, il existe un élément x non nul car H n'est pas réduit à $\{0\}$. Quitte à considérer $-x$, on peut affirmer qu'il existe dans H au moins un élément strictement positif. La partie

$$H^+ = \{h \in H \mid h > 0\}$$

est donc une partie de $\mathbb{R}$ non vide et minorée : on peut introduire sa borne inférieure.

(a) L'élément $2a$ ne minore pas H^+ car $2a > a$ et a est le plus grand des minorants de H^+. Il existe donc un élément $x \in H^+$ vérifiant $a \leqslant x < 2a$.

Si par l'absurde $x > a$, on peut encore introduire $y \in H^+$ tel que $a \leqslant y < x$. La différence $x - y$ détermine alors un élément du sous-groupe H strictement positif et strictement inférieur à a. Ceci contredit la définition de a comme borne inférieure. On en déduit $x = a$ puis $a \in H$.

méthode

On obtient l'égalité $H = a\mathbb{Z}$ par double inclusion. Écrire $a\mathbb{Z} = \langle a \rangle$ donne une première inclusion, la seconde se déduit d'une division euclidienne réelle[2].

Puisque a est élément du sous-groupe H, ce dernier contient le sous-groupe engendré par a : $a\mathbb{Z} = \langle a \rangle \subset H$. Inversement, soit $x \in H$. Par la division euclidienne de x par le réel $a > 0$, on écrit $x = aq + r$ avec $q \in \mathbb{Z}$ et $r \in [0\,;a[$. On a alors $r = x - aq \in H$ par opérations dans le sous-groupe H. En effet, x est élément de H et aq est élément de $a\mathbb{Z}$ donc de H. Si r est strictement positif, c'est un élément de H^+. Or ceci est impossible car $r < a$. On a donc nécessairement $r = 0$ puis $x = aq \in a\mathbb{Z}$. Par double inclusion, on conclut $H = a\mathbb{Z}$.

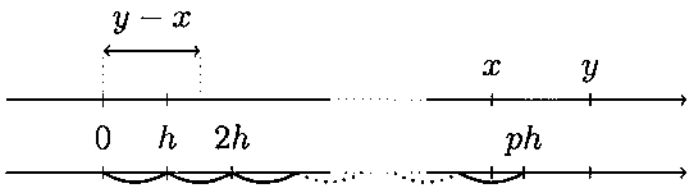
(b) **méthode**

On vérifie $]x\,;y[\cap H \neq \emptyset$ pour tous réels x et y tels que $x < y$.

1. Ici, la notation $\overline{A}$ désigne l'adhérence topologique d'une partie, voir le chapitre 3 de l'ouvrage *Exercices d'analyse MP* dans la même collection.

2. Pour $a, b \in \mathbb{R}$ avec $b > 0$, on peut écrire $a = bq + r$ avec $q \in \mathbb{Z}$ et $r \in [0\,;b[$: il suffit de prendre q égal à la partie entière de a/b et $r = a - bq$.

Soit $x, y \in \mathbb{R}$ avec $x < y$. Le réel $y - x$ est strictement positif, il ne minore donc pas la partie H^+. On peut alors introduire $h \in H$ tel que $0 < h < y - x$. Considérons ensuite p le plus petit entier tel que $ph > x$. Par définition de p, on a l'inégalité $(p-1)h \leqslant x$ ce qui entraîne $ph \leqslant x + h < y$. Ainsi, ph est un élément de $\langle h \rangle$, donc de H, strictement compris entre x et y.



(c) La partie $\{\cos n \mid n \in \mathbb{N}\}$ est incluse dans l'intervalle fermé $[-1\,;1]$ et son adhérence [1] est donc aussi incluse dans cet intervalle :

$$\overline{\{\cos n \mid n \in \mathbb{N}\}} \subset [-1\,;1].$$

méthode

‖ $\{\cos n \mid n \in \mathbb{N}\}$ est l'image par la fonction cos du sous-groupe $\mathbb{Z} + 2\pi\mathbb{Z}$.

On vérifie aisément [2] que $\mathbb{Z} + 2\pi\mathbb{Z}$ est un sous-groupe de $(\mathbb{R}, +)$. Si celui-ci est de la forme $a\mathbb{Z}$ avec $a > 0$, on peut écrire $1 = ak$ et $2\pi = a\ell$ avec $k, \ell \in \mathbb{Z}$ non nuls car 1 et 2π sont deux éléments de $\mathbb{Z} + 2\pi\mathbb{Z}$. On en déduit que π est le nombre rationnel $\ell/2k$ ce qui est absurde. Par conséquent, $\mathbb{Z} + 2\pi\mathbb{Z}$ est un sous-groupe dense de $(\mathbb{R}, +)$.

Considérons $x \in [-1\,;1]$ et $\theta = \arccos x$. Par densité de $\mathbb{Z} + 2\pi\mathbb{Z}$, on peut introduire deux suites d'entiers (k_n) et (ℓ_n) telles que

$$k_n + 2\pi\ell_n \xrightarrow[n\to+\infty]{} \theta.$$

Par parité, périodicité et continuité de la fonction cos, on obtient

$$\cos|k_n| = \cos k_n = \cos\bigl(k_n + 2\pi\ell_n\bigr) \xrightarrow[n\to+\infty]{} \cos\theta = x.$$

Ainsi, x est limite d'une suite d'éléments de $\{\cos n \mid n \in \mathbb{N}\}$.

Exercice 28 ***

À isomorphisme près, déterminer tous les groupes à 6 éléments.

Solution

Soit G un groupe à 6 éléments noté multiplicativement. Ses éléments sont d'ordres finis divisant 6 donc d'ordres possibles 1, 2, 3 ou 6.

1. Rappelons qu'un réel appartient à l'adhérence d'une partie A de $\mathbb{R}$ lorsqu'il est limite d'une suite d'éléments de A.

2. Éventuellement, consulter le sujet 11 p. 16 ou remarquer $\mathbb{Z} + 2\pi\mathbb{Z} = \langle 1, 2\pi \rangle$.

S'il existe un élément d'ordre 6 dans G, celui-ci engendre G qui est donc cyclique isomorphe à $\mathbb{Z}/6\mathbb{Z}$. On suppose désormais ce cas exclu.

méthode

Il existe [1] dans G un élément d'ordre 3.

Par l'absurde, si un tel élément n'existe pas, tous les éléments de G, sauf le neutre, sont d'ordre 2 : chaque élément est égal à son inverse et le groupe est commutatif car

$$xy = (xy)^{-1} = y^{-1}x^{-1} = yx \quad \text{pour tous } x, y \in G.$$

Soit a et b deux éléments de G distincts et distincts de 1. Le produit ab est nécessairement distinct de 1, a et b. Introduisons encore un nouvel élément c distinct des précédents. On connaît alors 5 éléments dans G : $1, a, b, ab$ et c. Les produits ac et bc sont forcément distincts des éléments précédents [2] et sont donc égaux car G n'a que six éléments. On en déduit $a = b$ ce qui est absurde [3].

Notons a un élément d'ordre 3 de G et b un élément n'appartenant pas au groupe engendré par a. Les deux éléments ab et a^2b complètent [4] alors le groupe G :

$$G = \left\{1, a, a^2, b, ab, a^2b\right\}.$$

méthode

On calcule b^2 et ba en se rappelant que G ne comporte pas d'éléments d'ordre 6.
On pourra ensuite déterminer la table d'opérations de G.

L'élément b^2 est immédiatement différent de b, ab et a^2b. L'élément b^2 est aussi différent de a car sinon b serait un élément d'ordre 6, situation que nous avons exclue. Le même argument assure que b^2 est différent de a^2. Il reste $b^2 = 1$.

L'élément ba ne peut être ni 1, ni a, ni b, ni a^2. Si $ba = ab$, les éléments a et b commutent et l'on vérifie que ab est d'ordre 6 ce qui est exclu. Il reste $ba = a^2b$.

On peut alors former la table d'opérations de G qui est identique [5] à celle du groupe symétrique $\mathcal{S}_3$.

	1	a	a^2	b	ab	a^2b
1	1	a	a^2	b	ab	a^2b
a	a	a^2	1	ab	a^2b	b
a^2	a^2	1	a	a^2b	b	ab
b	b	a^2b	ab	1	a^2	a
ab	ab	b	a^2b	a	1	a^2
a^2b	a^2b	ab	b	a^2	a	1

En résumé, à isomorphisme près, il existe deux groupes à 6 éléments : ce sont $(\mathbb{Z}/6\mathbb{Z}, +)$ et $(\mathcal{S}_3, \circ)$.

1. L'usage du lemme de Cauchy (sujet suivant) résout automatiquement cette question.
2. $ac \neq 1$ car c n'est pas $a^{-1} = a$, $ac \neq a$ car c n'est pas 1, $ac \neq b$ car $c \neq a^{-1}b = ab$, etc.
3. Un groupe fini dans lequel $x^2 = 1$ pour tout x a un cardinal qui est une puissance de 2 : voir le sujet 27 du chapitre 4 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI* dans la même collection.
4. Encore une fois, $ab \neq 1$ car $b \neq a^2$, $ab \neq a$ car $b \neq 1$, $ab \neq a^2$ car $b \neq a$, etc.
5. Et ce n'est même pas nécessaire de le vérifier ! Ce qui importe est que cette table soit entièrement calculable. Sachant que $\mathcal{S}_3$ est un groupe à 6 éléments non cyclique, il a forcément la même table.

Exercice 29 * (Lemme de Cauchy)**

Soit G un groupe fini noté multiplicativement. Soit p un nombre premier divisant le cardinal n du groupe G. On ambitionne de montrer qu'il existe dans G au moins un élément d'ordre p. On introduit pour cela l'ensemble

$$E = \left\{(g_1, g_2, \dots, g_p) \in G^p \mid g_1 g_2 \dots g_p = 1\right\}$$

et la relation d'équivalence $\mathcal{R}$ sur E déterminée par

$$(g_1, g_2, \dots, g_p)\ \mathcal{R}\ (h_1, h_2, \dots, h_p) \iff \exists k \in \mathbb{Z},\ \forall i \in [\![1\,;p]\!],\ h_i = g_{i+k}$$

où l'on adopte une notation circulaire : $g_{p+1} = g_1$, $g_{p+2} = g_2$, etc. Plus précisément, on pose $g_i = g_j$ pour $i \equiv j\ [p]$.

(a) Déterminer le cardinal de E.

(b) Montrer qu'une classe d'équivalence pour la relation $\mathcal{R}$ est de cardinal 1 ou p.

(c) En déduire l'existence d'un élément $g \neq 1$ vérifiant $g^p = 1$.

Solution

Deux éléments de E sont en relation par $\mathcal{R}$ si, et seulement si, ils se correspondent par permutation circulaire. Cette relation est évidemment réflexive, symétrique et transitive : c'est bien une relation d'équivalence.

(a) L'ensemble E est en bijection avec G^{p-1} via l'application qui à $(g_1, \dots, g_p)$ associe $(g_1, \dots, g_{p-1})$. En effet, la condition $g_1 g_2 \dots g_p = 1$ suffit à déterminer g_p en fonction des éléments $g_1, \dots, g_{p-1}$: $g_p = (g_1 \dots g_{p-1})^{-1}$. On en déduit

$$\mathrm{Card}(E) = \mathrm{Card}\left(G^{p-1}\right) = n^{p-1}.$$

(b) Étudions la classe d'équivalence de $(g_1, g_2, \dots, g_p)$ choisi dans E. Ses éléments sont les permutations circulaires $(g_{k+1}, g_{k+2}, \dots, g_{k+p})$ avec $k \in [\![0\,;p-1]\!]$ (pour les autres valeurs de $k \in \mathbb{Z}$, on obtient seulement des redites des éléments déjà listés).

méthode

On montre que si parmi les éléments ci-dessus, il y en a deux identiques alors $(g_1, g_2, \dots, g_p)$ est de la forme $(g, g, \dots, g)$.

Supposons qu'il existe $k, \ell \in [\![0\,;p-1]\!]$ avec $k < \ell$ tels que

$$(g_{k+1}, g_{k+2}, \dots, g_{k+p}) = (g_{\ell+1}, g_{\ell+2}, \dots, g_{\ell+p}).$$

On a donc $g_{k+i} = g_{\ell+i}$, d'abord pour tout $i \in [\![1\,;p]\!]$, puis pour tout $i \in \mathbb{Z}$ car $g_i = g_j$ dès que $i \equiv j\ [p]$. En notant $q = k - \ell \in [\![1\,;p-1]\!]$, on a alors $g_{i+q} = g_i$ pour tout i de $\mathbb{Z}$, puis, par récurrence, $g_{i+nq} = g_i$ pour tout $n \in \mathbb{N}$. L'entier p étant premier et q étant strictement inférieur à p, les entiers p et q sont premiers entre eux ce qui permet d'introduire $n \in \mathbb{N}^*$ tel que $nq \equiv 1\ [p]$. On obtient alors $g_{i+1} = g_i$ pour tout $i \in \mathbb{Z}$.

En résumé, soit les $(g_{k+1}, g_{k+2}, \dots, g_{k+p})$ pour k allant de 0 à $p-1$ sont deux à deux distincts, auquel cas il y en a p, soit $g_1 = g_2 = \cdots = g_p$ auquel cas la classe d'équivalence de $(g_1, g_2, \dots, g_p)$ se limite à un seul élément.

(c) **méthode**

> Le cardinal d'un ensemble muni d'une relation d'équivalence est la somme des cardinaux de ses classes d'équivalence.

Notons a le nombre de classes de d'équivalence de cardinal 1 et b le nombre de classes d'équivalence de cardinal p. Les classes d'équivalence réalisant une partition de E, on a l'égalité

$$\mathrm{Card}(E) = a \times 1 + b \times p.$$

L'entier p divisant le cardinal de E, il divise a. Or a est non nul car la classe d'équivalence de $(1, 1, \dots, 1) \in E$ est de cardinal 1. On en déduit $a \geqslant 2$ et donc l'existence de $(g_1, g_2, \dots, g_p) \neq (1, \dots, 1)$ tel que $g_1 = g_2 = \cdots = g_p$ et $g_1 g_2 \dots g_p = 1$. Ceci suffit à prouver l'existence[1] d'au moins un élément $g \in G$ tel que $g^p = 1$ et $g \neq 1$.

Exercice 30 * (Description des sous-groupes $\mathbb{Z}^2$)**

Dans ce sujet, on étudie les sous-groupes de $(\mathbb{Z}^2, +)$.

(a) Soit $e_1 = (x_1, y_1)$ et $e_2 = (x_2, y_2)$ deux éléments de $\mathbb{Z}^2$. Montrer

$$\langle e_1, e_2 \rangle = \mathbb{Z}^2 \iff \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \pm 1.$$

Soit H un sous-groupe de $\mathbb{Z}^2$ non réduit au neutre. On note d le plus petit PGCD de x et y pour (x, y) parcourant $H \setminus \{(0,0)\}$ et l'on introduit $(u_0, v_0) \in \mathbb{Z}^2$ et $(x_0, y_0) \in H$ tels que $d = u_0 x_0 + v_0 y_0$.

(b) Soit $(u, v) \in \mathbb{Z}^2$. Montrer qu'il existe $a_{u,v} \in \mathbb{Z}$ tel que

$$H_{u,v} = \{ux + vy \mid (x, y) \in H\} = a_{u,v}\mathbb{Z}.$$

(c) Déterminer a_{u_0,v_0}.

On introduit $e_1 = (x_1, y_1)$ et $e_2 = (x_2, y_2)$ avec

$$x_1 = \frac{x_0}{d}, \quad y_1 = \frac{y_0}{d}, \quad x_2 = -v_0 \quad \text{et} \quad y_2 = u_0.$$

(d) Montrer que $\langle e_1, e_2 \rangle = \mathbb{Z}^2$ et qu'il existe $n \in \mathbb{N}$ tel que $\langle de_1, ne_2 \rangle = H$.

(e) Vérifier que d divise n.

1. Plus précisément, il existe a éléments vérifiant $g^p = 1$ et, en mettant le neutre à part, on peut affirmer qu'il y a $a - 1 \equiv -1$ $[p]$ éléments d'ordre p dans G.

Solution

méthode

Dans un groupe commutatif[1] $\langle a, b\rangle = \{a^k b^\ell \mid k, \ell \in \mathbb{Z}\}$.

En notation additive $\langle a, b\rangle = \{ka + \ell b \mid k, \ell \in \mathbb{Z}\}$.

(a) Le groupe $\mathbb{Z}^2$ est engendré par $(1,0)$ et $(0,1)$. Les éléments e_1, e_2 engendrent $\mathbb{Z}^2$ si, et seulement si, il existe a, b, c, d entiers tels que

$$\begin{cases}(1,0) = ae_1 + be_2\\(0,1) = ce_1 + de_2.\end{cases}$$

Matriciellement, ce système s'écrit

$$\begin{pmatrix}x_1 & x_2\\y_1 & y_2\end{pmatrix}\begin{pmatrix}a & c\\b & d\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}.$$

On conclut en rappelant qu'une matrice à coefficients entiers est inversible d'inverse à coefficients entiers si, et seulement si, son déterminant vaut[2] ± 1.

(b) **méthode**

Les ensembles de la forme $a\mathbb{Z}$ sont les sous-groupes de $(\mathbb{Z}, +)$ (Th. 4 p. 5).

$H_{u,v}$ est une partie de $\mathbb{Z}$ non vide et stable par différence : c'est un sous-groupe de $(\mathbb{Z}, +)$ et l'on peut donc l'écrire $a_{u,v}\mathbb{Z}$ avec $a_{u,v} \in \mathbb{N}$.

(c) L'entier d appartient à H_{u_0,v_0} et est donc multiple de a_{u_0,v_0}. Inversement, on peut écrire $a_{u_0,v_0} = u_0x + v_0y$ avec $(x, y) \in H$, (x, y) non nul, et donc a_{u_0,v_0} est multiple du PGCD de x et y qui est supérieur à d. On en déduit $a_{u_0,v_0} = d$.

(d) Les couples e_1 et e_2 sont des éléments de $\mathbb{Z}^2$ vérifiant la condition de la première question car

$$x_1y_2 - x_2y_1 = \frac{u_0x_0 + v_0y_0}{d} = 1.$$

Les éléments e_1 et e_2 constituent une partie génératrice de $\mathbb{Z}^2$. Tout élément (x, y) de $\mathbb{Z}^2$ peut donc s'écrire $ae_1 + be_2$ avec $a, b \in \mathbb{Z}$. Après résolution, on obtient

$$\begin{cases}a = y_2x - x_2y\\b = -y_1x + x_1y.\end{cases}$$

Si $(x, y) \in H$, $a \in H_{y_2,-x_2} = H_{u_0,v_0} = d\mathbb{Z}$ et $b \in H_{-y_1,x_1} = n\mathbb{Z}$ en introduisant l'entier $n = a_{-y_1,x_1}$. Le sous-groupe H est donc inclus dans le groupe engendré par de_1 et ne_2.

1. Voir sujet 2 p. 9.
2. Voir sujet 25 du chapitre 10 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI.*

Inversement, de_1 appartient à H car on reconnaît (x_0, y_0). Pour conclure, il ne reste plus qu'à vérifier l'appartenance de ne_2 à H.

Puisque $n\mathbb{Z} = H_{-y_1,x_1}$, on peut introduire $(x, y) \in H$ tel que

$$n = -y_1 x + x_1 y.$$

Sachant $x_1 y_2 - x_2 y_1 = 1$, il vient après quelques calculs

$$ne_2 = (x, y) - (y_2 x - x_2 y)(x_1, y_1).$$

Or $y_2 x - x_2 y \in H_{y_2,-x_2} = H_{u_0,v_0} = d\mathbb{Z}$ et, en écrivant cet entier $d\ell$, on obtient

$$n.e_2 = (x, y) - \ell(x_0, y_0) \in H.$$

(e) **méthode**

> On introduit δ le PGCD de d et n et l'on détermine un élément (x, y) de H vérifiant $x \wedge y = \delta$.

Par une relation de Bézout, on écrit $\delta = ud + vn$ avec $u, v \in \mathbb{Z}$ puis on considère

$$(x, y) = u(\underbrace{de_1}_{\in H}) + v(\underbrace{ne_2}_{\in H}) = (udx_1 + vnx_2, udy_1 + vny_2) \in H.$$

Le couple (x, y) n'est pas nul car les vecteurs e_1 et e_2 sont linéairement indépendants et les coordonnées ud et vn ne sont pas toutes deux nulles puisque $\delta \neq 0$.

Le PGCD de x et y divise

$$x(y_2 - y_1) + y(x_1 - x_2) \underset{x_1 y_2 - x_2 y_1 = 1}{=} ud + vn = \delta.$$

Or δ divise aussi d qui est le plus petit PGCD des éléments (x, y) non nuls de H. On en déduit $x \wedge y = \delta = d$ et, en particulier, $d = \delta$ divise n.

CHAPITRE 2

Anneaux

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Structure d'anneau

2.1.1 Anneau

Définition

On appelle *anneau* tout triplet $(A, +, \times)$ formé d'un ensemble A et de deux lois de composition internes $+$ et $\times$ sur A vérifiant :

1) $(A, +)$ est un groupe abélien de neutre noté 0 (ou 0_A) ;
2) $\times$ est associative et possède un neutre dans A noté 1 (ou 1_A) ;
3) $\times$ est distributive sur $+$.

Si de plus, la loi $\times$ est commutative, on dit que l'anneau est *commutatif.*

$(\mathbb{Z}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$ et $(\mathbb{K}[X], +, \times)$ sont des anneaux commutatifs.

L'ensemble $\mathcal{F}(X, \mathbb{K})$ des fonctions numériques définies au départ d'une partie X est un anneau commutatif pour les opérations usuelles sur les fonctions numériques.

$\big(\mathcal{M}_n(\mathbb{K}), +, \times\big)$ est un anneau. Il n'est pas commutatif dès que $n \geqslant 2$.

Si E désigne un $\mathbb{K}$-espace vectoriel, $\big(\mathcal{L}(E), +, \circ\big)$ est un anneau. Il n'est pas commutatif dès que $\dim E \geqslant 2$.

Rappelons que l'ensemble $A^\times$ des éléments inversibles d'un anneau A est un groupe multiplicatif.

2.1.2 Anneau produit

Soit $(A_1,+,\times),\dots,(A_n,+,\times)$ des anneaux et $A = A_1 \times \cdots \times A_n$. On définit des lois $+$ et $\times$ sur A en posant (avec des notations entendues)

$$(x_1,\dots,x_n)+(y_1,\dots,y_n) \overset{\text{déf}}{=} (x_1+y_1,\dots,x_n+y_n),$$
$$(x_1,\dots,x_n)\times(y_1,\dots,y_n) \overset{\text{déf}}{=} (x_1\times y_1,\dots,x_n\times y_n).$$

Théorème 1
L'ensemble A muni des lois $+$ et $\times$ définies ci-dessus est un anneau de neutres

$$0_A = (0_{A_1},\dots,0_{A_n}) \quad \text{et} \quad 1_A = (1_{A_1},\dots,1_{A_n}).$$

Un élément $(a_1,\dots,a_n)$ de A est inversible si, et seulement si, les $a_1,\dots,a_n$ le sont tous et alors $(a_1,\dots,a_n)^{-1} = \left(a_1^{-1},\dots,a_n^{-1}\right)$.

2.1.3 Sous-anneaux

Définition
On appelle *sous-anneau* de $(A,+,\times)$ toute partie B de A contenant[1] 1_A et stable par différence et produit : $x-y \in B$ et $xy \in B$ pour tous x et y dans B.

Théorème 2
Si B est un sous-anneau d'un anneau $(A,+,\times)$ alors $(B,+,\times)$ est un anneau[2] de mêmes neutres que A.

2.1.4 Corps

Définition
On appelle corps tout anneau commutatif $(K,+,\times)$ non réduit à $\{0_K\}$ et dont tous les éléments, sauf le nul, sont inversibles.

$(\mathbb{Q},+,\times)$, $(\mathbb{R},+,\times)$, $(\mathbb{C},+,\times)$ et $\left(\mathbb{K}(X),+,\times\right)$ sont des corps.

Définition
On appelle *sous-corps* d'un corps $(K,+,\times)$ tout sous-anneau L de K contenant les inverses de ses éléments non nuls :

$$\forall x \in L, \quad x \neq 0_K \implies x^{-1} \in L.$$

Théorème 3
Si L est un sous-corps d'un corps $(K,+,\times)$ alors $(L,+,\times)$ est un corps.

1. On vérifie l'appartenance de 1_A et non seulement $B \neq \emptyset$: $2\mathbb{Z}$ est non vide, stable par différence et produit mais n'est pas un sous-anneau de $\mathbb{Z}$.
2. Les lois $+$ et $\times$ sur B sont les lois induites sur la partie B par les lois sur A.

2.1.5 Morphismes d'anneaux

Soit $(A, +, \times)$ et $(A', +, \times)$ deux anneaux.

Définition

On appelle *morphisme* de l'anneau A vers l'anneau A' toute application $\varphi : A \to A'$ vérifiant $\varphi(1_A) = 1_{A'}$ et, pour tous x et y de A,

$$\varphi(x+y) = \varphi(x) + \varphi(y) \quad \text{et} \quad \varphi(xy) = \varphi(x)\varphi(y).$$

Un morphisme d'anneaux est en particulier un morphisme de groupes additifs ce qui permet d'affirmer les propriétés calculatoires

$$\varphi(0_A) = 0_{A'}, \quad \varphi(-x) = -\varphi(x) \quad \text{et} \quad \varphi(nx) = n\varphi(x) \quad \text{pour tous } x \in A \text{ et } n \in \mathbb{Z}.$$

On a aussi $\varphi\left(x^p\right) = \varphi(x)^p$ pour tout $x \in A$ et tout $p \in \mathbb{N}$, et

$$x \text{ inversible} \implies \left(\varphi(x) \text{ inversible et } \varphi(x)^{-1} = \varphi\left(x^{-1}\right)\right).$$

En tant que morphisme de groupes additifs, on peut introduire l'image et le noyau d'un morphisme d'anneaux. L'image est un sous-anneau de l'anneau d'arrivée mais le noyau n'est généralement pas un sous-anneau de l'anneau de départ : c'est un idéal [1].

Définition

Un *isomorphisme* d'anneaux est un morphisme d'anneaux bijectif.

La composée de deux isomorphismes est un isomorphisme. Il en est de même pour la bijection réciproque d'un isomorphisme.

Lorsqu'il existe un isomorphisme entre deux anneaux, ceux-ci sont dits *isomorphes* : ils sont parfaitement identiques d'un point de vue calculatoire.

2.1.6 Intégrité

Dans un anneau $(A, +, \times)$, un produit est nul dès qu'un facteur est nul :

$$\forall (a,b) \in A^2, \quad \left(a = 0_A \text{ ou } b = 0_A\right) \implies ab = 0_A.$$

La réciproque peut cependant être fausse : le produit de deux matrices carrées non nulles peut être nul !

Définition

On dit qu'un élément non nul de l'anneau A est *diviseur* [2] *de zéro* s'il existe un élément non nul b de A vérifiant $ab = 0_A$ ou $ba = 0_A$.

Si l'on peut écrire $ab = 0_A$ avec a et b non nuls, a et b sont des diviseurs de zéro. On ne considère pas que 0_A soit un diviseur de zéro.

Les éléments inversibles d'un anneau ne sont jamais diviseurs de zéro.

1. Voir sujet 5 p. 44.
2. Ce vocabulaire ne doit pas être confondu avec celui de l'arithmétique.

Définition

On dit qu'un anneau $(A,+,\times)$ est *intègre* s'il n'est pas réduit à $\{0_A\}$ et s'il ne possède pas de diviseurs de zéros.

$(\mathbb{Z},+,\times)$ et $(\mathbb{K}[X],+,\times)$ sont des anneaux intègres.
Les corps sont aussi des anneaux intègres.

Théorème 4

Dans un anneau intègre $(A,+,\times)$, on dispose de l'*implication d'intégrité*

$$\forall (a,b)\in A^2,\quad ab=0_A \implies \big(a=0_A \text{ ou } b=0_A\big).$$

Dans un anneau intègre, les éléments non nuls sont *réguliers* : pour tous $a,b,c\in A$

$$\big(ab=ac \text{ et } a\neq 0_A\big) \implies b=c,$$
$$\big(ba=ca \text{ et } a\neq 0_A\big) \implies b=c.$$

2.2 Idéal d'un anneau commutatif

Soit $(A,+,\times)$ un anneau commutatif.

2.2.1 Idéal

Définition

On appelle *idéal* de l'anneau $(A,+,\times)$ toute partie I de A non vide, stable par addition et vérifiant la propriété d'*absorption* :

$$\forall a\in A,\ \forall x\in I,\quad ax\in I.$$

Les idéaux sont en particulier des sous-groupes additifs.
$\{0_A\}$ et A sont des idéaux de $(A,+,\times)$, ce sont ses *idéaux triviaux*.

Théorème 5

Si I et J sont deux idéaux de $(A,+,\times)$ alors

a) $I\cap J$ est un idéal inclus dans I et J et contenant tout idéal inclus dans I et J.

b) $I+J \overset{\text{déf}}{=} \{x+y \,|\, x\in I, y\in J\}$ est un idéal contenant I et J et inclus dans tout idéal contenant I et J.

2.2.2 Idéal engendré par un élément

Définition

On appelle *idéal engendré* par un élément x de A l'ensemble

$$xA \overset{\text{déf}}{=} \{xu \,|\, u\in A\}.$$

Lorsqu'un idéal peut être décrit comme engendré par un élément, on dit qu'il est *principal.*

Théorème 6
xA est un idéal contenant l'élément x et inclus dans tout idéal contenant x.

2.2.3 Idéaux de $\mathbb{Z}$ et de $\mathbb{K}[X]$

Théorème 7
Les idéaux de $(\mathbb{Z}, +, \times)$ sont les $n\mathbb{Z}$ avec $n \in \mathbb{N}$.
Les idéaux de $(\mathbb{K}[X], +, \times)$ sont les $P\,\mathbb{K}[X]$ avec $P \in \mathbb{K}[X]$.

Autrement dit, les idéaux de $\mathbb{Z}$ sont principaux. On dit que encore que l'anneau $\mathbb{Z}$ est *principal.* L'anneau $\mathbb{K}[X]$ est aussi principal [1].

2.2.4 Divisibilité

Soit $(A, +, \times)$ un anneau intègre et commutatif.

Définition
Soit a et b deux éléments de A. On dit que a *divise* b s'il existe $u \in A$ tel que $b = au$. On note alors $a \mid b$ et l'on dit que a est un *diviseur* de b ou encore que b est un *multiple* de a.

1_A divise a et a divise a. Tout élément de A divise 0_A mais 0_A ne divise que lui-même.

Signifier que b est un multiple de a s'écrit encore $b \in aA$. Une divisibilité peut alors s'interpréter en terme d'inclusion d'idéaux :

$$a \mid b \iff bA \subset aA.$$

2.2.5 PGCD, PPCM et idéaux

Soit a et b deux entiers. Par opérations sur les idéaux, $a\mathbb{Z} + b\mathbb{Z}$ et $a\mathbb{Z} \cap b\mathbb{Z}$ déterminent deux idéaux de $\mathbb{Z}$. Ceux-ci peuvent s'écrire $d\mathbb{Z}$ et $m\mathbb{Z}$ avec d et m naturels.

Théorème 8
Soit $a, b \in \mathbb{Z}$. Les entiers naturels d et m tels que

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z} \quad \text{et} \quad a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$$

sont respectivement les PGCD et PPCM de a et b.

On dispose de la même interprétation dans le cadre des polynômes :

1. Les deux anneaux $\mathbb{Z}$ et $\mathbb{K}[X]$ ont des propriétés communes du fait de l'existence d'une division euclidienne sur ceux-ci.

Théorème 9
Soit $P, Q \in \mathbb{K}[X]$. Des polynômes D et M de $\mathbb{K}[X]$ tels que

$$P\,\mathbb{K}[X] + Q\,\mathbb{K}[X] = D\,\mathbb{K}[X] \quad \text{et} \quad P\,\mathbb{K}[X] \cap Q\,\mathbb{K}[X] = M\,\mathbb{K}[X]$$

sont respectivement des[1] PGCD et PPCM de P et Q.

2.3 L'anneau $\mathbb{Z}/n\mathbb{Z}$

Soit $n, m \in \mathbb{N}^*$.

2.3.1 Présentation

Théorème 10
$(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif à n éléments.
Les éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ sont les $\overline{m}$ pour m entier premier avec n.

En particulier, $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps si, et seulement si, n est un nombre premier.

2.3.2 Théorème des restes chinois

Pour $k \in \mathbb{Z}$, on peut considérer les classes de k modulo mn, modulo m et modulo n. On notera distinctement celles-ci : $\bar{k}$, $\dot{k}$ et $\hat{k}$.

Théorème 11 (Théorème des restes chinois)
Si m et n sont deux entiers premiers entre eux, les anneaux $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes par l'application $\bar{k} \mapsto (\dot{k}, \hat{k})$.

Si l'on dispose de la relation de Bézout $mu + nv = 1$, l'isomorphisme réciproque est l'application

$$(\dot{a}, \hat{b}) \mapsto \overline{anv + bmu}.$$

2.3.3 Fonction indicatrice d'Euler

Définition
On appelle fonction *indicatrice d'Euler* l'application $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*$ définie par

$$\varphi(n) = \mathrm{Card}\big(\{k \in [\![1\,;n]\!] \mid k \wedge n = 1\}\big).$$

$\varphi(n)$ représente le nombre de générateurs du groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ ou, ce sont les mêmes éléments, le nombre d'éléments inversibles de l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$.

1. Tout polynôme associé à un PGCD de P et Q est encore un PGCD de P et Q et inversement. Lorsque P et Q ne sont pas tous deux nuls, on peut parler *du* PGCD de P et Q en privilégiant parmi les polynômes associés celui de coefficient dominant égal à 1. On a une remarque analogue pour le PPCM.

Si p est un nombre premier et $\alpha \in \mathbb{N}^*$, on a $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$.
Si $m, n \in \mathbb{N}^*$ sont premiers entre eux, on établit la formule $\varphi(mn) = \varphi(m)\varphi(n)$ à l'aide du théorème chinois. On en déduit le résultat suivant :

Théorème 12
Si la décomposition en facteurs premiers de $n \geqslant 2$ s'écrit $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ alors

$$\varphi(n) = n \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right).$$

Il suffit donc de déterminer les facteurs premiers de n pour savoir calculer $\varphi(n)$.

Théorème 13 (Théorème d'Euler)
Soit $a \in \mathbb{Z}$ et $n \in \mathbb{N}^*$.

$$a \wedge n = 1 \iff a^{\varphi(n)} \equiv 1 \ [n].$$

Si p est un nombre premier, $\varphi(p) = p - 1$ et l'on retrouve le petit théorème de Fermat :

$$a \not\equiv 0 \ [p] \implies a^{p-1} \equiv 1 \ [p].$$

2.4 Exercices d'apprentissage

2.4.1 Généralités sur les anneaux et les corps

Exercice 1
L'ensemble des nombres décimaux est

$$\mathbb{D} = \left\{ \frac{n}{10^k} \;\middle|\; n \in \mathbb{Z}, k \in \mathbb{N} \right\}.$$

(a) Montrer que $\mathbb{D}$ est un sous-anneau de $(\mathbb{Q}, +, \times)$.
(b) Déterminer ses éléments inversibles.

Solution

(a) **méthode**
‖ On vérifie que $\mathbb{D}$ contient 1 et est stable par différence et produit.

$\mathbb{D}$ est évidemment une partie de $\mathbb{Q}$ contenant 1 car on peut écrire $1 = 1/10^0$. Soit x et y deux éléments de $\mathbb{D}$. On écrit $x = n/10^k$ et $y = m/10^\ell$ avec $n, m \in \mathbb{Z}$ et $k, \ell \in \mathbb{N}$ et alors

$$x - y = \frac{n10^\ell - m10^k}{10^{k+\ell}} \in \mathbb{D} \quad \text{et} \quad xy = \frac{nm}{10^{k+\ell}} \in \mathbb{D}.$$

Ainsi, $\mathbb{D}$ est un sous-anneau de $(\mathbb{Q}, +, \times)$.

(b) **méthode**

Les éléments non nuls de $\mathbb{D}$ sont assurément inversibles dans $\mathbb{Q}$: on recherche ici ceux dont l'inverse est aussi dans $\mathbb{D}$.

Un élément $x \in \mathbb{D}$ est inversible dans l'anneau $\mathbb{D}$ si, et seulement si, il existe $y \in \mathbb{D}$ vérifiant $xy = 1$. En écrivant $x = n/10^k$ et $y = m/10^\ell$ avec $n, m \in \mathbb{Z}$ et $k, \ell \in \mathbb{N}$,

$$\begin{aligned} xy = 1 &\iff mn = 10^{k+\ell} \\ &\implies n \mid 10^{k+\ell}. \end{aligned}$$

Si x est inversible dans $\mathbb{D}$, les seuls facteurs premiers possibles de n sont 2 et 5. Au signe près, l'entier n s'exprime comme un produit de puissances de 2 et de 5. Le nombre x s'écrit alors $\pm 2^p 5^q$ avec $p, q \in \mathbb{Z}$. La réciproque est immédiate et, finalement,

$$\mathbb{D}^\times = \left\{ \pm 2^p 5^q \,\middle|\, p, q \in \mathbb{Z} \right\}.$$

Exercice 2

Soit $\alpha \in \mathbb{N}$ tel que $\sqrt{\alpha} \notin \mathbb{Q}$, on note

$$\mathbb{Q}[\sqrt{\alpha}] = \left\{ a + b\sqrt{\alpha} \,\middle|\, a, b \in \mathbb{Q} \right\}.$$

Montrer que $\mathbb{Q}[\sqrt{\alpha}]$ est un corps pour les opérations usuelles.

Solution

méthode

On vérifie que $\mathbb{Q}[\sqrt{\alpha}]$ est un sous-corps de $(\mathbb{R}, +, \times)$.

$\mathbb{Q}[\sqrt{\alpha}]$ est une partie de $\mathbb{R}$ contenant 1 car on peut écrire $1 = 1 + 0\sqrt{\alpha}$. Soit x et y deux éléments arbitraires de $\mathbb{Q}[\sqrt{\alpha}]$. On écrit $x = a + b\sqrt{\alpha}$ et $y = a' + b'\sqrt{\alpha}$ avec $a, b, a', b' \in \mathbb{Q}$ et alors

$$x - y = \underbrace{(a - a')}_{\in\mathbb{Q}} + \underbrace{(b - b')}_{\in\mathbb{Q}}\sqrt{\alpha} \in \mathbb{Q}[\sqrt{\alpha}] \quad \text{et} \quad xy = \underbrace{(aa' + bb'\alpha)}_{\in\mathbb{Q}} + \underbrace{(ab' + a'b)}_{\in\mathbb{Q}}\sqrt{\alpha} \in \mathbb{Q}[\sqrt{\alpha}].$$

Ainsi, $\mathbb{Q}[\sqrt{\alpha}]$ est un sous-anneau de $(\mathbb{R}, +, \times)$.

De plus, si $x \neq 0$, on peut multiplier numérateur et dénominateur par la quantité conjuguée[1] $a - b\sqrt{\alpha}$ et l'on obtient

$$\frac{1}{x} = \frac{1}{a + b\sqrt{\alpha}} = \frac{a - b\sqrt{\alpha}}{a^2 - \alpha b^2} = \underbrace{\frac{a}{a^2 - \alpha b^2}}_{\in\mathbb{Q}} - \underbrace{\frac{b}{a^2 - \alpha b^2}}_{\in\mathbb{Q}}\sqrt{\alpha} \in \mathbb{Q}[\sqrt{\alpha}].$$

Finalement, $\mathbb{Q}[\sqrt{\alpha}]$ est un sous-corps de $(\mathbb{R}, +, \times)$ et c'est donc un corps[2] pour les opérations usuelles.

1. Celle-ci est non nulle car $\sqrt{\alpha}$ est nombre irrationnel.
2. Lorsque $\sqrt{\alpha} \in \mathbb{Q}$, on remarque $\mathbb{Q}[\sqrt{\alpha}] = \mathbb{Q}$ et la conclusion reste vraie.

Exercice 3

Soit a un élément inversible d'un anneau A. Vérifier que l'application $f\colon x \mapsto axa^{-1}$ est un isomorphisme de l'anneau A vers lui-même [1].

Solution

L'application f est bien définie au départ de A et à valeurs dans A.

méthode

On vérifie que f transforme la somme en la somme, le produit en le produit mais aussi le 1_A en le 1_A.

On a immédiatement $f(1_A) = a1_Aa^{-1} = 1_A$. Soit $x, y \in A$. Par distributivité

$$f(x+y) = a(x+y)a^{-1} = (ax+ay)a^{-1} = axa^{-1} + aya^{-1} = f(x) + f(y).$$

Aussi, en écrivant « astucieusement » $1_A = a^{-1}a$, on obtient par associativité,

$$f(xy) = axya^{-1} = ax\big(a^{-1}a\big)ya^{-1} = \big(axa^{-1}\big)\big(aya^{-1}\big) = f(x)f(y).$$

L'application f est donc un morphisme de l'anneau A. Étudions sa bijectivité. Soit $y \in A$. On résout l'équation $f(x) = y$ d'inconnue $x \in A$:

$$\begin{aligned} f(x) = y &\iff axa^{-1} = y \\ &\iff xa^{-1} = a^{-1}y && \text{en multipliant à gauche par } a^{-1} \\ &\iff x = a^{-1}ya && \text{en multipliant à droite par } a. \end{aligned}$$

L'application f est donc bijective d'application réciproque $y \mapsto a^{-1}ya$.

Finalement, f est un isomorphisme de l'anneau A vers lui-même.

Exercice 4

Dans un anneau A, on étudie l'équation $x^2 = 1_A$ d'inconnue $x \in A$.

(a) On suppose l'anneau A intègre. Résoudre l'équation introduite.

(b) Observer que l'équation peut posséder d'autres solutions que les précédentes lorsque l'anneau A est l'un des anneaux non intègres suivants : $\mathbb{Z}^2$, $\mathbb{Z}/8\mathbb{Z}$ et $\mathcal{M}_2(\mathbb{R})$.

Solution

(a) **méthode**

Dans un anneau intègre, on peut résoudre une équation en factorisant une égalité à 0_A.

Soit $x \in A$.

$$\begin{aligned} x^2 = 1_A &\iff x^2 - 1_A = 0_A \\ &\iff (x - 1_A)(x + 1_A) = 0_A. \end{aligned}$$

1. On peut parler d'*automorphisme* de l'anneau A.

Dans un anneau intègre, un produit est nul si, et seulement si, l'un des facteurs est nul :

$$x^2 = 1_A \iff \big(x - 1_A = 0_A \text{ ou } x + 1_A = 0_A\big).$$

L'équation étudiée possède donc deux [1] solutions 1_A et -1_A.

(b) Dans $\mathbb{Z}^2$, l'équation $x^2 = 1_A$ se relit $(k,\ell)^2 = (1,1)$ avec $x = (k,\ell) \in \mathbb{Z}^2$. Celle-ci possède 4 solutions : $1_{\mathbb{Z}^2} = (1,1)$, $-1_{\mathbb{Z}^2} = (-1,-1)$ mais aussi $(1,-1)$ et $(-1,1)$.

Dans $\mathbb{Z}/8\mathbb{Z}$, on peut calculer les carrés des 8 éléments $\overline{0}, \overline{1}, \ldots, \overline{7}$. Les solutions de l'équation $x^2 = \overline{1}$ sont $\overline{1}$, $\overline{3}$, $\overline{5}$ et $\overline{7} = -\overline{1}$.

Dans $\mathcal{M}_2(\mathbb{R})$, les solutions de l'équation $M^2 = \mathrm{I}_2$ sont les matrices de symétrie. Hormis les matrices I_2 et $-\mathrm{I}_2$, ce sont les matrices [2]

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \quad \text{avec} \quad a^2 + bc = 1.$$

2.4.2 Idéaux

Exercice 5

Vérifier que le noyau d'un morphisme d'anneaux est un idéal.

Solution

Soit A et A' deux anneaux avec A commutatif [3] et $f : A \to A'$ un morphisme d'anneaux. On étudie l'ensemble $I = \mathrm{Ker}(f)$ des solutions de l'équation $f(x) = 0_{A'}$.

méthode

On vérifie que I est une partie non vide stable par addition et qu'elle satisfait la propriété d'absorption : $ax \in I$ pour tout $a \in A$ et tout $x \in I$.

L'ensemble I est une partie non vide de A car $f(0_A) = 0_{A'}$. Soit x et y deux éléments de I. Par morphisme on a $f(x+y) = f(x) + f(y) = 0_{A'}$ et donc $x + y \in I$. Soit de plus a un élément de A. Par morphisme on a aussi $f(ax) = f(a)f(x) = 0_{A'}$ car $f(x)$ est nul. On en déduit que ax est effectivement élément de I.

Finalement, I est un idéal de A.

Exercice 6

Soit A un anneau commutatif.

(a) Que dire d'un idéal de A qui contient le neutre 1_A ?

(b) Quels sont les idéaux d'un corps K ?

1. Ces solutions peuvent exceptionnellement être confondues : c'est le cas lorsque $A = \mathbb{Z}/2\mathbb{Z}$.
2. Voir sujet 10 du chapitre 9 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.
3. Le programme limite la notion d'idéal aux anneaux commutatifs.

Solution

(a) **méthode**

‖ Par la propriété d'absorption, un idéal I qui contient 1_A doit contenir A.

Soit I un idéal de A contenant l'élément 1_A. Pour tout $a \in A$, on peut écrire $a = a1_A$ et donc $a \in I$ car c'est le produit d'un élément de A par un élément de l'idéal I. Par double inclusion, on conclut $A = I$.

(b) Soit I un idéal d'un corps K. Si $I \neq \{0_K\}$ alors I contient un élément x non nul. Dans le corps K celui-ci est inversible ce qui permet d'introduire x^{-1}. La propriété d'absorption de l'idéal donne $ax \in I$ pour tout $a \in K$. En particulier, pour $a = x^{-1}$, on obtient $1_K \in I$ et donc $I = K$. En résumé, en dehors de $\{0_K\}$, le seul idéal d'un corps est lui-même.

2.4.3 Équations et systèmes en congruence

Exercice 7 (Équations modulaires)
Résoudre les équations suivantes d'inconnue $x \in \mathbb{Z}$.
(a) $6x + 2 \equiv 0\ [11]$ (b) $6x + 2 \equiv 0\ [10]$ (c) $6x + 2 \equiv 0\ [9]$.

Solution

(a) **méthode**

‖ Une équation modulo n a une expression équivalente dans $\mathbb{Z}/n\mathbb{Z}$.

On considère les classes d'équivalence dans $\mathbb{Z}/11\mathbb{Z}$

$$\begin{aligned} 6x + 2 \equiv 0\ [11] &\iff \overline{6}\overline{x} + \overline{2} = \overline{0} \\ &\iff \overline{6}\overline{x} = \overline{9} \qquad \text{en ajoutant } -\overline{2} = \overline{9} \text{ de part et d'autre.} \end{aligned}$$

Puisque 6 est premier avec 11, $\overline{6}$ est inversible dans $\mathbb{Z}/11\mathbb{Z}$ (Th. 10 p. 40) et son inverse est [1] $\overline{2}$.

$$\begin{aligned} 6x + 2 \equiv 0\ [11] &\iff \overline{x} = \overline{2} \times \overline{9} \\ &\iff \overline{x} = \overline{7} \qquad \text{car } 18 \equiv 7\ [11] \end{aligned}$$

Les solutions de l'équation étudiée sont donc [2] les $7 + 11k$ avec $k \in \mathbb{Z}$.

(b) **méthode**

‖ La démarche précédente ne peut être immédiatement reprise car $\overline{6}$ n'est pas inversible dans $\mathbb{Z}/10\mathbb{Z}$: on réduit l'équation en simplifiant par 2.

1. Cet inverse peut être déterminé en testant différentes valeurs ou en rappelant que, si m est premier avec n, l'inverse de m modulo n est le facteur u d'une relation de Bézout $mu + nv = 1$.

2. En congruence modulo 11, il n'y avait que 11 valeurs à tester pour trouver les solutions, on aurait pu les étudier toutes... En congruence supérieure, la démarche présentée ici devient efficace.

On a

$$6x+2\equiv 0\ [10] \iff 3x+1\equiv 0\ [5].$$

On considère les classes d'équivalence dans $\mathbb{Z}/5\mathbb{Z}$. $\overline{3}$ est inversible d'inverse $\overline{2}$.

$$\begin{aligned} 3x+1\equiv 0\ [5] &\iff \overline{3}\overline{x}+\overline{1}=\overline{0} \\ &\iff \overline{3}\overline{x}=\overline{4} \\ &\iff \overline{x}=\overline{3} \qquad \text{car} \quad 2\times 4=8\equiv 3\ [5]. \end{aligned}$$

Les solutions de l'équation sont les $3+5k$ avec $k\in\mathbb{Z}$.

(c) **méthode**

Encore une fois $\overline{6}$ n'est pas inversible dans $\mathbb{Z}/9\mathbb{Z}$. Cependant, ici aucune réduction n'est possible.

3 divise $6x$ et divise 9 mais ne divise pas 2 : il ne peut pas y avoir de solutions à l'équation étudiée.

Exercice 8 (Systèmes chinois)

Résoudre les systèmes suivants d'inconnue $x\in\mathbb{Z}$.

(a) $\begin{cases} x\equiv 2\ [5] \\ x\equiv 3\ [9] \end{cases}$ (b) $\begin{cases} 9x\equiv 3\ [21] \\ 5x\equiv 2\ [8] \end{cases}$ (c) $\begin{cases} x\equiv 7\ [9] \\ x\equiv 6\ [7] \\ x=3\ [5] \end{cases}$.

Solution

méthode

Lorsque les entiers m et n sont premiers entre eux, on peut écrire une relation de Bézout $mu+nv=1$ avec $u,v\in\mathbb{Z}$. Les entiers $y=nv$ et $z=mu$ sont alors solutions des systèmes

$$\begin{cases} y\equiv 1\ [m] \\ y\equiv 0\ [n] \end{cases} \quad \text{et} \quad \begin{cases} z\equiv 0\ [m] \\ z\equiv 1\ [n] \end{cases}.$$

(a) Les entiers 5 et 9 sont premiers entre eux et l'on peut écrire $-7\times 5+4\times 9=1$. L'entier

$$x=2\times \underbrace{4\times 9}_{nv}+3\times \underbrace{(-7)\times 5}_{mu}=-33$$

est alors solution du système et donc

$$\begin{cases} x\equiv 2\ [5] \\ x\equiv 3\ [9] \end{cases} \iff \begin{cases} x\equiv -33\ [5] \\ x\equiv -33\ [9] \end{cases}$$

Par le théorème des restes chinois (Th. 11 p. 40).

$$\begin{aligned} \begin{cases} x\equiv -33\ [5] \\ x\equiv -33\ [9] \end{cases} &\iff x\equiv -33\ [45] \\ &\iff x\equiv \ \ 12\ [45]. \end{aligned}$$

Finalement, les solutions du système sont donc les $12+45k$ avec $k\in\mathbb{Z}$.

(b) **méthode**

‖ On résout les deux équations du système avant d'exploiter le théorème chinois.

$$\begin{cases} 9x \equiv 3 \ [21] \\ 5x \equiv 2 \ [8] \end{cases} \iff \begin{cases} 3x \equiv 1 \ [7] \\ 5x \equiv 2 \ [8] \end{cases}$$
$$\iff \begin{cases} x \equiv 5 \ [7] \\ x \equiv 2 \ [8] \end{cases}$$
$$\iff x \equiv 5 \times 8 + 2 \times (-7) \ [56] \quad \text{car} \quad -7 + 8 = 1.$$

Les solutions du système sont les $26 + 56k$ avec $k \in \mathbb{Z}$.

(c) On utilise deux fois le théorème chinois :

$$\begin{cases} x \equiv 7 \ [9] \\ x \equiv 6 \ [7] \\ x \equiv 3 \ [5] \end{cases} \iff \begin{cases} x \equiv 7 \ [9] \\ x \equiv 13 \ [35] \end{cases}$$
$$\iff x \equiv 223 \ [315].$$

2.5 Exercices d'entraînement

2.5.1 Anneaux et corps

Exercice 9 *

Soit $(A, +, \times)$ un anneau intègre de cardinal fini.

(a) Soit $a \in A$ non nul. Montrer que $x \mapsto ax$ définit une permutation de A.

(b) En déduire que tout élément non nul de $(A, +, \times)$ est inversible.

Solution

(a) Notons $f : A \to A$ l'application déterminée par $f(x) = ax$.

méthode

‖ On établit la bijectivité de f par injectivité et un argument de cardinalité.

Soit x et y deux éléments de A. Supposons $f(x) = f(y)$ c'est-à-dire $ax = ay$. Par différence et factorisation, $a(x - y) = 0_A$. Or l'anneau A est intègre et a est non nul, on a donc $x - y = 0_A$ (Th. 4 p. 38) et par conséquent [1] $x = y$. L'application f est donc une injection de A dans A. Cependant, l'ensemble A est fini et l'injection f est nécessairement bijective.

1. On retient que dans un anneau intègre $ax = ay \implies x = y$ dès que a est non nul : on dit que les éléments non nuls d'un anneau intègre sont *réguliers*.

(b) Soit $a \in A$ différent de 0_A. Par la surjectivité de l'application f, on peut introduire un élément $b \in A$ tel que $ab = 1_A$. Il reste à vérifier [1] $ba = 1_A$ pour pouvoir conclure que a est inversible (et que b est son inverse).

méthode

‖ On compare $f(ba)$ et $f(1_A)$.

Par associativité, $f(ba) = a(ba) = (ab)a = a$ et donc $f(ba) = f(1_A)$. Par l'injectivité de f, on en déduit l'égalité $ba = 1_A$. On peut alors conclure que a est inversible.

Exercice 10 ** (Description des sous-anneaux de $\mathbb{Z}^2$)

Pour $d \in \mathbb{N}$, on note $A_d = \{(x, y) \in \mathbb{Z}^2 \mid d \text{ divise } y - x\}$.

(a) Montrer que A_d est un sous-anneau $(\mathbb{Z}^2, +, \times)$.

Inversement, soit A un sous-anneau de $(\mathbb{Z}^2, +, \times)$.

(b) Montrer qu'il existe $d \in \mathbb{N}$ tel que $\{x \in \mathbb{Z} \mid (x, 0) \in A\} = d\mathbb{Z}$.

(c) En déduire que $A = A_d$.

Solution

(a) A_d est un partie de $\mathbb{Z}^2$ contenant l'élément $1_{\mathbb{Z}^2} = (1, 1)$ car d divise 0. Soit (x, y) et (x', y') deux éléments de A_d. L'entier d divise $y - x$ et $y' - x'$ et l'on peut affirmer

$$(x, y) - (x', y') = (x - x', y - y') \in A_d \quad \text{et} \quad (x, y)(x', y') = (xx', yy') \in A_d.$$

En effet, d'une part, d divise

$$(y - y') - (x - x') = (y - x) - (y' - x')$$

et, d'autre part, d divise

$$(yy' - xx') = (y - x)y' + x(y' - x').$$

Ainsi, A_d est stable par différence et produit : c'est un sous-anneau de $\mathbb{Z}^2$.

(b) **méthode**

‖ On vérifie que $H = \{x \in \mathbb{Z} \mid (x, 0) \in A\}$ est un sous-groupe de $(\mathbb{Z}, +)$.

H est une partie non vide de $\mathbb{Z}$. En effet, $0 \in H$ car $(0, 0) = 0_{\mathbb{Z}^2} \in A$. Au surplus, pour tous x et $y \in H$, on a $(x - y, 0) = (x, 0) - (y, 0) \in A$ par différence d'éléments du sous-anneau A. On en déduit $x - y \in H$ et l'on peut affirmer que H est un sous-groupe de $(\mathbb{Z}, +)$. Il existe donc $d \in \mathbb{N}$ tel que $H = d\mathbb{Z}$ (Th. 4 p. 5).

(c) **méthode**

‖ Le sous-anneau A contient $1_{\mathbb{Z}^2}$ et donc le sous-groupe additif qu'il engendre.

1. Si l'anneau est commutatif, on peut immédiatement conclure et même affirmer que A est un corps.

On raisonne par double inclusion.

Soit $(x,y) \in A$, on a $(x-y,0) = (x,y) - (y,y)$ avec $(y,y) = y(1,1) \in \langle (1,1) \rangle \subset A$. Par différence de deux éléments de A, $(x-y,0)$ appartient à A et donc $x-y \in H = d\mathbb{Z}$. Ainsi, d divise $x-y$ et aussi $y-x$. On en déduit une première inclusion $A \subset A_d$.

Inversement, soit $(x,y) \in A_d$. L'entier d divise $y-x$ et donc $x-y$ est élément de $d\mathbb{Z}$, ce qui donne $(x-y,0) \in A$. On en déduit $(x,y) = (x-y,0) + y(1,1) \in A$ par opérations dans A. On a donc l'inclusion réciproque $A_d \subset A$ et l'on peut conclure à l'égalité $A = A_d$.

Exercice 11 ** (L'anneau des entiers de Gauss)

On note

$$\mathbb{Z}[\mathrm{i}] = \{a + \mathrm{i}b \mid (a,b) \in \mathbb{Z}^2\}.$$

(a) Montrer que $\mathbb{Z}[\mathrm{i}]$ est un anneau commutatif pour l'addition et la multiplication des nombres complexes.

(b) Déterminer les éléments inversibles de l'anneau $\mathbb{Z}[\mathrm{i}]$.

(c) Soit u et v deux éléments de $\mathbb{Z}[\mathrm{i}]$ avec $v \neq 0$. Montrer qu'il existe un couple (q,r) d'éléments de $\mathbb{Z}[\mathrm{i}]$ tel que $u = qv + r$ et $|r| < |v|$.

(d) Vérifier que les idéaux de $\mathbb{Z}[\mathrm{i}]$ sont de la forme $v\mathbb{Z}[\mathrm{i}]$ avec $v \in \mathbb{Z}[\mathrm{i}]$.

Solution

(a) **méthode**

‖ On vérifie que $\mathbb{Z}[\mathrm{i}]$ est un sous-anneau de l'anneau $(\mathbb{C},+,\times)$.

$\mathbb{Z}[\mathrm{i}]$ est une partie $\mathbb{C}$ contenant 1 puisque l'on peut écrire $1 = 1 + 0\mathrm{i}$. Soit $x,y \in \mathbb{Z}[\mathrm{i}]$. On écrit $x = a + \mathrm{i}b$ et $y = c + \mathrm{i}d$ avec $a,b,c,d \in \mathbb{Z}$ et alors

$$x - y = \underbrace{(a-c)}_{\in\mathbb{Z}} + \mathrm{i}\underbrace{(b-d)}_{\in\mathbb{Z}} \in \mathbb{Z}[\mathrm{i}] \quad \text{et} \quad xy = \underbrace{(ac-bd)}_{\in\mathbb{Z}} + \mathrm{i}\underbrace{(ad+bc)}_{\in\mathbb{Z}} \in \mathbb{Z}[\mathrm{i}].$$

Ainsi, $\mathbb{Z}[\mathrm{i}]$ est un sous-anneau de $(\mathbb{C},+,\times)$ donc un anneau pour les lois induites.

(b) *Analyse :* Soit $x \in \mathbb{Z}[\mathrm{i}]$ un élément inversible d'inverse $y \in \mathbb{Z}[\mathrm{i}]$. On écrit comme au-dessus $x = a + \mathrm{i}b$ et $y = c + \mathrm{i}d$ et l'on étudie l'égalité $xy = 1$.

méthode

‖ On obtient une relation remarquablement simple en considérant le module.

En calculant le carré du module, on obtient l'identité $(a^2+b^2)(c^2+d^2) = 1$. Les nombres a,b,c,d sont entiers et la seule façon d'écrire 1 comme le produit de deux naturels est $1 = 1 \times 1$. On peut alors affirmer $a^2 + b^2 = 1$ et donc

$$(a,b) = (1,0), \quad (-1,0), \quad (0,1) \quad \text{ou} \quad (0,-1).$$

L'élément x est alors 1, -1, i ou $-\mathrm{i}$.

Synthèse : Les éléments listés ci-dessus sont effectivement inversibles dans $\mathbb{Z}[\mathrm{i}]$.

$$\mathbb{Z}[\mathrm{i}]^\times = \{1, -1, \mathrm{i}, -\mathrm{i}\} = \mathbb{U}_4.$$

(c) **méthode**

‖ On introduit q élément de $\mathbb{Z}[\mathrm{i}]$ proche du complexe u/v.

Le quotient u/v détermine un complexe $x + \mathrm{i}y$ avec x et y réels. Considérons alors a et b des entiers au plus proche[1] de x et y et posons $q = a + \mathrm{i}b \in \mathbb{Z}[\mathrm{i}]$. On a $|x - a| \leqslant \frac{1}{2}$ et $|y - b| \leqslant \frac{1}{2}$ donc

$$\left|\frac{u}{v} - q\right| = \sqrt{(x-a)^2 + (y-b)^2} \leqslant \frac{1}{2}.$$

En posant $r = u - qv$, on a $q, r \in \mathbb{Z}[\mathrm{i}]$ et $|r| \leqslant \frac{1}{2}|v| < |v|$.

(d) Pour chaque $v \in \mathbb{Z}[\mathrm{i}]$, l'ensemble $v\mathbb{Z}[\mathrm{i}]$ est un idéal, plus précisément, c'est l'idéal engendré par v (Th. 6 p. 39).

Inversement, soit I un idéal de $\mathbb{Z}[\mathrm{i}]$. Si I est réduit à $\{0\}$, on peut écrire $I = v\mathbb{Z}[\mathrm{i}]$ avec $v = 0$. Sinon, il existe[2] dans I un élément v de module minimal parmi les éléments non nuls de I :

$$v \in I \quad \text{et} \quad \forall u \in I,\ u \neq 0 \implies |u| \geqslant |v|. \qquad (*)$$

Vérifions alors par double inclusion $I = v\mathbb{Z}[\mathrm{i}]$.

Puisque v est élément de l'idéal I, on sait déjà l'inclusion $v\mathbb{Z}[\mathrm{i}] \subset I$. Inversement, soit u un élément de I. On peut écrire $u = qv + r$ avec $q, r \in \mathbb{Z}[\mathrm{i}]$ et $|r| < |v|$. Par opérations dans l'idéal I, $r = u - qv$ est élément de I. Puisque $|r| < |v|$, la propriété $(*)$ assure que r est nul. On a donc $u = qv \in v\mathbb{Z}[\mathrm{i}]$ ce qui produit la deuxième inclusion et permet de conclure à l'égalité.

Exercice 12 ***

On se propose d'établir une correspondance bijective entre l'ensemble des sous-anneaux de l'anneau $(\mathbb{Q}, +, \times)$ et l'ensemble $\wp(\mathcal{P})$ des parties de l'ensemble $\mathcal{P}$ des nombres premiers. Pour A un sous-anneau de $(\mathbb{Q}, +, \times)$, on note

$$\mathcal{P}_A = \left\{p \in \mathcal{P} \,\middle|\, \frac{1}{p} \in A\right\}.$$

(a) Soit A et B deux sous-anneaux de $(\mathbb{Q}, +, \times)$. Établir

$$\mathcal{P}_A = \mathcal{P}_B \implies A = B.$$

(b) Soit $P \subset \mathcal{P}$. Déterminer un sous-anneau A de $(\mathbb{Q}, +, \times)$ vérifiant $\mathcal{P}_A = P$.

1. Pour $n \in \mathbb{Z}$, si $x \in [n\,;n + 1/2[$, on choisit $a = n$ et si $x \in]n + 1/2\,;n + 1[$, on choisit $a = n + 1$. Dans le cas restant $x = n + 1/2$, on prend l'une ou l'autre des deux valeurs précédentes. Dans tous les cas, $a = \lfloor x + 1/2 \rfloor$ convient. On procède de même pour y.

2. L'ensemble des $|u|^2$ pour u parcourant $I \setminus \{0\}$ est une partie non vide de $\mathbb{N}$: elle possède une valeur minimale et v est un élément réalisant celle-ci.

Solution

(a) Supposons $\mathcal{P}_A = \mathcal{P}_B$. Commençons par noter que les anneaux A et B contiennent chacun $\mathbb{Z} = \langle 1 \rangle$ car il s'agit de sous-groupes additifs contenant 1.

Soit $x \in A$ de représentant irréductible a/b.

méthode

‖ On montre que $1/b$, puis $1/p$ avec p facteur premier de b, appartiennent à A.

Puisque a et b sont premiers entre eux, il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 1$. Par opérations dans le sous-anneau A

$$\frac{1}{b} = \frac{au + bv}{b} = u \times \frac{a}{b} + v \in A \quad \text{car} \quad u, v, \frac{a}{b} \in A.$$

Soit p un diviseur premier de b. On peut écrire $b = pk$ avec $k \in \mathbb{N}^*$ et alors

$$\frac{1}{p} = k \times \frac{1}{b} \in A.$$

Par suite, les diviseurs premiers de b sont tous éléments de $\mathcal{P}_A$. Or $\mathcal{P}_A = \mathcal{P}_B$ et les inverses des diviseurs premiers de b sont aussi éléments de B. Puisque B est stable par produit, l'élément $1/b$ appartient à B et, finalement, $x = a \times \frac{1}{b}$ aussi.

Ainsi, A est inclus dans B. Par un raisonnement symétrique, on obtient l'inclusion réciproque donc l'égalité.

(b) Considérons

$$A = \left\{ \frac{a}{b} \;\middle|\; \text{les diviseurs premiers de } b \text{ sont éléments de } P \right\}.$$

L'ensemble A est une partie de $\mathbb{Q}$ contenant 1.

Soit x et y deux éléments de A que l'on écrit a/b et c/d. On a

$$x - y = \frac{ad - bc}{bd} \quad \text{et} \quad xy = \frac{ac}{bd}$$

avec bd dont les diviseurs premiers divisent b ou d et sont donc éléments de P. Ceci montre que A est stable par différence et produit et c'est donc un sous-anneau de $\mathbb{Q}$. De plus, l'inverse d'un nombre premier p est élément de A si, et seulement si, $p \in P$. Autrement dit, $\mathcal{P}_A = P$.

Finalement, l'application qui à un anneau A associe $\mathcal{P}_A$ est une bijection entre l'ensemble des sous-anneaux de $(\mathbb{Q}, +, \times)$ et l'ensemble des parties de l'ensemble des nombres premiers.

2.5.2 Idéaux

Exercice 13 * (Description des idéaux de $\mathbb{Z}^2$)
Soit I un idéal de l'anneau produit $(\mathbb{Z}^2, +, \times)$. On introduit

$$I_1 = \{x \in \mathbb{Z} \mid (x, 0) \in I\} \quad \text{et} \quad I_2 = \{y \in \mathbb{Z} \mid (0, y) \in I\}.$$

(a) Montrer que I_1 et I_2 sont des idéaux de $(\mathbb{Z}, +, \times)$.

(b) Établir $I = I_1 \times I_2$.

(c) Conclure que les idéaux de l'anneau $(\mathbb{Z}^2, +, \times)$ sont de la forme $x\mathbb{Z}^2$ avec $x \in \mathbb{Z}^2$.

Solution

(a) I_1 est une partie non vide de $\mathbb{Z}$ car 0 en est élément puisque $(0, 0) = 0_{\mathbb{Z}^2}$ appartient à I. Soit x et $x' \in I_1$. On a $x + x' \in I_1$ car $(x + x', 0) = (x, 0) + (x', 0) \in I$ par addition de deux éléments de I. Soit de plus $a \in \mathbb{Z}$. On a $ax \in I_1$ car $(ax, 0) = (a, 0) \times (x, 0) \in I$ par la propriété d'absorption de l'idéal I. Ainsi, I_1 est un idéal de $\mathbb{Z}$. De façon analogue on établit que I_2 est aussi est un idéal de $\mathbb{Z}$.

(b) Raisonnons par double inclusion.
Soit $(x, y) \in I_1 \times I_2$. On a $(x, 0) \in I$ et $(0, y) \in I$ donc $(x, y) = (x, 0) + (0, y) \in I$. Ainsi, on obtient une première inclusion $I_1 \times I_2 \subset I$.
Inversement, soit $(x, y) \in I$. Par absorption, $(x, 0) = (1, 0) \times (x, y) \in I$ donc $x \in I_1$. De même $y \in I_2$ et donc $(x, y) \in I_1 \times I_2$. Ainsi, $I \subset I_1 \times I_2$ puis $I = I_1 \times I_2$.

(c) **méthode**

‖ Les idéaux de $(\mathbb{Z}, +, \times)$ sont de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$ (Th. 7 p. 39).

On peut introduire a et b naturels tels que $I_1 = a\mathbb{Z}$ et $I_2 = b\mathbb{Z}$. L'idéal I apparaît alors comme étant celui engendré par $x = (a, b)$:

$$I = a\mathbb{Z} \times b\mathbb{Z} = \{(ak, b\ell) \mid (k, \ell) \in \mathbb{Z}^2\} = x\mathbb{Z}^2.$$

Exercice 14 **
Un idéal d'un anneau $(A, +, \times)$ est dit *principal* lorsqu'il est de la forme xA pour un certain $x \in A$. Montrer que les idéaux de tous les sous-anneaux de $\mathbb{Q}$ sont principaux.

Solution
Soit I un idéal d'un sous-anneau A de $(\mathbb{Q}, +, \times)$.

méthode

‖ On détermine x tel que $I = xA$ en étudiant $I \cap \mathbb{Z}$.

Par intersection de sous-groupes, $I \cap \mathbb{Z}$ est un sous-groupe de $(\mathbb{Z}, +)$, il est donc de la forme $x\mathbb{Z}$ pour un certain $x \in \mathbb{N}$ (Th. 4 p. 5). Vérifions alors que I est l'idéal engendré par x. Puisque $x \in I$, on sait déjà $xA \subset I$ (Th. 6 p. 39). Reste à établir l'inclusion inverse.

Soit $r \in I$. On peut écrire $r = p/q$ avec $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$ premiers entre eux. On a alors $qr = p \in \mathbb{Z}$ et, par addition, $qr = r + \cdots + r \in I$ (somme à q termes). Ainsi, qr est élément de $I \cap \mathbb{Z} = x\mathbb{Z}$ et l'on peut écrire $r = xk/q$ pour un certain $k \in \mathbb{Z}$.

Pour conclure $r \in xA$, il suffit d'établir que k/q est élément du sous-anneau A. Sachant les entiers p et q premiers entre eux, on peut écrire la relation de Bézout $pu + qv = 1$ avec $u, v \in \mathbb{Z}$. On a alors

$$\frac{k}{q} = ku\frac{p}{q} + kv = kur + kv1.$$

Les éléments r et 1 appartiennent à A et donc k/q appartient aussi à A par opérations dans le groupe $(A, +)$. On peut alors conclure $r \in xA$ et, finalement, $I = xA$ par double inclusion.

Exercice 15 ** (Idéaux premiers)

Un idéal I d'un anneau commutatif $(A, +, \times)$ est dit *premier* lorsque

$$\forall (x, y) \in A^2, \quad xy \in I \implies x \in I \text{ ou } y \in I.$$

(a) Déterminer les idéaux premiers de $\mathbb{Z}$.

(b) On suppose que A un anneau commutatif non réduit à $\{0_A\}$ dont tout idéal est premier. Établir que A est intègre puis que A est un corps.

Solution

(a) Les idéaux de l'anneau $(\mathbb{Z}, +, \times)$ sont les $n\mathbb{Z}$ avec $n \in \mathbb{N}$ (Th. 7 p. 39).
Si $n = 0$, $n\mathbb{Z} = \{0\}$ est un idéal premier. Si $n = 1$, $n\mathbb{Z} = \mathbb{Z}$ est aussi un idéal premier.
Si $n \geqslant 2$ est un nombre premier, le lemme d'Euclide donne

$$\forall (x, y) \in \mathbb{Z}^2, \quad n \mid xy \implies n \mid x \text{ ou } n \mid y.$$

L'idéal $n\mathbb{Z}$ est alors premier.

Enfin, si $n \geqslant 2$ est un nombre composé, on peut écrire $n = ab$ avec $1 < a, b < n$. Dans ce cas, $ab \in n\mathbb{Z}$ alors que $a \notin n\mathbb{Z}$ et $b \notin n\mathbb{Z}$. L'idéal $n\mathbb{Z}$ n'est alors pas premier.

(b) $I = \{0_A\}$ est un idéal. Affirmer qu'il est premier donne la propriété

$$\forall (x, y) \in A^2, \quad xy = 0_A \implies x = 0_A \text{ ou } y = 0_A.$$

On en déduit que l'anneau A est intègre.

méthode

‖ On montre que $x \in A \setminus \{0_A\}$ est inversible en considérant l'idéal x^2A.

Soit $x \in A$ tel que $x \neq 0_A$. L'idéal x^2A est premier et $x^2 = x \times x$ en est un élément. On en déduit $x \in x^2A$ et ainsi, il existe $y \in A$ tel que $x = x^2y$. Puisque $x \neq 0_A$ et puisque A est intègre, on peut simplifier par x et affirmer $xy = 1_A$. L'élément x est inversible et l'on peut conclure que A est un corps.

Exercice 16 ** (Radical d'un idéal)

Soit I un idéal d'un anneau commutatif A. On appelle *radical*[1] de l'idéal I l'ensemble $R(I)$ des éléments x de A pour lesquels il existe $q \in \mathbb{N}^*$ tel que $x^q \in I$.

(a) Montrer que $R(I)$ est un idéal de A contenant I.

(b) Soit I et J deux idéaux. Vérifier

$$R(I \cap J) = R(I) \cap R(J)$$

(c) On suppose que $A = \mathbb{Z}$. Déterminer le radical de $n\mathbb{Z}$ pour $n \in \mathbb{N}$.

Solution

(a) L'ensemble $R(I)$ est par définition une partie de A et celle-ci est non vide car elle contient I. En effet, pour tout $x \in I$, on a $x^1 = x \in I$ et donc $x \in R(I)$.

Soit x et y deux éléments de $R(I)$. Il existe $q, r \in \mathbb{N}^*$ tels que $x^q \in I$ et $y^r \in I$.

méthode

On détermine un exposant n suffisamment grand pour que le développement de $(x+y)^n$ fasse apparaître une somme d'éléments de I, soit parce que l'exposant de x est supérieur à q, soit parce que l'exposant de y est supérieur à r.

Pour $n = q + r - 1$, la formule du binôme permet d'écrire

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{q-1} \binom{n}{k} x^k \underbrace{y^{n-k}}_{\in I} + \sum_{k=q}^{n} \binom{n}{k} \underbrace{x^k}_{\in I} y^{n-k} \in I.$$

En effet, pour la seconde somme, on a $k \geqslant q$ et donc[2] $x^k = x^q x^{k-q} \in I$ tandis que pour la première somme, on a $k \leqslant q-1$ donc $n-k \geqslant r$ puis $y^{n-k} \in I$. Ainsi, on peut affirmer que $x+y$ est élément de $R(I)$.

Enfin, considérons de plus $a \in A$. On a $(ax)^q = a^q x^q \in I$ et donc $ax \in R(I)$. On peut alors conclure que $R(I)$ est un idéal de A.

(b) Si $x \in R(I \cap J)$, il existe $q \in \mathbb{N}^*$ tels que $x^q \in I \cap J$. On a alors $x^q \in I$ donc $x \in R(I)$ et de même $x \in R(J)$. Ainsi, $R(I \cap J) \subset R(I) \cap R(J)$. Inversement, soit $x \in R(I) \cap R(J)$. Il existe $q, r \in \mathbb{N}^*$ tel que $x^q \in I$ et $x^r \in J$. Considérons alors $n = \max(q, r)$. Par la propriété d'absorption, on a à la fois $x^n \in I$ et $x^n \in J$ donc $x^n \in I \cap J$. Ainsi, l'élément x appartient à $R(I \cap J)$ et l'on peut affirmer l'inclusion réciproque $R(I \cap J) \supset R(I) \cap R(J)$ puis l'égalité.

(c) **méthode**

Les éléments de $R(n\mathbb{Z})$ ont les mêmes facteurs premiers que n.

1. Lorsque $I = \{0\}$, le radical de I regroupe les éléments nilpotents de l'anneau A.
2. L'écriture $x^k = x^q x^{k-q}$ est possible car les exposants sont positifs.

On a immédiatement $R(0\mathbb{Z}) = \{0\}$ et $R(\mathbb{Z}) = \mathbb{Z}$. Supposons désormais $n \geqslant 2$ et écrivons sa décomposition en facteurs premiers

$$n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$$

avec $p_1, \dots, p_r$ nombres premiers deux à deux distincts et $\alpha_1, \dots, \alpha_r \in \mathbb{N}^*$.

Soit $x \in R(n\mathbb{Z})$. Il existe $q \in \mathbb{N}^*$ tel que $x^q \in n\mathbb{Z}$. Les facteurs premiers p_i de n divise alors x^q et donc divise x. L'entier x s'écrit alors

$$x = p_1 \dots p_r k \quad \text{avec} \quad k \in \mathbb{Z}.$$

Inversement, pour un entier de cette forme, on a $x^\alpha \in n\mathbb{Z}$ pour $\alpha = \max(\alpha_1, \dots, \alpha_r)$ et donc $x \in R(n\mathbb{Z})$. En résumé, $R(n\mathbb{Z}) = m\mathbb{Z}$ avec $m = p_1 \dots p_r$.

Exercice 17 ***

Soit p un nombre premier. On note Z_p l'ensemble des nombres rationnels dont le dénominateur n'est pas divisible par p.

(a) Vérifier que Z_p est un sous-anneau de $(\mathbb{Q}, +, \times)$. Quels en sont les éléments inversibles ?

On introduit J l'ensemble des éléments non inversibles de Z_p.

(b) Montrer que J est un idéal de Z_p. Que dire d'un idéal contenant J et distinct de J ?

(c) Déterminer tous les idéaux de Z_p.

Solution

(a) Z_p est une partie de $\mathbb{Q}$ contenant l'élément 1. Soit x et y deux éléments de Z_p. On peut écrire $x = a/b$ et $y = c/d$ avec $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{N}^*$ et p ne divisant ni b ni d. On a alors

$$x - y = \frac{ad - bc}{bd} \in Z_p \quad \text{et} \quad xy = \frac{ac}{bd} \in Z_p$$

car le nombre premier p ne divise pas le produit bd puisqu'il ne divise aucun des deux facteurs. Finalement, Z_p est un sous-anneau [1] de $(\mathbb{Q}, +, \times)$.

Tout élément non nul de Z_p est inversible dans le corps $\mathbb{Q}$. Il s'agit alors de déterminer ceux dont l'inverse est dans Z_p : ce sont les a/b tels que p ne divise ni a ni b.

(b) Les éléments de J sont les a/b tels que p divise a mais pas b. L'ensemble J est une partie non vide de Z_p. Pour x et y deux éléments J, on écrit $x = a/b$ et $y = c/d$ avec p divisant a et c mais ne divisant ni b ni d. On a alors

$$x + y = \frac{ad + bc}{bd} \in J \quad \text{car} \quad p \mid (ad + bc) \text{ et } p \nmid bd.$$

1. Cette étude est un cas particulier de celle du sujet 12 p. 50.

Si de plus $z = e/f$ désigne un élément de Z_p

$$zx = \frac{ae}{bf} \in J \quad \text{car} \quad p \mid ae \text{ et } p \not\mid bf.$$

L'ensemble J est donc un idéal de Z_p. En fait, J est l'idéal engendré[1] par p.

De plus, si I est un idéal contenant J et distinct de J, il contient un élément inversible de l'anneau et est donc égal[2] à Z_p.

(c) Pour $k \in \mathbb{N}$, introduisons J_k l'idéal engendré par p^k. En particulier, $J_0 = Z_p$ et $J_1 = J$. On convient aussi de noter J_∞ l'idéal nul. Montrons qu'il n'existe pas d'autres idéaux dans Z_p que ceux-ci. Soit I un idéal de Z_p.

méthode

On introduit $k = \max\{\ell \in \mathbb{N} \mid I \subset J_\ell\}$.

Supposons que l'ensemble des $\ell \in \mathbb{N}$ tels que $I \subset J_\ell$ est infini. Pour $x = a/b \in I$, il existe une infinité de ℓ tel que p^ℓ divise a et donc $a = 0$. Dans ce cas, l'idéal I est $\{0\}$ c'est-à-dire J_∞.

Sinon, l'ensemble des $\ell \in \mathbb{N}$ tels que $I \subset J_\ell$ est fini et possède donc un plus grand élément k. Pour celui-ci $I \subset J_k$ et $I \not\subset J_{k+1}$. En particulier, il existe dans I un élément x qui appartient à J_k mais pas à J_{k+1}. Cet élément s'écrit $x = a/b$ avec $a = p^k c$ et p ne divisant ni b, ni c. L'élément c/b est inversible dans Z_p et, si l'on introduit y son inverse, on obtient que $p^k = xy$ est élément de l'idéal I. On en déduit que J_k est inclus dans I et donc égal à I.

Finalement, les idéaux de Z_p sont les J_k avec $k \in \mathbb{N} \cup \{\infty\}$.

2.5.3 Calculs dans $\mathbb{Z}/p\mathbb{Z}$

Exercice 18 **

Soit p un nombre premier et k un entier naturel premier avec $p - 1$.

Montrer que l'application $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ définie par $\varphi(x) = x^k$ est bijective.

Solution

méthode

On exploite le petit théorème de Fermat pour exprimer une bijection réciproque à l'application φ.

Les entiers k et $p - 1$ étant premiers entre eux, on peut exprimer une relation[3] de Bézout $ku - (p-1)v = 1$ avec u et v entiers relatifs. Considérons ensuite l'applica-

1. Les idéaux des sous-anneaux de $\mathbb{Q}$ sont principaux : voir sujet 14 p. 52.
2. Tout idéal contenant un élément inversible est égal à l'anneau : voir sujet 6 p. 44.
3. Par commodité, on écrit celle-ci avec un signe moins en passant v à l'opposé.

tion $\psi\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ définie par [1]

$$\psi(x) = \begin{cases} x^u & \text{si } x \neq \overline{0} \\ 0 & \text{si } x = \overline{0}. \end{cases}$$

Soit x un élément de $\mathbb{Z}/p\mathbb{Z}$.

Cas : $x \neq \overline{0}$. On a

$$\varphi\big(\psi(x)\big) = \big(x^u\big)^k = x^{ku} = x \times x^{(p-1)v}.$$

Par le petit théorème de Fermat, on sait $x^{p-1} = \overline{1}$ et l'on en déduit $\varphi\big(\psi(x)\big) = x$.

Cas : $x = \overline{0}$. On vérifie immédiatement $\varphi\big(\psi(x)\big) = \overline{0} = x$.

Ainsi, $\varphi \circ \psi = \mathrm{Id}_{\mathbf{Z}/p\mathbf{Z}}$. L'application φ opérant de l'ensemble fini $\mathbb{Z}/p\mathbb{Z}$ vers lui-même, c'est une bijection et ψ est son application réciproque.

Exercice 19 ** (Théorème de Wilson)

Soit p un nombre premier.

(a) Quels sont les éléments de $\mathbb{Z}/p\mathbb{Z}$ qui sont égaux à leurs inverses ?

(b) En déduire que p divise $(p-1)! + 1$.

(c) Inversement, montrer que si un entier n supérieur à 2 divise $(n-1)! + 1$ alors celui-ci est premier.

Solution

(a) Dans $\mathbb{Z}/p\mathbb{Z}$, la condition $x = x^{-1}$ équivaut à l'équation $x^2 = \overline{1}$ qui s'écrit encore $(x - \overline{1})(x + \overline{1}) = \overline{0}$. Or $\mathbb{Z}/p\mathbb{Z}$ est un corps, il est donc intègre et cette équation a pour seules solutions $\overline{1}$ et $-\overline{1} = \overline{p-1}$.

(b) **méthode**

Dans le produit $(p-1)! = 1 \times 2 \times \cdots \times (p-1)$ on regroupe chaque facteur avec son inverse dans $\mathbb{Z}/p\mathbb{Z}$.

Lorsque $\overline{k}$ est différent de son inverse $\overline{\ell} = \overline{k}^{-1}$, ces deux facteurs se simplifient [2] dans le produit $\overline{1} \times \overline{2} \times \cdots \times \overline{p-1}$. Une fois ces simplifications réalisées, il ne reste dans le produit que les facteurs égaux à leur inverse, à savoir $\overline{1}$ et $\overline{p-1}$. On en déduit

$$\overline{(p-1)!} = \overline{1} \times \overline{p-1} = \overline{p-1} \quad \text{donc} \quad \overline{(p-1)! + 1} = \overline{0}.$$

(c) On suppose $n \geqslant 2$ et n diviseur de $(n-1)! + 1$.

méthode

On étudie les diviseurs de n.

1. Dans cette définition, on distingue le cas $x = \overline{0}$ pour la situation où l'exposant u serait négatif. Une alternative est aussi de remplacer u par $u + q(p-1)$ avec $q \in \mathbb{N}$ assez grand.

2. Dans $\mathbb{Z}/7\mathbb{Z}$, $\overline{6}! = \overline{1} \times \overline{2} \times \overline{3} \times \overline{4} \times \overline{5} \times \overline{6}$. Dans ce produit $\overline{1}$ et $\overline{6}$ sont égaux à leurs inverses tandis que $\overline{2}$ se simplifie avec $\overline{4}$ et $\overline{3}$ avec $\overline{5}$.

Soit d un diviseur positif de n différent de n. Cet entier d figure parmi les facteurs constituant $(n-1)!$ et divise donc ce nombre. Aussi, d divise n et donc divise $(n-1)!+1$. Par différence d divise 1 et donc $d = 1$. L'entier n n'est donc divisible que par 1 et lui-même, il est premier [1].

Exercice 20 ** (Sommes de Newton dans $\mathbb{Z}/p\mathbb{Z}$)

Soit p un entier premier. On admet que le groupe des inversibles du corps $\mathbb{Z}/p\mathbb{Z}$ est cyclique [2]. En discutant selon la valeur de $k \in \mathbb{N}$, calculer

$$S_k = \sum_{x \in \mathbb{Z}/p\mathbb{Z}} x^k.$$

Solution

méthode

‖ On réordonne les termes de S_k par une permutation $x \mapsto ax$ avec $a \neq \overline{0}$.

Considérons a un élément non nul de $\mathbb{Z}/p\mathbb{Z}$. Cet élément étant inversible, l'application $x \mapsto ax$ est une permutation sur l'ensemble $\mathbb{Z}/p\mathbb{Z}$.

Les termes sommés étant identiques

$$S_k = \sum_{x \in \mathbb{Z}/p\mathbb{Z}} (ax)^k = \sum_{x \in \mathbb{Z}/p\mathbb{Z}} a^k x^k = a^k \sum_{x \in \mathbb{Z}/p\mathbb{Z}} x^k.$$

On obtient donc la relation $S_k = a^k S_k$. Ceci amène à discuter selon que a^k vaut $\overline{1}$ ou non.

Cas : $k \equiv 0\ [p-1]$. Le petit théorème de Fermat assure que $a^k = \overline{1}$ pour tout a non nul de $\mathbb{Z}/p\mathbb{Z}$. Dans ce cas un calcul direct de S_k est possible :

$$S_k = \overline{0} + \sum_{k \in \mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}} \overline{1} = \overline{p-1}.$$

Cas : $k \not\equiv 0\ [p-1]$. Un générateur du groupe cyclique des inversibles de $\mathbb{Z}/p\mathbb{Z}$ définit [3] un élément a pour lequel $a^k \neq \overline{1}$. Dans ce cas $S_k = \overline{0}$.

On peut retrouver ce résultat en introduisant de nouveau a un générateur du groupe des inversibles de $\mathbb{Z}/p\mathbb{Z}$ et en menant le calcul d'une somme géométrique :

$$S_k = \overline{0} + \sum_{i=1}^{p-1} \left(a^i\right)^k = \sum_{i=1}^{p-1} \left(a^k\right)^i.$$

Ceci conduit de nouveau à discuter selon que la raison a^k est égale à $\overline{1}$ ou non.

1. Ce résultat n'est pas un bon test de primalité car le calcul de $(n-1)!$ est coûteux.
2. Plus généralement, le groupe des inversibles d'un corps fini est cyclique : voir sujet 29 p. 65.
3. Le groupe des inversibles de $\mathbb{Z}/p\mathbb{Z}$ comporte $p-1$ éléments, un générateur de celui-ci est donc un élément d'ordre exactement $p-1$ et sa puissance k-ième n'est alors pas égale au neutre.

Exercice 21 ***

Soit p un nombre premier supérieur à 3.

(a) Quel est le nombre de carrés dans $\mathbb{Z}/p\mathbb{Z}$?

(b) On suppose $p \equiv 1$ [4]. Justifier que $\overline{-1}$ est un carré dans $\mathbb{Z}/p\mathbb{Z}$ en calculant de deux façons la classe de congruence de $(p-1)!$.

(c) On suppose $p \equiv 3$ [4]. Montrer que $\overline{-1}$ n'est pas un carré dans $\mathbb{Z}/p\mathbb{Z}$.

Solution

(a) **méthode**

> En calculant les carrés dans $\mathbb{Z}/p\mathbb{Z}$ pour de petites valeurs de p, on observe que, en dehors de $\overline{0}$, chaque carré possède exactement deux antécédents.

Considérons l'application $\varphi\colon x \mapsto x^2$ définie de $\mathbb{Z}/p\mathbb{Z}$ vers lui-même. Le nombre de carrés dans $\mathbb{Z}/p\mathbb{Z}$ correspond au cardinal de l'image de φ. Dénombrons l'image de φ en étudiant les antécédents des valeurs prises par cette application.

Soit $x, y \in \mathbb{Z}/p\mathbb{Z}$. Dans le corps $\mathbb{Z}/p\mathbb{Z}$

$$\begin{aligned}\varphi(x) = \varphi(y) &\iff (x-y)(x+y) = \overline{0}\\ &\iff x = y \text{ ou } x = -y.\end{aligned}$$

Dans $\mathrm{Im}(\varphi)$, la valeur $\overline{0}$ possède un seul antécédent, les autres éléments possèdent deux antécédents distincts[1]. On en déduit

$$\mathrm{Card}(\mathbb{Z}/p\mathbb{Z}) = 1 + 2\Big(\mathrm{Card}\big(\mathrm{Im}(\varphi)\big) - 1\Big).$$

Il y a donc exactement $\frac{p+1}{2}$ carrés dans $\mathbb{Z}/p\mathbb{Z}$.

(b) Comme détaillé dans le sujet 19 p. 57, $\overline{(p-1)!} = \overline{-1}$ dans $\mathbb{Z}/p\mathbb{Z}$. Au surplus, en posant $n = \frac{p-1}{2}$, on peut séparer le produit exprimant la factorielle en son milieu :

$$\begin{aligned}\overline{(p-1)!} &= \overline{1} \times \cdots \times \overline{n} \times \overline{(n+1)} \times \cdots \times \overline{(p-1)}\\ &= \overline{1} \times \cdots \times \overline{n} \times \overline{(-n)} \times \cdots \times \overline{(-1)} = \overline{(-1)^n (n!)^2}.\end{aligned}$$

Or $p \equiv 1$ [4] donc n est pair et $\overline{-1} = \overline{(p-1)!} = \overline{(n!)}^2$ est un carré dans $\mathbb{Z}/p\mathbb{Z}$.

(c) Si $\overline{-1}$ est un carré de $\mathbb{Z}/p\mathbb{Z}$, l'application $\varphi : x \mapsto -x$ définit une involution sur l'ensemble des carrés de $\mathbb{Z}/p\mathbb{Z}$. En effet, on a $\varphi \circ \varphi = \mathrm{Id}_{\mathbb{Z}/p\mathbb{Z}}$ et un carré est transformé en un carré car le produit de deux carrés est un carré.

Puisque $\overline{0}$ est le seul point fixe de cette application, on peut affirmer qu'il y a un nombre impair de carrés dans $\mathbb{Z}/p\mathbb{Z}$. Or si $p \equiv 3$ [4], $(p+1)/2$ est un entier pair : $\overline{-1}$ ne peut donc être un carré[2] dans $\mathbb{Z}/p\mathbb{Z}$.

1. Les éléments y et $-y$ sont distincts car $2y = \overline{0}$ implique $y = \overline{0}$ puisque 2 est inversible dans $\mathbb{Z}/p\mathbb{Z}$. Cette propriété n'est plus vraie quand $p = 2$.

2. On peut retrouver ce résultat à l'aide du petit théorème de Fermat : si l'on peut écrire $\overline{-1} = a^2$

2.5.4 Fonction indicatrice d'Euler

Exercice 22 *

Montrer que, pour tout entier $n \geqslant 3$, $\varphi(n)$ est un nombre pair.

Solution

méthode

‖ On a $\varphi(mn) = \varphi(m)\varphi(n)$ lorsque m et n sont premiers entre eux.

Soit $n \in \mathbb{N}$ avec $n \geqslant 3$.

Cas : n possède un facteur premier impair p. On peut écrire $n = p^\alpha m$ avec m premier avec p et $\alpha \in \mathbb{N}^*$. On a alors

$$\varphi(n) = \varphi(p^\alpha)\varphi(m) = (p^\alpha - p^{\alpha-1})\varphi(m).$$

Puisque $p^\alpha - p^{\alpha-1}$ est la différence de nombres impairs, c'est un entier pair et $\varphi(n)$ est donc un nombre pair.

Cas : n ne possède pas de facteurs premiers impairs. On peut écrire $n = 2^\alpha$ avec $\alpha \geqslant 2$ auquel cas $\varphi(n) = 2^{\alpha-1}$ est encore un nombre pair.

Exercice 23 *

Soit $a \in \mathbb{Z}$ et $n \in \mathbb{N}$ avec $n \geqslant 2$.

(a) On suppose a et n premiers entre eux, montrer $a^{\varphi(n)} \equiv 1\ [n]$.

(b) On suppose $a^{n-1} \equiv 1\ [n]$ et $a^d \not\equiv 1\ [n]$ pour tout entier naturel d diviseur strict de $n-1$. Montrer que n est un nombre premier.

Solution

(a) L'ensemble des inversibles[1] de l'anneau $\mathbb{Z}/n\mathbb{Z}$ est un groupe multiplicatif de cardinal $\varphi(n)$ et $\overline{a}$ en est élément. On a donc $\overline{a}^{\varphi(n)} = \overline{1}$, c'est-à-dire $a^{\varphi(n)} \equiv 1\ [n]$.

(b) **méthode**

‖ On montre que $\varphi(n) = n - 1$ ce qui assure que n est un nombre premier.

Notons que la propriété $a^{n-1} \equiv 1\ [n]$ impose que a et n sont premiers entre eux. Introduisons $d = (n-1) \wedge \varphi(n)$. Par une relation de Bézout, on écrit $d = (n-1)u + \varphi(n)v$ avec u et v entiers. Par calculs[2] dans $\mathbb{Z}/n\mathbb{Z}$

$$\overline{a}^d = \left(\overline{a}^{n-1}\right)^u \left(\overline{a}^{\varphi(n)}\right)^v = \overline{1} \quad \text{donc} \quad a^d \equiv 1\ [n].$$

avec a dans $\mathbb{Z}/p\mathbb{Z}$, on a $(\overline{-1})^{(p-1)/2} = a^{p-1} = \overline{1}$ ce qui entraîne que $(p-1)/2$ est un entier pair car $\overline{-1}$ et $\overline{1}$ sont distincts.

1. On reproduit ici la démonstration du Th. 13 p. 41.

2. L'écriture des puissances u et v non nécessairement positives est possible car $\overline{a}$ est inversible dans l'anneau $\mathbb{Z}/n\mathbb{Z}$.

Compte tenu des hypothèses, l'entier d ne peut être un diviseur strict de $n-1$, il est donc égal à $n-1$. On en déduit que $n-1$ divise $\varphi(n)$ et donc $\varphi(n) = n-1$ car $\varphi(n)$ est toujours inférieur à $n-1$. On peut conclure que n est premier.

Exercice 24 **

Soit n un entier naturel non nul.

(a) Pour d diviseur positif de n, combien y a-t-il de $k \in [\![1\,;n]\!]$ vérifiant $k \wedge n = d$?

(b) En déduire

$$n = \sum_{d|n} \varphi(d)$$

où la somme s'étend sur les diviseurs positifs de n.

Solution

(a) On écrit $n = dn'$ avec $n' \in \mathbb{N}^*$.

méthode

On met en correspondance les entiers k tels que $k \wedge n = d$ avec les entiers k' tels que $k' \wedge n' = 1$.

Soit $k \in [\![1\,;n]\!]$ vérifiant $k \wedge n = d$. L'entier d divisant k, on peut écrire $k = dk'$ avec k' dans $[\![1\,;n']\!]$. De plus, $d = k \wedge n = (dk') \wedge (dn') = d(k' \wedge n')$ donne $k' \wedge n' = 1$.

Inversement, si $k = dk'$ avec $k' \in [\![1\,;n']\!]$ tel que $k' \wedge n' = 1$ alors $k \in [\![1\,;n]\!]$ et

$$k \wedge n = (dk') \wedge (dn') = d.$$

Ainsi, il y a autant de k cherchés que de k' éléments de $[\![1\,;n']\!]$ premiers avec n', à savoir, $\varphi(n') = \varphi(n/d)$.

(b) **méthode**

On dénombre les éléments $[\![1\,;n]\!]$ en discutant selon la valeur que constitue leur PGCD avec n.

Pour chaque k compris entre 1 et n, le PGCD d de k et n est un diviseur de n. On peut alors écrire l'ensemble $[\![1\,;n]\!]$ comme la réunion suivante

$$[\![1\,;n]\!] = \bigcup_{d|n} \left\{ k \in [\![1\,;n]\!] \;\middle|\; k \wedge n = d \right\}$$

où les ensembles sont deux à deux disjoints. On en déduit

$$\operatorname{Card} [\![1\,;n]\!] = \sum_{d|n} \operatorname{Card}\left(\left\{ k \in [\![1\,;n]\!] \;\middle|\; k \wedge n = d \right\} \right) \quad \text{puis} \quad n = \sum_{d|n} \varphi\left(\frac{n}{d}\right).$$

Enfin, lorsque d parcourt les diviseurs positifs de n, l'entier $\delta = n/d$ parcourt aussi cet ensemble ce qui permet de réorganiser la somme et conclure [1]

$$n = \sum_{\delta | n} \varphi(\delta).$$

Exercice 25 **

Soit $a, n \in \mathbb{N}$ au moins égaux à 2 et $N = a^n - 1$. Montrer que n divise $\varphi(N)$.

Solution

méthode

‖ On détermine l'ordre de $\overline{a}$ dans le groupe des inversibles de $\mathbb{Z}/N\mathbb{Z}$.

Par définition de la valeur de N, on peut affirmer

$$a^n \equiv 1 \ [N] \quad \text{et} \quad \forall 1 \leqslant k < n,\ a^k \not\equiv 1 \ [N].$$

On en déduit que $\overline{a}$ est un élément inversible de l'anneau $\mathbb{Z}/N\mathbb{Z}$ et, plus précisément, que $\overline{a}$ est un élément d'ordre exactement n dans le groupe multiplicatif des inversibles de $\mathbb{Z}/N\mathbb{Z}$. Or ce groupe est de cardinal $\varphi(N)$ et, puisque l'ordre des éléments divise le cardinal du groupe, on obtient que n divise $\varphi(N)$.

2.6 Exercices d'approfondissement

Exercice 26 *

Montrer que l'ensemble $\mathcal{S}$ des fonctions de $\mathbb{R}$ vers $\mathbb{C}$ développables en série entière sur $\mathbb{R}$ est un anneau intègre pour les opérations usuelles.

Solution

On montre que $\mathcal{S}$ est un sous-anneau de l'anneau $\left(\mathcal{F}(\mathbb{R}, \mathbb{C}), +, \times\right)$ des fonctions de $\mathbb{R}$ vers $\mathbb{C}$. Par définition, $\mathcal{S}$ est inclus dans $\mathcal{F}(\mathbb{R}, \mathbb{C})$ et il est entendu que la fonction constante égale à 1 est développable en série entière sur $\mathbb{R}$. De plus, la différence et le produit de deux fonctions développables en série entière sur $\mathbb{R}$ l'est aussi [2]. L'ensemble $\mathcal{S}$ est donc un sous-anneau de $\mathcal{F}(\mathbb{R}, \mathbb{C})$ et c'est donc un anneau pour les opérations usuelles. Il reste à montrer que celui-ci est intègre.

Soit f et g deux fonctions non nulles éléments de $\mathcal{S}$. On note $\sum a_n x^n$ et $\sum b_n x^n$ les séries entières associées, chacune de rayon de convergence $+\infty$.

méthode

‖ On introduit les plus petits entiers p et q tels que $a_p \neq 0$ et $b_q \neq 0$.

1. Une autre démonstration très élégante de cette identité est la suivante : parmi les n nombres rationnels $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$, il y en a exactement $\varphi(d)$ dont l'écriture irréductible présente un dénominateur égal à d et ce pour chaque d diviseur de n.

2. Voir section 9.3.1 du chapitre 9 de l'ouvrage *Exercices d'analyse MP*.

La fonction f n'étant pas nulle, la suite (a_n) n'est pas constante égale à 0 et l'on peut introduire le plus petit entier p tel que a_p est non nul. De même, on introduit le plus petit entier q tel que $b_q \neq 0$. Par produit de séries entières, on sait

$$\left(\sum_{n=0}^{+\infty} a_n x^n\right)\left(\sum_{n=0}^{+\infty} b_n x^n\right) = \left(\sum_{n=0}^{+\infty} c_n x^n\right) \quad \text{avec} \quad c_n = \sum_{k+\ell=n} a_k b_\ell.$$

Puisque

$$c_{p+q} = \underbrace{a_0 b_{p+q} + \cdots + a_{p-1} b_{q+1}}_{=0} + a_p b_q + \underbrace{a_{p+1} b_{q-1} + \cdots + a_{p+q} b_0}_{=0} = a_p b_q \neq 0$$

les coefficients de la série entière produit ne sont pas tous nuls. Par unicité des coefficients d'un développement en série entière, on peut affirmer que la fonction fg n'est pas nulle.

Finalement, le produit de deux éléments non nuls de $\mathcal{S}$ est non nul. Au surplus, $\mathcal{S}$ n'est pas réduit à $\{0\}$, c'est donc un anneau intègre.

Exercice 27 * (Déterminant de Smith)

Soit $T = (t_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ la matrice de coefficients

$$t_{i,j} = \begin{cases} 1 & \text{si } i \text{ divise } j \\ 0 & \text{sinon.} \end{cases}$$

Soit aussi $D \in \mathcal{M}_n(\mathbb{R})$ la matrice diagonale de coefficients diagonaux $\varphi(1), \ldots, \varphi(n)$ où φ désigne la fonction indicatrice d'Euler.

(a) Exprimer le coefficient d'indice (i,j) de la matrice ${}^tTDT$ en fonction de $i \wedge j$.

(b) En déduire la valeur du déterminant de la matrice de Smith

$$S = \begin{pmatrix} 1\wedge 1 & 1\wedge 2 & \cdots & 1\wedge n \\ 2\wedge 1 & 2\wedge 2 & \cdots & 2\wedge n \\ \vdots & \vdots & & \vdots \\ n\wedge 1 & n\wedge 2 & \cdots & n\wedge n \end{pmatrix}.$$

Solution

(a) **méthode**

$\| \; n \in \mathbb{N}^*$ est [1] la somme des $\varphi(d)$ pour d divisant n.

Le coefficient d'indice (i,j) du produit DT est $\varphi(i)t_{i,j}$ et celui de la matrice ${}^tTDT$ s'exprime alors

$$\sum_{k=1}^{n} t_{k,i}\varphi(k)t_{k,j} = \sum_{k|i \text{ et } k|j} \varphi(k).$$

1. Voir sujet 24 p. 61.

Or les diviseurs communs à i et j sont exactement les diviseurs de $i \wedge j$ et donc

$$\sum_{k=1}^{n} t_{k,i}\varphi(k)t_{k,j} = \sum_{k|i\wedge j} \varphi(k) = i \wedge j.$$

On a ainsi ${}^tTDT = S$ avec S la matrice introduite à la question suivante.

(b) La matrice T est triangulaire supérieure à coefficients diagonaux égaux à 1, elle est donc de déterminant égal à 1 tout comme sa transposée. On en déduit

$$\det(S) = \det({}^tT)\det(D)\det(T) = \det(D) = \prod_{k=1}^{n} \varphi(k).$$

Exercice 28 **

Déterminer les tables d'opérations sur $\mathbb{F}_4$ corps fini [1] à 4 éléments.

Solution

Hormis 0 et 1, le corps $\mathbb{F}_4$ possède deux éléments a et b. Son groupe des inversibles $\mathbb{F}_4 \setminus \{0\}$ possède 3 éléments 1, a et b, il est donc [2] isomorphe à $(\mathbb{Z}/3\mathbb{Z}, +)$. On en déduit que l'élément b se confond avec a^2 et il est donc facile de former la table de multiplication dans $\mathbb{F}_4$:

$\times$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

méthode

Pour déterminer la table d'addition, on observe $1 + 1 = 0$ dans $\mathbb{F}_4$.

$(\mathbb{F}_4, +)$ est un groupe à 4 éléments et donc $4\mathbb{1} = 1 + 1 + 1 + 1 = 0$ (Th. 14 p. 8). Par développement $(1 + 1)(1 + 1) = 4\mathbb{1}$ et donc $(1 + 1)^2 = 0$. Tout corps étant intègre, on en déduit $1 + 1 = 0$. Il en découle $a + a = (1 + 1)a = 0$ et de même $b + b = 0$. On peut alors former une première table d'addition que l'on complète

$+$	0	1	a	b
0	0	1	a	b
1	1	0		
a	a		0	
b	b			0

puis

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

En effet, $1 + a$ ne peut valoir 1, ni a, ni encore 0, on a donc $1 + a = b$. On complète les autres valeurs sachant que, sur chaque rangée, figurent tous les éléments du groupe une fois et une seule.

1. Un théorème hors-programme assure que, à isomorphisme près, il existe un unique corps de cardinal p^n pour tout nombre premier p et tout $n \in \mathbb{N}^*$. Au surplus, un corps fini a nécessairement un cardinal de cette forme.

2. Voir le sujet 6 p. 12, ou plus généralement, le sujet suivant.

Exercice 29 * (Groupe des inversibles d'un corps fini)**

On étudie le groupe $(\mathbb{F}^\times, \times)$ des inversibles d'un corps fini $\mathbb{F}$.

(a) Soit x un élément de $\mathbb{F}^\times$ d'ordre $d \in \mathbb{N}^*$ et y un élément de $\mathbb{F}^\times$ vérifiant $y^d = 1_\mathbb{F}$. Montrer que y appartient au groupe engendré par x.

On admettra que l'on peut étendre[1] *la théorie des polynômes à ceux dont les coefficients appartiennent au corps* $\mathbb{F}$.

(b) Pour $d \in \mathbb{N}^*$, on note $N(d)$ le nombre d'éléments d'ordre d dans $\mathbb{F}^\times$. Justifier $N(d) \leqslant \varphi(d)$ où φ désigne la fonction indicatrice d'Euler.

(c) En déduire que $(\mathbb{F}^\times, \times)$ est un groupe cyclique.

Solution

(a) **méthode**

‖ y est racine du polynôme $X^d - 1_\mathbb{F}$ dont on connaît d racines dans $\mathbb{F}$.

L'élément x étant d'ordre d, les x^k pour $k \in [\![0\,;d-1]\!]$ sont deux à deux distincts et tous racines du polynôme $X^d - 1_\mathbb{F}$. Celui-ci étant de degré d, il ne possède pas d'autres racines. Or l'élément y est racine de ce polynôme, c'est donc une puissance de x, c'est-à-dire un élément du groupe multiplicatif $\langle x \rangle$ engendré par x.

(b) S'il existe un élément x d'ordre d dans $\mathbb{F}^\times$, la résolution ci-dessus assure que tous les autres éléments d'ordre d appartiennent au groupe $\langle x \rangle$. Or celui-ci est cyclique de cardinal d donc isomorphe à $(\mathbb{Z}/d\mathbb{Z}, +)$. Ce dernier groupe possède exactement $\varphi(d)$ éléments d'ordre d. On peut donc affirmer que, si $N(d)$ est non nul, on a $N(d) = \varphi(d)$ et, si $N(d)$ est nul, on a évidemment $N(d) \leqslant \varphi(d)$.

(c) **méthode**

‖ On dénombre $\mathbb{F}^\times$ selon l'ordre de ses éléments.

Posons n le cardinal de $\mathbb{F}^\times$. Les éléments de ce groupe sont d'ordre d divisant n et il y en a exactement $N(d)$. On a donc une première égalité

$$n = \sum_{d|n} N(d).$$

Parallèlement, on sait[2]

$$N(d) \leqslant \varphi(d) \quad \text{et} \quad n = \sum_{d|n} \varphi(d).$$

On a donc $N(d) = \varphi(d)$ pour tout d diviseur de n. En particulier, $N(n) = \varphi(n) \neq 0$ et il existe dans $(\mathbb{F}^\times, \times)$ des éléments d'ordre n : ce groupe est cyclique.

1. En particulier, un polynôme à coefficients dans $\mathbb{F}$ ne peut avoir plus de racines dans $\mathbb{F}$ que son degré.

2. Voir sujet 24 p. 61.

Exercice 30 ***

Soit $(A, +, \times)$ un anneau.

(a) On suppose que 0_A est la seule solution à l'équation $x^2 = 0_A$ d'inconnue $x \in A$. Soit $e \in A$ vérifiant[1] $e^2 = e$. Montrer que e commute avec tout élément de A.

(b) On suppose $x^2 = x$ pour tout $x \in A$. Montrer que l'anneau A est commutatif.

(c) On suppose $x^3 = x$ pour tout $x \in A$. Montrer que $3x + 3x^2$ est nul puis que l'anneau A est commutatif.

(d) On suppose $x^4 = x$ pour tout $x \in A$. Montrer que $2x$ est nul puis que $x + x^2$ commute avec tout y de A. En déduire que x^2 commute avec y et conclure que l'anneau A est commutatif.

Solution

(a) Soit x un élément de l'anneau A.

méthode

On vérifie que ex et xe sont égaux à exe.

On observe

$$(ex - exe)^2 = \big(ex(1-e)\big)^2 = ex\underbrace{(1-e)e}_{=0_A}x(1-e) = 0_A.$$

On en déduit $ex = exe$ car seul 0_A est de carré nul. Un calcul semblable montre $xe = exe$.

Notons que l'hypothèse de cette question est remplie dans chacune des trois études qui suit.

(b) L'hypothèse $x^2 = x$ pour tout $x \in A$ signifie l'idempotence de tous les éléments de A : l'anneau est donc commutatif en vertu de l'étude ci-dessus.

(c) Soit $x \in A$. En développant l'égalité $(1_A + x)^3 = 1_A + x$ et en simplifiant on obtient[2] $3x + 3x^2 = 0_A$.

méthode

Les éléments x^2 et $(1+x)^2$ sont idempotents.

Pour tout $x \in A$, l'élément x^2 est idempotent car $\big(x^2\big)^2 = x^4 = x^3x = x^2$. En remplaçant x par $1_A + x$, on peut aussi affirmer l'idempotence de $(1_A + x)^2$. Ainsi, $1_A + 2x + x^2$ commute avec tout élément de A. Or 1_A et x^2 commutent aussi avec tout élément de A et donc $2x$ commutent avec tout élément de A. Enfin, $3x = -3x^2$ commute avec tout élément de A car x^2 commute avec tout élément de A. On peut alors conclure que $x = 3x - 2x$ commute avec tout élément de A.

1. On dit que l'élément e est *idempotent*. 0_A et 1_A sont des exemples d'éléments idempotents.
2. On ne peut pas *a priori* simplifier cette égalité par 3 : dans $\mathbb{Z}/3\mathbb{Z}$, on a $3x = \overline{0}$ pour tout x !

(d) Soit $x \in A$.

méthode

‖ On applique l'hypothèse $x^4 = x$ à l'élément $-x$.

On a $-x = (-x)^4 = x^4 = x$ et donc $2x = x + x = 0_A$.
L'élément $x + x^2$ commute alors avec tout élément de A car il est idempotent :

$$\left(x + x^2\right)^2 = x^2 + \underbrace{2x^3}_{=0_A} + \underbrace{x^4}_{=x} = x + x^2.$$

Soit $y \in A$. Puisque les éléments de la forme $x+x^2$ commutent avec tout élément de A, on peut affirmer que $(x + y) + (x + y)^2$ commute avec x. Sachant

$$x\big((x + y) + (x + y)^2\big) = x^2 + xy + x^3 + x^2y + xyx + xy^2 \text{ et}$$
$$\big((x + y) + (x + y)^2\big)x = x^2 + yx + x^3 + xyx + yx^2 + y^2x$$

on obtient $x^2y = yx^2$ après simplification et usage de l'identité $x\big(y + y^2\big) = \big(y + y^2\big)x$.
Ainsi, x^2 commute avec tout élément de A puis $x = \big(x + x^2\big) - x$ aussi.

Exercice 31 ***

Soit n un entier supérieur à 3 et U le groupe des inversibles de l'anneau $\mathbb{Z}/2^n\mathbb{Z}$.

(a) Montrer que $a^{2^{n-2}} \equiv 1\ [2^n]$ pour tout entier impair a.

(b) Le groupe $(U, \times)$ est-il cyclique ?

(c) Trouver le plus petit entier $k > 0$ tel que $3^k \equiv 1\ [2^n]$.

(d) Montrer que U est isomorphe au groupe produit $\big(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}, +\big)$.

Solution

(a) Par la factorisation $a^2 - b^2 = (a - b)(a + b)$ on écrit

$$a^{2^{n-2}} - 1 = \big(a^{2^{n-3}} + 1\big)\big(a^{2^{n-3}} - 1\big).$$

En répétant cette factorisation

$$a^{2^{n-2}} - 1 = \big(a^{2^{n-3}} + 1\big)\big(a^{2^{n-4}} + 1\big) \ldots \big(a^{2^0} + 1\big)\big(a^{2^0} - 1\big).$$

Il y a $n - 1$ facteurs dans ce produit et ceux-ci sont tous pairs car a est impair. De plus, les deux derniers facteurs $a + 1$ et $a - 1$ sont des entiers pairs consécutifs : l'un d'eux est divisible par 4. Au total, 2^n divise $a^{2^{n-2}} - 1$ et donc $a^{2^{n-2}} \equiv 1\ [2^n]$.

(b) Les éléments du groupe $U = \big(\mathbb{Z}/2^n\mathbb{Z}\big)^*$ sont les $\overline{k}$ avec $2 \wedge k = 1$, ce sont donc les classes des entiers impairs. Le calcul au-dessus assure que tous ces éléments sont d'ordres inférieurs à 2^{n-2}. Or le groupe U à 2^{n-1} éléments, il n'existe donc pas d'éléments engendrant U : le groupe U n'est pas cyclique.

(c) Puisque 2 et 3 sont premiers entre eux, $\overline{3}$ est inversible dans $\mathbb{Z}/2^n\mathbb{Z}$, il s'agit alors de déterminer l'ordre de $\overline{3}$ dans le groupe U. Puisque $3^{2^{n-2}} \equiv 1\ [2^n]$, cet ordre divise 2^{n-2} et c'est donc une puissance de 2 : on l'écrit 2^p. Par une écriture analogue à celle de la première question

$$3^{2^p} - 1 = \left(3^{2^{p-1}} + 1\right)\left(3^{2^{p-2}} + 1\right) \ldots \left(3^{2^0} + 1\right)\left(3^{2^0} - 1\right)$$

où le produit comporte $p+1$ facteurs en tout. Les deux derniers facteurs sont 4 et 2. Les précédents sont de la forme $3^{2^k} + 1$ avec $k \geqslant 1$.

méthode

‖ Pour $k \geqslant 1$, on vérifie $3^{2^k} \equiv 1\ [4]$.

On a $3^2 = 9 \equiv 1\ [4]$ donc $3^{2^k} = (3^2)^{2^{k-1}} \equiv 1^{2^{k-1}} \equiv 1\ [4]$ et donc $3^{2^k} + 1$ n'est pas divisible par 4. En conséquence, la plus grande puissance de 2 divisant $3^{2^p} - 1$ est 2^{p+2}. Pour que 2^n divise ce nombre, il faut $p \geqslant n - 2$. L'élément 3 est donc d'ordre exactement 2^{n-2} dans U.

(d) **méthode**

‖ $\overline{-1}$ est un élément d'ordre 2 n'appartenant pas au groupe engendré par $\overline{3}$.

L'élément $\overline{-1}$ appartient bien à U et c'est évidemment un élément d'ordre 2. Modulo 8, les puissances de 3 sont égales à 1 ou 3 mais jamais à -1. *A fortiori*, aucune puissance de 3 n'est égale à -1 modulo 2^n (rappelons que n est supérieur à 3).

Considérons ensuite l'application $\varphi\colon \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z} \to U$ définie par

$$\varphi(\dot{k}, \hat{\ell}) = \overline{(-1)^k 3^\ell}$$

en notant $\dot{k}$ et $\hat{\ell}$ les classes d'équivalence dans $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/2^{n-2}\mathbb{Z}$.

L'application φ est bien définie (car $\overline{-1}$ est d'ordre 2 et $\overline{3}$ d'ordre 2^{n-2}) et à valeurs dans U. L'application φ est clairement un morphisme du groupe $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}, +)$ vers $(U, \times)$. Étudions son noyau afin d'établir son injectivité.

Si $(\dot{k}, \hat{\ell})$ appartient à $\mathrm{Ker}(\varphi)$, on a $\overline{(-1)^k} = \overline{3}^{-\ell}$. Or $\overline{-1}$ n'appartient pas au groupe engendré par $\overline{3}$ et donc $k \equiv 0\ [2]$ puis $\overline{3}^\ell = \overline{1}$ ce qui entraîne $\ell \equiv 0\ \left[2^{n-2}\right]$.

Finalement, le noyau de φ est réduit au neutre $(\dot{0}, \hat{0})$ et le morphisme φ est injectif. Au surplus, φ opère entre deux ensembles finis ayant le même cardinal, c'est un isomorphisme de groupes.

CHAPITRE 3

Compléments d'algèbre linéaire

3.1 Extension du cours de première année

Le cours relatif aux polynômes, aux espaces vectoriels, aux applications linéaires, aux matrices et aux déterminants a été exposé en première année dans la situation où le corps $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$. Il se reprend dans les mêmes termes dans le cas plus général où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$.

3.2 Structure d'algèbre

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$.

3.2.1 Définition

Définition

On appelle $\mathbb{K}$-*algèbre* tout quadruplet $(A, +, \times, .)$ formé d'un ensemble A, de deux lois de composition internes $+$ et $\times$ sur A et d'un produit extérieur $(.)$ opérant de $\mathbb{K}$ sur A tels que $(A, +, .)$ est un $\mathbb{K}$-espace vectoriel, $(A, +, \times)$ est un anneau et, pour tout $\lambda \in \mathbb{K}$ et tous x et $y \in A$

$$(\lambda x)y = \lambda(xy) = x(\lambda y).$$

Si de plus la multiplication est commutative, la $\mathbb{K}$-algèbre est dite *commutative*.

Une algèbre est donc la combinaison d'une structure d'espace vectoriel et d'une structure d'anneau avec une propriété de compatibilité calculatoire engageant la multiplication et le produit extérieur.

$\mathbb{K}$, $\mathbb{K}^n$, $\mathbb{K}[X]$ et $\mathcal{F}(X,\mathbb{K})$ sont des $\mathbb{K}$-algèbres commutatives usuelles.
$\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre. Elle est non commutative dès que $n \geqslant 2$.
Si E désigne un $\mathbb{K}$-espace vectoriel, $\mathcal{L}(E)$ est une $\mathbb{K}$-algèbre. Celle-ci est non commutative dès que la dimension de E est supérieure à 2. Dans l'algèbre $\mathcal{L}(E)$ la multiplication correspond à la composition des endomorphismes.

Par restriction du produit extérieur, tout espace vectoriel complexe peut se comprendre comme un espace vectoriel[1] réel. De même, toute algèbre complexe peut s'interpréter comme une algèbre réelle. Plus généralement, si $\mathbb{L}$ est un sous-corps de $\mathbb{K}$, toute $\mathbb{K}$-algèbre est aussi une $\mathbb{L}$-algèbre.

3.2.2 Sous-algèbres

Définition

On appelle *sous-algèbre* d'une $\mathbb{K}$-algèbre $(A,+,\times,.)$ toute partie B de A contenant 1_A et vérifiant $\lambda x \in B$, $x+y \in B$ et $xy \in B$ pour tout $\lambda \in \mathbb{K}$ et tous x et $y \in B$.

Une sous-algèbre est donc à la fois un sous-espace vectoriel et un sous-anneau.

Théorème 1

Si B est une sous-algèbre d'une $\mathbb{K}$-algèbre $(A,+,\times,.)$ alors $(B,+,\times,.)$ est une $\mathbb{K}$-algèbre[2] de mêmes neutres que A.

Si X est un ensemble quelconque, l'ensemble des fonctions bornées de X vers $\mathbb{K}$ est une sous-algèbre de $\mathcal{F}(X,\mathbb{K})$, c'est donc une $\mathbb{K}$-algèbre.

3.2.3 Morphismes d'algèbres

Soit $(A,+,\times,.)$ et $(A',+,\times,.)$ deux $\mathbb{K}$-algèbres.

Définition

On appelle *morphisme* de l'algèbre A vers l'algèbre A' toute application $\varphi\colon A \to A'$ vérifiant $\varphi(1_A) = 1_{A'}$ et, pour tous $x,y \in A$ et tout $\lambda \in \mathbb{K}$,

$$\varphi(\lambda x) = \lambda\varphi(x), \quad \varphi(x+y) = \varphi(x)+\varphi(y) \quad \text{et} \quad \varphi(xy) = \varphi(x)\varphi(y).$$

Un morphisme d'algèbres est à la fois une application linéaire et un morphisme d'anneaux. En particulier, le noyau d'un morphisme d'algèbres est à la fois un sous-espace vectoriel et un idéal.

Lorsqu'un morphisme d'algèbres $\varphi\colon A \to A'$ est bijectif, on parle *d'isomorphisme d'algèbres* et l'on dit que les algèbres A et A' sont *isomorphes*.

1. Si E est un $\mathbb{C}$-espace vectoriel de dimension finie n, E est un $\mathbb{R}$-espace vectoriel de dimension $2n$.
2. Les lois $+$, $\times$ et $(.)$ sur B sont définies par restriction des lois correspondantes sur A.

3.3 Exercices d'apprentissage

3.3.1 Structure d'algèbre

Exercice 1

Soit u un endomorphisme d'un espace vectoriel E. On introduit le commutant de u

$$\mathcal{C}_u = \{v \in \mathcal{L}(E) \mid u \circ v = v \circ u\}.$$

Montrer que $\mathcal{C}_u$ est une sous-algèbre de l'algèbre $\mathcal{L}(E)$ des endomorphismes de E.

Solution

méthode

Une sous-algèbre est une partie contenant[1] le neutre pour la multiplication et stable par combinaison linéaire et produit.

Par définition, $\mathcal{C}_u$ est une partie de l'algèbre $(\mathcal{L}(E), +, \circ,)$. Elle contient le neutre pour le produit de composition car $u \circ \mathrm{Id}_E = u = \mathrm{Id}_E \circ u$. De plus, pour $\lambda, \mu \in \mathbb{K}$ et $v, w \in \mathcal{C}_u$, on vérifie par opérations dans l'algèbre des endomorphismes

$$u \circ (\lambda v + \mu w) = \lambda(u \circ v) + \mu(u \circ w) = \lambda(v \circ u) + \mu(w \circ u) = (\lambda v + \mu w) \circ u.$$

$$u \circ (v \circ w) = (u \circ v) \circ w = (v \circ u) \circ w = v \circ (u \circ w) = v \circ (w \circ u) = (v \circ w) \circ u.$$

On a donc $\lambda v + \mu w \in \mathcal{C}_u$ et $v \circ w \in \mathcal{C}_u$. On peut conclure que $\mathcal{C}_u$ est une sous-algèbre de $\mathcal{L}(E)$.

Exercice 2

Soit

$$E = \left\{ M(a,b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid (a,b) \in \mathbb{R}^2 \right\}.$$

(a) Montrer que E est une algèbre réelle commutative pour les lois usuelles.

(b) Vérifier que l'algèbre E est isomorphe à $\mathbb{C}$.

Solution

(a) **méthode**

Une sous-algèbre est une algèbre pour les lois restreintes (Th. 1 p. 70).

On montre que E est une sous-algèbre de l'algèbre réelle $\mathcal{M}_2(\mathbb{R})$ des matrices carrées de taille 2.

1. On vérifie l'appartenance du neutre pour la multiplication et non seulement le caractère non vide ou l'appartenance du neutre additif.

L'ensemble E est une partie de $\mathcal{M}_2(\mathbb{R})$ et la matrice $\mathrm{I}_2 = M(1,0)$ appartient à E. Soit $\lambda, \mu \in \mathbb{R}$ et $A, B \in E$. On peut écrire $A = M(a,b)$ et $B = M(c,d)$ avec a, b, c, d réels et l'on vérifie par calcul matriciel [1]

$$\lambda A + \mu B = M(\lambda a + \mu c, \lambda b + \mu d) \in E \quad \text{et} \quad AB = M(ac - bd, ad + bc) \in E.$$

La partie E est donc une sous-algèbre de $\mathcal{M}_2(\mathbb{R})$, c'est une algèbre réelle pour les lois usuelles. De plus, elle est commutative car, avec les notations précédentes, on observe l'égalité $AB = BA$.

(b) Pour que la question ait un sens, on doit comprendre $\mathbb{C}$ comme une algèbre réelle. On introduit l'application $\varphi\colon \mathbb{C} \to E$ définie par

$$\varphi(a + \mathrm{i}b) = M(a,b) \quad \text{pour} \quad a, b \in \mathbb{R}.$$

méthode

Une application opérant entre deux algèbres est un morphisme d'algèbres lorsqu'elle envoie le neutre multiplicatif sur le neutre multiplicatif et que l'image d'une combinaison linéaire (resp. d'un produit) est la combinaison linéaire des images (resp. le produit).

L'application φ opère entre les deux algèbres réelles $\mathbb{C}$ et E. On a $\varphi(1) = \mathrm{I}_2$ et, pour tous $\lambda, \mu \in \mathbb{R}$ et $z, z' \in \mathbb{C}$, on vérifie en écrivant $z = a + \mathrm{i}b$ et $z' = c + \mathrm{i}d$ avec a, b, c, d réels

$$\varphi(\lambda z + \mu z') = \lambda\varphi(z) + \mu\varphi(z') \quad \text{et} \quad \varphi(zz') = \varphi(z)\varphi(z').$$

L'application φ est donc un morphisme d'algèbres. De plus, φ est surjective car ses valeurs sont exactement les $M(a,b)$. Enfin, φ est aussi injective [2] car $M(a,b) = \mathrm{O}_2$ si, et seulement si, $a = b = 0$ et donc $\mathrm{Ker}(\varphi) = \{0\}$.

Finalement, φ est un isomorphisme d'algèbres réelles.

Exercice 3

Soit x un élément d'une $\mathbb{K}$-algèbre $(A, +, \times, .)$ et $P = a_0 + a_1 X + \cdots + a_p X^p$ un polynôme de $\mathbb{K}[X]$. On appelle *valeur* du polynôme P en x l'élément

$$P(x) = \sum_{n=0}^{p} a_n x^n = a_0 1_A + a_1 x + \cdots + a_p x^p \in A.$$

(a) Montrer que l'application $E_x\colon P \mapsto P(x)$ détermine un morphisme d'algèbres.

On dit qu'un polynôme P est *annulateur* de x si $P(x) = 0_A$.

(b) Que dire de l'ensemble des polynômes annulateurs de l'élément x ?

1. On peut aussi mener le calcul en écrivant $M(a,b) = a\mathrm{I}_2 + bJ$ avec $J = M(0,1)$ vérifiant $J^2 = -\mathrm{I}_2$.
2. L'application φ est linéaire, on peut établir son injectivité en étudiant son noyau.

Solution

(a) L'application E_x est définie au départ de la $\mathbb{K}$-algèbre $\mathbb{K}[X]$ et à valeurs dans la $\mathbb{K}$-algèbre A. Vérifions que E_x envoie le polynôme constant égal à 1 sur 1_A et que l'image d'une combinaison linéaire (resp. d'un produit) est la combinaison linéaire des images (resp. le produit).

Si P est le polynôme constant égal à 1, la formule définissant la valeur d'un polynôme en x donne

$$E_x(1) = 1x^0 = 1_A.$$

Soit $\lambda, \mu \in \mathbb{K}$ et $P = a_0 + a_1X + \cdots + a_pX^p$ et $Q = b_0 + b_1X + \cdots + b_qX^q$ des polynômes de $\mathbb{K}[X]$. Quitte à adjoindre des coefficients nuls à la description de ces deux polynômes, on peut écrire

$$\lambda P + \mu Q = \sum_{n=0}^{\max(p,q)} (\lambda a_n + \mu b_n)X^n \quad \text{et} \quad PQ = \sum_{n=0}^{p+q} c_n X^n \text{ avec } c_n = \sum_{k+\ell=n} a_k b_\ell$$

où la somme définissant c_n s'étend sur les couples (k, ℓ) vérifiant la condition $k + \ell = n$ mais aussi $k \in [\![0\,;p]\!]$ et $\ell \in [\![0\,;q]\!]$. Par opérations dans l'algèbre A, on a alors

$$\begin{aligned} E_x(\lambda P + \mu Q) &= \sum_{n=0}^{\max(p,q)} (\lambda a_n + \mu b_n)x^n = \lambda \sum_{n=0}^{\max(p,q)} a_n x^n + \mu \sum_{n=0}^{\max(p,q)} b_n x^n \\ &= \lambda \sum_{n=0}^{p} a_n x^n + \mu \sum_{n=0}^{q} b_n x^n = \lambda E_x(P) + \mu E_x(Q) \end{aligned}$$

et

$$\begin{aligned} E_x(PQ) &= \sum_{n=0}^{p+q} c_n x^n = \sum_{n=0}^{p+q} \left(\sum_{k+\ell=n} (a_k x^k)(b_\ell x^\ell) \right) = \sum_{k=0}^{p} \left(\sum_{\ell=0}^{q} (a_k x^k)(b_\ell x^\ell) \right) \\ &= \left(\sum_{k=0}^{p} a_k x^k \right) \left(\sum_{\ell=0}^{q} b_\ell x^\ell \right) = E_x(P)E_x(Q). \end{aligned}$$

Ainsi, on peut affirmer que E_x est un morphisme d'algèbres.

(b) **méthode**

> Le noyau d'un morphisme d'algèbres est un sous-espace vectoriel et un idéal de l'algèbre de départ.

L'ensemble des polynômes annulateurs d'un élément x est le noyau du morphisme d'algèbres E_x, c'est donc un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$.

Les idéaux de $\mathbb{K}[X]$ étant les $P\,\mathbb{K}[X]$ pour P polynôme, on peut affirmer que, lorsque l'ensemble des polynômes annulateurs de x n'est pas réduit au polynôme nul, il existe un unique polynôme unitaire [1] M_x tel que les polynômes annulateurs de x correspondent aux polynômes multiples de M_x.

1. Le polynôme M_x est alors appelé le *polynôme minimal* de x.

3.3.2 Matrices semblables

Exercice 4

Parmi les matrices suivantes, figure-t-il des matrices semblables [1] ?

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

Solution

On note respectivement A, B, C et D les quatre matrices introduites.

méthode

Deux matrices semblables ont le même rang, la même trace et le même déterminant (les réciproques sont fausses).

Les quatre matrices sont de rang 3 et de déterminant 1 mais la matrice A est de trace égale à 1 alors que les trois autres matrices sont de traces égales à 3 : la matrice A n'est semblable à aucune des autres matrices.

méthode

Après permutation des vecteurs de base, les matrices B et C figurent le même endomorphisme.

Soit E un espace de dimension 3 muni d'une base $e = (e_1, e_2, e_3)$ et u l'endomorphisme figuré par B dans e. Considérons la base $e' = (e_3, e_2, e_1)$ obtenue par permutation des vecteurs de e. La matrice de u dans e' est exactement [2] C. Les matrices B et C sont donc semblables.

méthode

Si deux matrices carrées M et N de taille n sont semblables, $\lambda \mathrm{I}_n + M$ et $\lambda \mathrm{I}_n + N$ le sont aussi.

Par l'absurde, supposons les matrices B et D semblables. Il existe P inversible telle que $B = PDP^{-1}$. On a alors $B - \mathrm{I}_3 = P(D - \mathrm{I}_3)P^{-1}$ et donc $B - \mathrm{I}_3$ et $D - \mathrm{I}_3$ sont semblables. Or $B - \mathrm{I}_3$ est une matrice de rang 2 alors que $D - \mathrm{I}_3$ est de rang 3 : c'est absurde. Les matrices B et D ne sont donc pas semblables. Par transitivité de la relation de similitude, les matrices C et D ne sont pas non plus semblables.

Finalement, seules les matrices B et C sont semblables.

1. Deux matrices carrées de même taille A et B sont semblables lorsqu'il existe une matrice P inversible vérifiant $A = PBP^{-1}$. Par la formule de changement de bases, cela revient à signifier qu'elles figurent le même endomorphisme.

2. Les colonnes de la matrice de u dans e' sont formées des coordonnées dans e' des images des vecteurs de e' : permuter les vecteurs de la base permute les colonnes et les lignes.

Exercice 5

Soit $A \in \mathcal{M}_3(\mathbb{R})$ une matrice non nulle vérifiant $A^2 = O_3$. Établir que A est semblable à la matrice

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Solution

La matrice B est une matrice non nulle de $\mathcal{M}_3(\mathbb{R})$ vérifiant elle $B^2 = O_3$: ce constat est une nécessité pour affirmer que A puisse être semblable à B mais ne démontre pas que ceci a lieu !

méthode

On introduit un endomorphisme figuré par A et l'on recherche (souvent en menant une analyse) une base dans laquelle cet endomorphisme est figuré par B.

Soit a l'endomorphisme canoniquement[1] associé à la matrice A. C'est un endomorphisme de $E = \mathbb{R}^3$ vérifiant $a \neq 0$ et $a^2 = 0$ car $A \neq O_3$ et $A^2 = O_3$.

Analyse : Supposons qu'il existe une base $e = (e_1, e_2, e_3)$ dans laquelle la matrice de a soit égale à B. Par les colonnes de la matrice B, on lit

$$a(e_1) = e_2, \quad a(e_2) = 0_E \quad \text{et} \quad a(e_3) = 0_E. \qquad (*)$$

Le choix du vecteur e_1 détermine le vecteur $e_2 = a(e_1)$. Ce dernier ne doit pas être nul et e_1 est à choisir en dehors de $\mathrm{Ker}(a)$. En revanche, le vecteur e_2 doit appartenir au noyau de a mais ceci est assuré car $a^2 = 0$. Enfin, le vecteur e_3 doit aussi appartenir au noyau de a et cet espace doit donc être de dimension au moins 2. Vérifions cette dernière propriété. Sachant $a^2 = 0$, on peut affirmer $\mathrm{Im}(a) \subset \mathrm{Ker}(a)$. Or la formule du rang donne $\mathrm{rg}(a) + \dim \mathrm{Ker}(a) = 3$ et donc $\mathrm{rg}(a) = 1$ et $\dim \mathrm{Ker}(a) = 2$ car a n'est pas nul.

Ces conditions nécessaires étant étudiées, vérifions maintenant qu'il est possible de construire une telle base.

Synthèse : Soit e_1 un vecteur de E n'appartenant pas à $\mathrm{Ker}(a)$. Un tel vecteur existe car l'endomorphisme a est non nul. Posons $e_2 = a(e_1)$ ce qui définit un vecteur non nul appartenant au noyau de a car $a^2 = 0$. Enfin, complétons la famille libre (e_2) en une base (e_2, e_3) de $\mathrm{Ker}(a)$ ce qui est possible car cet espace est de dimension 2. Par construction, les égalités $(*)$ sont vérifiées. Il reste à justifier que la famille $e = (e_1, e_2, e_3)$ est une base de E. Il s'agit d'une famille de longueur 3 dans un espace de dimension 3, il suffit de vérifier sa liberté.

Soit $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ tel que

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0_E. \qquad (\triangle)$$

1. L'application linéaire canoniquement associée à une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est l'application linéaire de $\mathbb{K}^p$ vers $\mathbb{K}^n$ figurée par la matrice A dans les bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$. C'est aussi l'application qui à $x \in \mathbb{K}^p$ associe $y = Ax$ où, dans ce calcul, on identifie vecteur de $\mathbb{K}^p$ (resp. $\mathbb{K}^n$) et colonne de $\mathcal{M}_{p,1}(\mathbb{K})$ (resp. de $\mathcal{M}_{n,1}(\mathbb{K})$) formée des mêmes coefficients.

En appliquant a aux deux membres de l'égalité $(\triangle)$ on obtient $\lambda_1 e_2 = 0_E$ et donc $\lambda_1 = 0$. La relation $(\triangle)$ se simplifie alors en $\lambda_2 e_2 + \lambda_3 e_3 = 0_E$ et donc $\lambda_2 = \lambda_3 = 0$ car la famille (e_2, e_3) est libre. La famille e est donc libre.

Finalement, la famille $e = (e_1, e_2, e_3)$ est une base de E dans laquelle l'endomorphisme a est figuré par B : les matrices A et B sont semblables.

3.4 Exercices d'entraînement

3.4.1 Lorsque le corps de base est $\mathbb{Q}$...

Exercice 6 **

(a) Le polynôme $X^3 - 3X - 1$ est-il irréductible dans $\mathbb{Q}[X]$? dans $\mathbb{R}[X]$?

(b) Mêmes questions avec $X^4 + 1$.

Solution

(a) **méthode**

> Un polynôme est irréductible dans $\mathbb{K}[X]$ lorsque ses seuls diviseurs sont les polynômes constants non nuls et ses polynômes associés.

Par l'absurde, si le polynôme $P = X^3 - 3X + 1$ n'est pas irréductible dans $\mathbb{Q}[X]$, il est possible de l'écrire comme produit de deux polynômes à coefficients rationnels non constants. L'un d'eux est de degré 1 et détermine donc une racine $r \in \mathbb{Q}$ de P. On écrit r sous forme irréductible p/q (avec $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$ et $p \wedge q = 1$) et l'égalité $P(r) = 0$ donne après réduction au même dénominateur

$$p^3 - 3pq^2 - q^3 = 0.$$

L'entier p divise $p^3 - 3pq^2$ et donc divise $q^3 = q^3 \times 1$. Or p est premier avec q et donc (par le lemme de Gauss) p divise 1, c'est-à-dire $p = \pm 1$. Aussi, q divise $3pq^2 + q^3 = p^3$ mais est premier avec p et donc $q = 1$. Ainsi, $r = \pm 1$. Cependant, ni 1, ni -1, ne sont racines de P. C'est absurde. Le polynôme P est donc irréductible dans $\mathbb{Q}[X]$.

En revanche, il n'est pas irréductible dans $\mathbb{R}[X]$ car il est de degré impair et possède donc une racine réelle a : le polynôme $X - a$ est alors un facteur non trivial de P.

(b) Bien qu'il ne possède pas de racines réelles, le polynôme $Q = X^4 + 1$ n'est pas irréductible [1] dans $\mathbb{R}[X]$: on peut le factoriser en faisant apparaître une différence de deux carrés [2]

$$X^4 + 1 = \left(X^2 + 1\right)^2 - 2X^2 = \left(X^2 - \sqrt{2}X + 1\right)\left(X^2 + \sqrt{2}X + 1\right). \qquad (*)$$

1. Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 sans racines réelles.

2. On peut aussi factoriser le polynôme dans $\mathbb{C}[X]$ et recombiner les facteurs conjugués pour former la factorisation dans $\mathbb{R}[X]$.

En revanche, le polynôme Q est irréductible dans $\mathbb{Q}[X]$. En effet, si D est un diviseur non trivial de Q dans $\mathbb{Q}[X]$, c'est aussi un diviseur dans $\mathbb{R}[X]$. Cependant, la factorisation $(*)$ fournit la décomposition en facteurs irréductibles de Q dans $\mathbb{R}[X]$ et les diviseurs non triviaux de Q dans $\mathbb{R}[X]$ sont donc les polynômes associés à $X^2 - \sqrt{2}X + 1$ et à $X^2 + \sqrt{2}X + 1$. Le polynôme D ne peut être de cette forme car $\sqrt{2}$ est un nombre irrationnel.

Exercice 7 **

On munit $\mathbb{R}$ de sa structure[1] de $\mathbb{Q}$-espace vectoriel.

(a) Soit $d \in \mathbb{N}^*$. À quelle condition la famille $\left(1, \sqrt{d}\right)$ est-elle libre ?

(b) Établir la liberté de la famille $\left(1, \sqrt{2}, \sqrt{3}, \sqrt{6}\right)$.

Solution

(a) **méthode**

> Une famille est libre lorsqu'une combinaison linéaire nulle des vecteurs de la famille est nécessairement écrite avec des scalaires tous nuls. Ici, on sera attentif à ce que les scalaires sont des nombres rationnels.

Cas : d est le carré d'un entier n. On peut simplifier la racine et affirmer que la famille $\left(1, \sqrt{d}\right)$ est liée en vertu de la relation linéaire suivante

$$\underbrace{n}_{\in\mathbb{Q}} 1 + \underbrace{(-1)}_{\in\mathbb{Q}}\sqrt{d} = 0.$$

Cas : d n'est pas le carré d'un entier. Soit $(\lambda, \mu) \in \mathbb{Q}^2$ tel que $\lambda 1 + \mu\sqrt{d} = 0$. En réduisant λ et μ à un dénominateur commun q, on obtient l'écriture

$$a + b\sqrt{d} = 0 \quad \text{avec} \quad a, b \in \mathbb{Z},\ \lambda = \frac{a}{q} \text{ et } \mu = \frac{b}{q}.$$

On en tire $a = -b\sqrt{d}$, puis, en élevant au carré, $a^2 = b^2 d$.

Si a ou b est non nul, l'autre est aussi non nul et les facteurs premiers de a^2 et b^2 s'écrivent tous avec des exposants pairs. Il en est alors de même pour d qui est donc le carré d'un entier. Ceci étant exclu, on obtient $(a, b) = (0, 0)$ puis $(\lambda, \mu) = (0, 0)$. La famille $\left(1, \sqrt{d}\right)$ est alors libre.

Finalement, la famille $\left(1, \sqrt{d}\right)$ est libre[2] dans le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$ si, et seulement si, d n'est pas le carré d'un entier.

1. Les vecteurs sont les nombres réels et les scalaires exprimant les combinaisons linéaires sont des nombres rationnels. Le produit extérieur est la multiplication usuelle. Cet espace est de dimension infinie (voir sujet suivant).

2. Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}$ cette famille est assurément liée car comporte deux vecteurs en dimension 1.

(b) Soit $(\alpha, \beta, \gamma, \delta) \in \mathbb{Q}^4$ tel que $\alpha + \beta\sqrt{2} + \gamma\sqrt{3} + \delta\sqrt{6} = 0$. Après réduction des rationnels $\alpha, \beta, \gamma, \delta$ à un dénominateur commun q, on obtient l'écriture

$$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0 \quad \text{avec} \quad a, b, c, d \in \mathbb{Z},\ \alpha = \frac{a}{q},\ \beta = \frac{b}{q},\ \gamma = \frac{c}{q} \text{ et } \delta = \frac{d}{q}.$$

méthode

On forme des relations sur a, b, c, d en isolant deux termes et en élevant au carré.

D'une part, $\left(a + b\sqrt{2}\right)^2 = \left(c\sqrt{3} + d\sqrt{6}\right)^2$ et, en développant les carrés,

$$a^2 + 2ab\sqrt{2} + 2b^2 = 3c^2 + 6cd\sqrt{2} + 6d^2.$$

La famille $\left(1, \sqrt{2}\right)$ étant libre, on peut identifier les rationnels en facteur de 1 et de $\sqrt{2}$:

$$\begin{cases} a^2 + 2b^2 = 3c^2 + 6d^2 & (1) \\ 2ab = 3cd. \end{cases}$$

D'autre part, $\left(a + c\sqrt{3}\right)^2 = \left(b\sqrt{2} + d\sqrt{6}\right)^2$ et le même raisonnement qu'au-dessus donne

$$\begin{cases} a^2 + 3c^2 = 2b^2 + 6d^2 & (2) \\ 2ac = 4bd. \end{cases}$$

Après simplification, la somme des équations (1) et (2) donne $a^2 = 6d^2$ ce qui entraîne $a = d = 0$ car 6 n'est pas le carré d'un entier. Aussi, la différence des équations (1) et (2) conduit à $b = c = 0$.

Finalement, la famille $\left(1, \sqrt{2}, \sqrt{3}, \sqrt{6}\right)$ est libre dans le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$.

Exercice 8 **

On peut énumérer [1] l'infinité des nombres premiers en ordre croissant afin de former une suite $(p_n)_{n\geqslant 1}$: $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc.

(a) Montrer que la famille $(\ln p_n)_{n\geqslant 1}$ est une famille libre du $\mathbb{Q}$-espace vectoriel $\mathbb{R}$.

(b) Que dire de la dimension du $\mathbb{Q}$-espace vectoriel $\mathbb{R}$?

Solution

(a) **méthode**

La liberté d'une famille infinie équivaut à la liberté de toutes ses sous-familles finies.

Soit $n \in \mathbb{N}^*$ arbitrairement grand. Il nous suffit d'établir la liberté de la sous-famille finie $(\ln p_1, \ldots, \ln p_n)$.

1. L'ensemble des nombres premiers est dénombrable.

Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{Q}^n$ tel que $\lambda_1 \ln p_1 + \cdots + \lambda_n \ln p_n = 0$. En réduisant les nombres rationnels $\lambda_1, \dots, \lambda_n$ à un même dénominateur q, on obtient l'écriture

$$\alpha_1 \ln p_1 + \cdots + \alpha_n \ln p_n = 0 \quad \text{avec} \quad \alpha_k \in \mathbb{Z} \text{ et } \lambda_k = \frac{\alpha_k}{q} \text{ pour tout } k \in [\![1\,;n]\!].$$

Par la propriété de morphisme du logarithme, on obtient l'égalité

$$\ln\left(\prod_{k=1}^{n} p_k^{\alpha_k}\right) = 0 \quad \text{puis} \quad \prod_{k=1}^{n} p_k^{\alpha_k} = 1.$$

Afin d'interpréter cette égalité dans le cadre de l'arithmétique des entiers, on sépare les facteurs du produit selon le signe de α_k

$$\prod_{\substack{1\leqslant k\leqslant n\\ \alpha_k\geqslant 0}} p_k^{\alpha_k} = \prod_{\substack{1\leqslant k\leqslant n\\ \alpha_k< 0}} p_k^{-\alpha_k}.$$

Les deux membres de l'égalité correspondent à l'écriture d'un entier en produit de facteurs premiers, l'unicité de la décomposition en facteurs premiers des entiers entraîne $\alpha_k = 0$, puis $\lambda_k = 0$, pour tout $k \in [\![1\,;n]\!]$.

La famille étudiée est donc libre.

(b) Le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$ contient une famille libre infinie, il est donc de dimension infinie.

3.4.2 Applications linéaires

Exercice 9 *

Soit f un endomorphisme d'un espace E de dimension finie vérifiant $\operatorname{rg}(f^2) = \operatorname{rg}(f)$.

(a) Établir $\operatorname{Im}(f^2) = \operatorname{Im}(f)$ et $\operatorname{Ker}(f^2) = \operatorname{Ker}(f)$.

(b) Montrer que les espaces $\operatorname{Im}(f)$ et $\operatorname{Ker}(f)$ sont supplémentaires dans E.

Solution

(a) **méthode**

> On transforme les inclusions $\operatorname{Im}(f^2) \subset \operatorname{Im}(f)$ et $\operatorname{Ker}(f) \subset \operatorname{Ker}(f^2)$ en égalité par un argument de dimension.

Soit $y \in \operatorname{Im}(f^2)$. Il existe un antécédent $x \in E$ tel que $y = f^2(x)$ et donc y est élément de $\operatorname{Im}(f)$ car on peut écrire $y = f(f(x)) = f(a)$ avec $a = f(x)$ élément de E. Ainsi, on a l'inclusion[1] $\operatorname{Im}(f^2) \subset \operatorname{Im}(f)$. De plus, l'hypothèse $\operatorname{rg}(f) = \operatorname{rg}(f^2)$ fournit l'égalité des dimensions et l'on peut affirmer $\operatorname{Im}(f^2) = \operatorname{Im}(f)$.

1. Plus généralement, on peut affirmer $\operatorname{Im}(g \circ f) \subset \operatorname{Im}(g)$ pour tous f et g endomorphismes de E.

Aussi, pour $x \in \mathrm{Ker}(f)$, on a $f(x) = 0_E$ donc $f^2(x) = f(0_E) = 0_E$. Ainsi, on peut écrire[1] $\mathrm{Ker}(f) \subset \mathrm{Ker}(f^2)$. De plus, la formule du rang appliquée à f et f^2 donne

$$\dim E = \mathrm{rg}(f) + \dim \mathrm{Ker}(f) \quad \text{et} \quad \dim E = \mathrm{rg}(f^2) + \dim \mathrm{Ker}(f^2).$$

On en déduit $\dim \mathrm{Ker}(f) = \dim \mathrm{Ker}(f^2)$. Par inclusion et égalité des dimensions, on conclut $\mathrm{Ker}(f^2) = \mathrm{Ker}(f)$.

(b) **méthode**

Il suffit de vérifier que $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont en somme directe avant de conclure avec un argument de dimension.

Soit $x \in \mathrm{Ker}(f) \cap \mathrm{Im}(f)$. On a $f(x) = 0_E$ et l'on peut introduire un antécédent $a \in E$ tel que $x = f(a)$. On a alors $f\big(f(a)\big) = f(x) = 0_E$ et donc a appartient au noyau de f^2. Or celui-ci se confond avec le noyau de f et donc $x = f(a) = 0_E$. Ainsi, les espaces $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont en somme directe. De plus, la formule du rang donne

$$\dim E = \dim \mathrm{Im}(f) + \dim \mathrm{Ker}(f)$$

et l'on peut conclure que les espaces $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont supplémentaires.

Exercice 10 *

Soit f un endomorphisme d'un espace E de dimension finie. Montrer

$$\mathrm{Ker}(f) = \mathrm{Im}(f) \iff \big(f^2 = 0 \text{ et } \dim E = 2\,\mathrm{rg}(f)\big).$$

Solution

méthode

Pour $f, g \in \mathcal{L}(E)$, on sait

$$g \circ f = 0 \iff \mathrm{Im}(f) \subset \mathrm{Ker}(g).$$

On raisonne par double implication.

($\Longrightarrow$) Supposons $\mathrm{Ker}(f) = \mathrm{Im}(f)$. D'une part, $f^2 = 0$ car[2] $\mathrm{Im}(f) \subset \mathrm{Ker}(f)$. D'autre part, la formule du rang donne $\dim E = \mathrm{rg}(f) + \dim \mathrm{Ker}(f) = 2\,\mathrm{rg}(f)$ car les dimensions de $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont égales.

($\Longleftarrow$) Supposons $f^2 = 0$ et $\dim E = 2\,\mathrm{rg}(f)$. D'une part, $\mathrm{Im}(f) \subset \mathrm{Ker}(f)$ car[3] $f \circ f = 0$. D'autre part, la formule du rang donne $2\,\mathrm{rg}(f) = \dim E = \mathrm{rg}(f) + \dim \mathrm{Ker}(f)$ et donc $\dim \mathrm{Im}(f) = \dim \mathrm{Ker}(f)$. Par inclusion et égalité des dimensions, on peut conclure $\mathrm{Im}(f) = \mathrm{Ker}(f)$.

1. Aussi, on peut affirmer $\mathrm{Ker}(g) \subset \mathrm{Ker}(f \circ g)$ pour tous f et g endomorphismes de E.
2. Pour tout $x \in E$, $f(x)$ est élément de $\mathrm{Im}(f)$ donc de $\mathrm{Ker}(f)$ et par conséquent $f\big(f(x)\big) = 0_E$.
3. Tout y de $\mathrm{Im}(f)$ peut s'écrire $f(x)$ avec $x \in E$ et alors $f(y) = f\big(f(x)\big) = 0_E$ donc $y \in \mathrm{Ker}(f)$.

Exercice 11 **

Soit f et g deux endomorphismes d'un espace vectoriel E de dimension finie.

(a) Montrer

$$\mathrm{rg}(g \circ f) = \mathrm{rg}(g) \iff E = \mathrm{Im}(f) + \mathrm{Ker}(g).$$

(b) Montrer

$$\mathrm{rg}(g \circ f) = \mathrm{rg}(f) \iff \mathrm{Im}(f) \cap \mathrm{Ker}(g) = \{0_E\}.$$

Solution

(a) **méthode**

Sachant l'inclusion $\mathrm{Im}(g \circ f) \subset \mathrm{Im}(g)$, l'égalité des rangs de $g \circ f$ et g signifie l'égalité des espaces images.

On raisonne par double implication.

($\Longleftarrow$) Supposons $E = \mathrm{Im}(f) + \mathrm{Ker}(g)$. Soit y élément de $\mathrm{Im}(g)$. On peut introduire un antécédent $x \in E$ tel que $y = g(x)$ et écrire $x = a + b$ avec $a \in \mathrm{Im}(f)$ et $b \in \mathrm{Ker}(g)$. On peut aussi écrire $a = f(c)$ avec $c \in E$ et alors

$$y = g(x) = g(a) + g(b) = g(a) = g\big(f(c)\big) \in \mathrm{Im}(g \circ f).$$

Par conséquent, $\mathrm{Im}(g) \subset \mathrm{Im}(g \circ f)$ puis $\mathrm{Im}(g) = \mathrm{Im}(g \circ f)$ donc $\mathrm{rg}(g \circ f) = \mathrm{rg}(g)$.

($\Longrightarrow$) Supposons $\mathrm{rg}(g \circ f) = \mathrm{rg}(g)$ et donc $\mathrm{Im}(g) = \mathrm{Im}(g \circ f)$: la force de cette hypothèse réside dans l'inclusion $\mathrm{Im}(g) \subset \mathrm{Im}(g \circ f)$ qui signifie que n'importe quel vecteur $g(x)$ peut s'écrire sous la forme $g\big(f(a)\big)$ pour un certain vecteur a de E.

Soit $x \in E$. Introduisons, un vecteur a tel que $g(x) = g\big(f(a)\big)$ et considérons le vecteur $b = x - f(a)$. On a $x = f(a) + b$ avec $f(a) \in \mathrm{Im}(f)$ et $b \in \mathrm{Ker}(g)$ car $g(x) = g\big(f(a)\big)$. Ainsi, $E \subset \mathrm{Im}(f) + \mathrm{Ker}(g)$ et donc $E = \mathrm{Im}(f) + \mathrm{Ker}(g)$ puisque l'inclusion réciproque est entendue.

(b) **méthode**

Sachant l'inclusion $\mathrm{Ker}(f) \subset \mathrm{Ker}(g \circ f)$, l'égalité des rangs de f et $g \circ f$ signifient l'égalité des noyaux.

On raisonne par double implication.

($\Longleftarrow$) Supposons $\mathrm{Im}(f) \cap \mathrm{Ker}(g) = \{0_E\}$. Soit x un élément de $\mathrm{Ker}(g \circ f)$. Pour celui-ci, on a à la fois $f(x)$ dans $\mathrm{Im}(f)$ et dans $\mathrm{Ker}(g)$. On a donc $f(x) = 0_E$. On en déduit $\mathrm{Ker}(g \circ f) \subset \mathrm{Ker}(f)$. L'autre inclusion étant immédiate, on obtient l'égalité des noyaux, puis, par la formule du rang, l'égalité des rangs de f et de $g \circ f$.

($\Longrightarrow$) Supposons $\mathrm{rg}(g \circ f) = \mathrm{rg}(f)$. À l'inverse du raisonnement précédent, on peut affirmer $\mathrm{Ker}(g \circ f) = \mathrm{Ker}(f)$. Soit $x \in \mathrm{Im}(f) \cap \mathrm{Ker}(g)$. On peut écrire $x = f(a)$ avec $a \in E$ et l'on a $g(x) = 0_E$. On en déduit $(g \circ f)(a) = 0_E$ et donc a est élément de $\mathrm{Ker}(g \circ f)$, c'est-à-dire de $\mathrm{Ker}(f)$. On peut alors conclure $x = f(a) = 0_E$. Ainsi, $\mathrm{Im}(f) \cap \mathrm{Ker}(g) \subset \{0_E\}$ et l'on conclut $\mathrm{Im}(f) \cap \mathrm{Ker}(g) = \{0_E\}$ car l'inclusion réciproque est entendue.

Exercice 12 ** (Noyaux et images itérés)

Soit f un endomorphisme d'un espace vectoriel E et $p \in \mathbb{N}$.

(a) Montrer que si $\mathrm{Ker}(f^p) = \mathrm{Ker}(f^{p+1})$ alors, pour $k \in \mathbb{N}$, $\mathrm{Ker}(f^p) = \mathrm{Ker}(f^{p+k})$.

(b) Établir la même propriété avec les espaces images.

(c) Donner un exemple d'endomorphisme f pour lequel il n'existe pas d'entiers p tels que $\mathrm{Ker}(f^p) = \mathrm{Ker}(f^{p+1})$.

(d) Même question avec les espaces images.

Solution

(a) On connaît [1] la croissance (au sens de l'inclusion) de la suite des $\mathrm{Ker}(f^p)$. Il s'agit ici d'établir que la suite devient constante dès que deux noyaux successifs sont égaux.

Supposons $\mathrm{Ker}(f^{p+1}) = \mathrm{Ker}(f^p)$.

méthode

On propage l'égalité $\mathrm{Ker}(f^{p+1}) = \mathrm{Ker}(f^p)$ en écrivant

$$f^{p+k+1}(x) = f^{p+1}(f^k(x)).$$

Soit $x \in E$ et $k \in \mathbb{N}$. On a

$$\begin{aligned} x \in \mathrm{Ker}(f^{p+k+1}) &\iff f^k(x) \in \mathrm{Ker}(f^{p+1}) \\ &\iff f^k(x) \in \mathrm{Ker}(f^p) \\ &\iff x \in \mathrm{Ker}(f^{p+k}). \end{aligned}$$

On en déduit $\mathrm{Ker}(f^{p+k+1}) = \mathrm{Ker}(f^{p+k})$. Par une récurrence facile, on peut conclure à l'égalité $\mathrm{Ker}(f^{p+k}) = \mathrm{Ker}(f^p)$ pour tout $k \in \mathbb{N}$.

(b) On connaît la décroissance [2] de la suite des $\mathrm{Im}(f^p)$. On montre ici que la suite devient constante dès que deux images successives sont égales.

Supposons $\mathrm{Im}(f^{p+1}) = \mathrm{Im}(f^p)$.

méthode

On propage l'égalité $\mathrm{Im}(f^{p+1}) = \mathrm{Im}(f^p)$ par l'écriture

$$f^{p+k}(a) = f^k(f^p(a)).$$

Soit $y \in E$ et $k \in \mathbb{N}$. On a

$$\begin{aligned} y \in \mathrm{Im}(f^{p+k}) &\iff \exists x \in \mathrm{Im}(f^p),\ y = f^k(x) \\ &\iff \exists x \in \mathrm{Im}(f^{p+1}),\ y = f^k(x) \\ &\iff y \in \mathrm{Im}(f^{p+k+1}). \end{aligned}$$

1. Si $f^p(x) = 0_E$ alors $f^{p+1}(x) = f(0_E) = 0_E$: Voir le sujet 24 du chapitre 8 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

2. Si $y = f^{p+1}(x)$ alors $y = f^p(a)$ avec $a = f(x)$.

On en déduit $\mathrm{Im}(f^{p+k+1}) = \mathrm{Im}(f^{p+k})$ et l'on peut alors conclure $\mathrm{Im}(f^{p+k}) = \mathrm{Im}(f^p)$ pour tout $k \in \mathbb{N}$.

(c) Dans un espace de dimension finie, les dimensions des noyaux itérés forment une suite d'entiers majorée donc constante à partir d'un certain rang. Les noyaux sont alors égaux à partir de ce rang. Un exemple tel que demandé ne peut alors être trouvé qu'en dimension infinie. Dans l'espace $\mathbb{K}[X]$, on peut proposer l'endomorphisme D de dérivation pour lequel $\mathrm{Ker}(\mathrm{D}^p) = \mathbb{K}_{p-1}[X]$ pour tout $p \geqslant 1$.

(d) De la même façon, un exemple doit être recherché en dimension infinie : l'endomorphisme $\psi : P \mapsto XP$ dans $\mathbb{K}[X]$ convient [1] puisque $\mathrm{Im}(\psi^p) = X^p\mathbb{K}[X]$.

Exercice 13 **

Soit $f, g, h \in \mathcal{L}(E)$ tels que $f \circ g = h$, $g \circ h = f$ et $h \circ f = g$.

(a) Montrer que f, g et h ont même noyau et même image.

(b) Vérifier $f^5 = f$.

(c) En déduire que l'image et le noyau de f sont supplémentaires dans E.

Solution

(a) **méthode**

On emploie l'égalité $f \circ g = h$ pour affirmer que le noyau de g est inclus dans celui de h et l'image de h incluse dans celle de f.

Soit x élément de $\mathrm{Ker}(g)$. On a $h(x) = (f \circ g)(x) = f\big(g(x)\big) = f(0_E) = 0_E$ donc x appartient à $\mathrm{Ker}(h)$. Ainsi, on a l'inclusion $\mathrm{Ker}(g) \subset \mathrm{Ker}(h)$. De même, l'égalité $g \circ h = f$ donne $\mathrm{Ker}(h) \subset \mathrm{Ker}(f)$ et $h \circ f = g$ donne $\mathrm{Ker}(f) \subset \mathrm{Ker}(g)$. On a alors les inclusions

$$\mathrm{Ker}(g) \subset \mathrm{Ker}(h) \subset \mathrm{Ker}(f) \subset \mathrm{Ker}(g)$$

et l'on peut conclure que les trois noyaux sont égaux.

Aussi, $h = f \circ g$ entraîne que les valeurs prises par h sont des valeurs prises par f. On peut ainsi affirmer $\mathrm{Im}(h) \subset \mathrm{Im}(f)$ mais aussi $\mathrm{Im}(f) \subset \mathrm{Im}(g)$ et $\mathrm{Im}(g) \subset \mathrm{Im}(h)$ car on a respectivement $f = g \circ h$ et $g = h \circ f$. Les trois espaces images sont égaux.

(b) **méthode**

On vérifie $f^2 = g^2 = h^2$.

Par associativité $f^2 = (g \circ h) \circ f = g \circ (h \circ f) = g^2$. De même, on a $g^2 = h^2$ et alors

$$f^5 = g^2 \circ h^2 \circ f = g \circ (g \circ h) \circ (h \circ f) = g \circ (f \circ g) = g \circ h = f.$$

1. Les deux exemples précédents sont à rapprocher des endomorphismes qui transforment une suite numériques $(u_0, u_1, \ldots)$ en une suite obtenue par décalage, soit en perdant le premier terme $(u_1, u_2, \ldots)$, soit en insérant un premier terme nul $(0, u_0, u_1, \ldots)$.

(c) **méthode**

On vérifie [1] par analyse-synthèse que tout vecteur de E s'écrit de façon unique comme la somme d'un vecteur de l'image et d'un vecteur du noyau.

Soit $x \in E$.

Analyse : Supposons $x = a + b$ avec $a \in \mathrm{Im}(f)$ et $b \in \mathrm{Ker}(f)$. On peut écrire $a = f(c)$ avec $c \in E$ et l'on a $f^4(x) = f^5(c) = f(c) = a$. Ceci détermine a puis $b = x - a$ de façon unique.

Synthèse : Posons $a = f^4(x)$ et $b = x - a$. On a immédiatement $a \in \mathrm{Im}(f)$ et $x = a + b$. Reste à vérifier l'appartenance de b au noyau de f : ce qui résulte du calcul suivant :

$$f(b) = f(x) - f(a) = f(x) - f^5(x) = 0_E.$$

On peut conclure que les espaces $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont supplémentaires.

Exercice 14 **

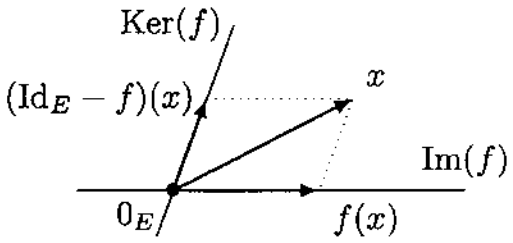
Soit f un endomorphisme d'un espace vectoriel E de dimension $n \in \mathbb{N}$. Montrer

$$f \text{ est un projecteur} \iff \mathrm{rg}(f) + \mathrm{rg}(\mathrm{Id}_E - f) = n.$$

Solution

Raisonnons par double implication.

($\Longrightarrow$) Si f est un projecteur, les espaces $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont supplémentaires et f est la projection sur $\mathrm{Im}(f)$ parallèlement à $\mathrm{Ker}(f)$. L'endomorphisme $\mathrm{Id}_E - f$ correspond alors à la projection complémentaire de f, c'est-à-dire la projection sur $\mathrm{Ker}(f)$ parallèlement à $\mathrm{Im}(f)$. On en déduit



$$\mathrm{rg}(f) + \mathrm{rg}(\mathrm{Id}_E - f) = \dim \mathrm{Im}(f) + \dim \mathrm{Ker}(f) = n.$$

($\Longleftarrow$) Supposons $\mathrm{rg}(f) + \mathrm{rg}(\mathrm{Id}_E - f) = n$.

méthode

On vérifie que les espaces $F = \mathrm{Im}(f)$ et $G = \mathrm{Im}(\mathrm{Id}_E - f)$ sont supplémentaires.

Pour tout $x \in E$, on peut écrire $x = a + b$ avec $a = f(x) \in F$ et $b = x - f(x) \in G$. Ainsi, $E \subset F + G$ puis $E = F + G$. Or $\dim F + \dim G = \mathrm{rg}(f) + \mathrm{rg}(\mathrm{Id}_E - f) = \dim E$ et donc $E = F \oplus G$.

L'écriture $x = f(x) + \big(x - f(x)\big)$ apparaît comme l'unique façon de décomposer un vecteur x en la somme d'un vecteur de F et d'un vecteur de G. Puisque l'application f associe à x le premier vecteur de cette écriture, elle se comprend comme la projection sur F parallèlement à G.

1. S'il y avait eu une hypothèse de dimension finie, on aurait pu établir que les espaces sont supplémentaires en étudiant leur intersection et en employant la formule du rang : $\dim E = \mathrm{rg}(f) + \dim \mathrm{Ker}(f)$.

Exercice 15 **

Soit u et v deux endomorphismes d'un espace vectoriel E de dimension finie. Montrer

$$\dim \operatorname{Ker}(u \circ v) \leqslant \dim \operatorname{Ker}(u) + \dim \operatorname{Ker}(v).$$

Solution

méthode

On applique la formule du rang à la restriction de v au départ de $\operatorname{Ker}(u \circ v)$.

Notons v' la restriction de v au départ de $\operatorname{Ker}(u \circ v)$. Celle-ci est une application linéaire et la formule du rang appliquée à v' permet d'écrire

$$\dim \underbrace{\operatorname{Ker}(u \circ v)}_{\text{espace de départ}} = \operatorname{rg}(v') + \dim \operatorname{Ker}(v'). \qquad (*)$$

Cependant,

$$\operatorname{Ker}(v') = \operatorname{Ker}(v) \cap \operatorname{Ker}(u \circ v)$$

et donc $\operatorname{Ker}(v') \subset \operatorname{Ker}(v)$ ce qui entraîne $\dim \operatorname{Ker}(v') \leqslant \dim \operatorname{Ker}(v)$.

Aussi, les valeurs prises par v' appartiennent à $\operatorname{Ker}(u)$ car, pour tout $x \in \operatorname{Ker}(u \circ v)$, on a $u\big(v'(x)\big) = u\big(v(x)\big) = 0_E$. On a ainsi $\operatorname{Im}(v') \subset \operatorname{Ker}(u)$ et donc $\operatorname{rg}(v') \leqslant \dim \operatorname{Ker}(u)$. L'équation $(*)$ donne alors la relation voulue.

Exercice 16 **

Soit F un sous-espace vectoriel d'un espace vectoriel E de dimension finie.

(a) Déterminer la dimension de l'espace $A = \big\{f \in \mathcal{L}(E) \;\big|\; \operatorname{Im}(f) \subset F\big\}$.

(b) Déterminer la dimension de l'espace $B = \big\{f \in \mathcal{L}(E) \;\big|\; F \subset \operatorname{Ker}(f)\big\}$.

Solution

On vérifie dans les deux études que les parties A et B sont des sous-espaces vectoriels de $\mathcal{L}(E)$ car non vides et stables par combinaisons linéaires.

(a) **méthode**

On peut calculer la dimension d'un espace en déterminant un isomorphisme entre celui-ci et un espace de dimension connue.

Les endomorphismes dont l'image est incluse dans F peuvent s'identifier aux applications linéaires de E vers F. Plus précisément, l'application $\Phi\colon A \to \mathcal{L}(E,F)$ qui associe à $f \in A$ sa restriction [1] $f|_E^F$ au départ de E et à valeurs dans F est un isomorphisme d'espaces vectoriels. On en déduit

$$\dim A = \dim \mathcal{L}(E,F) = \dim E \times \dim F.$$

1. Cette restriction est bien définie car l'application f prend ses valeurs dans F.

(b) **méthode**

Une application linéaire est entièrement déterminée par ses restrictions linéaires sur des espaces supplémentaires.

Introduisons G un espace supplémentaire de F et, pour $f \in \mathcal{L}(E)$, notons $f|_F^E$ et $f|_G^E$ les restrictions de f au départ de F et G et à valeurs dans E. Les éléments de B sont exactement les endomorphismes dont la restriction au départ de F est nulle : ils sont entièrement déterminés par leur restriction au départ de G qui est une application linéaire quelconque de G vers E. Ainsi, l'application $\Phi\colon B \to \mathcal{L}(G,E)$ qui à $f \in B$ associe $f|_G^E$ est un isomorphisme d'espaces vectoriels. On en déduit

$$\dim B = \dim \mathcal{L}(G,E) = (\dim E - \dim F)\dim E.$$

méthode

Une résolution matricielle de ce sujet est aussi possible.

Si l'on introduit une base de E dont les premiers vecteurs forment une base de F, les endomorphismes des espaces A et B correspondent respectivement à ceux figurés dans e par les matrices par blocs

$$\begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$$

où les blocs diagonaux sont carrés de tailles r et $n-r$ avec $n = \dim E$ et $r = \dim F$.

Exercice 17 **

Soit f un endomorphisme d'un espace vectoriel E tel que la famille $\big(x, f(x)\big)$ est liée pour tout vecteur x de E.

(a) Justifier que, pour tout $x \in E$, il existe un scalaire α tel que $f(x) = \alpha x$.

(b) En déduire que f est une homothétie vectorielle.

Solution

(a) Si $x = 0_E$, n'importe quel scalaire α convient.

Sinon, la famille $\big(x, f(x)\big)$ étant liée, il existe $(\lambda,\mu) \neq (0,0)$ tel que $\lambda x + \mu f(x) = 0_E$. Le scalaire μ ne peut pas être nul car x n'est pas le vecteur nul et le couple (λ,μ) n'est pas le couple nul. On peut alors diviser par μ et écrire $f(x) = \alpha x$ avec $\alpha = -\lambda/\mu$.

(b) *A priori*, le scalaire α introduit [1] à la question précédente dépend de la valeur de x. Pour souligner cette dépendance, notons le α_x lorsque x est non nul.

méthode

On montre que $\alpha_x = \alpha_y$ en étudiant $f(x+y)$.

Soit x et y deux vecteurs non nuls de E. Par linéarité, on a $f(x+y) = f(x) + f(y)$ et donc

$$\alpha_{x+y}(x+y) = \alpha_x x + \alpha_y y.$$

1. Lorsque x est un vecteur non nul, le scalaire α est déterminé de façon unique. Lorsque x est le vecteur nul, le scalaire α peut être choisi de façon arbitraire.

Poursuivons [1] en discutant selon que la famille (x, y) est libre ou non.

Cas : (x, y) est libre. L'égalité précédente donne la combinaison linéaire nulle

$$(\alpha_{x+y} - \alpha_x)x + (\alpha_{x+y} - \alpha_y)y = 0_E.$$

On en déduit $\alpha_{x+y} - \alpha_x = \alpha_{x+y} - \alpha_y = 0$ donc $\alpha_x = \alpha_y$.

Cas : (x, y) est liée. On peut écrire $y = \lambda x$ avec $\lambda \neq 0$ car x et y sont deux vecteurs non nuls. On vérifie alors directement l'égalité de α_x et α_y car

$$f(y) = \alpha_y y = \alpha_y \lambda x \quad \text{et} \quad f(y) = f(\lambda x) = \lambda . f(x) = \lambda \alpha_x x.$$

Sachant $\lambda \neq 0$ et $x \neq 0_E$, on conclut : $\alpha_x = \alpha_y$.

Ainsi, l'application $x \mapsto \alpha_x$ est constante sur $E \setminus \{0_E\}$. Notons α la valeur de cette constante. Pour tout vecteur x non nul, on a $f(x) = \alpha x$ et cette égalité est aussi valable si x est le vecteur nul. Finalement, f est une homothétie.

Exercice 18 ***

Soit u et v deux endomorphismes d'un espace vectoriel E de dimension finie.
Résoudre l'équation $u \circ f = v$ d'inconnue $f \in \mathcal{L}(E)$.

Solution

S'il existe un endomorphisme f tel que $u \circ f = v$, les valeurs prises par v sont nécessairement des valeurs prises par u, c'est-à-dire $\mathrm{Im}(v) \subset \mathrm{Im}(u)$. Supposons cette condition remplie pour la suite. Pour résoudre l'équation $u \circ f = v$, on souhaite pouvoir inverser l'endomorphisme u.

méthode

Une application linéaire induit un isomorphisme entre tout supplémentaire de son noyau et son image.

Soit S un espace supplémentaire de $\mathrm{Ker}(u)$ dans E et φ l'isomorphisme réalisé par la restriction de u au départ de S et à valeurs dans $\mathrm{Im}(u)$. Considérons ensuite $f_1 = \varphi^{-1} \circ v$. Cette application est bien définie car v prend ses valeurs dans $\mathrm{Im}(u)$ qui est l'espace de définition de φ^{-1}. Par composition d'applications linéaires, on peut aussi affirmer que f_1 est linéaire et l'on peut comprendre f_1 comme un endomorphisme de E. Enfin, puisque φ^{-1} prend ses valeurs dans S et que u se confond avec φ sur S, on vérifie :

$$u \circ f_1 = \underbrace{\varphi \circ \varphi^{-1}}_{=\mathrm{Id}_{\mathrm{Im}(u)}} \circ v = v.$$

Ainsi, f_1 détermine une solution de l'équation $u \circ f = v$.

méthode

L'obtention d'une solution particulière à une équation linéaire permet de ramener l'étude à la résolution d'une équation homogène.

1. Il n'est possible d'identifier les scalaires en facteurs de x et y que si la famille (x, y) est libre.

Pour $f \in \mathcal{L}(E)$,

$$\begin{aligned} u \circ f = v &\iff u \circ f = u \circ f_1 \\ &\iff u \circ (f - f_1) = 0 \\ &\iff \mathrm{Im}(f - f_1) \subset \mathrm{Ker}(u). \end{aligned}$$

Les solutions de l'équation $u \circ f = v$ sont alors les endomorphismes

$$f = f_1 + g \quad \text{avec} \quad g \in \mathcal{L}(E) \text{ tel que } \mathrm{Im}(g) \subset \mathrm{Ker}(u).$$

3.4.3 Produit matriciel

Exercice 19 *

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$. Établir $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Solution

méthode

Le produit MN de deux matrices M et N est possible seulement si le nombre n de colonnes de M est égal au nombre de lignes de N. Le coefficient général de la matrice produit est alors [1]

$$[MN]_{i,j} = \sum_{k=1}^{n} [M]_{i,k} [N]_{k,j}.$$

Les produits matriciels AB et BA sont tous deux possibles et déterminent des matrices carrées. On peut donc calculer les traces étudiées.

D'une part, la matrice AB étant carrée de taille n,

$$\mathrm{tr}(AB) = \sum_{i=1}^{n} [AB]_{i,i} = \sum_{i=1}^{n} \left(\sum_{j=1}^{p} [A]_{i,j} [B]_{j,i} \right). \qquad (*)$$

D'autre part, la matrice BA étant carrée de taille p,

$$\mathrm{tr}(BA) = \sum_{j=1}^{p} [BA]_{j,j} = \sum_{j=1}^{p} \left(\sum_{i=1}^{n} [B]_{j,i} [A]_{i,j} \right). \qquad (**)$$

Les termes sommés dans $(*)$ et $(**)$ sont identiques et il suffit d'échanger les deux sommes pour pouvoir affirmer $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Exercice 20 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Pour $M \in \mathcal{M}_n(\mathbb{R})$, on pose $\varphi(M) = AM - MA$.

(a) Vérifier que φ définit un endomorphisme de $\mathcal{M}_n(\mathbb{R})$.

(b) Calculer la trace de φ.

1. $[A]_{i,j}$ est une notation possible pour désigner le coefficient d'indice (i,j) d'une matrice A.

Solution

(a) L'application φ est bien définie de $\mathcal{M}_n(\mathbb{R})$ vers lui-même. Pour tous $\lambda_1, \lambda_2 \in \mathbb{R}$ et $M_1, M_2 \in \mathcal{M}_n(\mathbb{R})$ on vérifie $\varphi(\lambda_1 M_1 + \lambda_2 M_2) = \lambda_1 \varphi(M_1) + \lambda_2 \varphi(M_2)$ car

$$A(\lambda_1 M_1 + \lambda_2 M_2) - (\lambda_1 M_1 + \lambda_2 M_2)A = \lambda_1(AM_1 - M_1 A) + \lambda_2(AM_2 - M_2 A).$$

(b) **méthode**

La trace d'un endomorphisme est la trace d'une matrice figurant celui-ci.

Étudions la matrice représentant φ dans la base canonique de $\mathcal{M}_n(\mathbb{R})$ constituée des matrices élémentaires [1] $E_{i,j}$. En introduisant les coefficients $a_{i,j}$ de la matrice A, on peut écrire

$$A = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} E_{i,j} \right).$$

méthode

On obtient le produit de deux matrices élémentaires [2] par la formule [3] :

$$E_{i,j} E_{k,\ell} = \delta_{j,k} E_{i,\ell} \quad \text{avec} \quad \delta_{j,k} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

Pour $(k, \ell) \in [\![1\,;n]\!]^2$

$$\varphi(E_{k,\ell}) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} \underbrace{E_{i,j} E_{k,\ell}}_{=\delta_{j,k} E_{i,\ell}} \right) - \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} \underbrace{E_{k,\ell} E_{i,j}}_{=\delta_{\ell,i} E_{k,j}} \right)$$

$$= \sum_{i=1}^{n} a_{i,k} E_{i,\ell} - \sum_{j=1}^{n} a_{\ell,j} E_{k,j}.$$

Les coefficients diagonaux de la matrice figurant φ dans la base des matrices élémentaires sont les coordonnées des $\varphi(E_{k,\ell})$ selon $E_{k,\ell}$ à savoir les $a_{k,k} - a_{\ell,\ell}$. La trace de φ est donc [4]

$$\operatorname{tr}(\varphi) = \sum_{k=1}^{n} \left(\sum_{\ell=1}^{n} (a_{k,k} - a_{\ell,\ell}) \right) = n \operatorname{tr}(A) - n \operatorname{tr}(A) = 0$$

car

$$\sum_{k=1}^{n} \left(\sum_{\ell=1}^{n} a_{k,k} \right) = \sum_{k=1}^{n} n a_{k,k} = n \operatorname{tr}(A) \quad \text{et} \quad \sum_{k=1}^{n} \left(\sum_{\ell=1}^{n} a_{\ell,\ell} \right) = \sum_{k=1}^{n} \operatorname{tr}(A) = n \operatorname{tr}(A).$$

1. Une matrice élémentaire $E_{i,j}$ est une matrice dont tous les coefficients sont nuls sauf celui d'indice (i,j) qui vaut 1.

2. Ces matrices élémentaires doivent être de types compatibles : $E_{i,j}$ de type (n,p), $E_{k,\ell}$ de type (p,q), le produit étant quant à lui de type (n,q). Voir sujet 1 du chapitre 9 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

3. $\delta j,k$ est un *symbole de Kronecker*.

4. En fait, les endomorphismes $M \mapsto AM$ et $M \mapsto MA$ sont tous deux de trace $n \operatorname{tr}(A)$. Par des calculs semblables, on peut aussi établir que la trace de l'endomorphisme $M \mapsto AMA$ est $\left(\operatorname{tr}(A)\right)^2$.

3.4.4 Ensemble de matrices

Exercice 21 * (Le « corps » des quaternions)
On note $\mathbb{H}$ l'ensemble des matrices

$$M(a,b) = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \quad \text{avec} \quad a, b \in \mathbb{C}.$$

(a) Montrer que $\mathbb{H}$ est une algèbre réelle de dimension 4 pour les opérations usuelles.

(b) Vérifier que tout élément non nul de $\mathbb{H}$ est inversible dans $\mathbb{H}$.

Solution

(a) **méthode**

‖ On vérifie que $\mathbb{H}$ est une sous-algèbre de l'algèbre réelle $\mathcal{M}_2(\mathbb{C})$.

$\mathcal{M}_2(\mathbb{C})$ est une $\mathbb{C}$-algèbre de dimension 4 donc aussi une $\mathbb{R}$-algèbre de dimension 8.
L'ensemble $\mathbb{H}$ est une partie de $\mathcal{M}_2(\mathbb{C})$ contenant le neutre multiplicatif I_2 (obtenu pour $a = 1$ et $b = 0$). De plus, pour $a, b, c, d \in \mathbb{C}$ et $\lambda, \mu \in \mathbb{R}$, on vérifie par le calcul

$$\lambda M(a,b) + \mu M(c,d) = M(\lambda a + \mu c, \lambda b + \mu d) \in \mathbb{H}$$
$$M(a,b)M(c,d) = M(ac - b\overline{d}, ad + b\overline{c}) \in \mathbb{H}.$$

Ainsi, $\mathbb{H}$ est une sous-algèbre de la $\mathbb{R}$-algèbre $\mathcal{M}_2(\mathbb{C})$ et c'est donc une $\mathbb{R}$-algèbre. Enfin, en introduisant les parties réelles et imaginaires des complexes a et b, on peut écrire

$$M(a,b) = t\mathrm{I}_2 + xJ + yK + zL$$

avec $t = \mathrm{Re}(a)$, $x = \mathrm{Im}(a)$, $y = \mathrm{Re}(b)$, $z = \mathrm{Im}(b)$ et

$$J = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{et} \quad L = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$$

$\mathbb{H}$ est l'ensemble des combinaisons linéaires réelles des quatre matrices[1] I_2, J, K et L. Ces dernières étant linéairement indépendantes, $\mathbb{H}$ est un espace réel de dimension 4.

(b) Soit $(a,b) \in \mathbb{C}^2$. Étudions l'inversibilité de $M(a,b)$ dans $\mathbb{H}$.
Si $M(a,b) \neq \mathrm{O}_2$, on a $(a,b) \neq (0,0)$ et $\det\big(M(a,b)\big) = |a|^2 + |b|^2 \neq 0$. On en déduit que la matrice $M(a,b)$ est inversible dans $\mathcal{M}_2(\mathbb{C})$. De plus, en introduisant sa comatrice, on peut exprimer l'inverse de $M(a,b)$ puis vérifier son appartenance à $\mathbb{H}$:

$$\big(M(a,b)\big)^{-1} = \frac{1}{\det\big(M(a,b)\big)}\,{}^t\Big(\mathrm{Com}\big(M(a,b)\big)\Big) = \frac{1}{|a|^2 + |b|^2}\begin{pmatrix} \overline{a} & -b \\ \overline{b} & a \end{pmatrix}$$
$$= M\left(\frac{\overline{a}}{|a|^2 + |b|^2}, \frac{-b}{|a|^2 + |b|^2}\right) \in \mathbb{H}.$$

1. On a les relations remarquables $J^2 = K^2 = L^2 = -\mathrm{I}_2$, $JK = L$, $KL = J$ et $LJ = K$.

Finalement, tout élément non nul de l'algèbre $\mathbb{H}$ est inversible [1].

Exercice 22 ** (Matrices de permutation)

Soit n un entier au moins égal à 2. On appelle *matrice de permutation* associée à une permutation σ de $\mathcal{S}_n$, la matrice déterminée par

$$P(\sigma) = \left(\delta_{i,\sigma(j)}\right)_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R}) \quad \text{avec} \quad \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

(a) Vérifier $P(\sigma \circ \sigma') = P(\sigma)P(\sigma')$ pour tous σ et σ' dans $\mathcal{S}_n$.

(b) Calculer le déterminant d'une matrice de permutation.

(c) Justifier ${}^t\big(P(\sigma)\big) = P\big(\sigma^{-1}\big)$ pour toute permutation $\sigma \in \mathcal{S}_n$.

Solution

(a) **méthode**

‖ On interprète [2] $P(\sigma)$ comme la matrice d'un endomorphisme simple à décrire.

Soit $\sigma \in \mathcal{S}_n$ et u_σ l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à $P(\sigma)$. En notant $e = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$, on a $u_\sigma(e_j) = e_{\sigma(j)}$ pour tout $1 \leqslant j \leqslant n$.

Pour σ et $\sigma' \in \mathcal{S}_n$, l'endomorphisme canoniquement associé à $P(\sigma)P(\sigma')$ est $u_\sigma \circ u_{\sigma'}$. Or, pour tout $j \in [\![1\,;n]\!]$,

$$u_\sigma \circ u_{\sigma'}(e_j) = u_\sigma\big(e_{\sigma'(j)}\big) = e_{\sigma(\sigma'(j))} = e_{\sigma\circ\sigma'(j)}.$$

Une application linéaire étant complètement caractérisée par l'image d'une base, on obtient $u_\sigma \circ u_{\sigma'} = u_{\sigma\circ\sigma'}$ puis $P(\sigma)P(\sigma') = P(\sigma \circ \sigma')$.

(b) Soit $\sigma \in \mathcal{S}_n$. Le déterminant de la matrice $P(\sigma)$ est aussi celui de u_σ. Par la propriété d'antisymétrie du déterminant d'une famille de vecteurs, on obtient

$$\det\big(P(\sigma)\big) = \det(u_\sigma) = \det_e\big(e_{\sigma(1)}, \ldots, e_{\sigma(n)}\big) = \varepsilon(\sigma) \underbrace{\det_e(e_1, \ldots, e_n)}_{=1} = \varepsilon(\sigma).$$

(c) Soit $\sigma \in \mathcal{S}_n$.

méthode

‖ L'inverse d'une matrice orthogonale est sa transposée.

La matrice $P(\sigma)$ est orthogonale car ses colonnes sont unitaires et deux à deux orthogonales. La transposée de $P(\sigma)$ est donc son inverse. Aussi, on a $P(\sigma)P\big(\sigma^{-1}\big) = P(\mathrm{Id}) = \mathrm{I}_n$ et $P\big(\sigma^{-1}\big)$ est l'inverse [3] de $P(\sigma)$. On en déduit l'égalité proposée.

1. La multiplication sur $\mathbb{H}$ n'est pas commutative et donc $\mathbb{H}$ n'est pas un corps dans le sens où ce concept est défini dans le cours.

2. On peut aussi calculer le coefficient général de la matrice produit $P(\sigma)P(\sigma')$ et y simplifier le produit des symboles de Kronecker.

3. L'application qui à $\sigma \in \mathcal{S}_n$ associe $P(\sigma) \in \mathrm{GL}_n(\mathbb{R})$ est un morphisme de groupes : il transforme l'inverse en l'inverse.

Exercice 23 * (Matrice stochastique)**

On dit qu'une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ est *stochastique* si tous ses coefficients sont positifs et si

$$\sum_{j=1}^{n} a_{i,j} = 1 \quad \text{pour tout } i \in [\![1\,;n]\!].$$

(a) Montrer que l'ensemble des matrices stochastiques de taille n est stable pour la multiplication des matrices.

(b) À quelle condition une matrice stochastique est-elle inversible tout en ayant pour inverse une matrice stochastique ?

Solution

(a) Soit $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices stochastiques de taille n. Vérifions que la matrice $C = AB = (c_{i,j})$ est aussi stochastique.

Pour tous i et j dans $[\![1\,;n]\!]$, on a

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}.$$

Par somme de produits de réels positifs, on peut affirmer que les coefficients de C sont tous positifs. Au surplus, pour tout $i \in [\![1\,;n]\!]$, on peut écrire en échangeant les deux sommes

$$\sum_{j=1}^{n} c_{i,j} = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,j} \right) = \sum_{k=1}^{n} \left(a_{i,k} \underbrace{\sum_{j=1}^{n} b_{k,j}}_{=1} \right) = \sum_{k=1}^{n} a_{i,k} = 1.$$

La matrice C est donc stochastique[1].

(b) *Analyse* : Supposons la matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ stochastique et inversible d'inverse $B = (b_{i,j})$ elle aussi stochastique. Pour tous i et $j \in [\![1\,;n]\!]$, on a

$$\sum_{k=1}^{n} a_{i,k} b_{k,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

méthode

‖ On montre que la matrice A comporte un et un seul 1 sur chaque ligne.

Soit $i \in [\![1\,;n]\!]$. Pour tout $k \in [\![1\,;n]\!]$, on a $a_{i,k} \geqslant 0$ et $b_{k,i} \leqslant 1$ donc $a_{i,k} b_{k,i} \leqslant a_{i,k}$. En sommant ces inégalités pour k allant de 1 à n, on obtient

$$1 = \sum_{k=1}^{n} a_{i,k} b_{k,i} \leqslant \sum_{k=1}^{n} a_{i,k} = 1.$$

1. En notant J la colonne de hauteur n dont tous les coefficients sont égaux à 1, la condition sur la somme des coefficients de A se relit $AJ = J$ ce qui permet une justification rapide : $ABJ = AJ = J$.

Les inégalités $a_{i,k}b_{k,i} \leqslant a_{i,k}$ sont donc toutes des égalités et l'on peut affirmer que, pour tous les indices i et k de $[\![1\,;n]\!]$, on a $a_{i,k} = 1$ ou $b_{k,i} = 0$. Puisque la matrice B est inversible, elle ne possède pas de colonne nulle et, donc, pour chaque i de $[\![1\,;n]\!]$, il existe k dans $[\![1\,;n]\!]$ tel que $b_{k,i} \neq 0$ ce qui entraîne $a_{i,k} = 1$. De plus, les coefficients de la i-ème ligne de A sont positifs et de somme 1 : en dehors de l'élément d'indice k, tous les autres éléments de la i-ème ligne de A sont nuls.

Résumons, pour chaque indice i de $[\![1\,;n]\!]$, il existe k dans $[\![1\,;n]\!]$ tel que $a_{i,k} = 1$ et $a_{i,j} = 0$ pour tout $j \in [\![1\,;n]\!]$ tel que $j \neq k$. Posons alors $\sigma(i) = k$ ce qui détermine une application σ de $[\![1\,;n]\!]$ vers lui-même telle que $a_{i,j} = \delta_{\sigma(i),j}$ pour tous les indices i et j. De plus, l'application σ est injective car la matrice A ne peut posséder deux lignes identiques puisqu'elle est inversible. L'application σ est donc une permutation de $[\![1\,;n]\!]$ vérifiant $A = \left(\delta_{\sigma(i),j}\right)$.

Synthèse : Inversement, une telle matrice est inversible et son inverse[1] est $B = \left(\delta_{i,\sigma(j)}\right)$ qui est une matrice stochastique.

Finalement, les matrices stochastiques inversibles d'inverse stochastique sont les matrices de permutation.

3.4.5 Rang d'une matrice

Exercice 24 *

Soit G une partie de $\mathcal{M}_n(\mathbb{K})$ telle que $(G, \times)$ soit un groupe[2].

Que peut-on dire du rang des matrices de G ?

Solution

méthode

‖ Le rang commun des éléments de G est le rang du neutre de $(G, \times)$.

Notons J l'élément neutre du groupe $(G, \times)$ et considérons A un élément quelconque de G. Puisque le rang d'un produit de deux matrices est inférieur aux rangs des facteurs qui le constituent, on a

$$\mathrm{rg}(A) = \mathrm{rg}(AJ) \leqslant \mathrm{rg}(J).$$

De plus, la matrice A est inversible[3] dans le groupe $(G, \times)$. On peut donc introduire une matrice $B \in G$ telle que $AB = J$ et alors

$$\mathrm{rg}(J) = \mathrm{rg}(AB) \leqslant \mathrm{rg}(A).$$

On en déduit $\mathrm{rg}(A) = \mathrm{rg}(J)$: tous les éléments de $(G, \times)$ ont le même rang.

1. Voir le sujet précédent, A est la matrice de la permutation σ^{-1}.

2. On ne suppose pas *a priori* que G soit un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$. En particulier, le neutre du groupe $(G, \times)$ peut être différent de la matrice I_n.

3. On ignore si la matrice A est inversible dans $\mathcal{M}_n(\mathbb{K})$ et elle ne l'est pas quand $J \neq \mathrm{I}_n$.

Exercice 25 *

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice de rang r. Montrer qu'il existe des matrices B et C, respectivement dans $\mathcal{M}_{n,r}(\mathbb{K})$ et $\mathcal{M}_{r,p}(\mathbb{K})$, telles que $A = BC$.

Solution

Puisque de rang r, la matrice A est équivalente[1] à la matrice

$$J_r = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

où les 0 désignent des blocs nuls de tailles appropriées. On peut donc introduire des matrices $P \in \mathrm{GL}_p(\mathbb{K})$ et $Q \in \mathrm{GL}_n(\mathbb{K})$ telles que $A = QJ_rP^{-1}$.

méthode

On décompose la matrice J_r sous la forme demandée et l'on adapte cette écriture à la matrice A.

Introduisons les matrices

$$K = \begin{pmatrix} \mathrm{I}_r \\ 0 \end{pmatrix} \in \mathcal{M}_{n,r}(\mathbb{K}) \quad \text{et} \quad L = \begin{pmatrix} \mathrm{I}_r & 0 \end{pmatrix} \in \mathcal{M}_{r,p}(\mathbb{K}).$$

Par produit par blocs, on vérifie $J_r = KL$ et, en introduisant $B = QK \in \mathcal{M}_{n,r}(\mathbb{K})$ et $C = LP^{-1} \in \mathcal{M}_{r,p}(\mathbb{K})$, on obtient l'écriture $A = BC$ avec B et C qui ont les types voulus[2].

Exercice 26 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice de rang r. Déterminer la dimension de l'espace

$$\left\{B \in \mathcal{M}_n(\mathbb{R}) \;\middle|\; ABA = \mathrm{O}_n\right\}.$$

Solution

L'ensemble étudié est le noyau de l'endomorphisme $B \mapsto ABA$ défini sur $\mathcal{M}_n(\mathbb{R})$, il s'agit bien d'un espace vectoriel comme l'affirme l'énoncé du sujet.

méthode

On transforme l'étude en une étude équivalente où la matrice A devient la matrice J_r.

On introduit des matrices P et $Q \in \mathrm{GL}_n(\mathbb{R})$ vérifiant $A = QJ_rP^{-1}$ et alors

$$\begin{aligned} ABA = \mathrm{O}_n &\iff QJ_rP^{-1}BQJ_rP^{-1} = \mathrm{O}_n \\ &\iff J_r \underbrace{P^{-1}BQ}_{=M} J_r = \mathrm{O}_n. \end{aligned}$$

1. Deux matrices A, B de même type sont équivalentes dans $\mathcal{M}_n(\mathbb{K})$ lorsqu'il existe des matrices inversibles P et Q de $\mathcal{M}_n(\mathbb{K})$ telles que $A = QBP^{-1}$. Il revient au même de dire que les matrices A et B ont le même rang.

2. Lorsque la matrice A est de rang 1, on obtient l'écriture $A = Y\,{}^tX$ avec X et Y colonnes : on pourra comparer cette résolution à celle du sujet 25 du chapitre 9 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

L'application $B \mapsto P^{-1}BQ$ définit un isomorphisme de $\{B \in \mathcal{M}_n(\mathbb{R}) \mid ABA = \mathrm{O}_n\}$ vers $\{M \in \mathcal{M}_n(\mathbb{R}) \mid J_r M J_r = \mathrm{O}_n\}$.

méthode

On détermine les matrices $M \in \mathcal{M}_n(\mathbb{R})$ vérifiant $J_r M J_r = \mathrm{O}_n$ en raisonnant par blocs.

On écrit la matrice M par blocs

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{pmatrix} \quad \text{avec} \quad M_{1,1} \in \mathcal{M}_r(\mathbb{R}).$$

On obtient alors après calculs par blocs

$$J_r M J_r = \mathrm{O}_r \iff M_{1,1} = \mathrm{O}_r.$$

Les matrices M vérifiant $J_r M J_r = \mathrm{O}_n$ sont donc celles de la forme

$$\begin{pmatrix} \mathrm{O}_r & * \\ * & * \end{pmatrix}.$$

Ces matrices constituent un espace de dimension [1] $n^2 - r^2$ et par isomorphisme

$$\dim\{B \in \mathcal{M}_n(\mathbb{R}) \mid ABA = \mathrm{O}_n\} = n^2 - r^2.$$

3.4.6 Matrices semblables

Exercice 27 *

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice non nulle telle que les espaces $\mathrm{Im}(A)$ et $\mathrm{Ker}(A)$ sont supplémentaires. Montrer que la matrice A est semblable à une matrice de la forme

$$\begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix} \quad \text{avec} \quad A' \in \mathrm{GL}_r(\mathbb{R})$$

(où les 0 désignent des blocs nuls de tailles appropriées).

Solution

Soit a l'endomorphisme de $E = \mathbb{R}^n$ canoniquement associé à la matrice A. Par hypothèse, on a $E = \mathrm{Im}(a) \oplus \mathrm{Ker}(a)$.

méthode

On introduit [2] une base adaptée à la supplémentarité de $\mathrm{Im}(a)$ et $\mathrm{Ker}(a)$.

1. On détermine la dimension de cet espace en dénombrant les matrices élémentaires qui en constituent une base.

2. Le cadre hypothétique invite directement à introduire une telle base. Cependant, une « petite » analyse est aussi possible et fait observer que r doit être le rang de A, que l'image de a est incluse dans l'espace engendré par les r premiers vecteurs de la base et que les derniers vecteurs doivent appartenir au noyau de a.

Posons r le rang de a, $(e_1, \dots, e_r)$ une base de l'image de a et $(e_{r+1}, \dots, e_n)$ une base de $\mathrm{Ker}(a)$. La famille $e = (e_1, \dots, e_n)$ est une base de E adaptée à la supplémentarité de $\mathrm{Im}(a)$ et $\mathrm{Ker}(a)$. Formons la matrice de a dans cette base.

Pour tout $j \in [\![1\,;r]\!]$, les vecteurs $a(e_j)$ appartiennent à l'image de a et sont donc combinaisons linéaires des vecteurs $e_1, \dots, e_r$: leurs coordonnées selon les vecteurs $e_{r+1}, \dots, e_n$ sont toutes nulles.

Pour tout $j \in [\![r+1\,;n]\!]$, les vecteurs $a(e_j)$ sont nuls donc de coordonnées nulles.

La matrice de a dans la base e est donc de la forme

$$\begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix} \quad \text{avec} \quad A' \in \mathcal{M}_r(\mathbb{R}).$$

Enfin, cette matrice est de rang $r = \mathrm{rg}(a)$ et donc[1] $\mathrm{rg}(A') = r$ ce qui assure que la matrice A' est inversible[2].

Finalement, A est semblable à une matrice de la forme proposée avec $r = \mathrm{rg}(A)$.

Exercice 28 **

Soit $A \in \mathcal{M}_n(\mathbb{K})$ vérifiant $A^{n-1} \neq \mathrm{O}_n$ et $A^n = \mathrm{O}_n$. Établir que A est semblable à la matrice

$$B = \begin{pmatrix} 0 & & & (0) \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ (0) & & 1 & 0 \end{pmatrix}.$$

Solution

Soit a l'endomorphisme de $E = \mathbb{K}^n$ canoniquement associé à la matrice A. Les hypothèses $A^n = \mathrm{O}_n$ et $A^{n-1} \neq \mathrm{O}_n$ se traduisent par $a^n = 0$ et $a^{n-1} \neq 0$.

Analyse : Supposons qu'il existe une base $e = (e_1, \dots, e_n)$ dans laquelle la matrice de a soit égale à B. Sur les colonnes de B on lit

$$a(e_1) = e_2, \quad a(e_2) = e_3, \ \dots, \ a(e_{n-1}) = e_n \quad \text{et} \quad a(e_n) = 0_E. \tag{$*$}$$

Ainsi, le choix du vecteur e_1 détermine l'ensemble des autres vecteurs de la base

$$\begin{cases} e_2 = a(e_1) \\ e_3 = a^2(e_1) \\ \quad \vdots \\ e_n = a^{n-1}(e_1). \end{cases} \tag{$**$}$$

méthode

Le vecteur e_1 doit être choisi en dehors[3] de $\mathrm{Ker}\left(a^{n-1}\right)$ car aucun des vecteurs $e_2, \dots, e_n$ ne doit être nul pour que la famille e puisse être libre.

1. On ne modifie pas le rang d'une matrice lorsque l'on retire des rangées nulles de celle-ci.

2. Puisque $\mathrm{Im}(a)$ est un supplémentaire de $\mathrm{Ker}(a)$, l'endomorphisme a induit un isomorphisme de $\mathrm{Im}(a)$ vers lui-même : la matrice A' figure cet isomorphisme.

Le vecteur e_1 doit aussi être choisi tel que $a(e_n)$ soit nul mais cette condition est automatiquement remplie car a^n est l'endomorphisme nul.

Vérifions maintenant qu'il est possible de construire une telle base.

Synthèse : Soit e_1 un vecteur de E n'appartenant pas à $\mathrm{Ker}(a^{n-1})$ (ce qui est possible car a^{n-1} est non nul). Introduisons ensuite les vecteurs $e_2, \ldots, e_n$ définis par (**). Par construction, les égalités (*) sont toutes satisfaites. Il reste à justifier que la famille $e = (e_1, \ldots, e_n)$ est une base de $\mathbb{K}^n$. Il s'agit d'une famille de longueur n dans un espace de dimension n, il suffit d'établir sa liberté.

Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ tel que

$$\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n = 0_E.$$

Ceci revient à écrire

$$\lambda_1 e_1 + \lambda_2 a(e_1) + \cdots + \lambda_n a^{n-1}(e_1) = 0_E. \qquad (\triangle)$$

En appliquant a aux deux membres de l'identité $(\triangle)$, cette équation se réduit à

$$\lambda_1 a(e_1) + \lambda_2 a^2(e_1) + \cdots + \lambda_{n-1} a^{n-1}(e_1) = 0_E \quad \text{car} \quad a^n(e_1) = 0_E.$$

En répétant plusieurs fois cette manipulation, on forme le système suivant :

$$\begin{cases} \lambda_1 e_1 + \lambda_2 a(e_1) + \cdots + \lambda_{n-1} a^{n-2}(e_1) + \lambda_n a^{n-1}(e_1) = 0_E \\ \lambda_1 a(e_1) + \cdots + \lambda_{n-2} a^{n-2}(e_1) + \lambda_{n-1} a^{n-1}(e_1) = 0_E \\ \ddots \qquad\qquad \vdots \\ \lambda_1 a^{n-2}(e_1) + \lambda_2 a^{n-1}(e_1) = 0_E \\ \lambda_1 a^{n-1}(e_1) = 0_E. \end{cases}$$

Sachant $a^{n-1}(e_1)$ non nul, la dernière équation donne $\lambda_1 = 0$. Ceci permet de simplifier l'équation précédente et d'obtenir $\lambda_2 = 0$. Ainsi, de proche en proche, on acquiert la nullité de tous les λ_i : la famille e est libre et c'est donc une base de E.

Finalement, la famille e est une base dans laquelle l'endomorphisme a est figuré par B, on peut affirmer que les matrices A et B sont semblables.

Exercice 29 **

Soit $A \in \mathcal{M}_3(\mathbb{R})$ une matrice non nulle vérifiant $A^3 + A = O_3$. Montrer que A est semblable à la matrice

$$B = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

3. On sait $\mathrm{Ker}(a) \subset \mathrm{Ker}(a^2) \subset \ldots \mathrm{Ker}(a^{n-1})$: choisir $e_1 \notin \mathrm{Ker}(a^{n-1})$, assure non seulement la non-nullité de e_n mais aussi celles de e_2, e_3, etc.

Solution

Soit a l'endomorphisme de $E = \mathbb{R}^3$ canoniquement associé à la matrice A. Cet endomorphisme est non nul et vérifie $a^3 + a = 0$.

Analyse : Supposons qu'il existe une base $e = (e_1, e_2, e_3)$ dans laquelle B est la matrice de a. Par les colonnes de la matrice B, on lit

$$a(e_1) = e_2, \quad a(e_2) = -e_1 \quad \text{et} \quad a(e_3) = 0_E. \qquad (*)$$

On en déduit

$$e_1 \in \operatorname{Ker}\left(a^2 + \mathrm{Id}_E\right), \quad e_2 = a(e_1) \quad \text{et} \quad e_3 \in \operatorname{Ker}(a).$$

méthode

‖ On vérifie que les noyaux de $a^2 + \mathrm{Id}_E$ et a ne sont pas réduits au vecteur nul.

D'une part, $a^3 + a = \left(a^2 + \mathrm{Id}_E\right) \circ a = 0$ et donc $\operatorname{Im}(a) \subset \operatorname{Ker}\left(a^2 + \mathrm{Id}_E\right)$. Or l'endomorphisme a n'est pas nul et par conséquent $\operatorname{Ker}\left(a^2 + \mathrm{Id}_E\right) \neq \{0_E\}$.

D'autre part, si par l'absurde $\operatorname{Ker}(a)$ est réduit au vecteur nul, l'endomorphisme a est bijectif et l'on peut simplifier l'identité $a^3 + a = 0$ en composant par a^{-1} pour obtenir l'égalité $a^2 + \mathrm{Id}_E = 0$. En exploitant le déterminant, on obtient l'absurdité

$$\det\left(a^2\right) = \det(-\mathrm{Id}_E) \underset{E=\mathbb{R}^3}{=} (-1)^3 = -1 \quad \text{et} \quad \det\left(a^2\right) = \left(\det(a)\right)^2 \geqslant 0.$$

On peut donc affirmer que $\operatorname{Ker}(a)$ n'est pas réduit au vecteur nul.

Synthèse : Les espaces $\operatorname{Ker}\left(a^2 + \mathrm{Id}_E\right)$ et $\operatorname{Ker}(a)$ n'étant pas réduits au vecteur nul, on peut introduire des vecteurs non nuls e_1 et e_3 leur appartenant. Posons de plus $e_2 = a(e_1)$. Ces trois vecteurs vérifient les relations $(*)$. Il reste à justifier que la famille $e = (e_1, e_2, e_3)$ constitue une base de E. Puisqu'il s'agit d'une famille de trois vecteurs en dimension 3, il suffit d'en étudier la liberté.

Soit $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ tel que

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0_E. \qquad (\triangle)$$

En appliquant deux fois a à cette relation, on obtient

$$\begin{cases} \lambda_1 e_2 - \lambda_2 e_1 = 0_E & (1) \\ -\lambda_1 e_1 - \lambda_2 e_2 = 0_E & (2). \end{cases}$$

La combinaison d'équations $\lambda_2(1) + \lambda_1(2)$ donne $-\left(\lambda_1^2 + \lambda_2^2\right)e_1 = 0_E$. Puisque le vecteur e_1 a été choisi non nul, on obtient $\lambda_1^2 + \lambda_2^2 = 0$ puis [1] $\lambda_1 = \lambda_2 = 0$. La relation $(\triangle)$ se simplifie alors en $\lambda_3 e_3 = 0_E$ ce qui donne $\lambda_3 = 0$ puisque le vecteur e_3 a été choisi non nul.

Finalement, la famille $e = (e_1, e_2, e_3)$ est une base de E et la matrice de a dans celle-ci est celle voulue : les matrices A et B sont semblables.

Exercice 30 ***

Soit A et B dans $\mathcal{M}_n(\mathbb{R})$ telles qu'il existe $P \in \mathrm{GL}_n(\mathbb{C})$ vérifiant $A = PBP^{-1}$.
Montrer [2] qu'il existe $Q \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = QBQ^{-1}$.

1. Car λ_1 et λ_2 sont des réels : le sujet n'est pas valable dans le cadre des matrices complexes, la matrice iI_3 vérifie l'hypothèse sans vérifier la conclusion.

2. Les matrices réelles A et B sont semblables sur $\mathbb{C}$, il s'agit ici de montrer qu'elles sont aussi semblables sur $\mathbb{R}$.

Solution

L'égalité $A = PBP^{-1}$ donne $AP = PB$ dans laquelle nous allons identifier les parties réelles et imaginaires. Notons R la matrice formée des parties réelles des coefficients de P et S celle formée des parties imaginaires. L'égalité $AP = PB$ avec $P = R + \mathrm{i}S$ donne

$$AR + \mathrm{i}AS = BR + \mathrm{i}BS.$$

Les matrices A et B étant réelles, on peut identifier les parties réelles et imaginaires des deux membres pour obtenir $AR = BR$ et $AS = BS$ avec R et S matrices réelles. Cependant, bien que la matrice P soit inversible, rien n'assure que R ou S le soit [1].

méthode

‖ On définit Q inversible de la forme $Q = R + xS$ avec $x \in \mathbb{R}$ bien choisi.

Pour tout $x \in \mathbb{R}$, on a $AR + xAS = RB + xSB$ ce qui donne $AQ = BQ$. Aussi, l'application $x \mapsto \det(R + xS)$ est une fonction polynomiale [2]. Celle-ci n'est pas nulle car on peut l'évaluer le polynôme associé en $x = \mathrm{i}$ et l'on sait $\det(R+\mathrm{i}S) = \det(P) \neq 0$. Il existe donc un réel x tel que $\det(R + xS) \neq 0$. On peut alors conclure : la matrice $Q = R + xS$ est inversible, réelle et permet d'écrire $A = QBQ^{-1}$.

Exercice 31 ***

Soit $A \in \mathcal{M}_n(\mathbb{R})$ de trace nulle. Montrer que A est semblable à une matrice de diagonale nulle.

Solution

méthode

‖ On raisonne par récurrence sur la taille de la matrice.

Montrons par récurrence sur $n \in \mathbb{N}^*$ que toute matrice de $\mathcal{M}_n(\mathbb{R})$ de trace nulle est semblable à une matrice de diagonale nulle.

Pour $n = 1$, A est une matrice carrée de taille 1 et de trace nulle, c'est la matrice nulle et elle est donc de la forme voulue. Supposons la propriété vraie au rang $n \geqslant 1$. Soit $A \in \mathcal{M}_{n+1}(\mathbb{R})$ une matrice de trace nulle et f l'endomorphisme de $E = \mathbb{R}^{n+1}$ canoniquement associé. Cet endomorphisme est de trace nulle.

Si f est une homothétie vectorielle, c'est-à-dire s'il existe $\lambda \in \mathbb{R}$ tel que $f = \lambda \mathrm{Id}_E$, alors $\mathrm{tr}(f) = (n+1)\lambda$ et donc $\lambda = 0$ puis $f = 0$. La matrice A est alors nulle et la propriété voulue est immédiate.

Sinon, il existe [3] au moins un vecteur $e_1 \in E$ tel que $f(e_1)$ n'est pas colinéaire à e_1. Posons alors $e_2 = f(e_1)$. La famille (e_1, e_2) est libre et peut être complétée en une base de E. La matrice de f dans celle-ci s'exprime

$$B = \begin{pmatrix} 0 & * & \cdots & * \\ 1 & & & \\ 0 & & A' & \\ \vdots & & & \\ 0 & & & \end{pmatrix} \quad \text{avec} \quad A' \in \mathcal{M}_n(\mathbb{R}).$$

1. La matrice $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{smallmatrix}\right)$ est inversible mais ni sa partie réelle ni sa partie imaginaire le sont.
2. On s'en convainc par la formule définissant le déterminant ou en calculant celui-ci par développements successifs selon une rangée.
3. Voir sujet 17 p. 86.

La matrice A est semblable à B et $\mathrm{tr}(A') = \mathrm{tr}(B) = \mathrm{tr}(A) = 0$. Par hypothèse de récurrence, la matrice A' est semblable à une matrice de diagonale nulle et l'on peut écrire

$$P^{-1}A'P = \begin{pmatrix} 0 & & (*) \\ & \ddots & \\ (*) & & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}) \quad \text{avec} \quad P \in \mathrm{GL}_n(\mathbb{R}).$$

On introduit alors la matrice Q carrée de taille $n+1$ définie par blocs :

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}.$$

Puisque P est inversible, la matrice Q est inversible avec

$$Q^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix}.$$

Un produit par blocs permet de vérifier

$$Q^{-1}BQ = \begin{pmatrix} 0 & & (*) \\ & \ddots & \\ (*) & & 0 \end{pmatrix}.$$

La matrice B est donc semblable à une matrice de diagonale nulle et par conséquent A aussi.

La récurrence est établie.

Exercice 32 **

Soit $D \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonale à coefficients diagonaux deux à deux distincts et φ l'endomorphisme de $\mathcal{M}_n(\mathbb{R})$ défini par $\varphi(X) = DX - XD$.

(a) Déterminer le noyau et l'image de l'endomorphisme φ.

(b) En exploitant le résultat du sujet précédent, montrer que, si $M \in \mathcal{M}_n(\mathbb{R})$ est une matrice de trace nulle, il existe deux matrices A et B dans $\mathcal{M}_n(\mathbb{R})$ telles que

$$M = AB - BA.$$

Solution

(a) Notons $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux de la matrice D. Soit X dans $\mathcal{M}_n(\mathbb{R})$ une matrice de coefficients $x_{i,j}$.

méthode

‖ On exprime le coefficient général de la matrice $\varphi(X)$.

D'une part [1], $DX = (\lambda_i x_{i,j})$ et, d'autre part, $XD = (\lambda_j x_{i,j})$. On a donc

$$\varphi(X) = \big((\lambda_i - \lambda_j)x_{i,j}\big).$$

1. Lorsque l'on multiplie X par D à gauche, on opère sur les lignes de X, lorsque l'on multiplie à droite, on opère sur les colonnes.

Déterminons le noyau de φ.

$$\begin{aligned} X \in \mathrm{Ker}(\varphi) &\iff \forall (i,j) \in [\![1\,;n]\!]^2,\ (\lambda_i - \lambda_j)x_{i,j} = 0 \\ &\iff \forall (i,j) \in [\![1\,;n]\!]^2,\ \big(i \neq j \implies x_{i,j} = 0\big) \end{aligned}$$

car les $\lambda_1, \ldots, \lambda_n$ sont supposés deux à deux distincts. Le noyau de φ est donc constitué des matrices diagonales.

Déterminons l'image de φ. Les coefficients diagonaux des matrices $\varphi(X)$ sont nuls et l'on a donc l'inclusion

$$\mathrm{Im}(\varphi) \subset \underbrace{\big\{A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R}) \,\big|\, \forall i \in [\![1\,;n]\!], a_{i,i} = 0\big\}}_{\text{espace des matrices de diagonale nulle}}.$$

De plus, la formule du rang donne

$$\dim \mathrm{Im}(\varphi) = \dim \mathcal{M}_n(\mathbb{R}) - \dim \mathrm{Ker}(\varphi) = n^2 - n.$$

Or l'espace des matrices de diagonale nulle est aussi de dimension [1] $n^2 - n$ et donc, par inclusion et égalité des dimensions,

$$\mathrm{Im}(\varphi) = \big\{A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R}) \,\big|\, \forall i \in [\![1\,;n]\!], a_{i,i} = 0\big\}.$$

(b) Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice de trace nulle. La matrice M est semblable à une matrice de diagonale nulle et cette dernière appartient à l'image de φ. On peut donc introduire $P \in \mathrm{GL}_n(\mathbb{R})$ et $X \in \mathcal{M}_n(\mathbb{R})$ telles que

$$M = P\varphi(X)P^{-1} = P(DX - XD)P^{-1} = PDXP^{-1} - PXDP^{-1}.$$

En posant $A = PDP^{-1}$ et $B = PXP^{-1}$, on obtient l'écriture voulue.

3.4.7 Déterminants

Exercice 33 * (Déterminant de Vandermonde)
Soit $n \in \mathbb{N}^*$ et $a_1, \ldots, a_n \in \mathbb{C}$. Calculer le déterminant

$$V_n(a_1, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}_{[n]}.$$

1. On forme une base de cet espace en considérant les matrices élémentaires $E_{i,j}$ avec $1 \leqslant i \neq j \leqslant n$.

Solution

méthode

Par opérations élémentaires [1], on fait apparaître des zéros afin d'obtenir après développement une relation de récurrence. Les opérations élémentaires choisies ne doivent pas trop « perturber » la structure du déterminant.

Soit $n \in \mathbb{N}$ avec $n \geqslant 2$. On retranche à chaque colonne a_1 fois la précédente en commençant par la dernière

$$V_n(a_1,\dots,a_n) \underset{\substack{C_n \leftarrow C_n - a_1 C_{n-1} \\ \cdots \\ C_2 \leftarrow C_2 - a_1 C_1}}{=} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2-a_1 & a_2(a_2-a_1) & \cdots & a_2^{n-2}(a_2-a_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n-a_1 & a_n(a_n-a_1) & \cdots & a_n^{n-2}(a_n-a_1) \end{vmatrix}_{[n]}.$$

On développe le déterminant selon sa première ligne et l'on factorise un terme $a_j - a_1$ sur chaque ligne

$$\begin{aligned} V_n(a_1,\dots,a_n) &= \begin{vmatrix} a_2-a_1 & a_2(a_2-a_1) & \cdots & a_2^{n-2}(a_2-a_1) \\ \vdots & \vdots & & \vdots \\ a_n-a_1 & a_n(a_n-a_1) & \cdots & a_n^{n-2}(a_n-a_1) \end{vmatrix}_{[n-1]} \\ &= \prod_{j=2}^{n}(a_j-a_1) \times \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix}_{[n-1]}. \end{aligned}$$

On obtient ainsi la formule de récurrence

$$V_n(a_1,\dots,a_n) = \prod_{j=2}^{n}(a_j-a_1) \times V_{n-1}(a_2,\dots,a_n).$$

De proche en proche, il vient

$$V_n(a_1,\dots,a_n) = \prod_{j=2}^{n}(a_j-a_1) \times \prod_{j=3}^{n}(a_j-a_2) \times \cdots \times \prod_{j=n}^{n}(a_j-a_{n-1}) \times V_1(a_n).$$

Enfin, $V_1(a_1) = 1$ et l'on peut conclure

$$V_n(a_1,\dots,a_n) = \prod_{1 \leqslant i < j \leqslant n}(a_j-a_i).$$

1. Une alternative élégante consiste à établir que la fonction $x \mapsto V_n(a_1,\dots,a_{n-1},x)$ est une fonction polynôme de degré au plus $n-1$ qui s'annule en $a_1, a_2, \dots, a_{n-1}$ et dont le coefficient de x^{n-1} est $V_{n-1}(a_1,\dots,a_{n-1})$. On a donc $V_n(a_1,\dots,a_{n-1},x) = (x-a_1)\dots(x-a_{n-1})V_{n-1}(a_1,\dots,a_{n-1})$ lorsque les $a_1,\dots,a_{n-1}$ sont deux à deux distincts.

Exercice 34 **

Soit $n \geqslant 2$, $\lambda_1, \ldots, \lambda_n$ et x des nombres complexes deux à deux distincts. Calculer :

$$D_n(x) = \begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-2} & P_1(x) \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-2} & P_2(x) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-2} & P_n(x) \end{vmatrix}_{[n]} \quad \text{avec} \quad P_i(x) = \prod_{\substack{1 \leqslant k \leqslant n \\ k \neq i}} (x - \lambda_k).$$

Solution

Les coefficients de la dernière colonne sont des polynômes de degré $n-1$ de la variable x. En développant le déterminant selon sa dernière colonne, on peut affirmer que $D_n(x)$ est une expression polynomiale de degré au plus $n-1$ en la variable x.

méthode

‖ Connaître n valeurs suffit à déterminer un polynôme de degré au plus $n-1$.

Soit $j \in [\![1\,;n]\!]$. On a $P_i(\lambda_j) = 0$ pour tout indice i différent de j et donc, en développant le déterminant selon la dernière colonne,

$$D_n(\lambda_j) = \begin{vmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-2} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_j & \cdots & \lambda_j^{n-2} & P_j(\lambda_j) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-2} & 0 \end{vmatrix}_{[n]} = (-1)^{n+j} P_j(\lambda_j) \begin{vmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_{j-1} & \cdots & \lambda_{j-1}^{n-2} \\ 1 & \lambda_{j+1} & \cdots & \lambda_{j+1}^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-2} \end{vmatrix}_{[n-1]}.$$

Le dernier déterminant est un déterminant de Vandermonde que l'on sait calculer [1]

$$D_n(\lambda_j) = (-1)^{n+j} \underbrace{\prod_{\substack{1 \leqslant k \leqslant n \\ k \neq j}} (\lambda_j - \lambda_k)}_{=P_j(\lambda_j)} \underbrace{\prod_{\substack{1 \leqslant i < k \leqslant n \\ i,k \neq j}} (\lambda_k - \lambda_i)}_{=V_{n-1}(\lambda_1, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots, \lambda_n)}.$$

On réorganise ensuite les facteurs du premier produit

$$\begin{aligned} \prod_{\substack{1 \leqslant k \leqslant n \\ k \neq j}} (\lambda_j - \lambda_k) &= \prod_{1 \leqslant k < j} (\lambda_j - \lambda_k) \times \prod_{j < k \leqslant n} (-1) \times (\lambda_k - \lambda_j) \\ &= \prod_{1 \leqslant i < j} (\lambda_j - \lambda_i) \times (-1)^{n-j} \prod_{j < k \leqslant n} (\lambda_k - \lambda_j). \end{aligned}$$

On obtient alors

$$D_n(\lambda_j) = \prod_{1 \leqslant i < k \leqslant n} (\lambda_k - \lambda_i) = V_n(\lambda_1, \ldots, \lambda_n).$$

1. Voir sujet précédent.

Finalement, la fonction polynomiale $x \mapsto D_n(x)$, qui est de degré au plus $n-1$, coïncide en n points distincts avec la constante $V_n(\lambda_1, \ldots, \lambda_n)$, elle est donc égale à cette constante et l'on peut conclure

$$D_n(x) = V_n(\lambda_1, \ldots, \lambda_n) \quad \text{pour tout } x \in \mathbb{C}.$$

Exercice 35 **

Soit x un nombre complexe. Calculer pour tout $n \geqslant 1$

$$D_n = \begin{vmatrix} 1+x^2 & x & & (0) \\ x & \ddots & \ddots & \\ & \ddots & \ddots & x \\ (0) & & x & 1+x^2 \end{vmatrix}_{[n]}.$$

Solution

méthode

On forme une relation de récurrence linéaire double par développement de déterminant.

Soit $n \in \mathbb{N}$ avec $n \geqslant 3$. On développe le déterminant selon la première ligne

$$D_n = (1+x^2) \begin{vmatrix} 1+x^2 & x & & (0) \\ x & \ddots & \ddots & \\ & \ddots & \ddots & x \\ (0) & & x & 1+x^2 \end{vmatrix}_{[n-1]} - x \begin{vmatrix} x & x & 0 & \cdots & 0 \\ 0 & 1+x^2 & x & & (0) \\ \vdots & x & 1+x^2 & \ddots & \\ \vdots & & \ddots & \ddots & x \\ 0 & (0) & & x & 1+x^2 \end{vmatrix}_{[n-1]}$$

puis on développe le second déterminant par rapport à sa première colonne

$$D_n = (1+x^2) \underbrace{\begin{vmatrix} 1+x^2 & x & & (0) \\ x & \ddots & \ddots & \\ & \ddots & \ddots & x \\ (0) & & x & 1+x^2 \end{vmatrix}_{[n-1]}}_{=D_{n-1}} - x^2 \underbrace{\begin{vmatrix} 1+x^2 & x & & (0) \\ x & \ddots & \ddots & \\ & \ddots & \ddots & x \\ (0) & & x & 1+x^2 \end{vmatrix}_{[n-2]}}_{=D_{n-2}}.$$

La suite $(D_n)_{n \geqslant 1}$ satisfait la relation de récurrence

$$D_n - (1+x^2) D_{n-1} + x^2 D_{n-2} = 0.$$

C'est une relation de récurrence linéaire double d'équation caractéristique

$$r^2 - (1+x^2) r + x^2 = 0$$

de racines 1 et x^2. On poursuit la résolution en discutant selon que ces racines sont distinctes ou non.

Cas : $x^2 = 1$. Le terme général de la suite $(D_n)_{n\geqslant 1}$ s'exprime

$$D_n = (\lambda n + \mu) \quad \text{avec} \quad (\lambda, \mu) \in \mathbb{C}^2.$$

On détermine λ et μ par les valeurs initiales $D_1 = 2$ et $D_2 = 3$. On obtient $\lambda = \mu = 1$ et l'on conclut

$$D_n = n + 1 \quad \text{pour tout } n \geqslant 1.$$

Cas : $x^2 \neq 1$. Le terme général de la suite s'exprime

$$D_n = \lambda + \mu x^{2n} \quad \text{avec} \quad (\lambda, \mu) \in \mathbb{C}^2.$$

Sachant $D_1 = 1 + x^2$ et $D_2 = 1 + x^2 + x^4$, on obtient après quelques calculs

$$D_n = \frac{1 - x^{2n+2}}{1 - x^2} \quad \text{pour tout } n \geqslant 1.$$

Exercice 36 **

Soit $n \in \mathbb{N}^*$. Calculer

$$D_n = \begin{vmatrix} a+b & & (b) \\ & \ddots & \\ (a) & & a+b \end{vmatrix}_{[n]}$$

Solution

méthode

Il On forme une relation de récurrence sur les termes de la suite $(D_n)_{n\geqslant 1}$.

Soit $n \in \mathbb{N}$ avec $n \geqslant 2$. On décompose la première colonne en somme de deux colonnes

$$\begin{pmatrix} a+b \\ a \\ \vdots \\ a \end{pmatrix} = \begin{pmatrix} a \\ a \\ \vdots \\ a \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Par multilinéarité du déterminant en la famille des colonnes, on décompose D_n

$$D_n = \begin{vmatrix} a & b & \cdots & b \\ a & a+b & & (b) \\ \vdots & & \ddots & \\ a & (a) & & a+b \end{vmatrix}_{[n]} + \begin{vmatrix} b & b & \cdots & b \\ 0 & a+b & & (b) \\ \vdots & & \ddots & \\ 0 & (a) & & a+b \end{vmatrix}_{[n]}.$$

Dans le premier déterminant, on retranche la première ligne à chaque ligne et dans le second on développe selon la première colonne

$$D_n = \begin{vmatrix} a & b & \cdots & b \\ 0 & a & & (0) \\ \vdots & & \ddots & \\ 0 & (a-b) & & a \end{vmatrix}_{[n]} + b \begin{vmatrix} a+b & & (b) \\ & \ddots & \\ (a) & & a+b \end{vmatrix}_{[n-1]}.$$

Le premier déterminant se calcule en développant selon la première colonne tandis que le second correspond à D_{n-1}

$$D_n = a^n + bD_{n-1}.$$

On en déduit

$$\begin{aligned} D_n &= a^n + a^{n-1}b + b^2 D_{n-2} \\ &= a^n + a^{n-1}b + a^{n-2}b^2 + b^3 D_{n-3} \\ &= a^n + a^{n-1}b + a^{n-2}b^2 + a^{n-3}b^3 + b^4 D_{n-4} \\ &= \cdots \end{aligned}$$

De proche en proche et sachant $D_1 = a + b$, on obtient

$$D_n = a^n + a^{n-1}b + a^{n-2}b^2 + \cdots + ab^{n-1} + b^n = \sum_{k=0}^{n} a^{n-k}b^k.$$

Par sommation géométrique, on conclut [1]

$$D_n = \begin{cases} \dfrac{a^{n+1} - b^{n+1}}{a-b} & \text{si } a \neq b \\ (n+1)a^n & \text{si } a = b. \end{cases}$$

Exercice 37 **

Soit $A, B, C, D \in \mathcal{M}_n(\mathbb{R})$ telles que C et D commutent.

(a) On suppose que D est inversible, établir

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC).$$

On introduit $D_x = D + x\mathrm{I}_n$ avec $x \in \mathbb{R}$.

(b) Généraliser la formule précédente au cas où D n'est plus supposée inversible.

1. Cette étude est un cas particulier du sujet 10 du chapitre 10 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

Solution

(a) **méthode**

On multiplie la matrice étudiée par une matrice triangulaire par blocs afin que le produit obtenu soit lui aussi triangulaire par blocs.

On a

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D & \mathrm{O}_n \\ -C & \mathrm{I}_n \end{pmatrix} = \begin{pmatrix} AD - BC & B \\ \mathrm{O}_n & D \end{pmatrix}.$$

Le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux. En calculant le déterminant des deux membres

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} \det(D) = \det(AD - BC) \det(D).$$

On conclut en simplifiant par $\det(D)$ ce qui est possible car $\det(D) \neq 0$.

(b) **méthode**

On vérifie que la matrice D_x est inversible pour une infinité de valeurs de x.

Le déterminant de D_x est une fonction polynomiale[1] en la variable x. Celle-ci n'est pas constante puisque

$$\det(D_x) = \underbrace{x^n}_{\to +\infty} \det\left(\frac{1}{x} D + \mathrm{I}_n\right) \xrightarrow[x \to +\infty]{} +\infty \quad \text{car} \quad \det\left(\frac{1}{x} D + \mathrm{I}_n\right) \xrightarrow[x \to +\infty]{} \det(\mathrm{I}_n) = 1.$$

En effet, le déterminant d'une matrice se calcule par somme de produits de ses coefficients ce qui permet de réaliser le passage à la limite ci-dessus. On en déduit que la matrice D_x est inversible pour une infinité de valeurs de x.

La matrice D_x commute avec C et pour les valeurs de x pour lesquelles D_x est inversible on peut écrire

$$\det \begin{pmatrix} A & B \\ C & D_x \end{pmatrix} = \det(AD_x - BC).$$

Les deux membres de cette équation correspondent à des fonctions polynomiales en la variable x. Puisque celles-ci prennent les mêmes valeurs une infinité de fois, elles sont égales. En particulier, elles sont égales en $x = 0$ ce qui donne

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC).$$

1. Plus précisément, c'est une fonction polynôme de degré n qui est liée au polynôme caractéristique de la matrice $-D$. Cette dernière notion est présentée dans le chapitre qui suit.

Exercice 38 **

(a) Soit $M \in \mathcal{M}_n(\mathbb{R})$ vérifiant :

$$\det(M+X) = \det(X) \quad \text{pour tout } X \in \mathcal{M}_n(\mathbb{R}).$$

Déterminer M.

(b) Soit A et B de $\mathcal{M}_n(\mathbb{R})$ vérifiant

$$\det(A+X) = \det(B+X) \quad \text{pour tout } X \in \mathcal{M}_n(\mathbb{R}).$$

Montrer que les matrices A et B sont égales.

Solution

(a) On montre que la matrice M est nulle en étudiant son rang.

méthode

‖ La matrice M est équivalente à la matrice J_r de même type avec $r = \mathrm{rg}(M)$.

Posons $r = \mathrm{rg}(M)$. On peut écrire $M = QJ_rP^{-1}$ avec P et Q des matrices inversibles de taille n et

$$J_r = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$$

où les 0 désignent des blocs nuls de tailles appropriées.

Posons alors $X = QJ'_{n-r}P^{-1}$ avec

$$J'_{n-r} = \begin{pmatrix} 0 & 0 \\ 0 & \mathrm{I}_{n-r} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

La matrice $M+X$ est inversible car

$$M + X = Q(\underbrace{J_r + J'_{n-r}}_{=\mathrm{I}_n})P^{-1} = QP^{-1} \in \mathrm{GL}_n(\mathbb{R}).$$

On a donc $\det(X) = \det(M+X) \neq 0$. On en déduit que la matrice J'_{n-r} est la matrice I_n et donc $r = 0$ puis $M = \mathrm{O}_n$.

(b) Quand X parcourt $\mathcal{M}_n(\mathbb{R})$, la matrice $Y = B + X$ parcourt aussi $\mathcal{M}_n(\mathbb{R})$. En posant $M = A - B$, l'hypothèse donne

$$\det(M+Y) = \det(Y) \quad \text{pour tout } Y \in \mathcal{M}_n(\mathbb{R}).$$

On en déduit $M = \mathrm{O}_n$ puis $A = B$.

Exercice 39 **

Soit $n \in \mathbb{N}$ avec $n \geqslant 2$. Résoudre l'équation [1] $\mathrm{Com}(M) = M$ d'inconnue $M \in \mathcal{M}_n(\mathbb{R})$.

1. $\mathrm{Com}(A)$ désigne la comatrice de $A \in \mathcal{M}_n(\mathbb{K})$, c'est-à-dire la matrice des cofacteurs de A.

Solution

méthode

On analyse les solutions de l'équation en employant l'identité

$$^t\big(\mathrm{Com}(M)\big)M = M\,{}^t\big(\mathrm{Com}(M)\big) = \det(M)\mathrm{I}_n.$$

Soit M une matrice solution de l'équation étudiée. L'égalité $^t\big(\mathrm{Com}(M)\big)M = \det(M)\mathrm{I}_n$ donne

$$^tMM = \det(M)\mathrm{I}_n. \qquad (*)$$

Nous allons étudier la valeur de $\det(M)$.

En appliquant la fonction déterminant aux deux membres de $(*)$, on obtient

$$\big(\det(M)\big)^2 = \big(\det(M)\big)^n.$$

De plus, en calculant la trace des deux membres de $(*)$, il vient

$$\mathrm{tr}\big({}^tMM\big) = n\det(M).$$

Cependant [1]

$$\mathrm{tr}\big({}^tMM\big) = \sum_{j=1}^{n}\big[{}^tMM\big]_{j,j} = \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\big[{}^tM\big]_{j,i}\,[M]_{i,j}\right) = \sum_{i,j=1}^{n} m_{i,j}^2 \qquad (**)$$

en notant $m_{i,j}$ le coefficient général de la matrice M. On en déduit que le déterminant de M est positif.

On poursuit l'étude en traitant séparément le cas des matrices de taille 2.

Cas : $n \geqslant 3$. L'égalité $\big(\det(M)\big)^2 = \big(\det(M)\big)^n$ avec $\det(M) \geqslant 0$ entraîne $\det(M) = 0$ ou 1. Lorsque $\det(M) = 0$, on obtient ${}^tMM = \mathrm{O}_n$ et M est donc la matrice nulle en vertu de $(**)$. Lorsque $\det(M) = 1$, $(*)$ se relit ${}^tMM = \mathrm{I}_n$ ce qui caractérise une matrice orthogonale. De plus, $\det(M) = 1$ et M est donc une matrice du groupe spécial orthogonal $\mathrm{SO}_n(\mathbb{R})$.

Inversement, la matrice nulle est solution de l'équation étudiée et, si M est une matrice orthogonale de déterminant 1, l'égalité $^t\big(\mathrm{Com}(M)\big)M = \mathrm{I}_n = {}^tMM$ entraîne $\mathrm{Com}(M) = M$ car M est inversible.

Cas : $n = 2$. On étudie les coefficients de la comatrice de M en fonction de ceux de M :

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{et} \quad \mathrm{Com}(M) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

On en déduit que la comatrice de M est égale à M si, et seulement si, M est de la forme

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{avec} \quad (a,b) \in \mathbb{R}^2.$$

1. La trace de ${}^tMM$ correspond au carré de la norme euclidienne de la matrice M pour le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$: $\langle A, B\rangle = \mathrm{tr}({}^tAB)$.

3.5 Exercices d'approfondissement

Exercice 40 *

Soit $n \in \mathbb{N}$ avec $n \geqslant 2$. Montrer que les comatrices de deux matrices semblables de $\mathcal{M}_n(\mathbb{C})$ sont aussi semblables.

Solution

Soit A et P deux matrices de $\mathcal{M}_n(\mathbb{C})$ avec P inversible.

méthode

On étudie la comatrice de $P^{-1}AP$ en commençant par le cas où la matrice A est inversible avant de généraliser par un argument de densité et continuité.

On sait ${}^t\big(\mathrm{Com}(A)\big)A = \det(A)\mathrm{I}_n$. Si la matrice A est inversible, on peut exprimer la comatrice de A à l'aide de son inverse :

$$\mathrm{Com}(A) = \det(A)\,{}^t\big(A^{-1}\big).$$

En appliquant ce résultat à la matrice $P^{-1}AP$, il vient

$$\mathrm{Com}\big(P^{-1}AP\big) = \underbrace{\det\big(P^{-1}AP\big)}_{=\det(A)}\,{}^t\big(P^{-1}A^{-1}P\big) = {}^tP\,\mathrm{Com}(A)\,{}^t\big(P^{-1}\big).$$

Réalisons l'extension de cette égalité à toute matrice A de $\mathcal{M}_n(\mathbb{C})$.

La fonction $M \mapsto \mathrm{Com}(M)$ est continue sur $\mathcal{M}_n(\mathbb{C})$ car les coefficients de la comatrice de M sont ses cofacteurs et sont donc des sommes de produits de coefficients de M. Pour $P \in \mathrm{GL}_n(\mathbb{C})$, l'application $M \mapsto P^{-1}MP$ est continue car linéaire au départ d'un espace de dimension finie. Par opérations sur les fonctions continues, on peut affirmer que les deux applications $A \mapsto \mathrm{Com}\big(P^{-1}AP\big)$ et $A \mapsto {}^tP\,\mathrm{Com}(A)\,{}^t\big(P^{-1}\big)$ sont continues sur $\mathcal{M}_n(\mathbb{C})$. On vient d'établir qu'elles sont égales sur $\mathrm{GL}_n(\mathbb{C})$ qui est une partie dense[1] de $\mathcal{M}_n(\mathbb{C})$, ces deux fonctions sont donc égales sur l'intégralité de $\mathcal{M}_n(\mathbb{C})$. Ainsi, pour toute matrice A de $\mathcal{M}_n(\mathbb{C})$ et toute matrice P de $\mathrm{GL}_n(\mathbb{C})$, on a l'égalité[2]

$$\mathrm{Com}\big(P^{-1}AP\big) = {}^tP\,\mathrm{Com}(A)\,{}^t\big(P^{-1}\big).$$

On en déduit que si deux matrices sont semblables, leurs comatrices le sont aussi puisque, lorsque P est inversible, ${}^tP$ l'est aussi et $({}^tP)^{-1} = {}^t\big(P^{-1}\big)$.

1. Voir sujet 16 du chapitre 5 de l'ouvrage *Exercices d'analyse MP*.
2. Celle-ci découle aussi de l'égalité $\mathrm{Com}(AB) = \mathrm{Com}(A)\,\mathrm{Com}(B)$ qui s'établit par un procédé analogue.

Exercice 41 **

Soit E un espace vectoriel de dimension finie et G un sous-groupe de $\mathrm{GL}(E)$ de cardinal fini n. Montrer

$$\dim\left(\bigcap_{g\in G}\mathrm{Ker}(g-\mathrm{Id}_E)\right)=\frac{1}{n}\sum_{g\in G}\mathrm{tr}(g).$$

Solution

méthode

‖ La trace d'un projecteur est la dimension de son image.

On introduit l'endomorphisme

$$p=\frac{1}{n}\sum_{g\in G}g.$$

Vérifions que p est un projecteur de E. On a

$$p\circ p=\left(\frac{1}{n}\sum_{g\in G}g\right)\circ p=\frac{1}{n}\sum_{g\in G}(g\circ p).$$

Pour $g\in G$ fixé, l'application $h\mapsto g\circ h$ réalise une permutation du groupe G. Ainsi, le terme $k=g\circ h$ parcourt G lorsque h décrit G et l'on peut écrire

$$g\circ p=\frac{1}{n}\sum_{h\in G}(g\circ h)=\frac{1}{n}\sum_{k\in G}k=p \qquad (*)$$

puis

$$p\circ p=\frac{1}{n}\sum_{g\in G}p=p.$$

Ainsi, l'endomorphisme p est un projecteur et la dimension de son image est égale sa trace

$$\dim\mathrm{Im}(p)=\mathrm{tr}(p)=\frac{1}{n}\sum_{g\in G}\mathrm{tr}(g).$$

méthode

‖ L'image d'un projecteur est l'ensemble de ses vecteurs invariants.

Puisque p est un projecteur, on a $\mathrm{Im}(p)=\mathrm{Ker}(p-\mathrm{Id}_E)$. Vérifions par double inclusion que ce noyau se confond avec l'intersection des noyaux de $g-\mathrm{Id}_E$ pour g parcourant G.

Soit $g\in G$, on sait $g\circ p=p$ par $(*)$. Si x est invariant par p, on a $p(x)=x$ et donc $g(x)=g\big(p(x)\big)=p(x)=x$. On en déduit que x est élément de $\mathrm{Ker}(g-\mathrm{Id}_E)$ ce qui établit une première inclusion

$$\mathrm{Ker}(p-\mathrm{Id}_E)\subset\bigcap_{g\in G}\mathrm{Ker}(g-\mathrm{Id}_E).$$

L'inclusion réciproque est immédiate car, si x est un vecteur de l'intersection, il satisfait $g(x) = x$ pour tout $g \in G$ et l'on vérifie alors $p(x) = x$ par un simple calcul.

On a donc l'égalité

$$\bigcap_{g\in G} \operatorname{Ker}(g - \mathrm{Id}_E) = \operatorname{Ker}(p - \mathrm{Id}_E).$$

Finalement, en considérant les dimensions de ces espaces,

$$\dim\left(\bigcap_{g\in G} \operatorname{Ker}(g - \mathrm{Id}_E)\right) = \operatorname{rg}(p) = \operatorname{tr}(p) = \frac{1}{n}\sum_{g\in G} \operatorname{tr}(g).$$

Exercice 42 **

Soit F et G deux sous-espaces vectoriels supplémentaires d'un espace E et

$$\Gamma = \{u \in \mathcal{L}(E) \mid \operatorname{Im}(u) = F \text{ et } \operatorname{Ker}(u) = G\}.$$

(a) Soit u appartenant à Γ. Justifier que la restriction de u au départ de F et à valeurs dans F est un automorphisme.

(b) Vérifier que Γ est stable pour le produit de composition des applications.

(c) Établir que $(\Gamma, \circ)$ est un groupe dont on précisera l'élément neutre.

Solution

(a) **méthode**

La restriction d'une application linéaire au départ d'un espace supplémentaire de son noyau réalise un isomorphisme vers son image.

Si u appartient à Γ, l'espace $F = \operatorname{Im}(u)$ est un supplémentaire du noyau de u et la restriction de u au départ de celui-ci définit un isomorphisme de F vers $\operatorname{Im}(u)$, c'est-à-dire un automorphisme de F.

(b) Soit u et v deux endomorphismes éléments de Γ. On a

$$\operatorname{Im}(u \circ v) = u\big(v(E)\big) = u(F) = F$$

car u réalise un isomorphisme de F vers F.

Aussi, pour $x \in \operatorname{Ker}(u \circ v)$, on a $v(x)$ élément de $\operatorname{Im}(v) = F$ et de $\operatorname{Ker}(u) = G$ donc $v(x)$ est nul car F et G sont en somme directe. Ainsi, $x \in \operatorname{Ker}(v) = G$. L'inclusion réciproque étant immédiate, on a $\operatorname{Ker}(u \circ v) = G$ et l'on peut conclure que $u \circ v$ est élément de Γ.

(c) Il n'existe pas de groupe « référencé » dont Γ soit sous-groupe : on revient alors la définition axiomatique de groupe. La loi $\circ$ définit bien une loi de composition interne associative sur Γ. Vérifions l'existence d'un neutre et l'inversibilité des éléments.

méthode

On introduit une bijection entre Γ et $\mathrm{GL}(F)$ compatible avec le produit de composition.

Notons φ l'application de Γ vers $\mathrm{GL}(F)$ qui à l'endomorphisme u associe sa restriction au départ de F et à l'arrivée dans F. La question initiale assure la bonne définition de l'application φ. Vérifions sa bijectivité.

Soit $v \in \mathrm{GL}(F)$. Déterminons un endomorphisme u dans Γ vérifiant $\varphi(u) = v$. On sait qu'une application linéaire est entièrement déterminée par ses restrictions au départ de deux espaces supplémentaires. Les éléments de Γ étant par définition nuls sur l'espace G, on peut affirmer qu'il existe au plus un seul endomorphisme u pouvant appartenir à Γ et vérifier $\varphi(u) = v$, c'est celui déterminé par

$$\forall x \in F,\ u(x) = v(x) \quad \text{et} \quad \forall x \in G,\ u(x) = 0_E.$$

Inversement, cet endomorphisme convient car d'image et de noyau exactement égaux à F et G. Ainsi, l'application φ est bijective. Enfin, celle-ci est aussi compatible avec le produit de composition des applications dans le sens où

$$\varphi(u_1 \circ u_2) = \varphi(u_1) \circ \varphi(u_2) \quad \text{pour tous } u_1 \text{ et } u_2 \in \Gamma.$$

Par l'application φ on peut alors transporter les propriétés de groupe de $\big(\mathrm{GL}(F), \circ\big)$.

L'antécédent par φ du neutre Id_F, à savoir la projection p sur F parallèlement à G, est neutre pour la structure $(\Gamma, \circ)$. En effet, pour tout $u \in \Gamma$,

$$\varphi(p \circ u) = \underbrace{\varphi(p)}_{=\mathrm{Id}_F} \circ \varphi(u) = \varphi(u) \quad \text{et} \quad \varphi(u \circ p) = \varphi(u) \circ \underbrace{\varphi(p)}_{=\mathrm{Id}_F} = \varphi(u)$$

donc $p \circ u = u$ et $u \circ p = u$ par injectivité de φ.

Aussi, tout élément u de Γ est inversible d'inverse $u' = \varphi^{-1}\big(\varphi(u)^{-1}\big)$ car

$$\varphi(u \circ u') = \varphi(u) \circ \varphi(u') = \varphi(u) \circ \varphi(u)^{-1} = \mathrm{Id}_F \quad \text{et de même} \quad \varphi(u' \circ u) = \mathrm{Id}_F$$

donc $u \circ u' = u' \circ u = p$.

Finalement, $(\Gamma, \circ)$ est un groupe de neutre la projection p (et φ est un isomorphisme du groupe Γ vers $\mathrm{GL}(F)$).

Exercice 43 **

Soit $a < b$ deux réels et E l'espace des fonctions continues et affines par morceaux du segment $[a\,;b]$ vers $\mathbb{R}$. Pour $\alpha \in [a\,;b]$, on note f_α la fonction de E définie par

$$f_\alpha(x) = |x - \alpha| \quad \text{pour tout } x \in [a\,;b].$$

Montrer que la famille $(f_\alpha)_{\alpha \in [a;b]}$ est une base de E.

Solution

méthode

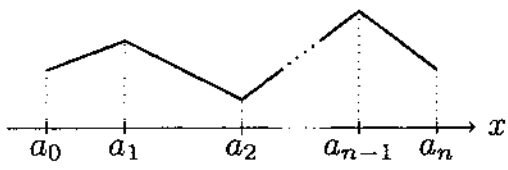
On vérifie la liberté d'une famille infinie en observant la liberté de toutes ses sous-familles finies.

Soit $n \in \mathbb{N}^*$, $a_1, \ldots, a_n$ des réels deux à deux distincts de $[a\,;b]$ et $\lambda_1, \ldots, \lambda_n$ des réels tels que $\lambda_1 f_{a_1} + \cdots + \lambda_n f_{a_n} = 0$.

Soit $i \in [\![1\,;n]\!]$. Par l'absurde, si λ_i est non nul, on peut écrire f_{a_i} comme combinaison linéaire des f_{a_j} pour $j \neq i$. Or ces dernières fonctions sont dérivables en a_i alors que f_{a_i} ne l'est pas. C'est absurde et donc $\lambda_i = 0$.

Ainsi, la sous-famille $(f_{a_i})_{1 \leqslant i \leqslant n}$ est libre et l'on peut conclure à la liberté de la famille infinie $(f_\alpha)_{\alpha \in [a;b]}$.

Montrons maintenant que cette famille est génératrice. Soit f une fonction réelle continue et affine par morceaux définie sur $[a\,;b]$. Soit aussi $(a_0, a_1, \ldots, a_n)$ une subdivision adaptée[1] à f.



méthode

On vérifie que $(f_{a_i})_{0 \leqslant i \leqslant n}$ est une base du sous-espace E_n constitué des fonctions de E pour lesquelles la subdivision $(a_0, a_1, \ldots, a_n)$ est adaptée.

On vérifie aisément que E_n est un sous-espace vectoriel et que les fonctions f_{a_i} en sont éléments. De plus, une fonction f de E_n est entièrement déterminée par la connaissance de ses valeurs en chaque a_i. L'application

$$\varphi : \begin{cases} E_n \to \mathbb{R}^{n+1} \\ f \mapsto \big(f(a_0), f(a_1), \ldots, f(a_n)\big) \end{cases}$$

détermine donc un isomorphisme entre E_n et $\mathbb{R}^{n+1}$. On en déduit que l'espace E_n est de dimension $n+1$. La famille $(f_{a_i})_{0 \leqslant i \leqslant n}$ étant libre et constituée de $n+1$ éléments qui appartiennent tous à E_n, c'est une base de cet espace. En particulier, la fonction f initiale est combinaison linéaire des éléments de cette famille.

On peut alors conclure que la famille $(f_\alpha)_{\alpha \in [a;b]}$ est génératrice de E et, finalement, c'est une base de E.

Exercice 44 ***

Soit $A, B \in \mathcal{M}_n(\mathbb{C})$ vérifiant $A^2B = A$ et $\mathrm{rg}(A) = \mathrm{rg}(B)$. Montrer $B^2A = B$.

Solution

Par l'égalité $A^2B = A$, on peut affirmer que le noyau de B est inclus dans celui de A. L'égalité des rangs de A et B entraîne celle des dimensions des noyaux de A et B et donc $\mathrm{Ker}(A) = \mathrm{Ker}(B)$.

méthode

On considère des matrices semblables à A et B permettant de visualiser simplement l'égalité des noyaux des deux matrices.

On introduit une matrice de passage P dont les dernières colonnes constituent une base du noyau commun de A et B. Les matrices $A' = P^{-1}AP$ et $B' = P^{-1}BP$ sont alors

1. La famille $(a_0, a_1, \ldots, a_n)$ est une famille de réels strictement croissante commençant par $a_0 = a$ et finissant par $a_n = b$ telle que f soit affine sur chaque intervalle $[a_{i-1}\,;a_i]$.

de la forme

$$A' = \begin{pmatrix} A_1 & 0 \\ A_2 & 0 \end{pmatrix} \quad \text{et} \quad B' = \begin{pmatrix} B_1 & 0 \\ B_2 & 0 \end{pmatrix} \quad \text{avec} \quad A_1, B_1 \in \mathcal{M}_r(\mathbb{C})$$

où r est le rang commun des matrices A et B.

L'égalité $A^2B = A$ donne $A'^2B' = A'$ ce qui entraîne par produit par blocs

$$A_1^2 B_1 = A_1 \quad \text{et} \quad A_2 A_1 B_1 = A_2$$

et ainsi

$$A_1(A_1B_1 - \mathrm{I}_r) = \mathrm{O}_r \quad \text{et} \quad A_2(A_1B_1 - \mathrm{I}_r) = \mathrm{O}_{n-r,r}.$$

Introduisons C la matrice constituée des blocs A_1 et A_2

$$C = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in \mathcal{M}_{n,r}(\mathbb{C}).$$

On a $C(A_1B_1 - \mathrm{I}_r) = \mathrm{O}_{n,r}$ et donc $\mathrm{Im}(A_1B_1 - \mathrm{I}_r) \subset \mathrm{Ker}(C)$. Or, par adjonction de colonnes nulles, le rang de C est celui de la matrice A'. Cette dernière est semblable à A et donc $\mathrm{rg}(C) = \mathrm{rg}(A) = r$. La formule du rang donne alors

$$\dim \mathrm{Ker}(C) = \underbrace{r}_{\text{nombre de colonnes}} - \mathrm{rg}(C) = r - r = 0.$$

On en déduit que la matrice $A_1B_1 - \mathrm{I}_r$ est nulle. La matrice carrée A_1 est donc inversible d'inverse B_1 et alors

$$B'^2A' = \begin{pmatrix} B_1^2A_1 & 0 \\ B_2B_1A_1 & 0 \end{pmatrix} = \begin{pmatrix} B_1 & 0 \\ B_2 & 0 \end{pmatrix} = B'.$$

Finalement, on obtient $B^2A = B$.

Exercice 45 ***

Soit $(A, +, \times)$ une $\mathbb{R}$-algèbre intègre de dimension finie $n \geqslant 2$ et a un élément de A. On assimile $\mathbb{R}$ à $\mathbb{R}.1_A$ où 1_A est l'élément de A neutre pour le produit.

(a) En observant que l'application $x \mapsto ax$ est un endomorphisme de A, montrer que a est inversible si, et seulement si, a est non nul.

(b) Montrer qu'il existe des réels $\lambda_0, \lambda_1, \ldots, \lambda_n$ non tous nuls tels que

$$\lambda_0 + \lambda_1 a + \cdots + \lambda_n a^n = 0_A.$$

(c) Montrer que, à isomorphisme près, $\mathbb{C}$ est la seule $\mathbb{R}$-algèbre commutative intègre de dimension finie $n \geqslant 2$.

Solution

(a) L'application $\varphi\colon x \mapsto ax$ est bien définie de A vers A. Elle est linéaire car, pour tous $\lambda, \mu \in \mathbb{R}$ et $a, b \in A$, les règles d'opérations dans une algèbre permettent d'écrire

$$\varphi(\lambda x + \mu y) = a(\lambda x + \mu y) = \lambda ax + \mu ay = \lambda\varphi(x) + \mu\varphi(y).$$

Si a est nul, il est entendu que a n'est pas inversible. Supposons maintenant a non nul et étudions la bijectivité de φ.

Soit $x \in A$ tel que $\varphi(x) = 0_A$. On a $ax = 0_A$ et donc $x = 0_A$ car l'algèbre A est intègre et $a \neq 0_A$. On en déduit que l'application φ est injective donc bijective car c'est un endomorphisme en dimension finie. Il existe donc $b \in A$ tel que $\varphi(b) = ab = 1_A$. Il reste à vérifier l'égalité $ba = 1_A$ pour pouvoir affirmer que a est inversible.

D'une part, $\varphi(ba) = aba = a$ et, d'autre part, $\varphi(1_A) = a$. Par injectivité de φ, on obtient $ba = 1_A$ et l'on peut conclure que a est inversible d'inverse b.

(b) La famille $\left(1_A, a, \ldots, a^n\right)$ comporte $n+1$ vecteurs en dimension n, c'est donc une famille liée. Une relation linéaire sur les éléments de cette famille permet d'écrire

$$\lambda_0 + \lambda_1 a + \cdots + \lambda_n a^n = 0_A \quad \text{avec} \quad (\lambda_0, \lambda_1, \ldots, \lambda_n) \neq 0_{\mathbb{R}^{n+1}}.$$

(c) On suppose l'algèbre A commutative et l'on introduit un élément a non réel de A ce qui est possible car $n = \dim A \geqslant 2$.

méthode

‖ On vérifie qu'il existe des réels μ, ν tels que $a^2 + \mu a + \nu = 0_A$ avec $\mu^2 - 4\nu < 0$.

Comme au-dessus, on écrit $\lambda_0 + \lambda_1 a + \cdots + \lambda_n a^n = 0_A$ avec $\lambda_0, \lambda_1, \ldots, \lambda_n$ non tous nuls. Considérons alors le polynôme $P = \lambda_0 + \lambda_1 X + \cdots + \lambda_n X^n$ de $\mathbb{R}[X]$. La propriété précédente s'interprète en affirmant que le polynôme P est un polynôme non nul *annulateur*[1] de a. En factorisant celui-ci dans $\mathbb{R}[X]$, l'intégrité de l'anneau permet d'affirmer l'existence d'un polynôme irréductible unitaire réel annulant a. Ce polynôme ne peut être de degré 1 car a n'est pas réel, c'est donc un polynôme de la forme $X^2 + \mu X + \nu$ sans racines réelles donc de discriminant $\Delta = \mu^2 - 4\nu < 0$.

méthode

‖ On introduit un élément i de $\mathrm{Vect}(1_A, a)$ vérifiant $i^2 = -1_A$.

Partant de l'équation $a^2 + \mu a + \nu = 0_A$ avec $\mu^2 - 4\nu < 0$, on peut introduire

$$i = \frac{a + \mu/2}{\sqrt{4\nu - \mu^2}} \in \mathrm{Vect}(1_A, a) \quad \text{vérifiant } i^2 = -1_A.$$

Observons alors que l'espace A est exactement $\mathrm{Vect}(1_A, i)$. Supposons par l'absurde, qu'il existe $b \in A$ extérieur à $\mathrm{Vect}(1_A, i)$. Comme au-dessus on peut introduire un élément j de

1. Voir sujet 3 p. 72.

$\mathrm{Vect}(1_A, b)$ vérifiant $j^2 = -1_A$ et donc $j^2 = i^2$. Par commutativité de l'anneau A, cette équation se relit $(j-i)(j+i) = 0_A$ ce qui donne $j = i$ ou $j = -i$ par intégrité. Comme b est combinaison linéaire de 1_A et de j, on obtient $b \in \mathrm{Vect}(1_A, i)$: c'est absurde.

Il reste à établir que A est isomorphe à $\mathbb{C}$. Considérons pour cela l'application linéaire $\varphi\colon A \to \mathbb{C}$ déterminée par $\varphi(1_A) = 1$ et $\varphi(i) = \mathrm{i}$. Celle-ci est un isomorphisme d'espaces vectoriels car transforme une base en une base. C'est aussi un morphisme d'algèbres car on vérifie $\varphi(xy) = \varphi(x)\varphi(y)$ pour tous $x, y \in A$ en exploitant $i^2 = -1_A$.

Finalement, l'algèbre A est isomorphe à $\mathbb{C}$.

Exercice 46 ***

On note $\mathrm{SL}_n(\mathbb{Z})$ l'ensemble des matrices carrées de taille $n \geqslant 2$ à coefficients entiers et de déterminants 1.

(a) Montrer que $\mathrm{SL}_n(\mathbb{Z})$ est un groupe multiplicatif.

On note H le sous-groupe de $\mathrm{SL}_n(\mathbb{Z})$ engendré par les matrices $\mathrm{I}_n + E_{i,j}$ avec i, j éléments distincts de $[\![1\,;n]\!]$.

(b) Vérifier l'appartenance à H des matrices $\mathrm{I}_n + kE_{i,j}$ pour tous indices i, j distincts dans $[\![1\,;n]\!]$ et tout $k \in \mathbb{Z}$.

(c) Soit $M \in \mathrm{SL}_n(\mathbb{Z})$. Montrer qu'il existe une matrice $A \in H$ telle que la première colonne de AM est constituée d'un 1 suivi de 0.

(d) Montrer $H = \mathrm{SL}_n(\mathbb{Z})$.

Solution

(a) Les matrices de $\mathrm{SL}_n(\mathbb{Z})$ sont inversibles car de déterminants non nuls. Montrons alors que $\mathrm{SL}_n(\mathbb{Z})$ est un sous-groupe du groupe $\big(\mathrm{GL}_n(\mathbb{R}), \times\big)$.

La partie $\mathrm{SL}_n(\mathbb{Z})$ est non vide car la matrice I_n en est élément. La partie $\mathrm{SL}_n(\mathbb{Z})$ est stable par produit car le produit de deux matrices à coefficients entiers est à coefficients entiers et le produit de deux matrices de déterminants 1 est de déterminant 1. Il reste à étudier les inverses des éléments de $\mathrm{SL}_n(\mathbb{Z})$.

méthode

‖ On peut exprimer l'inverse d'une matrice à l'aide de sa comatrice.

Soit $M \in \mathrm{SL}_n(\mathbb{Z})$. Sachant $\det(M) = 1$, l'inverse de M s'écrit

$$M^{-1} = \frac{1}{\det(M)}\,{}^t\big(\mathrm{Com}(M)\big) = {}^t\big(\mathrm{Com}(M)\big).$$

Les coefficients de la comatrice sont les cofacteurs de M et ceux-ci sont au signe près égaux aux mineurs de M. Chacun de ces mineurs est un déterminant d'une matrice à coefficients entiers, c'est donc un entier. On en déduit que la comatrice de M, et donc l'inverse de M, est à coefficients entiers.

(b) Soit $i, j \in [\![1\,;n]\!]$ avec $i \neq j$. Pour $k \in \mathbb{N}$, on a

$$(\mathrm{I}_n + kE_{i,j})(\mathrm{I}_n + E_{i,j}) = \mathrm{I}_n + (k+1)E_{i,j} \quad \text{car} \quad E_{i,j}^2 = \mathrm{O}_n.$$

On vérifie alors la propriété $\mathrm{I}_n + kE_{i,j} \in H$ par récurrence sur $k \in \mathbb{N}$. De plus, cette propriété s'étend à $k \in \mathbb{Z}$ car $\mathrm{I}_n - kE_{i,j}$ est l'inverse de $\mathrm{I}_n + kE_{i,j}$.

(c) **méthode**

‖ Multiplier à gauche[1] par $\mathrm{I}_n + kE_{i,j}$ réalise l'opération $L_i \leftarrow L_i + kL_j$.

Soit $M \in \mathrm{SL}_n(\mathbb{Z})$. Cette matrice étant inversible, il existe au moins un coefficient non nul sur sa première colonne. Parmi ceux-ci considérons m celui qui est le plus petit en valeur absolue. Par des opérations élémentaires sur les lignes du type $L_i \leftarrow L_i + kL_j$ avec k entier, on peut remplacer chaque autre coefficient de la première colonne par son reste lors de la division euclidienne par m. En répétant si besoin cette manipulation avec un nouveau coefficient de la première colonne, on peut transformer M en une matrice où il ne figure plus qu'un seul coefficient non nul sur la première colonne. Si l'on note j l'indice de ligne où il figure, l'opération $L_1 \leftarrow L_1 + L_j$ suivie de $L_j \leftarrow L_j - L_1$ suffit à positionner ce coefficient en première ligne si ce n'est pas déjà le cas. On peut aussi le transformer en un coefficient positif si besoin par $L_2 \leftarrow L_2 - L_1$, $L_1 \leftarrow L_1 + 2L_2$ et $L_2 \leftarrow L_2 - L_1$.

Ces différentes opérations élémentaires sur les lignes déterminent une matrice $A \in H$ telle que

$$AM = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \quad \text{avec} \quad a_{i,j} \in \mathbb{Z} \text{ et } a_{1,1} > 0.$$

Enfin, le nombre $a_{1,1}$ divise le déterminant de AM et ce dernier vaut 1. On a donc $a_{1,1} = 1$ et la matrice AM est de la forme voulue.

(d) **méthode**

‖ Multiplier à droite par $\mathrm{I}_n + kE_{i,j}$ réalise l'opération élémentaire $C_j \leftarrow C_j + kC_i$.

Par opérations sur les colonnes, on peut disposer des zéros à droite du coefficient en position $(1,1)$. Ces opérations sur les colonnes déterminent une matrice $A' \in H$ telle que

$$AMA' = \begin{pmatrix} 1 & 0 \\ 0 & M' \end{pmatrix} \quad \text{avec} \quad M' \in \mathrm{SL}_{n-1}(\mathbb{Z}).$$

Tant que la matrice M' ainsi formée n'est pas de taille 1, on peut répéter les opérations précédentes. Lorsque la matrice M' obtenue est de taille 1, c'est la matrice (1). On peut ainsi définir des matrices B et B' appartenant à H telles que $BMB' = \mathrm{I}_n$ et donc $M = B^{-1}B'^{-1} \in H$.

Finalement, toute matrice de $\mathrm{SL}_n(\mathbb{Z})$ appartient à H. L'inclusion réciproque étant entendue, on conclut $H = \mathrm{SL}_n(\mathbb{Z})$.

1. Une matrice $\mathrm{I}_n + \lambda E_{i,j}$ se nomme une *matrice de transvection*. Multiplier par celle-ci à gauche opère sur les lignes : $L_i \leftarrow L_i + \lambda L_j$. Multiplier par celle-ci à droite opère sur les colonnes : $C_j \leftarrow C_j + \lambda C_i$.

CHAPITRE 4

Réduction géométrique

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$, E un $\mathbb{K}$-espace vectoriel de dimension quelconque et n un entier naturel non nul.

4.1 Sous-espaces stables

4.1.1 Définition

Définition

Un sous-espace vectoriel F de E est dit *stable* par $u \in \mathcal{L}(E)$ si $u(F) \subset F$, c'est-à-dire si $u(x) \in F$ pour tout vecteur x de F.

Si F et G sont des sous-espaces vectoriels stables par un endomorphisme u, les espaces $F + G$ et $F \cap G$ le sont aussi.

4.1.2 Endomorphismes induits

Définition

Si F est un sous-espace vectoriel de E stable par un endomorphisme u de E, on peut considérer l'application restreinte [1]

$$u_F : \begin{cases} F \to F \\ x \mapsto u_F(x) = u(x). \end{cases}$$

Celle-ci définit un endomorphisme de F que l'on appelle l'*endomorphisme induit* par u sur F.

Si l'endomorphisme u est injectif, les endomorphismes qu'il induit le sont aussi. S'il est surjectif, on ignore si les endomorphismes induits le sont encore[2].

Théorème 1

Si F est un sous-espace vectoriel de E stable par des endomorphismes u et v de E, F est aussi stable par λu pour tout $\lambda \in \mathbb{K}$ et stable par $u+v$ et $u \circ v$.
De plus

$$(\lambda u)_F = \lambda u_F, \quad (u+v)_F = u_F + v_F \quad \text{et} \quad (u \circ v)_F = u_F \circ v_F.$$

L'ensemble des endomorphismes stabilisant F est une sous-algèbre de $\mathcal{L}(E)$ et l'application qui à u associe u_F y définit un morphisme d'algèbres à valeurs dans $\mathcal{L}(F)$.

4.1.3 Stabilité et représentation matricielle par blocs

Ici, E désigne un $\mathbb{K}$-espace vectoriel de dimension finie n.

Théorème 2

Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E de dimension p muni d'une base $(e_1, \dots, e_p)$ complétée en une base $e = (e_1, \dots, e_n)$ de E. On a équivalence entre :
(i) F est stable par u ;
(ii) la matrice de u dans e est de la forme

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \text{ avec } A \in \mathcal{M}_p(\mathbb{K}).$$

De plus, si tel est le cas, A est la matrice figurant u_F dans la base $(e_1, \dots, e_p)$.

Si l'on considère une base de E dont les derniers vecteurs forment une base de F, la stabilité de F par u se visualise par une matrice de la forme

$$\begin{pmatrix} B & 0 \\ C & A \end{pmatrix}$$

avec A matrice carrée de taille p figurant u_F.

Plus généralement, si des sous-espaces $E_1, \dots, E_m$ réalisent une décomposition en somme directe de l'espace E, ceux-ci sont stables par un endomorphisme u de E si, et seulement si, la matrice de u dans une base adaptée[3] à l'écriture $E = E_1 \oplus \cdots \oplus E_m$ est de la forme

$$\begin{pmatrix} A_1 & & (0) \\ & \ddots & \\ (0) & & A_m \end{pmatrix} \quad \text{avec} \quad A_k \in \mathcal{M}_{\alpha_k}(\mathbb{K}) \text{ et } \alpha_k = \dim E_k.$$

1. Il s'agit d'une restriction au départ et à l'arrivée : l'hypothèse de stabilité de F assure la bonne définition de cette application restreinte.
2. La dérivation sur $\mathbb{K}[X]$ est surjective mais l'endomorphisme qu'elle induit sur $\mathbb{K}_n[X]$ ne l'est pas.
3. Une telle base est déterminée en accolant des bases de chaque espace E_k.

4.2 Éléments propres

4.2.1 Valeurs propres et vecteurs propres d'un endomorphisme

Théorème 3 (Droite vectorielle stable)
Une droite vectorielle D engendrée par un vecteur x non nul de E est stable par un endomorphisme u de E si, et seulement si, il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.

Définition
On dit qu'un vecteur x de E est un *vecteur propre* d'un endomorphisme u de E s'il engendre une droite vectorielle stable par u.

Un vecteur propre est donc un vecteur non nul tel que $u(x) = \lambda x$ pour une certaine valeur $\lambda \in \mathbb{K}$. Celle-ci est unique, on l'appelle *valeur propre associée* au vecteur propre x.

Définition
On dit qu'un scalaire $\lambda \in \mathbb{K}$ est *valeur propre* d'un endomorphisme u de E s'il existe un vecteur $x \in E$ non nul vérifiant $u(x) = \lambda x$.

Un tel vecteur x est alors dit *vecteur propre associé* à la valeur propre λ.

4.2.2 Sous-espaces propres d'un endomorphisme

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Les vecteurs x de E solutions de l'équation $u(x) = \lambda x$ constituent l'espace $E_\lambda(u) = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)$. Le scalaire λ est valeur propre de E si, et seulement si, cet espace n'est pas réduit au vecteur nul.

Définition
Si λ est une valeur propre de u, l'ensemble $E_\lambda(u)$ s'appelle le *sous-espace propre*[1] associé à la valeur propre λ.

Théorème 4
Les sous-espaces propres d'un endomorphisme u sont stables par celui-ci et l'on a l'égalité $u_{E_\lambda(u)} = \lambda \mathrm{Id}_{E_\lambda(u)}$ pour toute valeur propre λ de u.

Plus généralement, si u et v commutent, les sous-espaces propres de u sont stables par v.

Théorème 5
La somme d'une famille finie de sous-espaces propres d'un endomorphisme u de E associés à des valeurs propres deux à deux distinctes est directe.

Théorème 6
Une famille (finie ou infinie) de vecteurs propres associés à des valeurs propres deux à deux distinctes est assurément libre.

1. Les vecteurs de $E_\lambda(u)$ sont les vecteurs propres de valeur propre λ auxquels on adjoint le vecteur nul (qui lui n'est pas vecteur propre).

Dans un espace de dimension finie n, un endomorphisme u possède au plus n valeurs propres. L'ensemble constitué de celles-ci se nomme le *spectre* de u, on le note $\mathrm{Sp}(u)$.

4.2.3 Éléments propres d'une matrice carrée

On définit les éléments propres d'une matrice carrée A de $\mathcal{M}_n(\mathbb{K})$ comme étant les éléments propres de l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé. Par l'identification usuelle des éléments de $\mathbb{K}^n$ avec les colonnes de $\mathcal{M}_{n,1}(\mathbb{K})$ formées des mêmes coefficients, on obtient le vocabulaire suivant :

Définition

Une colonne X de $\mathcal{M}_{n,1}(\mathbb{K})$ est un *vecteur propre* de la matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ si X est non nulle et s'il existe $\lambda \in \mathbb{K}$ tel que $AX = \lambda X$. Ce scalaire λ est appelé *valeur propre* associée au vecteur propre X.

Pour $\lambda \in \mathbb{K}$, on note $E_\lambda(A) = \mathrm{Ker}(A - \lambda \mathrm{I}_n)$ l'espace des solutions $X \in \mathcal{M}_{n,1}(\mathbb{K})$ de l'équation $AX = \lambda X$. Le scalaire λ est valeur propre de la matrice A si, et seulement si, cet espace n'est pas réduit à l'élément nul, auquel cas on le nomme le *sous-espace propre* associé à la valeur propre λ de la matrice A.

L'ensemble des valeurs propres de la matrice $A \in \mathcal{M}_n(\mathbb{K})$ est de cardinal fini inférieur à sa taille n. Cet ensemble constitue le *spectre* de la matrice A.

Théorème 7

Si u est un endomorphisme d'un espace E de dimension n représenté par une matrice A de $\mathcal{M}_n(\mathbb{K})$ dans une base e de E, l'endomorphisme u et la matrice A ont les mêmes valeurs propres.

De plus, leurs espaces propres se correspondent dans le sens où, pour tout $\lambda \in \mathbb{K}$ et tout $x \in E$ figuré par une colonne X dans la base e, on a

$$x \in E_\lambda(u) \iff X \in E_\lambda(A).$$

Deux matrices semblables ont alors le même spectre car figurent le même endomorphisme.

4.3 Polynôme caractéristique

4.3.1 Polynôme caractéristique d'une matrice carrée

Définition

On appelle *polynôme caractéristique* d'une matrice A de $\mathcal{M}_n(\mathbb{K})$, le polynôme χ_A de $\mathbb{K}[X]$ déterminé par l'expression polynomiale

$$\chi_A(\lambda) = \det(\lambda \mathrm{I}_n - A) \quad \text{pour tout } \lambda \in \mathbb{K}.$$

Théorème 8
Le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ est unitaire, de degré exactement n et possède les coefficients remarquables [1] de l'égalité suivante :

$$\chi_A = X^n - \mathrm{tr}(A)X^{n-1} + \cdots + (-1)^n \det(A).$$

Lorsque A est une matrice triangulaire de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$, on a

$$\chi_A = \prod_{i=1}^{n} (X - \lambda_i).$$

4.3.2 Polynôme caractéristique et valeurs propres

Théorème 9
Les valeurs propres d'une matrice A de $\mathcal{M}_n(\mathbb{K})$ sont exactement les racines [2] dans $\mathbb{K}$ de son polynôme caractéristique χ_A.

Les matrices A et ${}^tA$ ont le même polynôme caractéristique et donc les mêmes valeurs propres.

Le théorème de d'Alembert-Gauss assure l'existence d'une racine à tout polynôme complexe non constant. On en déduit le théorème d'existence de valeur propre qui suit :

Théorème 10
Toute matrice de $\mathcal{M}_n(\mathbb{C})$ possède au moins une valeur propre complexe.

Aussi, toute matrice réelle de taille impaire possède au moins une valeur propre réelle.

4.3.3 Polynôme caractéristique d'un endomorphisme

Deux matrices carrées semblables ont le même polynôme caractéristique : ceci permet d'introduire la définition suivante.

Définition
On appelle *polynôme caractéristique* d'un endomorphisme u d'un espace E de dimension finie non nulle [3], le polynôme caractéristique commun aux matrices représentant l'endomorphisme u. On le note χ_u et celui-ci vérifie :

$$\chi_u(\lambda) = \det(\lambda \mathrm{Id}_E - u) \quad \text{pour tout } \lambda \in \mathbb{K}.$$

1. Lorsque $n = 2$, le polynôme caractéristique de A est $\chi_A = X^2 - \mathrm{tr}(A)X + \det(A)$.
2. Le polynôme caractéristique de A étant de degré n, on retrouve que la matrice A possède au plus n valeurs propres.
3. Le seul endomorphisme de l'espace nul est l'endomorphisme nul, on convient que son polynôme caractéristique est constant égal à 1.

Si l'espace E est de dimension n, le polynôme caractéristique de u est unitaire, de degré exactement n et possède les coefficients remarquables de l'égalité suivante :

$$\chi_u = X^n - \mathrm{tr}(u)X^{n-1} + \cdots + (-1)^n \det(u).$$

De plus, les valeurs propres de u sont exactement les racines de χ_u.

Théorème 11

Si u est un endomorphisme d'un espace vectoriel complexe de dimension finie non nulle, il possède au moins une valeur propre.

4.3.4 Multiplicité d'une valeur propre

Définition

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. On appelle *multiplicité* de λ en tant que valeur propre de u, l'ordre de multiplicité [1] de λ en tant que racine du polynôme caractéristique χ_u. Cette multiplicité est notée $m_\lambda(u)$.

De façon analogue, pour $A \in \mathcal{M}_n(\mathbb{K})$, on définit la multiplicité $m_\lambda(A)$ de λ en tant que valeur propre de A.
Une valeur propre est dite *simple* lorsqu'elle est de multiplicité 1, elle est dite *multiple* si elle est de multiplicité au moins égale à 2. Le vocabulaire qui précède permet aussi de parler de valeur propre de multiplicité 0 pour signifier qu'un scalaire n'est pas valeur propre.

Théorème 12

La somme des multiplicités des valeurs propres d'un endomorphisme u de E est inférieure à la dimension de l'espace E. De plus, il y a égalité si, et seulement si, le polynôme caractéristique χ_u est scindé sur $\mathbb{K}$.

Lorsque $\mathbb{K} = \mathbb{C}$, tout endomorphisme u de E possède exactement $\dim E$ valeurs propres *comptées avec multiplicité* [2].

Ces résultats se transposent aux matrices carrées. En particulier, une matrice $A \in \mathcal{M}_n(\mathbb{C})$ possède exactement n valeurs propres comptées avec multiplicité.

4.3.5 Multiplicité et dimension des sous-espaces propres

Théorème 13

Si F est un sous-espace vectoriel de E stable par $u \in \mathcal{L}(E)$, le polynôme caractéristique de l'endomorphisme induit par u sur F divise le polynôme caractéristique de u.

1. L'ordre de multiplicité de $a \in \mathbb{K}$ en tant que racine d'un polynôme P non nul de $\mathbb{K}[X]$ est le plus grand entier α tel que $(X-a)^\alpha$ divise P. Lorsque cet ordre est nul, a n'est pas racine de P.
2. Une valeur propre simple est comptée une fois, une valeur propre double deux fois, etc.

De ce résultat découle une comparaison de la dimension [1] du sous-espace propre et de la multiplicité d'une valeur propre :

Théorème 14
Si $\lambda \in \mathbb{K}$ est une valeur propre de $u \in \mathcal{L}(E)$ alors $1 \leqslant \dim E_\lambda(u) \leqslant m_\lambda(u)$.

Le sous-espace propre associé à une valeur propre simple est de dimension 1.
Ces résultats se transposent aux matrices carrées.

4.3.6 Éléments propres complexes d'une matrice réelle

Une matrice carrée réelle A peut se comprendre comme une matrice carrée à coefficients complexes. On peut donc à la fois parler de ses valeurs propres réelles, qui constituent son *spectre réel* noté $\mathrm{Sp}_{\mathbb{R}}(A)$, et de ses valeurs propres complexes, qui forment son *spectre complexe* noté $\mathrm{Sp}_{\mathbb{C}}(A)$.

On a l'inclusion $\mathrm{Sp}_{\mathbb{R}}(A) \subset \mathrm{Sp}_{\mathbb{C}}(A)$ et la propriété $\mathrm{Sp}_{\mathbb{C}}(A) \neq \emptyset$. Plus précisément, une matrice réelle de taille n possède exactement n valeurs propres complexes comptées avec multiplicité.

On a aussi le résultat suivant :

Théorème 15
Les valeurs propres complexes d'une matrice réelle sont deux à deux conjuguées.
De plus, deux valeurs propres complexes conjuguées ont même multiplicité et leurs sous-espaces propres ont même dimension et se correspondent par conjugaison.

Plus généralement, si A est une matrice carrée à coefficients dans $\mathbb{L}$ sous-corps de $\mathbb{K}$, le spectre de A dans $\mathbb{L}$ est inclus dans le spectre de A dans $\mathbb{K}$.

4.4 Diagonalisabilité

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie.

4.4.1 Endomorphisme diagonalisable

Définition

Un endomorphisme u de l'espace E est dit *diagonalisable* s'il existe une base de E dans laquelle sa matrice est diagonale. Une telle base est appelée *base de diagonalisation* de l'endomorphisme u.

Une base de diagonalisation est une base constituée de vecteurs propres [2] de l'endomorphisme.

1. La dimension du sous-espace propre définit la multiplicité *géométrique* de la valeur propre par opposition à la multiplicité, quelquefois dite *algébrique*, introduite à partir du polynôme caractéristique.
2. On parle quelquefois de *base propre* pour désigner une base formée de vecteurs propres.

Une condition simple et suffisante de diagonalisabilité est donnée par le résultat suivant :

Théorème 16
Si un endomorphisme u de E possède exactement $\dim E$ valeurs propres, celui-ci est diagonalisable et ses sous-espaces propres sont des droites vectorielles.

Des conditions nécessaires et suffisantes de diagonalisabilité sont fournies par le théorème qui suit :

Théorème 17
Soit u un endomorphisme de l'espace E. On a équivalence entre :
(i) u est diagonalisable ;
(ii) E est la somme directe des sous-espaces propres de u ;
(iii) la somme des dimensions des sous-espaces propres de u égale la dimension de E ;
(iv) χ_u est scindé sur $\mathbb{K}$ et $\dim E_\lambda(u) = m_\lambda(u)$ pour tout $\lambda \in \mathrm{Sp}(u)$.
De plus, les matrices diagonales figurant u sont alors celles dont les coefficients diagonaux sont les valeurs propres de u comptées avec multiplicité.

Lorsque u est diagonalisable, une base adaptée à la décomposition de E en la somme directe des sous-espaces propres de u est une base de diagonalisation dans laquelle la matrice de u est diagonale par blocs avec des blocs diagonaux de la forme $\lambda \mathrm{I}_{m_\lambda(u)}$ pour chaque valeur propre λ de u.

4.4.2 Matrice diagonalisable

Définition
Une matrice A de $\mathcal{M}_n(\mathbb{K})$ est dite *diagonalisable* si l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé est diagonalisable.

Ceci signifie encore que la matrice A est semblable à une matrice diagonale, c'est-à-dire qu'il existe une matrice P inversible et une matrice D diagonale permettant d'écrire

$$A = PDP^{-1}.$$

Les critères de diagonalisabilité énoncés pour les endomorphismes se transposent aux matrices :

Théorème 18
Si une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ admet n valeurs propres distinctes, celle-ci est diagonalisable et ses sous-espaces propres sont des droites vectorielles.

Théorème 19

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a équivalence entre :

(i) A est diagonalisable ;

(ii) $\mathcal{M}_{n,1}(\mathbb{K})$ est [1] la somme directe des sous-espaces propres de A ;

(iii) la somme des dimensions des sous-espaces propres de A est égale à sa taille n ;

(iv) χ_A est scindé sur $\mathbb{K}$ et $\dim E_\lambda(A) = m_\lambda(A)$ pour tout $\lambda \in \mathrm{Sp}(A)$.

De plus, les matrices diagonales semblables à A sont celles dont les coefficients diagonaux sont les valeurs propres de A comptées avec multiplicité.

Si une matrice A de $\mathcal{M}_n(\mathbb{K})$ figure un endomorphisme u de E dans une certaine base, dire que la matrice A est diagonalisable équivaut à dire que l'endomorphisme u l'est.

4.5 Trigonalisabilité

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie.

4.5.1 Endomorphisme trigonalisable

Définition

Un endomorphisme u de l'espace E est dit *trigonalisable* s'il existe une base de E dans laquelle sa matrice est triangulaire supérieure. Une telle base est appelée *base de trigonalisation* de l'endomorphisme u.

Les endomorphismes diagonalisables sont *a fortiori* trigonalisables.

Lorsqu'un endomorphisme est trigonalisable, le premier vecteur d'une base de trigonalisation est un vecteur propre de celui-ci. On peut approfondir cette affirmation par le résultat géométrique qui suit :

Théorème 20

Soit u un endomorphisme d'espace E de dimension n et $e = (e_1, \ldots, e_n)$ une base de E. On a équivalence entre :

(i) la matrice de u dans e est triangulaire supérieure ;

(ii) $u(e_j) \in \mathrm{Vect}(e_1, \ldots, e_j)$ pour tout $j \in [\![1\,;n]\!]$;

(iii) l'espace $\mathrm{Vect}(e_1, \ldots, e_j)$ est stable par u pour tout $j \in [\![1\,;n]\!]$.

4.5.2 Matrice trigonalisable

Définition

Une matrice A de $\mathcal{M}_n(\mathbb{K})$ est dite *trigonalisable* si l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé est trigonalisable.

1. $\mathcal{M}_{n,1}(\mathbb{K})$ ou $\mathbb{K}^n$ selon que l'on considère les sous-espaces propres de A comme sous-espace vectoriel de l'un ou l'autre.

Ceci signifie aussi que la matrice A est semblable à une matrice triangulaire supérieure. En particulier, les matrices diagonalisables sont trigonalisables.

Si une matrice A de $\mathcal{M}_n(\mathbb{K})$ figure un endomorphisme u de E dans une certaine base, dire que la matrice A est trigonalisable équivaut à dire que l'endomorphisme u l'est.

4.5.3 Caractérisation

Théorème 21

Un endomorphisme u de E est trigonalisable si, et seulement si, son polynôme caractéristique χ_u est scindé sur $\mathbb{K}$.

De plus, une matrice triangulaire figurant u a pour coefficients diagonaux les valeurs propres de u comptées avec multiplicité.

En particulier, tout endomorphisme d'un espace vectoriel complexe E de dimension finie est trigonalisable car tout polynôme complexe non nul est scindé sur $\mathbb{C}$.

Ces résultats se transposent aux matrices carrées :

Théorème 22

Une matrice A de $\mathcal{M}_n(\mathbb{K})$ est trigonalisable si, et seulement si, son polynôme caractéristique χ_A est scindé sur $\mathbb{K}$.

De plus, les coefficients diagonaux d'une matrice triangulaire semblable à A sont les valeurs propres de A comptées avec multiplicité.

Les matrices de $\mathcal{M}_n(\mathbb{C})$ sont assurément trigonalisables.

4.5.4 Somme et produit des valeurs propres

Théorème 23

Soit u un endomorphisme de E. Si le polynôme caractéristique χ_u est scindé sur $\mathbb{K}$, la trace et le déterminant de u sont respectivement la somme et le produit[1] des valeurs propres de u comptées avec multiplicité.

On énonce un résultat semblable pour les matrices carrées.

Théorème 24

La trace et le déterminant d'une matrice complexe sont la somme et le produit de ses valeurs propres comptées avec multiplicité.

Ce résultat s'étend aux matrices réelles sous réserve de considérer leurs valeurs propres complexes.

1. Lorsqu'un polynôme unitaire est scindé de degré n, le coefficient de X^{n-1} et son coefficient constant déterminent au signe près la somme et le produit de ses racines. Pour le polynôme caractéristique, ces coefficients sont liés à la trace et au déterminant de la matrice.

4.5.5 Nilpotence

Définition

Un endomorphisme u de l'espace E est dit *nilpotent* s'il existe $p \in \mathbb{N}$ vérifiant $u^p = 0$.
Le plus petit entier naturel p vérifiant cette identité définit l'*indice de nilpotence* de u.

Ce vocabulaire se transpose aux matrices. En particulier, les matrices triangulaires supérieures strictes [1] sont nilpotentes.

Théorème 25

Un endomorphisme u de E est nilpotent si, et seulement si, il est trigonalisable et 0 est sa seule valeur propre.

Un endomorphisme nilpotent u d'un espace E de dimension n vérifie $u^n = 0$ car il peut être figuré par une matrice triangulaire supérieure stricte.

Théorème 26

Une matrice A de $\mathcal{M}_n(\mathbb{K})$ est nilpotente si, et seulement si, elle est trigonalisable et 0 est sa seule valeur propre.

Cela revient encore à dire que la matrice est semblable à une matrice triangulaire supérieure stricte. Son indice de nilpotence est alors assurément inférieur à sa taille n.

4.6 Exercices d'apprentissage

4.6.1 Éléments propres

Exercice 1

Déterminer les valeurs propres et les vecteurs propres des endomorphismes suivants :

(a) $\varphi : P \mapsto XP'$ endomorphisme de $\mathbb{R}[X]$.

(b) $\psi : P \mapsto XP$ endomorphisme de $\mathbb{C}[X]$.

(c) $T : (u_n) \mapsto (u_{n+1})$ endomorphisme de l'espace $\mathcal{B}(\mathbb{N}, \mathbb{R})$ des suites réelles bornées.

(d) $u : M \mapsto M + \mathrm{tr}(M)\mathrm{I}_n$ endomorphisme de $\mathcal{M}_n(\mathbb{R})$ (avec $n \geqslant 2$).

Solution

On vérifie dans chaque cas que les applications considérées sont bien des endomorphismes des espaces proposés.

méthode

On détermine les valeurs propres d'un endomorphisme u de E en étudiant pour quels scalaires λ l'équation [2] $u(x) = \lambda x$ d'inconnue $x \in E$ possède d'autres solutions que le vecteur nul.

1. Une matrice triangulaire supérieure A de taille n est dite *stricte* lorsque les coefficients de sa diagonale sont tous nuls. On vérifie alors $A^n = \mathrm{O}_n$.

(a) Soit $\lambda \in \mathbb{R}$. On étudie l'équation $\varphi(P) = \lambda P$, c'est-à-dire $XP' = \lambda P$, d'inconnue $P \in \mathbb{R}[X]$.

méthode

L'identification des plus grandes puissances de X dans les deux membres révèle les valeurs possibles de λ.

Analyse : Supposons que cette équation possède une solution P non nulle. On peut introduire le degré n de P et écrire

$$P = a_n X^n + \cdots + a_1 X + a_0 \quad \text{avec} \quad a_0, \ldots, a_n \in \mathbb{R} \text{ et } a_n \neq 0.$$

L'équation $\varphi(P) = \lambda P$ s'exprime alors

$$\sum_{k=0}^{n} k a_k X^k = \sum_{k=0}^{n} \lambda a_k X^k$$

L'identification des coefficients des polynômes des deux membres donne $ka_k = \lambda a_k$ pour tout $k \in [\![0\,;n]\!]$. En particulier, $na_n = \lambda a_n$ et donc $\lambda = n$ car $a_n \neq 0$. Les égalités $ka_k = na_k$ donnent alors $a_k = 0$ pour tout $k \in [\![0\,;n-1]\!]$ puis $P = a_n X^n$.

Résumons : si l'équation $\varphi(P) = \lambda P$ possède une solution P non nulle, il existe $n \in \mathbb{N}$ tel que $\lambda = n$ et P est alors $a_n X^n$ avec $a_n \neq 0$.

Synthèse : Pour $P = a_n X^n$ avec $a_n \neq 0$, on vérifie $\varphi(P) = nP$ avec P non nul.

Finalement, les valeurs propres de φ sont les entiers naturels et les vecteurs propres associés à la valeur propre $n \in \mathbb{N}$ sont les $a_n X^n$ avec a_n réel non nul[3].

(b) Soit $\lambda \in \mathbb{C}$. Résolvons l'équation $\psi(P) = \lambda P$ d'inconnue $P \in \mathbb{C}[X]$.

$$\begin{aligned} \psi(P) = \lambda P &\iff XP = \lambda P \\ &\iff (X - \lambda)P = 0 \\ &\iff P = 0 \qquad \text{car } X - \lambda \text{ n'est pas le polynôme nul}^4 \end{aligned}$$

Ainsi, pour tout $\lambda \in \mathbb{C}$, seul le polynôme nul est solution de l'équation $\psi(P) = \lambda P$: l'endomorphisme ψ ne possède pas de valeurs propres[5].

(c) Soit $\lambda \in \mathbb{R}$. Étudions l'équation $T(u) = \lambda u$ d'inconnue $u = (u_n) \in \mathcal{B}(\mathbb{N}, \mathbb{R})$. Par résolution d'une relation de récurrence géométrique, on a

$$\begin{aligned} T(u) = \lambda u &\iff \forall n \in \mathbb{N},\ u_{n+1} = \lambda u_n \\ &\iff \exists a \in \mathbb{R},\ \forall n \in \mathbb{N},\ u_n = \lambda^n a \\ &\iff \exists a \in \mathbb{R},\ u = \left(\lambda^n a\right)_{n \in \mathbb{N}}. \end{aligned}$$

2. L'équation $u(x) = \lambda x$ d'inconnues $x \in E$ et $\lambda \in \mathbb{K}$ s'appelle *l'équation aux éléments propres.*

3. Le sous-espace propre associé à la valeur propre n est $E_n(\varphi) = \{a_n X^n \mid a_n \in \mathbb{R}\}$ c'est-à-dire $\mathrm{Vect}(X^n)$: il réunit les vecteurs propres et le vecteur nul.

4. L'anneau $\mathbb{C}[X]$ est intègre : le produit de deux polynômes est nul si, et seulement si, un des facteurs est nul.

5. C'est seulement en dimension finie que l'on peut affirmer qu'un endomorphisme d'un espace complexe possède au moins une valeur propre.

méthode

‖ On se limite aux solutions bornées.

Cas : $|\lambda| > 1$. La suite $(\lambda^n a)$ est bornée si, et seulement si, $a = 0$ et c'est alors la suite nulle.

Cas : $|\lambda| \leqslant 1$. La suite $(\lambda^n a)$ est bornée et est non nulle[1], dès que $a \neq 0$.

Par conséquent, les valeurs propres de T sont les éléments[2] de $[-1\,;1]$ et les vecteurs propres associés à $\lambda \in [-1\,;1]$ sont les suites $(\lambda^n a)$ avec a réel non nul.

(d) Soit $\lambda \in \mathbb{R}$. Étudions l'équation $u(M) = \lambda M$ d'inconnue $M \in \mathcal{M}_n(\mathbb{R})$:

$$M + \operatorname{tr}(M)\mathrm{I}_n = \lambda M. \qquad (*)$$

Dans cette équation l'inconnue M apparaît sous deux formes : M et $\operatorname{tr}(M)$.

méthode

‖ On étudie la trace des matrices solutions.

Soit M une solution de $(*)$. En appliquant la fonction trace aux deux membres de cette équation, on obtient

$$(n+1)\operatorname{tr}(M) = \lambda \operatorname{tr}(M).$$

Cette équation détermine la valeur de la trace de M lorsque $\lambda \neq n+1$. On distingue alors deux cas.

Cas : $\lambda \neq n+1$. On a nécessairement $\operatorname{tr}(M) = 0$ et l'équation $(*)$ devient $M = \lambda M$.

Si $\lambda \neq 1$, seule la matrice nulle est solution et λ n'est pas valeur propre.

Si $\lambda = 1$, toutes les matrices de trace nulle sont solutions de $(*)$. Parmi celles-ci, il figure des matrices non nulles (car $n \geqslant 2$) ce qui assure que 1 est valeur propre.

Cas : $\lambda = n+1$. L'équation $(*)$ se simplifie en

$$nM = \operatorname{tr}(M)\mathrm{I}_n \quad \text{soit encore} \quad M = \frac{\operatorname{tr}(M)}{n}\mathrm{I}_n.$$

Une solution de cette équation est de la forme $\alpha \mathrm{I}_n$ avec $\alpha \in \mathbb{R}$ et, inversement, une matrice de cette forme est solution . Parmi celles-ci, il figure des matrices non nulles et $n+1$ est donc valeur propre.

En résumé, u admet deux valeurs propres :

— la valeur $n+1$ dont les vecteurs propres associés sont les matrices non nulles de trace nulle ;

— la valeur 1 dont les vecteurs propres associés sont les $\alpha \mathrm{I}_n$ avec $\alpha \neq 0$.

Exercice 2

En introduisant l'endomorphisme de dérivation sur l'espace $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, montrer la liberté de la famille de fonctions $(e_\lambda)_{\lambda \in \mathbb{C}}$ avec $e_\lambda(t) = \mathrm{e}^{\lambda t}$ pour tout $t \in \mathbb{R}$.

1. Même lorsque λ vaut 0, la suite est non nulle car son terme initial est $a \neq 0$ sachant $0^0 = 1$.
2. Ceci détermine une infinité de valeurs propres : cela n'est possible qu'en dimension infinie.

Solution

La dérivation est une opération linéaire qui transforme une fonction de classe $\mathcal{C}^\infty$ en une fonction de classe $\mathcal{C}^\infty$, on peut donc introduire l'endomorphisme D de dérivation sur l'espace E.

méthode

Une famille de vecteurs propres associés à des valeurs propres deux à deux distinctes est une famille libre (Th. 6 p. 121).

Pour tout $\lambda \in \mathbb{C}$, la fonction e_λ est une fonction non nulle de l'espace E vérifiant $\mathrm{D}(e_\lambda) = \lambda e_\lambda$: cette fonction est vecteur propre associé à la valeur propre λ pour l'endomorphisme D. On en déduit la liberté [1] de la famille $(e_\lambda)_{\lambda\in\mathbb{C}}$.

Exercice 3

Déterminer les valeurs propres et les sous-espaces propres de la matrice

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

Solution

méthode

Les valeurs propres d'une matrice sont les racines de son polynôme caractéristique (Th. 9 p. 123).

On calcule le polynôme caractéristique de A en $\lambda \in \mathbb{R}$.

$$\chi_A(\lambda) = \det(\lambda \mathrm{I}_3 - A) = \begin{vmatrix} \lambda-2 & 1 & -1 \\ 1 & \lambda-2 & 1 \\ 1 & -1 & \lambda \end{vmatrix}.$$

méthode

Puisque ce sont les racines du polynôme caractéristique qui nous intéresse, on exprime dans la mesure du possible celui-ci sous forme factorisée.

On fait apparaître des facteurs $\lambda - 1$ en ajoutant la deuxième colonne à la première puis la troisième à la seconde

$$\begin{vmatrix} \lambda-2 & 1 & -1 \\ 1 & \lambda-2 & 1 \\ 1 & -1 & \lambda \end{vmatrix} \underset{\substack{C_1\leftarrow C_1+C_2 \\ C_2\leftarrow C_2+C_3}}{=} \begin{vmatrix} \lambda-1 & 0 & -1 \\ \lambda-1 & \lambda-1 & 1 \\ 0 & \lambda-1 & \lambda \end{vmatrix} = (\lambda-1)^2 \begin{vmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & \lambda \end{vmatrix}.$$

En développant le dernier déterminant selon une rangée ou en faisant encore apparaître des zéros, on obtient à la fin des calculs $\chi_A(\lambda) = (\lambda-1)^2(\lambda-2)$. Cette égalité valable pour tout scalaire λ détermine le polynôme caractéristique de A :

$$\chi_A = (X-1)^2(X-2).$$

1. On pourra rapprocher cet argument de la démarche suivie lors de la résolution du sujet 15 du chapitre 7 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

La matrice A possède alors deux valeurs propres : 1 (valeur propre double) et 2 (valeur propre simple). Ceci est conforme à la valeur de la trace de A : $4 = 1 + 1 + 2$.

méthode

On obtient le sous-espace propre associé à une valeur propre λ d'une matrice carrée A en résolvant l'équation $AX = \lambda X$ d'inconnue la colonne X.

Commençons par déterminer le sous-espace propre associé à la valeur propre 2 (nous pouvons anticiper que ce sous-espace propre est une droite vectorielle car 2 est une valeur propre simple).

Pour X colonne de coefficients x_1, x_2, x_3, on a [1]

$$\begin{aligned} AX = 2X &\iff (A - 2\mathrm{I}_3)X = 0 \\ &\iff \begin{cases} -x_2 + x_3 = 0 \\ -x_1 - x_3 = 0 \\ -x_1 + x_2 - 2x_3 = 0 \end{cases} \\ &\iff \begin{cases} x_2 = x_3 \\ x_1 = -x_3. \end{cases} \end{aligned}$$

Le sous-espace propre associé à la valeur propre 2 est donc [2]

$$E_2(A) = \left\{ \begin{pmatrix} -x_3 \\ x_3 \\ x_3 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\} = \left\{ x_3 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\} = \mathrm{Vect} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Déterminons ensuite le sous-espace vectoriel associé à la valeur propre 1. Avec les mêmes notations

$$\begin{aligned} AX = X &\iff (A - \mathrm{I}_3)X = 0 \\ &\iff \begin{cases} x_1 - x_2 + x_3 = 0 \\ -x_1 + x_2 - x_3 = 0 \\ -x_1 + x_2 - x_3 = 0 \end{cases} \\ &\iff x_1 - x_2 + x_3 = 0. \end{aligned}$$

Cette dernière équation se résout en exprimant, par exemple, x_3 en fonction de x_1 et de x_2. Le sous-espace propre associé à la valeur propre 1 est donc

$$\begin{aligned} E_1(A) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_2 - x_1 \end{pmatrix} \;\middle|\; x_1, x_2 \in \mathbb{R} \right\} = \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; x_1, x_2 \in \mathbb{R} \right\} \\ &= \mathrm{Vect} \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

1. Puisque 2 est valeur propre de la matrice A, le système étudié doit posséder des solutions non nulles : ce n'est pas un système de Cramer et des équations doivent s'y simplifier.

2. On peut aussi décrire $E_2(A)$ comme un sous-espace vectoriel de $\mathbb{R}^3$: $E_2(A) = \mathrm{Vect}((-1, 1, 1))$.

Exercice 4

Déterminer les valeurs propres et les sous-espaces propres complexes de la matrice de rotation

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \not\equiv 0 \ [\pi].$$

Solution

méthode

Le polynôme caractéristique d'une matrice carrée A de taille 2 est (Th. 8 p. 123)

$$\chi_A = X^2 - \mathrm{tr}(A)X + \det(A).$$

Le polynôme caractéristique de R_θ est $X^2 - 2\cos(\theta)X + 1$. Les racines de ce trinôme de discriminant $\Delta = -4\sin^2\theta$, c'est-à-dire les valeurs propres de la matrice $R(\theta)$, sont les complexes distincts[1] $\mathrm{e}^{\mathrm{i}\theta}$ et $\mathrm{e}^{-\mathrm{i}\theta}$.

Pour déterminer le sous-espace propre associé à la valeur propre $\mathrm{e}^{\mathrm{i}\theta}$, on résout l'équation $AX = \mathrm{e}^{\mathrm{i}\theta}X$ avec X colonne complexe de coefficients x et y. Cela produit le système

$$\begin{cases} \cos(\theta)x - \sin(\theta)y = \mathrm{e}^{\mathrm{i}\theta}x \\ \sin(\theta)x + \cos(\theta)y = \mathrm{e}^{-\mathrm{i}\theta}y. \end{cases}$$

En écrivant $\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$ et en simplifiant par $\sin\theta$ qui est non nul, le système se résume[2] à la seule équation $x = \mathrm{i}y$. L'espace propre associé à la valeur propre $\mathrm{e}^{\mathrm{i}\theta}$ est alors

$$E_{\mathrm{e}^{\mathrm{i}\theta}}(R_\theta) = \left\{ \begin{pmatrix} \mathrm{i}y \\ y \end{pmatrix} \;\middle|\; y \in \mathbb{C} \right\} = \mathrm{Vect}\begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}.$$

méthode

Les valeurs propres complexes d'une matrice réelle sont deux à deux conjuguées et les sous-espaces propres associés se correspondent par conjugaison.

Par conjugaison, on détermine le sous-espace propre associé à la valeur propre $\mathrm{e}^{-\mathrm{i}\theta} = \overline{\mathrm{e}^{\mathrm{i}\theta}}$

$$E_{\mathrm{e}^{-\mathrm{i}\theta}}(R_\theta) = \mathrm{Vect}\begin{pmatrix} -\mathrm{i} \\ 1 \end{pmatrix}.$$

4.6.2 Diagonalisabilité

Exercice 5

Soit f un endomorphisme d'un espace vectoriel E de dimension finie n non nulle ne possédant qu'une seule valeur propre.

À quelle condition cet endomorphisme est-il diagonalisable ?

1. Lorsque $\theta \equiv 0 \ [\pi]$, la matrice $R(\theta)$ est égale à I_2 ou $-\mathrm{I}_2$: elle admet une unique valeur propre et l'espace propre associé est l'espace des colonnes.

2. Puisque $\mathrm{e}^{\mathrm{i}\theta}$ est valeur propre, il est attendu qu'une équation du système se simplifie afin de proposer d'autres solutions que la solution nulle.

Solution

méthode

Lorsqu'un endomorphisme est diagonalisable, sa matrice dans une base de diagonalisation a pour coefficients diagonaux ses valeurs propres.

Notons λ l'unique valeur propre de f. Si l'endomorphisme f est diagonalisable, sa matrice dans une base de diagonalisation est diagonale avec des λ pour seuls coefficients diagonaux. Il s'agit de la matrice λI_n que l'on sait figurer l'endomorphisme $\lambda \mathrm{Id}_E$. Par unicité de l'endomorphisme représenté par une matrice dans une base donnée, on peut affirmer que l'endomorphisme f est $\lambda \mathrm{Id}_E$. La réciproque est immédiate[1].

Exercice 6

Soit u un endomorphisme diagonalisable d'un espace E de dimension finie $n \geqslant 1$. Montrer que le noyau et l'image de u sont supplémentaires.

Solution

méthode

Les espaces $\mathrm{Im}(u)$ et $\mathrm{Ker}(u)$ se déduisent des vecteurs d'une base de diagonalisation de u en discutant selon la nullité des valeurs propres associées[2].

Soit $e = (e_1, \ldots, e_n)$ une base de diagonalisation de u

$$\mathrm{Mat}_e(u) = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \ldots, \lambda_n$ les valeurs propres associées aux vecteurs propres $e_1, \ldots, e_n$. On partitionne l'ensemble des indices i selon que la valeur correspondante est nulle ou non :

$$I = \{i \in [\![1\,;n]\!] \mid \lambda_i = 0\} \quad \text{et} \quad J = \{i \in [\![1\,;n]\!] \mid \lambda_i \neq 0\}.$$

Soit $i \in [\![1\,;n]\!]$. Si $i \in I$, on a $\lambda_i = 0$ donc $u(e_i) = 0_E$ puis $e_i \in \mathrm{Ker}(u)$. Ainsi,

$$F = \mathrm{Vect}\{e_i \mid i \in I\} \subset \mathrm{Ker}(u). \qquad (*)$$

Si $i \in J$, on peut écrire $e_i = \frac{1}{\lambda_i} u(e_i) = u\left(\frac{1}{\lambda_i} e_i\right)$ car $\lambda_i \neq 0$ et donc $e_i \in \mathrm{Im}(u)$. On en déduit

$$G = \mathrm{Vect}\{e_i \mid i \in J\} \subset \mathrm{Im}(u). \qquad (**)$$

Les inclusions $(*)$ et $(**)$ sont des égalités car

$$\dim F + \dim G = \mathrm{Card}(I) + \mathrm{Card}(J) = n = \dim \mathrm{Ker}(u) + \dim \mathrm{Im}(u).$$

On peut alors conclure que les espaces $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont supplémentaires[3] car engendrés par deux familles de vecteurs obtenues par partition d'une base.

1. Ce résultat se transpose aux matrices : si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable de seule valeur propre λ, elle est semblable à la matrice λI_n et donc égale à celle-ci.
2. On peut aussi constater $\mathrm{Im}(u) \cap \mathrm{Ker}(u) = \{0_E\}$ en observant $\mathrm{Ker}(u) = \mathrm{Ker}(u^2)$ (voir sujet suivant).
3. L'image de u est en fait la somme des sous-espaces propres associés aux valeurs propres non nulles.

Exercice 7

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice diagonalisable. Montrer $\mathrm{Ker}(A) = \mathrm{Ker}(A^2)$.

Solution

On sait $\mathrm{Ker}(A) \subset \mathrm{Ker}(A^2)$ car une colonne X vérifiant $AX = 0$ vérifie aussi $A^2X = 0$.

méthode

On montre $\mathrm{rg}(A) = \mathrm{rg}(A^2)$ par diagonalisation de A.

Puisque la matrice A est diagonalisable, on peut écrire $A = PDP^{-1}$ avec $P \in \mathrm{GL}_n(\mathbb{R})$ et

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

où $\lambda_1, \ldots, \lambda_n$ désignent les valeurs propres comptées avec multiplicité de la matrice A. Par élévation au carré

$$A^2 = (PDP^{-1})^2 = PD\underbrace{P^{-1}P}_{=I_n}DP^{-1} = PD^2P^{-1}$$

avec

$$D^2 = \begin{pmatrix} \lambda_1^2 & & (0) \\ & \ddots & \\ (0) & & \lambda_n^2 \end{pmatrix}.$$

Les matrices D et D^2 ont le même rang correspondant au nombre de coefficients non nuls sur la diagonale. Les matrices A et A^2 leur étant respectivement semblables, ont aussi le même rang. Enfin, par la formule du rang, on obtient

$$\dim \mathrm{Ker}(A) = n - \mathrm{rg}(A) = n - \mathrm{rg}(A^2) = \dim \mathrm{Ker}(A^2).$$

Par inclusion et égalité des dimensions, on conclut $\mathrm{Ker}(A) = \mathrm{Ker}(A^2)$.

Exercice 8

Étudier la diagonalisabilité des matrices de $\mathcal{M}_3(\mathbb{R})$ suivantes :

(a) $A = \begin{pmatrix} 1 & 4 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix}$ (b) $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

(c) $C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ (d) $D = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution

méthode

L'étude la diagonalisabilité d'une matrice peut [1] se résoudre par l'étude de ses éléments propres (Th. 19 p. 127).

(a) Le polynôme caractéristique de A est [2] $(X-1)(X-2)(X-3)$. La matrice A est de taille 3 et possède 3 valeurs propres distinctes [3], elle est donc diagonalisable (Th. 18 p. 126).

(b) Le polynôme caractéristique de B est $(X-1)^2(X-2)$. La matrice B est de taille 3 et possède deux valeurs propres.

méthode

On étudie les dimensions des sous-espaces propres associés.

La valeur propre 2 est simple et l'on peut affirmer sans calculs que le sous-espace propre associé est de dimension 1. La valeur propre 1 est double et son sous-espace propre est de dimension 1 ou 2 (Th. 14 p. 125).

méthode

Il n'est pas nécessaire de calculer exactement un sous-espace propre pour en déterminer la dimension : un sous-espace propre est un noyau, sa dimension se déduit d'un calcul de rang.

L'espace propre associé à la valeur propre 1 est $E_1(B) = \mathrm{Ker}(B - \mathrm{I}_3)$ avec

$$\mathrm{rg}(B-\mathrm{I}_3) = \mathrm{rg}\begin{pmatrix}0&1&0\\0&0&1\\0&0&1\end{pmatrix} \underset{L_3 \leftarrow L_3 - L_2}{=} \mathrm{rg}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix} = 2.$$

Par la formule du rang, on obtient $\dim E_1(B) = 3-2 = 1$ et donc $\dim E_1(B) < m_1(B)$. La matrice B n'est pas diagonalisable.

(c) Le polynôme caractéristique de C est $(X+1)^2(X-1)$. La valeur propre 1 est simple et le sous-espace propre associé est de dimension 1. La valeur propre -1 est double et le sous-espace propre associé est de dimension $3-1=2$ car

$$\mathrm{rg}(C+\mathrm{I}_3) = \mathrm{rg}\begin{pmatrix}1&1&0\\1&1&0\\0&0&0\end{pmatrix} \underset{L_2 \leftarrow L_2 - L_1}{=} \mathrm{rg}\begin{pmatrix}1&1&0\\0&0&0\\0&0&0\end{pmatrix} = 1.$$

1. On présentera dans le chapitre 5 une démarche alternative basée sur le concept de polynômes annulateurs.

2. De façon générale, le polynôme caractéristique d'une matrice triangulaire est le produit des $(X-\lambda_i)$ avec λ_i ses coefficients diagonaux.

3. Il importe de souligner cette distinction. Toute matrice de $\mathcal{M}_n(\mathbb{C})$ n'est pas nécessairement diagonalisable mais possède pourtant n valeurs propres, sous réserve de les décompter avec multiplicité!

La somme des dimensions des sous-espaces propres de C est égale à sa taille, la matrice C est diagonalisable [1].

(d) Le polynôme caractéristique de D est $(X-1)(X^2+1)$. Il n'est pas scindé sur $\mathbb{R}$: la matrice D n'est donc pas diagonalisable dans [2] $\mathcal{M}_3(\mathbb{R})$.

méthode

Pour l'étudier la diagonalisabilité de $M \in \mathcal{M}_n(\mathbb{K})$, on retient les démarches :

- n valeurs propres distinctes $\Longrightarrow$ M diagonalisable ;
- $\sum \dim E_\lambda(M) = n \iff M$ diagonalisable ;
- χ_M non scindé sur $\mathbb{K} \Longrightarrow M$ non diagonalisable dans $\mathcal{M}_n(\mathbb{K})$;
- $\exists \lambda \in \mathrm{Sp}(M)$, $\dim E_\lambda(M) < m_\lambda(M) \Longrightarrow M$ non diagonalisable.

Exercice 9

Soit u un endomorphisme d'un espace vectoriel réel E représenté dans une base $e = (e_1, e_2, e_3)$ par la matrice

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

(a) Justifier que u est diagonalisable.

(b) Déterminer une base de diagonalisation de u.

Solution

(a) On calcule le polynôme caractéristique de u en $\lambda \in \mathbb{R}$.

$$\chi_u(\lambda) = \chi_A(\lambda) = \begin{vmatrix} \lambda & -1 & 1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} \underset{\substack{C_1 \leftarrow C_1 - C_2 \\ C_2 \leftarrow C_2 + C_3}}{=} \begin{vmatrix} \lambda+1 & 0 & 1 \\ -\lambda-1 & \lambda-1 & -1 \\ 0 & \lambda-1 & \lambda \end{vmatrix}$$

$$= (\lambda+1)(\lambda-1)\begin{vmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & \lambda \end{vmatrix} \underset{C_3 \leftarrow C_3 - C_1}{=} (\lambda+1)(\lambda-1)\begin{vmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & \lambda \end{vmatrix}$$

$$= \lambda(\lambda+1)(\lambda-1).$$

Le polynôme caractéristique de u est donc $\chi_u = X(X-1)(X+1)$ et ses valeurs propres sont les racines 0, 1 et -1. L'endomorphisme u possède 3 valeurs propres et opère dans un espace de dimension 3, il est donc diagonalisable (Th. 16 p. 126).

1. Il s'agit d'une matrice symétrique réelle, on verra dans le Th. 6 p. 291 que celles-ci sont toujours diagonalisables.

2. En revanche, le polynôme caractéristique s'écrit $(X-1)(X-\mathrm{i})(X+\mathrm{i})$ dans $\mathbb{C}[X]$: la matrice D est de taille 3 et possède 3 valeurs propres complexes distinctes, elle est diagonalisable dans $\mathcal{M}_3(\mathbb{C})$. En substance, la propriété pour une matrice d'être diagonalisable dépend du corps d'étude.

(b) Une base de diagonalisation est une base de vecteurs propres.

méthode

On forme une base de diagonalisation en considérant une base adaptée à l'écriture

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u).$$

Déterminons une base de chaque sous-espace propre de u.

Soit x un vecteur de E et X la colonne de ses coordonnées x_1, x_2, x_3 dans la base e. Étudions le sous-espace propre associé à la valeur propre -1.

$$\begin{aligned} u(x) = -x &\iff AX = -X \\ &\iff (A + \mathrm{I}_3)X = 0 \\ &\iff \begin{cases} x_1 + x_2 - x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \\ &\iff \begin{cases} x_2 = -x_1 \\ x_3 = 0. \end{cases} \end{aligned}$$

On a donc[1]

$$E_{-1}(u) = \{x_1(e_1 - e_2) \mid x_1 \in \mathbb{R}\} = \mathrm{Vect}(\underbrace{e_1 - e_2}_{=e'_1}).$$

La famille (e'_1) est une base de l'espace $E_{-1}(u)$. Par des calculs analogues, on obtient

$$E_0(u) = \mathrm{Ker}(u) = \mathrm{Vect}(\underbrace{e_1 - e_2 - e_3}_{=e'_2}) \quad \text{et} \quad E_1(u) = \mathrm{Vect}(\underbrace{e_2 + e_3}_{=e'_3}).$$

Les familles (e'_2) et (e'_3) sont des bases respectives des espaces propres $E_0(u)$ et $E_1(u)$. La famille $e' = (e'_1, e'_2, e'_3)$ est alors une base de E adaptée[2] à l'écriture

$$E = E_{-1}(u) \oplus E_0(u) \oplus E_1(u).$$

Il s'agit donc d'une base diagonalisant u avec[3]

$$\mathrm{Mat}_{e'}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1. Il apparaît que ce sous-espace propre est une droite vectorielle ce qui est attendu puisque -1 est une valeur propre simple.

2. La famille e' a été obtenue en accolant des bases des sous-espaces propres.

3. Si l'on choisit le vecteur $\alpha e'_1$ (avec $\alpha \neq 0$) au lieu du vecteur e'_1 comme premier vecteur de base, la matrice diagonale obtenue est identique : sa première colonne traduit simplement un vecteur changé en son opposé. Si l'on permute les vecteurs de e', les coefficients diagonaux de la matrice diagonale figurant u sont permutés de la même façon.

Exercice 10

Montrer que les matrices réelles suivantes sont diagonalisables et les diagonaliser[1].

(a) $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ (b) $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

Solution

(a) On calcule le polynôme caractéristique de A. Soit $\lambda \in \mathbb{R}$. On développe le déterminant selon la dernière ligne

$$\chi_A(\lambda) = \begin{vmatrix} \lambda-1 & -1 & 0 \\ -1 & \lambda-1 & -1 \\ 0 & 0 & \lambda+1 \end{vmatrix} = (\lambda+1)\begin{vmatrix} \lambda-1 & -1 \\ -1 & \lambda-1 \end{vmatrix} = \lambda(\lambda+1)(\lambda-2).$$

Le polynôme caractéristique de A est donc $\chi_A = X(X+1)(X-2)$ et ses valeurs propres sont les racines 0, -1 et 2. La matrice A est de taille 3 et possède 3 valeurs propres, elle est diagonalisable (Th. 18 p. 126).

méthode

Une base diagonalisant l'endomorphisme canoniquement associé à A détermine les colonnes d'une matrice de passage P diagonalisant A.

On détermine les sous-espaces propres de A. Soit $X \in \mathcal{M}_{3,1}(\mathbb{R})$ de coefficients x_1, x_2, x_3. La résolution des équations $AX = 0$, $AX = -X$ et $AX = 2X$ conduit à l'étude des trois systèmes suivants

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 + x_3 = 0 \\ -x_3 = 0 \end{cases} \quad \begin{cases} 2x_1 + x_2 = 0 \\ x_1 + 2x_2 + x_3 = 0 \\ 0 = 0 \end{cases} \quad \text{et} \quad \begin{cases} -x_1 + x_2 = 0 \\ x_1 - x_2 + x_3 = 0 \\ -3x_3 = 0. \end{cases}$$

Après résolution, on obtient

$$E_0(A) = \text{Vect}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad E_{-1}(A) = \text{Vect}\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \quad \text{et} \quad E_2(A) = \text{Vect}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Les colonnes ainsi obtenues déterminent une base de vecteurs propres[2] de l'endomorphisme canoniquement associé à A. Considérons alors la matrice P constituée par ces colonnes.

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 3 & 0 \end{pmatrix}.$$

1. *Diagonaliser* une matrice carrée A signifie déterminer une matrice inversible P telle que $P^{-1}AP$ soit diagonale. Une telle matrice P est alors appelée *matrice de diagonalisation* de A.

2. Par l'identification usuelle de l'espace des colonnes $\mathcal{M}_{n,1}(\mathbb{K})$ avec $\mathbb{K}^n$, les colonnes introduites définissent des vecteurs propres de l'endomorphisme canoniquement associé.

La matrice P est inversible car c'est la matrice de passage de la base canonique à une base de vecteurs propres. Au surplus, la formule de changement de base, donne [1]

$$A = PDP^{-1} \quad \text{avec} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

méthode

En résumé, on forme les colonnes d'une matrice de passage P diagonalisant A à l'aide de bases des sous-espaces propres de A. La matrice diagonale D obtenue est alors constituée des valeurs propres associées aux colonnes de P en ordre respectif.

(b) On calcule le polynôme caractéristique de B. Soit $\lambda \in \mathbb{R}$. On développe le déterminant selon la deuxième ligne

$$\chi_B(\lambda) = \begin{vmatrix} \lambda & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ -1 & 0 & \lambda \end{vmatrix} = (\lambda - 1) \begin{vmatrix} \lambda & -1 \\ -1 & \lambda \end{vmatrix} = (\lambda - 1)(\lambda^2 - 1) = (\lambda - 1)^2(\lambda + 1).$$

Le polynôme caractéristique de B est donc $\chi_B = (X-1)^2(X+1)$ et ses valeurs propres sont 1 (valeur propre double) et -1 (valeur propre simple). On détermine les sous-espaces propres associés en résolvant les équations $BX = X$ et $BX = -X$. Avec des notations entendues, ceci conduit à l'étude des systèmes

$$\begin{cases} -x_1 + x_3 = 0 \\ 0 = 0 \\ x_1 - x_3 = 0 \end{cases} \quad \text{et} \quad \begin{cases} x_1 + x_3 = 0 \\ 2x_2 = 0 \\ x_1 + x_3 = 0. \end{cases}$$

Après résolution, on obtient

$$E_1(B) = \text{Vect}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \quad \text{et} \quad E_{-1}(B) = \text{Vect}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Les colonnes introduites définissent les colonnes d'une matrice de passage inversible

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

et la formule de changement de base donne

$$B = PDP^{-1} \quad \text{avec} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

1. La matrice D est diagonale car figure l'endomorphisme canoniquement associé à A dans une base de vecteurs propres. Ses coefficients diagonaux sont les valeurs propres respectives associées aux vecteurs propres constituant les colonnes de P.

4.6.3 Trigonalisabilité

Exercice 11

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 2$ et $k \in \mathbb{N}$.

(a) On suppose que $\lambda \in \mathbb{K}$ est une valeur propre de u. Vérifier que λ^k est valeur propre de u^k.

(b) On suppose $\mathbb{K} = \mathbb{C}$. Montrer que les valeurs propres de u^k sont exactement les λ^k avec λ valeur propre de u.

(c) On suppose $\mathbb{K} = \mathbb{R}$. Donner un exemple illustrant que la propriété précédente n'est plus vraie.

Solution

(a) Soit x un vecteur propre associé à la valeur propre λ de l'endomorphisme u. Le vecteur x est non nul et vérifie $u(x) = \lambda x$. On a alors

$$u^2(x) = u\big(u(x)\big) = u(\lambda x) = \lambda u(x) = \lambda^2 x.$$

Par récurrence, on vérifie $u^k(x) = \lambda^k x$ pour tout $k \in \mathbb{N}$. Puisque le vecteur x est non nul[1], on peut affirmer que λ^k est valeur propre de u^k associée au vecteur propre x.

(b) **méthode**

L'endomorphisme u peut être figuré par une matrice triangulaire où figurent sur la diagonale ses valeurs propres comptées avec multiplicité.

Puisque $\mathbb{K} = \mathbb{C}$, le polynôme caractéristique de u est assurément scindé sur $\mathbb{C}$ et l'endomorphisme u est trigonalisable (Th. 21 p. 128). Il existe donc $e = (e_1, \ldots, e_n)$ base de E dans laquelle

$$\mathrm{Mat}_e(u) = T = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}.$$

Les valeurs propres de u sont celles de T et, la matrice T étant triangulaire, ses valeurs propres comptées avec multiplicité sont ses coefficients diagonaux $\lambda_1, \ldots, \lambda_n$.

Aussi, la matrice de u^k dans e est T^k avec

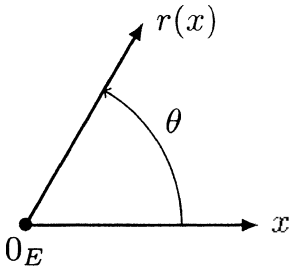
$$T^k = \begin{pmatrix} \lambda_1^k & & (*') \\ & \ddots & \\ (0) & & \lambda_n^k \end{pmatrix}.$$

Les valeurs propres de u^k comptées avec multiplicité sont donc les $\lambda_1^k, \ldots, \lambda_n^k$, c'est-à-dire les puissances d'exposant k des valeurs propres de u. Par élévation à la puissance k, les multiplicités de certaines valeurs propres peuvent fusionner.

1. Il est important d'insister sur la non nullité du vecteur x car, pour tout $\mu \in \mathbb{K}$, le vecteur nul est solution de l'équation $u(x) = \mu x$ sans pour autant pouvoir affirmer que μ est valeur propre de u.

(c) Dans un plan euclidien orienté E, une rotation r d'angle $\theta \not\equiv 0\ [\pi]$ ne possède pas de valeurs propres réelles. En effet, une rotation différente de l'identité et de la rotation d'angle π ne possède aucune droite vectorielle stable. Cependant, r^k correspond à la rotation d'angle $k\theta$. En choisissant $\theta = \pi/k$ pour $k \geqslant 2$, on obtient $r^k = -\mathrm{Id}_E$ donc

$$\mathrm{Sp}(r) = \emptyset \quad \text{et} \quad \mathrm{Sp}\left(r^k\right) = \{-1\}.$$



Exercice 12

Montrer que les matrices réelles suivantes sont trigonalisables et les trigonaliser[1].

$$\text{(a) } A = \begin{pmatrix} 1 & -3 & 1 \\ 1 & -2 & 0 \\ 0 & 1 & -2 \end{pmatrix} \qquad \text{(b) } B = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & 1 \\ -3 & -2 & 2 \end{pmatrix} \qquad \text{(c) } C = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & -3 & 3 \end{pmatrix}.$$

Solution

(a) **méthode**

> On montre qu'une matrice A de $\mathcal{M}_n(\mathbb{K})$ est trigonalisable en vérifiant que son polynôme caractéristique est scindé sur $\mathbb{K}$ (Th. 22 p. 128).

Après quelques calculs[2], le polynôme caractéristique de A est $\chi_A = (X+1)^3$. Ce polynôme est scindé sur $\mathbb{R}$ et la matrice A est donc trigonalisable[3]. Après résolution de l'équation $AX = -X$, on obtient que l'espace propre associé à la valeur propre -1 est

$$E_{-1}(A) = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Pour trigonaliser A, on forme une base dans laquelle l'endomorphisme a canoniquement associé à A est représenté par une matrice triangulaire supérieure.

méthode

> On commence par figurer l'endomorphisme a dans une base dont le premier vecteur est vecteur propre de a.

On pose $u_1 = (1,1,1)$ vecteur propre de a associé à la valeur propre -1. On complète celui-ci de deux vecteurs de la base canonique afin de former une base de $\mathbb{R}^3$, par exemple[4], en choisissant $u_2 = (0,1,0)$ et $u_3 = (0,0,1)$. On forme la matrice de a dans la

1. *Trigonaliser* une matrice carrée A signifie déterminer une matrice inversible P telle que $P^{-1}AP$ soit triangulaire supérieure. Une telle matrice P est alors appelée *matrice de trigonalisation* de A.

2. On pourra amorcer ces calculs par l'opération $C_1 \leftarrow C_1 + C_2 + C_3$.

3. Elle n'est cependant pas diagonalisable car une matrice diagonalisable dont la seule valeur propre est λ est semblable à $\lambda \mathrm{I}_n$ donc nécessairement égale à $\lambda \mathrm{I}_n$ (voir sujet 5 p. 134).

4. Durant cette démarche de trigonalisation des choix arbitraires sont effectués : il n'y a pas unicité de la matrice de passage ni de la matrice triangulaire à laquelle on parvient.

base (u_1, u_2, u_3) en calculant simplement les images de ses vecteurs

$$\begin{aligned} a(u_1) &= -u_1, \\ a(u_2) &= (3, -2, 1) = -3u_1 + u_2 + 4u_3 \\ a(u_3) &= (1, 0, -2) = u_1 - u_2 - 3u_3. \end{aligned}$$

Ainsi,

$$\mathrm{Mat}_{(u_1,u_2,u_3)}(a) = \begin{pmatrix} -1 & -3 & 1 \\ 0 & 1 & -1 \\ 0 & 4 & -3 \end{pmatrix}.$$

On considère ensuite l'endomorphisme a' du plan $\mathrm{Vect}(u_2, u_3)$ figuré dans la base (u_2, u_3) par le bloc

$$A' = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix}.$$

méthode

La détermination d'une base de $\mathrm{Vect}(u_2, u_3)$ trigonalisant a' permet [1] de former une base réalisant la trigonalisation de a.

Sans calculs, on peut affirmer que le polynôme caractéristique de a' est $(X+1)^2$ car $\chi_A = (X+1)\chi_{A'}$. Après quelques calculs, on obtient que $v_2 = u_2 + 2u_3 = (0, 1, 2)$ est vecteur propre de a'. On complète ce vecteur afin de former une base du plan $\mathrm{Vect}(u_2, u_3)$, par exemple en prenant le vecteur u_3.

On forme la matrice de a dans la base (u_1, v_2, u_3) de $\mathbb{R}^3$ en calculant les images

$$\begin{aligned} a(u_1) &= (-1, -1, -1) = -u_1, \\ a(v_2) &= (-1, -2, -3) = -u_1 - v_2 \\ a(u_3) &= (1, 0, -2) = u_1 - v_2 - u_3. \end{aligned}$$

On a donc

$$\mathrm{Mat}_{(u_1,v_2,u_3)}(a) = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Finalement, la formule de changement de base donne la trigonalisation

$$A = PTP^{-1} \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \text{ et } T = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix}.$$

(b) Cette étude est globalement semblable à la précédente. On présente rapidement les éléments de résolution de celle-ci. Le polynôme caractéristique est [2] $\chi_B = (X-1)^3$,

1. En effet, les images par a des vecteurs de $\mathrm{Vect}(u_2, u_3)$ se déduisent de leur image par a' en ajoutant un vecteur colinéaire à u_1.

2. On amorce le calcul de celui-ci par l'opération $C_1 \leftarrow C_1 + C_2$.

il est scindé sur $\mathbb{R}$ et la matrice B est trigonalisable. On introduit l'endomorphisme b canoniquement associé à B. Le vecteur $u_1 = (1,-1,1)$ est vecteur propre de b associé à la valeur propre 1. On complète ce vecteur avec $u_2 = (0,1,0)$ et $u_3 = (0,0,1)$ et l'on a

$$\mathrm{Mat}_{(u_1,u_2,u_3)}(b) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

On introduit l'endomorphisme b' du plan $\mathrm{Vect}(u_2,u_3)$ figuré dans la base (u_2,u_3) par le bloc

$$B' = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

Le vecteur $v_2 = u_2 + u_3$ est vecteur propre de b' et la matrice de l'endomorphisme b dans (u_1,v_2,u_3) donne la trigonalisation

$$B = PTP^{-1} \quad \text{avec} \quad T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(c) Cette fois-ci l'étude est un peu différente et beaucoup plus simple... Le polynôme caractéristique de C est $\chi_C = (X-1)(X-2)^2$ et les sous-espaces propres associés aux valeurs propres 1 et 2 sont [1]

$$E_1(A) = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{et} \quad E_2(A) = \mathrm{Vect}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Posons alors $u_1 = (1,1,1)$, $u_2 = (1,0,-1)$ et complétons la famille libre (u_1,u_2) en une base de $\mathbb{R}^3$ en prenant $u_3 = (0,0,1)$. On forme la matrice dans (u_1,u_2,u_3) de l'endomorphisme c canoniquement associé à la matrice C en calculant les images

$$\begin{aligned} c(u_1) &= u_1 \\ c(u_2) &= 2u_2 \\ c(u_3) &= (-1,0,3) = -u_2 + 2u_3. \end{aligned}$$

On a donc

$$\mathrm{Mat}_{(u_1,u_2,u_3)}(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Finalement,

$$C = PTP^{-1} \quad \text{avec} \quad P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

1. L'espace propre associé à la valeur propre 2 n'est que de dimension 1, la matrice C n'est pas diagonalisable.

4.7 Exercices d'entraînement

4.7.1 Sous-espaces vectoriels stables

Exercice 13 *

Soit u et v deux endomorphismes d'un espace vectoriel E vérifiant $u \circ v = v \circ u$.

(a) Montrer que les espaces $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont stables par v.

(b) Montrer que les sous-espaces propres de u sont stables par v.

(c) Soit f un endomorphisme de E et p un projecteur de E. Montrer

$$p \circ f = f \circ p \iff \mathrm{Im}(p) \text{ et } \mathrm{Ker}(p) \text{ sont stables par } f.$$

Solution

(a) **méthode**

> Un sous-espace vectoriel F est stable par un endomorphisme u lorsque, pour tout $x \in F$, on vérifie $u(x) \in F$.

Soit $x \in \mathrm{Ker}(u)$. On vérifie que $v(x)$ appartient au noyau de u par le calcul qui suit

$$u\big(v(x)\big) = (u \circ v)(x) = (v \circ u)(x) = v\big(u(x)\big) = v(0_E) = 0_E.$$

Soit $y \in \mathrm{Im}(u)$. On écrit $y = u(x)$ avec $x \in E$ et l'on vérifie $v(y) \in \mathrm{Im}(u)$ par le calcul

$$v(y) = v\big(u(x)\big) = (v \circ u)(x) = (u \circ v)(x) = u\big(\underbrace{v(x)}_{\in E}\big) \in \mathrm{Im}(u).$$

Ainsi, $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont stables par v.

(b) Soit λ une valeur propre de u. Étudions la stabilité par v de l'espace propre $E_\lambda(u)$. Soit $x \in E_\lambda(u)$. On a $u(x) = \lambda x$ et donc, par commutation de u et v,

$$u\big(v(x)\big) = (u \circ v)(x) = (v \circ u)(x) = v\big(u(x)\big) = v(\lambda x) = \lambda v(x).$$

Ainsi, $v(x)$ appartient[1] à $E_\lambda(u)$ et l'on peut affirmer que $E_\lambda(u)$ est stable[2] par v.

(c) L'implication directe est résolue par la première question. Étudions sa réciproque.

méthode

> On vérifie que les endomorphismes $p \circ f$ et $f \circ p$ sont égaux sur deux espaces supplémentaires[3].

1. Sans savoir si $v(x) \neq 0_E$, on n'affirme pas que $v(x)$ est vecteur propre de u.
2. Aussi, $u - \lambda \mathrm{Id}_E$ et v commutent donc $\mathrm{Ker}(u - \lambda \mathrm{Id}_E)$ est stable par v.
3. En dimension finie, une résolution matricielle est aussi possible en étudiant la commutation avec une matrice $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$ figurant p dans un base adaptée, $r = \mathrm{rg}(p)$.

On sait $E = \text{Ker}(p) \oplus \text{Im}(p)$ car p est un projecteur. Vérifions que les applications linéaires $p \circ f$ et $f \circ p$ sont égales sur chacun des espaces de cette écriture.

Soit $x \in \text{Ker}(p)$.

D'une part, $p(x) = 0_E$ et donc $f\big(p(x)\big) = f(0_E) = 0_E$.

D'autre part, $f(x) \in \text{Ker}(p)$ car l'espace $\text{Ker}(p)$ est stable par f et donc $p\big(f(x)\big) = 0_E$. Ainsi, les applications $p \circ f$ et $f \circ p$ sont égales sur $\text{Ker}(p)$.

Soit $x \in \text{Im}(p)$. On sait $p(x) = x$ car les vecteurs de l'image d'une projection sont invariants[1] par celle-ci. Aussi, on a $f(x)$ élément de $\text{Im}(p)$ car l'espace $\text{Im}(p)$ est stable par f et donc $p\big(f(x)\big) = f(x) = f\big(p(x)\big)$. Ainsi, les applications $p \circ f$ et $f \circ p$ sont aussi égales sur $\text{Im}(p)$.

Finalement, les applications linéaires $p \circ f$ et $f \circ p$ sont égales car égales sur deux espaces supplémentaires.

Exercice 14 **

Déterminer les sous-espaces vectoriels stables pour l'endomorphisme D de dérivation dans $\mathbb{K}[X]$.

Solution

Il est immédiat de vérifier que les espaces $\mathbb{K}[X]$, $\{0\}$ et $\mathbb{K}_n[X]$ (avec $n \in \mathbb{N}$) sont stables par l'endomorphisme de dérivation D. Établissons qu'il n'en existe pas d'autres. Soit F un sous-espace vectoriel stable par D non réduit à l'espace nul.

méthode

Si P est un polynôme de degré n, $\big(P^{(n)}, P^{(n-1)}, \ldots, P\big)$ est une famille de polynômes étagée en degré.

Soit P un polynôme de F et n son degré. L'espace F étant stable par D, les polynômes

$$P' = \text{D}(P),\ P'' = \text{D}^2(P),\ \ldots,\ P^{(n)} = \text{D}^n(P)$$

appartiennent tous à F. Or la famille $\big(P^{(n)}, P^{(n-1)}, \ldots, P\big)$ est une famille de polynômes de degrés étagés[2] et donc une base de $\mathbb{K}_n[X]$. Tout polynôme de $\mathbb{K}_n[X]$ peut donc s'écrire comme une combinaison linéaire de polynômes qui appartiennent à l'espace F et donc $\mathbb{K}_n[X] \subset F$. Résumons : $\mathbb{K}_n[X]$ est inclus dans F dès qu'il existe un polynôme de degré n appartenant à F. On peut alors conclure en distinguant deux cas :

Cas : Les polynômes de F sont de degrés majorés. On peut introduire un polynôme dans F de degré maximal N et alors $F = \mathbb{K}_N[X]$ par double inclusion.

Cas : Les polynômes de F ne sont pas de degrés majorés. Il existe des entiers $n \in \mathbb{N}$ arbitrairement grands tels que $\mathbb{K}_n[X] \subset F$ et donc $F = \mathbb{K}[X]$.

1. $\text{Im}(p) = \text{Ker}(p - \text{Id}_E)$ est le sous-espace propre associé à la valeur propre 1.

2. On dit qu'une famille $(P_0, \ldots, P_n)$ de polynômes de $\mathbb{K}[X]$ est une famille de *polynômes de degrés étagés* si $\deg(P_k) = k$ pour tout indice k. Une telle famille est une base de $\mathbb{K}_n[X]$ voir le sujet 24 du chapitre 7 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI.*

Exercice 15 **

Soit u endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie $n \geqslant 2$. On suppose que E et $\{0_E\}$ sont les seuls[1] sous-espaces vectoriels stables par u.

(a) L'endomorphisme u possède-t-il des valeurs propres ?

Soit x un vecteur non nul de E.

(b) Montrer que la famille $e_x = \big(x, u(x), \dots, u^{n-1}(x)\big)$ est une base de E.

(c) On note $a_0, a_1, \dots, a_{n-1}$ les coordonnées de $u^n(x)$ dans la base e_x. Établir

$$u^n = a_0 \mathrm{Id}_E + a_1 u + \cdots + a_{n-1} u^{n-1}.$$

(d) Exprimer la matrice de u dans la base e_x .

Solution

(a) **méthode**

Les vecteurs propres sont les vecteurs engendrant une droite vectorielle stable (Th. 3 p. 121).

Par l'absurde, si l'endomorphisme u possède un vecteur propre x, la droite vectorielle $D = \mathrm{Vect}(x)$ est stable par u. Ceci contredit l'hypothèse de travail car D est un sous-espace vectoriel non nul distinct de E. On en déduit que u ne possède ni vecteurs propres ni valeurs propres.

(b) **méthode**

On introduit le plus grand entier tel que la famille $\big(x, u(x), \dots, u^{p-1}(x)\big)$ est libre.

Puisque le vecteur x est supposé non nul, la famille (x) est libre. L'ensemble

$$A = \big\{m \in \mathbb{N}^* \;\big|\; \big(x, u(x), \dots, u^{m-1}(x)\big) \text{ est libre}\big\}$$

est donc une partie non vide de $\mathbb{N}$. Au surplus, la partie A est majorée car en dimension n une famille libre ne peut comporter plus de n éléments. La partie A possède donc un plus grand élément, autrement dit, il existe un plus grand entier $p \geqslant 1$ tel que la famille $\big(x, u(x), \dots, u^{p-1}(x)\big)$ est libre.

Par définition de l'entier p, la famille $\big(x, u(x), \dots, u^{p-1}(x), u^p(x)\big)$ est liée. Il existe donc des scalaires $\lambda_0, \dots, \lambda_p$ non tous nuls permettant d'écrire

$$\lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{p-1} u^{p-1}(x) + \lambda_p u^p(x) = 0_E. \qquad (*)$$

Si λ_p est nul, l'équation $(*)$ se simplifie en

$$\lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{p-1} u^{p-1}(x) = 0_E$$

1. En dimension finie, un endomorphisme d'un espace complexe admet au moins une valeur propre et donc une droite vectorielle stable, un endomorphisme d'un espace réel admet quant à lui au moins une droite ou un plan vectoriel stable (voir sujet 33 p. 232). Pour $n \geqslant 3$, l'hypothèse de ce sujet pourra être rencontrée si $\mathbb{K} = \mathbb{Q}$.

ce qui entraîne $\lambda_0 = \cdots = \lambda_{p-1} = 0$ par la liberté de la famille $\big(x, u(x), \dots, u^{p-1}(x)\big)$. C'est absurde car les scalaires λ_k ne sont pas tous nuls. On en déduit $\lambda_p \neq 0$ et l'égalité $(*)$ donne

$$u^p(x) = -\frac{1}{\lambda_p}\big(\lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{p-1}u^{p-1}(x)\big) \in \mathrm{Vect}\big(x, u(x), \dots, u^{p-1}(x)\big).$$

Le sous-espace vectoriel[1] $F = \mathrm{Vect}\big(x, u(x), \dots, u^{p-1}(x)\big)$ est alors stable par u car l'image par u d'une combinaison linéaire des vecteurs $x, u(x), \dots, u^{p-1}(x)$ est une combinaison linéaire des vecteurs $u(x), u^2(x), \dots, u^p(x)$ donc une combinaison linéaire des vecteurs $x, u(x), \dots, u^{p-1}(x)$. L'espace F n'étant pas nul, il est égal à l'espace E et la famille $\big(x, u(x), \dots, u^{p-1}(x)\big)$ est donc une base de E. En substance, $p = n$.

(c) La question présente une difficulté : les coordonnées $a_0, a_1, \dots, a_{n-1}$ dépendent *a priori* du choix du vecteur x !

méthode

On montre que les applications linéaires

$$v = a_0\mathrm{Id}_E + a_1 u + \cdots + a_{n-1}u^{n-1} \quad \text{et} \quad u^n$$

sont égales sur les vecteurs d'une base.

Par définition des scalaires a_i, les applications v et u^n sont égales sur le vecteur x. De plus, elles commutent[2] toutes deux avec u et donc, pour tout $k \in [\![0\,;n-1]\!]$,

$$v\big(u^k(x)\big) = u^k\big(v(x)\big) = u^k\big(u^n(x)\big) = u^n\big(u^k(x)\big).$$

Les applications linéaires v et u^n sont donc égales en tout point car égales en chaque vecteur de la base e_x.

(d) La matrice de u dans la base $e_x = \big(x, u(x), \dots, u^{n-1}(x)\big)$ est la matrice figurant les coordonnées dans e_x de la famille de vecteurs $\big(u(x), u^2(x), \dots, u^n(x)\big)$ ce qui donne[3]

$$\mathrm{Mat}_{e_x}(u) = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & & (0) & a_1 \\ & \ddots & & \vdots \\ (0) & & 1 & a_{n-1} \end{pmatrix}.$$

Notons que l'expression de cette matrice ne dépend pas du choix du vecteur x non nul.

1. Le sous-espace vectoriel F se nomme l'*espace cyclique* engendré par x : il correspond à l'ensemble des valeurs prises par les polynômes en u sur le vecteur x.
2. Dans le prochain chapitre, on dira que v et u^n sont des polynômes en u.
3. La matrice formée est une matrice compagnon (Voir sujet 32 p. 166).

Exercice 16 **

On munit l'espace $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique défini par $\langle X, Y\rangle = {}^tXY$ pour toutes colonnes X et Y.

(a) Montrer qu'un sous-espace vectoriel F de $\mathcal{M}_{n,1}(\mathbb{R})$ est stable par A si, et seulement si, $F^\perp$ est stable par ${}^tA$.

(b) Déterminer les sous-espaces vectoriels stables par l'endomorphisme u de $\mathbb{R}^3$ figuré dans la base canonique par la matrice

$$A = \begin{pmatrix} -3 & 2 & 2 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution

(a) **méthode**

> En identifiant matrice carrée et endomorphisme canoniquement associé, affirmer qu'un sous-espace vectoriel F de $\mathcal{M}_{n,1}(\mathbb{K})$ est stable par $A \in \mathcal{M}_n(\mathbb{K})$ signifie $AX \in F$ pour tout $X \in F$.

Raisonnons par double implication.

($\Longrightarrow$) Supposons le sous-espace vectoriel F stable par A. Pour tout X de F et tout Y de $F^\perp$, on a

$$\langle X, {}^tAY\rangle = {}^tX\,{}^tAY = {}^t(AX)Y = \langle AX, Y\rangle = 0$$

car AX est élément de F tandis que Y appartient à $F^\perp$. Ainsi, ${}^tAY$ est élément de $F^\perp$ et cet espace est donc stable par ${}^tA$.

($\Longleftarrow$) Supposons le sous-espace vectoriel $F^\perp$ stable par ${}^tA$. L'implication ci-dessus donne directement $F = (F^\perp)^\perp$ stable par $A = {}^t({}^tA)$.

(b) **méthode**

> On étudie les sous-espaces vectoriels stables en discutant [1] selon leur dimension.

Soit F un sous-espace vectoriel stable par u.

Si $\dim F = 0$ ou 3, l'espace F est $\{0\}$ ou $\mathbb{R}^3$ que l'on sait être tous deux stables par u.

Si $\dim F = 1$, F est une droite vectorielle et l'on sait que les droites vectorielles stables par un endomorphisme sont celles engendrées par les vecteurs propres. On est donc conduit à déterminer les vecteurs propres de u.

Enfin, si $\dim F = 2$, F est un plan et l'étude qui précède assure qu'un plan vectoriel P est stable par u si, et seulement si, sa droite normale $D = P^\perp$ (pour le produit scalaire canonique de $\mathbb{R}^3$) est stable par l'endomorphisme u' canoniquement associé à la matrice ${}^tA$. Il s'agit alors de déterminer les vecteurs propres de u'.

1. L'endomorphisme u étant diagonalisable, on verra dans le sujet 8 p. 207 que les sous-espaces vectoriels stables par u sont ceux admettant une base de vecteurs propres.

Reste à déterminer les éléments propres mentionnés. Après quelques calculs, on obtient que le polynôme caractéristique de la matrice A est

$$\chi_A = (X-1)(X+1)(X+2).$$

Les valeurs propres de A sont donc $1, -1$ et -2. Sans calculs, on peut affirmer que les valeurs propres de ${}^tA$ sont identiques car $\chi_A = \chi_{{}^tA}$. Il suffit ensuite de déterminer les sous-espaces propres associés à chaque valeur propre. Après résolutions,

$$E_1(A) = \text{Vect}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad E_{-1}(A) = \text{Vect}\begin{pmatrix}2\\1\\0\end{pmatrix}, \quad E_{-2}(A) = \text{Vect}\begin{pmatrix}1\\1\\0\end{pmatrix},$$

$$E_1({}^tA) = \text{Vect}\begin{pmatrix}0\\0\\1\end{pmatrix}, \quad E_{-1}({}^tA) = \text{Vect}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \quad E_{-2}({}^tA) = \text{Vect}\begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

Les droites vectorielles stables par u sont alors

$$\text{Vect}(1,1,1), \quad \text{Vect}(2,1,0) \quad \text{et} \quad \text{Vect}(1,1,0).$$

Les plans vectoriels stables par u sont déterminés par un vecteur normal vecteur propre de u', ils ont pour équation [1]

$$z = 0, \quad x - y = 0 \quad \text{et} \quad x - 2y + z = 0.$$

Exercice 17 **

Soit u un endomorphisme nilpotent d'un espace vectoriel E non réduit au vecteur nul et S un sous-espace vectoriel de E stable par u et tel que $E = S + \text{Im}(u)$.
Montrer $S = E$.

Solution

méthode

On vérifie $E = S + \text{Im}(u^k)$ par récurrence sur $k \in \mathbb{N}^*$.

La propriété est vraie par hypothèse pour $k = 1$.

Supposons la propriété vraie au rang $k \geqslant 1$. On a immédiatement $S + \text{Im}(u^{k+1}) \subset E$. Inversement, soit $x \in E$. Par hypothèse de récurrence, on peut écrire $x = s + u^k(a)$ avec $s \in S$ et $a \in E$. Or, par hypothèse, on peut aussi écrire $a = s' + u(b)$ avec $s' \in S$ et $b \in E$. On en déduit

$$x = \underbrace{s + u^k(s')}_{\in S} + u^{k+1}(b) \in S + \text{Im}(u^{k+1})$$

car le sous-espace vectoriel S est stable par u. Ainsi, $E \subset S + \text{Im}(u^{k+1})$ puis on peut affirmer l'égalité par double inclusion.

La récurrence est établie.

En appliquant cette propriété avec k l'indice de nilpotence de u, on conclut $E = S$.

1. Pour le produit scalaire canonique, un vecteur normal $(a,b,c) \neq (0,0,0)$ détermine dans $\mathbb{R}^3$ le plan d'équation $ax + by + cz = 0$.

4.7.2 Éléments propres d'un endomorphisme

Exercice 18 *

Soit u un endomorphisme de rang 1 d'un espace vectoriel réel E. Montrer qu'il existe un réel λ valeur propre de u tel que $u^2 = \lambda u$.

Solution

méthode

La droite $\mathrm{Im}(u)$ est stable par u.

L'endomorphisme u étant de rang 1 son image est une droite vectorielle. De plus, celle-ci est stable par u et donc engendrée par un vecteur propre (Th. 3 p. 121). Notons λ la valeur propre[1] associée. On a $\mathrm{Im}(u) \subset \mathrm{Ker}(u - \lambda \mathrm{Id}_E)$ donc $(u - \lambda \mathrm{Id}_E) \circ u = 0$. On peut alors conclure $u^2 = \lambda u$.

Exercice 19 *

Soit u et v deux endomorphismes d'un espace vectoriel E de dimension finie non nulle. Montrer que si λ est valeur propre de $u \circ v$, λ est aussi valeur propre de $v \circ u$.

Solution

Soit λ une valeur propre de $u \circ v$ et x un vecteur propre associé. Le vecteur x est non nul et satisfait l'égalité $(u \circ v)(x) = \lambda x$. En appliquant v à cette relation, il vient

$$(v \circ u)\big(v(x)\big) = v\big((u \circ v)(x)\big) = v(\lambda x) = \lambda v(x).$$

méthode

Il faut vérifier $v(x) \neq 0_E$ pour affirmer que $v(x)$ est vecteur propre de $v \circ u$.

Si $\lambda \neq 0$, l'égalité initiale $u\big(v(x)\big) = \lambda x$ avec $\lambda \neq 0$ et $x \neq 0_E$ entraîne $v(x) \neq 0_E$. On peut alors affirmer que $v(x)$ est vecteur propre de $v \circ u$ associé à la valeur propre λ.

Si $\lambda = 0$, on ne peut pas conclure aussi directement[2].

méthode

0 est valeur propre d'un endomorphisme d'un espace de dimension finie si, et seulement si, celui-ci n'est pas bijectif.

Si 0 est valeur propre de $u \circ v$, on a $\det(u \circ v) = 0$ et donc $\det(v \circ u) = 0$ car

$$\det(v \circ u) = \det(v) \times \det(u) = \det(u \circ v).$$

On en déduit que 0 est aussi valeur propre de $v \circ u$.

Finalement, u et v jouant des rôles symétriques, on peut conclure que les endomorphismes $u \circ v$ et $v \circ u$ ont exactement les mêmes[3] valeurs propres.

1. On verra dans le sujet 41 p. 175 que celle-ci correspond à la trace de u lorsque l'espace est de dimension finie.

2. Le résultat est même faux en dimension infinie. Si D désigne l'opérateur de dérivation sur $\mathbb{R}[X]$ et si I est l'opérateur déterminant le polynôme primitif s'annulant en 0, le réel 0 est valeur propre de I $\circ$ D mais n'est pas valeur propre de D $\circ$ I.

3. Ce résultat découle aussi de l'égalité $\chi_{AB} = \chi_{BA}$ établie dans le sujet 33 p. 168.

Exercice 20 **

Soit $A, B \in \mathcal{M}_n(\mathbb{R})$ vérifiant $AB - BA = A$.

(a) Calculer $A^k B - BA^k$ pour $k \in \mathbb{N}$.

On considère l'endomorphisme φ de $\mathcal{M}_n(\mathbb{R})$ défini par $\varphi(M) = MB - BM$.

(b) À quelle condition la matrice A^k est-elle vecteur propre de φ ?

(c) En déduire que la matrice A est nilpotente.

Solution

(a) **méthode**

‖ On commence par l'étude du cas $n = 2$ pour anticiper la relation attendue.

$$A^2B - BA^2 = A(\underbrace{AB}_{=A+BA}) - BA^2 = A^2 + ABA - BA^2 = A^2 + \underbrace{(AB - BA)}_{=A}A = 2A^2.$$

Par récurrence sur $k \in \mathbb{N}$, montrons $A^k B - BA^k = kA^k$.

Pour $k = 0$, l'identité est vraie car $A^0 = \mathrm{I}_n$.

Supposons l'égalité vraie au rang $k \geqslant 0$. Pour l'établir au rang suivant, on adapte le calcul vu ci-dessus

$$\begin{aligned} A^{k+1}B - BA^{k+1} &= A\big(A^kB\big) - BA^{k+1} = A\big(kA^k + BA^k\big) - BA^{k+1} \\ &= kA^{k+1} + ABA^k - BA^{k+1} = kA^{k+1} + (AB - BA)A^k \\ &= (k+1)A^k. \end{aligned}$$

La récurrence est établie.

(b) Soit $k \in \mathbb{N}$. L'étude ci-dessus donne $\varphi\big(A^k\big) = kA^k$. La matrice A^k est donc vecteur propre de φ si, et seulement si, $A^k \neq \mathrm{O}_n$.

(c) **méthode**

‖ En dimension finie, un endomorphisme n'admet qu'un nombre fini de valeurs propres.

Par l'absurde, si la matrice A n'est pas nilpotente, tout entier $k \in \mathbb{N}$ est valeur propre de l'endomorphisme φ. Or celui-ci n'admet qu'un nombre fini de valeurs propres. C'est absurde.

Exercice 21 **

Soit A une matrice élément de $E = \mathcal{M}_n(\mathbb{R})$. On introduit l'endomorphisme u de E défini par $u(M) = AM$ pour tout $M \in E$.

(a) Montrer que A et u ont les mêmes valeurs propres et préciser les sous-espaces propres de u en fonction de ceux de A.

(b) Que dire des éléments propres l'endomorphisme v de E défini par $v(M) = MA$?

Solution

(a) Soit $\lambda \in \mathbb{R}$. Étudions l'équation $u(M) = \lambda M$ d'inconnue $M \in \mathcal{M}_n(\mathbb{R})$.

méthode

‖ On étudie l'égalité $AM = \lambda M$ colonne par colonne.

Notons $C_1, \dots, C_n$ les colonnes de la matrice M. Les colonnes de la matrice AM sont alors $AC_1, \dots, AC_n$ et donc

$$AM = \lambda M \iff \forall j \in [\![1\,;n]\!],\ AC_j = \lambda C_j.$$

Si λ n'est pas valeur propre de la matrice A, seule la colonne X nulle vérifie $AX = \lambda X$. On a alors

$$AM = \lambda M \iff M = \mathrm{O}_n.$$

Le réel λ n'est pas valeur propre de u.

En revanche, si λ est une valeur propre de A, il existe des colonnes X non nulles vérifiant $AX = \lambda X$. Une matrice M constituée de telles colonnes est non nulle et vérifie $AM = \lambda M$: le réel λ est valeur propre de u.

Finalement, les valeurs propres de u sont exactement celles de A. De plus, si λ est une telle valeur, le sous-espace propre associé à la valeur propre λ est constitué des matrices dont les colonnes appartiennent [1] au sous-espace propre de A associé à la valeur propre λ.

(b) Soit $\lambda \in \mathbb{R}$. Étudions l'équation $v(M) = \lambda M$ d'inconnue $M \in \mathcal{M}_n(\mathbb{R})$. Par transposition

$$v(M) = \lambda M \iff {}^tA\,{}^tM = \lambda\,{}^tM \quad \text{avec} \quad M \neq \mathrm{O}_n \iff {}^tM \neq \mathrm{O}_n.$$

On retrouve l'équation aux éléments propres de la question précédente avec la matrice transposée de A au lieu de A et l'inconnue ${}^tM$ au lieu de M. On en déduit que les valeurs propres de v sont les valeurs propres de ${}^tA$, donc celles de A, et les vecteurs propres associés sont les matrices dont les lignes sont les transposées des colonnes appartenant au sous-espace propre de ${}^tA$ pour la même valeur propre.

Exercice 22 **

Soit E l'espace vectoriel des fonctions continues de $[0\,;+\infty[$ vers $\mathbb{R}$. Pour tout $f \in E$, on définit une fonction $\varphi(f)\colon [0\,;+\infty[\to \mathbb{R}$ par

$$\varphi(f)(0) = f(0) \quad \text{et} \quad \varphi(f)(x) = \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t \quad \text{pour tout } x > 0.$$

(a) Montrer que φ est un endomorphisme de E.

(b) Déterminer les valeurs propres et les vecteurs propres de φ.

1. Ces colonnes peuvent être nulles et ne sont donc pas forcément vecteurs propres de A.

Solution

(a) Soit f une fonction élément de E. La fonction $F : x \mapsto \int_0^x f(t)\,\mathrm{d}t$ est la primitive de la fonction continue f s'annulant en 0. Cette fonction est donc continue sur $[0\,;+\infty[$ et la fonction $\varphi(f)$ est continue sur $]0\,;+\infty[$ par produit de fonctions qui le sont. Reste à justifier la continuité de $\varphi(f)$ en 0.

méthode

On interprète $\varphi(f)(x)$ comme un taux d'accroissement.

Pour $x > 0$, on peut écrire

$$\varphi(f)(x) = \frac{1}{x}F(x) = \frac{F(x) - F(0)}{x}.$$

On en déduit

$$\varphi(f)(x) \xrightarrow[x\to 0^+]{} F'(0) = f(0) = \varphi(f)(0).$$

La fonction $\varphi(f)$ est donc continue en 0. Ainsi, l'application φ est bien définie de E vers E. Enfin, la linéarité de φ est immédiate car, pour $\lambda, \mu \in \mathbb{K}$ et $f, g \in E$, on observe $\varphi(\lambda f + \mu g)(x) = \lambda\varphi(f)(x) + \mu\varphi(g)(x)$ pour tout $x > 0$ et aussi pour $x = 0$.

(b) Soit $\lambda \in \mathbb{R}$. Étudions l'équation $\varphi(f) = \lambda f$ d'inconnue $f \in E$.

Analyse : Supposons f solution de cette équation. On a

$$\frac{1}{x}\int_0^x f(t)\,\mathrm{d}t = \lambda f(x) \quad \text{pour tout } x > 0. \tag{$*$}$$

méthode

On souhaite dériver cette relation. Afin de pouvoir justifier que la fonction f est dérivable, on traite le cas $\lambda = 0$ séparément.

Cas : $\lambda = 0$. L'équation $(*)$ se simplifie en

$$\int_0^x f(t)\,\mathrm{d}t = 0.$$

Par dérivation, on obtient $f(x) = 0$, d'abord pour $x > 0$, puis aussi pour $x = 0$ par continuité de f. Une fonction solution est nécessairement la fonction nulle : 0 n'est pas valeur propre de φ.

Cas : $\lambda \neq 0$.

méthode

On exprime une équation différentielle vérifiée par f.

En divisant les deux membres de $(*)$ par λ, on peut exprimer $f(x)$ à l'aide du terme intégrale et affirmer que f est de classe $\mathcal{C}^1$ sur $]0\,;+\infty[$. En multipliant $(*)$ par x puis en dérivant, on obtient

$$f(x) = \lambda x f'(x) + \lambda f(x) \quad \text{pour tout } x > 0$$

donc

$$xf'(x) = \alpha f(x) \quad \text{avec} \quad \alpha = \frac{1-\lambda}{\lambda}.$$

Ceci conduit à résoudre l'équation différentielle $xy' = \alpha y$ sur $]0;+\infty[$. Il s'agit d'une équation différentielle linéaire d'ordre 1 et, puisque

$$\int \frac{\alpha}{x}\,\mathrm{d}x = \alpha \ln x$$

sa solution générale est $y(x) = Ce^{\alpha \ln x} = Cx^\alpha$ avec $C \in \mathbb{R}$. La fonction f est donc de cette forme. La fonction f est aussi continue en 0 ce qui conduit à discuter selon le signe de α qui est encore celui de $\lambda(1-\lambda)$.

Si $\lambda < 0$ ou $\lambda > 1$, on a $\alpha < 0$ et l'expression Cx^α admet une limite finie en 0 si, et seulement si, $C = 0$. Dans ce cas la fonction f est nulle : un tel λ n'est pas valeur propre de φ.

Si $\lambda \in]0;1]$, l'expression Cx^α admet une limite finie en 0 qui n'est autre que[1] $C0^\alpha$.

Synthèse : Pour $\lambda \in]0;1]$, la fonction f donnée par

$$f(x) = Cx^\alpha \quad \text{avec} \quad C \in \mathbb{R} \text{ et } \alpha = \frac{1-\lambda}{\lambda}$$

est définie et continue sur $[0;+\infty[$. Par un rapide calcul intégral, on vérifie $\varphi(f) = \lambda f$ et l'on peut affirmer que la fonction f est non nulle dès que $C \neq 0$.

Finalement, les valeurs propres de φ sont les éléments de $]0;1]$ et les vecteurs propres associés à $\lambda \in]0;1]$ sont les fonctions suivantes définies sur $[0;+\infty[$

$$x \mapsto Cx^{\frac{1-\lambda}{\lambda}} \quad \text{avec} \quad C \in \mathbb{R}^*.$$

Exercice 23 **

Soit $n \in \mathbb{N}$. Déterminer les valeurs propres et les sous-espaces propres de l'endomorphisme[2] φ de $\mathbb{R}_n[X]$ défini par

$$\varphi(P) = (X^2-1)P' - nXP.$$

Solution

Soit $\lambda \in \mathbb{R}$. Résolvons l'équation $\varphi(P) = \lambda P$ d'inconnue $P \in \mathbb{R}_n[X]$.

$$\varphi(P) = \lambda P \iff (X^2-1)P' - (nX+\lambda)P = 0.$$

méthode

On recherche[3] les solutions polynomiales de degrés inférieurs à n à l'équation différentielle

$$(E_\lambda)\colon (x^2-1)y' - (nx+\lambda)y = 0.$$

1. Rappelons que x^α pour un exposant réel α est défini par l'égalité $x^\alpha = e^{\alpha \ln x}$ pour tout $x > 0$. On prolonge la fonction de x correspondante en 0 lorsque $\alpha \geqslant 0$ en posant $0^\alpha = 0$ pour $\alpha > 0$ et $0^0 = 1$.

2. On observe aisément que φ est un endomorphisme de $\mathbb{R}_n[X]$ car cette application est linéaire et l'on vérifie que φ transforme un polynôme de degré inférieur à n en un autre après une éventuelle simplification des termes X^{n+1}.

Il s'agit d'une équation différentielle linéaire d'ordre 1 que l'on sait résoudre sur chacun des intervalles $]-\infty\,;-1[$, $]-1\,;1[$ et $]1\,;+\infty[$ pour lesquels le facteur de y' ne s'annule pas. À l'aide d'une décomposition en éléments simples, on a

$$\int \frac{nx+\lambda}{x^2-1}\,\mathrm{d}x = \frac{1}{2}\int\left(\frac{n+\lambda}{x-1}+\frac{n-\lambda}{x+1}\right)\mathrm{d}x = \ln\left(|x-1|^{\frac{n+\lambda}{2}}\,|x+1|^{\frac{n-\lambda}{2}}\right).$$

La solution générale de l'équation (E_λ) sur chacun des intervalles précédents s'exprime

$$y(x) = C\,|x-1|^{\frac{n+\lambda}{2}}\,|x+1|^{\frac{n-\lambda}{2}} \quad \text{avec} \quad C\in\mathbb{R}.$$

Cette résolution invite à introduire, pour chaque $\lambda = n-2k$ avec $k\in[\![0\,;n]\!]$, le polynôme

$$P_\lambda = (X-1)^{n-k}(X+1)^k \in \mathbb{R}_n[X].$$

On vérifie par le calcul l'identité $\varphi(P_\lambda) = \lambda P_\lambda$ avec $P_\lambda \neq 0$ ce qui assure que λ est valeur propre de φ et P_λ est vecteur propre associé. On détermine ainsi $n+1$ valeurs propres distinctes pour l'endomorphisme φ. Or l'espace $\mathbb{R}_n[X]$ est de dimension $n+1$, il ne peut donc y avoir d'autres valeurs propres et celles-ci sont toutes simples (Th. 12 p. 124). De plus, chaque sous-espace propre de φ est de dimension 1.

Finalement, les valeurs propres de φ sont les $n-2k$ avec $k\in[\![0\,;n]\!]$ et le sous-espace propre associé à la valeur $n-2k$ est la droite

$$E_{n-2k}(\varphi) = \mathrm{Vect}\big((X-1)^{n-k}(X+1)^{n+k}\big).$$

Exercice 24 **

Soit u, v deux endomorphismes d'un espace vectoriel complexe de dimension finie non nulle.

(a) On suppose $u\circ v = v\circ u$. Montrer que u et v ont un vecteur propre en commun.

(b) On suppose $u\circ v = 0$. Montrer que u et v ont un vecteur propre en commun.

Solution

(a) **méthode**

Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie non nulle[4] admet au moins une valeur propre (Th. 11 p. 124).

Soit λ une valeur propre de u. L'espace $E_\lambda(u)$ n'est pas réduit au vecteur nul et tout vecteur non nul de celui-ci est vecteur propre de l'endomorphisme u. De plus, cet espace propre est stable[5] par v car u et v commutent. L'endomorphisme induit par v sur l'espace complexe $E_\lambda(u)$ de dimension finie non nulle admet donc une valeur propre. Le vecteur propre associé est alors un vecteur propre commun à u et v.

3. On peut aussi résoudre cette équation par l'unicité de la décomposition en éléments simples de P'/P qui s'exprime comme une somme de termes $\alpha/(X-x)$ avec $x\in\mathbb{C}$ racine de P et $\alpha\in\mathbb{N}^*$ sa multiplicité.

4. En dimension infinie, un endomorphisme d'un espace complexe peut ne pas admettre de valeur propre comme l'endomorphisme ψ du sujet 1 p. 129.

5. Voir sujet 13 p. 146.

(b) **méthode**

‖ L'espace $\mathrm{Im}(v)$ est stable[1] par v.

On a $u \circ v = 0$ et donc $\mathrm{Im}(v) \subset \mathrm{Ker}(u)$. Par conséquent, tout vecteur non nul de $\mathrm{Im}(v)$ est vecteur propre de u associé à la valeur propre 0.

Si v n'est pas l'endomorphisme nul, l'espace complexe $\mathrm{Im}(v)$ est de dimension finie non nulle et l'endomorphisme induit par v sur celui-ci admet au moins un vecteur propre qui est alors vecteur propre commun à u et v.

Si v est l'endomorphisme nul, n'importe quel vecteur propre de u (et il en existe) est vecteur propre commun à u et v.

4.7.3 Éléments propres d'une matrice

Exercice 25 *

Soit $n \geqslant 2$.

(a) Déterminer les valeurs propres de la matrice de $\mathcal{M}_n(\mathbb{C})$ suivante

$$M = \begin{pmatrix} a & & (b) \\ & \ddots & \\ (b) & & a \end{pmatrix} \quad \text{avec} \quad (a,b) \in \mathbb{C} \times \mathbb{C}^*.$$

(b) Cette matrice est-elle diagonalisable ?

Solution

(a) **méthode**

‖ On calcule le polynôme caractéristique de M.

Soit $\lambda \in \mathbb{C}$. Par l'opération $C_1 \leftarrow C_1 + C_2 + \cdots + C_n$

$$\chi_M(\lambda) = \begin{vmatrix} \lambda - a & & (-b) \\ & \ddots & \\ (-b) & & \lambda - a \end{vmatrix}_{[n]} = \begin{vmatrix} \lambda - \big(a + (n-1)b\big) & -b & \cdots & -b \\ \lambda - \big(a + (n-1)b\big) & \lambda - a & & (-b) \\ \vdots & & \ddots & \\ \lambda - \big(a + (n-1)b\big) & (-b) & & \lambda - a \end{vmatrix}_{[n]}.$$

On retranche ensuite la première ligne à chacune des suivantes afin d'obtenir un déterminant triangulaire

$$\chi_M(\lambda) = \begin{vmatrix} \lambda - \big(a + (n-1)b\big) & -b & \cdots & -b \\ 0 & \lambda - a + b & & (0) \\ \vdots & & \ddots & \\ 0 & (0) & & \lambda - a + b \end{vmatrix}_{[n]}$$

$$= \Big(\lambda - \big(a + (n-1)b\big)\Big)\big(\lambda - (a-b)\big)^{n-1}.$$

1. Un raisonnement analogue à celui proposé est possible en observant $\mathrm{Ker}(u)$ stable par v.

Les valeurs propres de M sont donc $a+(n-1)b$ et $a-b$ de multiplicités respectives 1 et $n-1 \geqslant 1$ (ces valeurs propres sont distinctes car $b \neq 0$).

(b) **méthode**

On étudie les dimensions des sous-espaces propres de M par un calcul de rang.

La valeur propre $a+(n-1)b$ est simple, le sous-espace propre associé est donc de dimension 1.

Par la formule du rang, la dimension du sous-espace propre associé à la valeur $a-b$ est

$$\dim \operatorname{Ker}\big(M-(a-b)\mathrm{I}_n\big) = n - \operatorname{rg}\big(M-(a-b)\mathrm{I}_n\big) = n - \operatorname{rg}\begin{pmatrix} b & \cdots & b \\ \vdots & & \vdots \\ b & \cdots & b \end{pmatrix} \underset{b\neq 0}{=} n-1.$$

La somme des dimensions des sous-espaces propres est égale à n, la matrice M est donc diagonalisable (Th. 19 p. 127).

Exercice 26 *

Soit $n \geqslant 3$. Déterminer les valeurs propres de la matrice de $\mathcal{M}_n(\mathbb{R})$ suivante

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & & (0) \\ \vdots & & \ddots & \\ 1 & (0) & & 1 \end{pmatrix}.$$

Solution

méthode

Il est quelquefois plus commode de calculer les valeurs propres d'une matrice carrée en étudiant l'équation $AX = \lambda X$ qu'en calculant son polynôme caractéristique [1].

Notons A la matrice étudiée et considérons X une colonne de hauteur n et de coefficients $x_1, \ldots, x_n$. L'équation $AX = \lambda X$ équivaut au système

$$\begin{cases} x_1 + x_2 + \cdots + x_n = \lambda x_1 \\ x_1 + x_2 \qquad\qquad = \lambda x_2 \\ \vdots \qquad\quad \ddots \qquad\quad \vdots \\ x_1 \qquad\qquad + x_n = \lambda x_n \end{cases} \quad \text{soit encore} \quad \begin{cases} x_1 + \cdots + x_n = \lambda x_1 \\ x_1 = (\lambda-1)x_2 \\ \vdots \\ x_1 = (\lambda-1)x_n. \end{cases}$$

1. Le calcul du polynôme caractéristique χ_n est néanmoins possible : en développant le déterminant selon la dernière ligne, on obtient $\chi_n(\lambda) = (\lambda-1)\chi_{n-1}(\lambda) - (\lambda-1)^{n-2}$. Le calcul des premiers termes de cette suite permet de conjecturer $\chi_n(\lambda) = (\lambda-1)^n - (n-1)(\lambda-1)^{n-2}$ ce que l'on vérifie par récurrence.

Cas : $\lambda = 1$. On obtient une solution non nulle en vérifiant les conditions [1]

$$x_1 = 0 \quad \text{et} \quad x_2 + \cdots + x_n = 0.$$

Il suffit par exemple de prendre $x_2 = 1$, $x_3 = -1$ et les autres valeurs nulles.

Cas : $\lambda \neq 1$. Le système équivaut au suivant :

$$\begin{cases} (n-1)x_1 = (\lambda-1)^2 x_1 \\ x_2 = \frac{x_1}{\lambda-1} \\ \vdots \\ x_n = \frac{x_1}{\lambda-1}. \end{cases}$$

Pour $x_1 = 0$, la solution de ce système est nulle. Pour $x_1 \neq 0$, on forme une solution non nulle au système à condition que $(\lambda-1)^2 = n-1$.

Finalement, la matrice admet trois valeurs propres

$$1, \quad 1+\sqrt{n-1} \quad \text{et} \quad 1-\sqrt{n-1}.$$

On peut aussi retrouver ce résultat en commençant par observer que $\operatorname{rg}(A - \mathrm{I}_n) = 2$ ce qui assure que 1 est valeur propre de multiplicité au moins $n-2$. Les deux autres valeurs propres (*a priori* complexes) α et β se déduisent de la résolution des équations

$$\operatorname{tr}(A) = (n-2) \times 1 + \alpha + \beta \quad \text{et} \quad \operatorname{tr}(A^2) = (n-2) \times 1^2 + \alpha^2 + \beta^2.$$

Exercice 27 **

Soit $n \in \mathbb{N}^*$ et χ_n le polynôme caractéristique de

$$A_n = \begin{pmatrix} 0 & 1 & & (0) \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ (0) & & 1 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

(a) Pour $\theta \in]0\,;\pi[$, calculer $u_n = \chi_n(2\cos\theta)$.

(b) Déterminer les valeurs propres de A_n. La matrice A_n est-elle diagonalisable ?

(c) Déterminer les sous-espaces propres de A_n.

Solution

(a) Le déterminant définissant $u_n = \chi_n(2\cos\theta)$ est un déterminant tridiagonal.

méthode

On forme une relation de récurrence linéaire double vérifiée par les éléments de la suite (u_n).

1. L'espace propre associé à la valeur propre 0 est l'intersection de deux hyperplans distincts, c'est un espace de dimension $n - 2 \geqslant 1$.

Pour $n \geqslant 3$, on développe selon la première ligne

$$u_n = 2\cos(\theta)\underbrace{\begin{vmatrix} 2\cos\theta & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2\cos\theta \end{vmatrix}_{[n-1]}}_{=u_{n-1}} + \begin{vmatrix} -1 & -1 & 0 & \cdots & 0 \\ 0 & 2\cos\theta & -1 & & (0) \\ \vdots & -1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & -1 \\ 0 & (0) & & -1 & 2\cos\theta \end{vmatrix}_{[n-1]}.$$

On développe ensuite le second déterminant selon sa première colonne et l'on obtient

$$u_n = 2\cos(\theta)u_{n-1} - u_{n-2}.$$

La suite $(u_n)_{n\geqslant 1}$ est une suite récurrente linéaire d'ordre 2 d'équation caractéristique

$$r^2 - 2\cos(\theta)r + 1 = 0$$

de discriminant $\Delta = -4\sin^2\theta < 0$ et de racines distinctes $e^{i\theta}$ et $e^{-i\theta}$. Il existe donc deux réels λ et μ tels que, pour tout $n \geqslant 1$,

$$u_n = \lambda\cos(n\theta) + \mu\sin(n\theta).$$

Sachant $u_1 = 2\cos\theta$ et $u_2 = 4\cos^2\theta - 1$, on détermine les valeurs[1] de λ et μ et l'on propose une expression simplifiée de u_n :

$$\lambda = 1, \quad \mu = \frac{\cos\theta}{\sin\theta} \quad \text{et} \quad u_n = \frac{\sin\big((n+1)\theta\big)}{\sin\theta}.$$

(b) Pour $\theta \in]0\,;\pi[$,

$$\begin{aligned} \chi_n(2\cos\theta) = 0 &\iff \sin\big((n+1)\theta\big) = 0 \\ &\iff (n+1)\theta \equiv 0 \ [\pi]. \end{aligned}$$

Pour $k \in [\![1\,;n]\!]$, les angles $\theta_k = \frac{k\pi}{n+1}$ sont distincts dans $]0\,;\pi[$. Par stricte décroissance de la fonction cosinus sur $[0\,;\pi]$, les réels

$$\lambda_k = 2\cos\left(\frac{k\pi}{n+1}\right)$$

sont des racines distinctes du polynôme caractéristique χ_n : ce sont des valeurs propres de A_n. Cependant, la matrice A_n est de taille n, elle admet donc au plus n valeurs propres. Les valeurs $\lambda_1, \ldots, \lambda_n$ précédentes sont donc exactement les valeurs propres de A_n. Au surplus, cette matrice est diagonalisable et ses sous-espaces propres sont de dimension 1 (Th. 18 p. 126).

1. La relation de récurrence vérifiée par (u_n) est compatible avec la valeur $u_0 = 1$: il est plus facile de calculer λ et μ à partir de u_0 et u_1 qu'à partir de u_1 et u_2.

(c) Étudions le sous-espace propre de A_n associé à la valeur propre $\lambda_k = 2\cos(\theta_k)$ pour $k \in [\![1\,;n]\!]$. Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$ une colonne de coefficients $x_1, \dots, x_n$.

$$A_nX = \lambda_k X \iff \begin{cases} x_2 = 2\cos(\theta_k)x_1 \\ x_{j-1} + x_{j+1} = 2\cos(\theta_k)x_j \quad \text{pour} \quad j \in [\![2\,;n-1]\!]. \\ x_{n-1} = 2\cos(\theta_k)x_n \end{cases}$$

méthode

‖ On homogénéise le système en introduisant x_0 et x_{n+1} égaux à 0.

$$A_nX = \lambda_k X \iff x_0 = x_{n+1} = 0 \quad \text{et} \quad \forall j \in [\![1\,;n]\!],\ x_{j-1} + x_{j+1} = 2\cos(\theta_k)x_j.$$

La suite finie $(x_j)_{0\leqslant j\leqslant n+1}$ est alors une suite récurrente linéaire double d'équation caractéristique

$$r^2 - 2\cos(\theta_k)r + 1 = 0.$$

Les racines de cette équation sont $e^{i\theta_k}$ et $e^{-i\theta_k}$. Il existe donc deux réels α et β tels que, pour tout $j \in [\![0\,;n+1]\!]$,

$$x_j = \alpha\cos(j\theta_k) + \beta\sin(j\theta_k) = \alpha\cos\left(\frac{jk\pi}{n+1}\right) + \beta\sin\left(\frac{jk\pi}{n+1}\right).$$

Les conditions $x_0 = x_{n+1} = 0$ correspondent à $\alpha = 0$ et, finalement,

$$x_j = \beta\sin\left(\frac{jk\pi}{n+1}\right) \quad \text{pour tout } j \in [\![1\,;n]\!].$$

Le sous-espace propre de A_n associé à la valeur propre λ_k est donc

$$E_{\lambda_k}(A_n) = \operatorname{Vect}\left(\sin\left(\frac{jk\pi}{n+1}\right)\right)_{1\leqslant j\leqslant n} = \operatorname{Vect}\begin{pmatrix}\sin\left(\frac{k\pi}{n+1}\right)\\ \vdots \\ \sin\left(\frac{nk\pi}{n+1}\right)\end{pmatrix}.$$

Exercice 28 ***

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ une matrice[1] à coefficients strictement positifs vérifiant

$$\sum_{j=1}^{n} a_{i,j} = 1 \quad \text{pour tout } i \in [\![1\,;n]\!].$$

(a) Montrer que 1 est valeur propre de A et que toute autre valeur propre complexe λ vérifie $|\lambda| \leqslant 1$.

(b) Établir que si $\lambda \in \mathbb{C}$ est une valeur propre de A vérifiant $|\lambda| = 1$ alors $\lambda = 1$.

(c) Montrer que l'espace propre associé à la valeur propre 1 est une droite.

1. La matrice A est stochastique, voir sujet 23 p. 92.

Solution

(a) Considérons le vecteur colonne $J = {}^t(1 \quad \dots \quad 1)$. On vérifie $AJ = J$ avec $J \neq 0$ et donc 1 est valeur propre de A et J est vecteur propre associé.

Soit λ une valeur propre de A et $X = {}^t(x_1 \quad \dots \quad x_n)$ un vecteur propre associé. L'égalité matricielle $AX = \lambda X$ donne, pour tout $i \in [\![1\,;n]\!]$,

$$\sum_{j=1}^{n} a_{i,j}x_j = \lambda x_i. \qquad (*)$$

méthode

On considère l'indice pour lequel $|x_i|$ est maximal.

Soit i_0 l'indice vérifiant

$$|x_{i_0}| = \max_{1\leqslant i\leqslant n} |x_i|.$$

On a $|x_{i_0}| \neq 0$ car X n'est pas la colonne nulle et l'équation $(*)$ pour $i = i_0$ donne

$$|\lambda|\,|x_{i_0}| = \left|\sum_{j=1}^{n} a_{i_0,j}x_j\right| \leqslant \sum_{j=1}^{n} \underbrace{a_{i_0,j}}_{\geqslant 0}\,|x_j| \leqslant \sum_{j=1}^{n} a_{i_0,j}\,|x_{i_0}| = |x_{i_0}| \qquad (**)$$

car la somme des coefficients de la ligne d'indice i_0 de A est égale à 1. En simplifiant par $|x_{i_0}|$ qui est strictement positif, on conclut $|\lambda| \leqslant 1$.

(b) On reprend les notations précédentes avec de plus $|\lambda| = 1$. Il y a donc égalité dans l'inégalité $(**)$.

méthode

Il y a égalité [1] dans l'inégalité triangulaire complexe

$$|z_1 + \cdots + z_n| \leqslant |z_1| + \cdots + |z_n|$$

si, et seulement si, les points d'affixes $z_1, \dots, z_n$ figurent sur une même droite issue de l'origine. Cela signifie l'existence d'un réel θ tel que $z_j = |z_j|\,\mathrm{e}^{\mathrm{i}\theta}$ pour tout $j \in [\![1\,;n]\!]$.

L'égalité dans la première inégalité de $(**)$ donne l'existence d'un réel θ tel que

$$a_{i_0,j}x_j = \underbrace{|a_{i_0,j}x_j|}_{a_{i_0,j}|x_j|}\,\mathrm{e}^{\mathrm{i}\theta} \quad \text{pour tout } j \in [\![1\,;n]\!].$$

L'égalité dans la deuxième inégalité de $(**)$ entraîne

$$a_{i_0,j}\,|x_j| = a_{i_0,j}\,|x_{i_0}| \quad \text{pour tout } j \in [\![1\,;n]\!].$$

1. Voir sujet 32 du chapitre 3 de l'ouvrage *Exercices d'analyse MPSI* dans la même collection.

Par hypothèse, le réel $a_{i_0,j}$ est non nul pour tout indice $j \in [\![1\,;n]\!]$ et donc

$$x_j = |x_{i_0}|\,\mathrm{e}^{\mathrm{i}\theta} \quad \text{pour tout } j \in [\![1\,;n]\!].$$

Les $x_1, \ldots, x_n$ sont donc tous égaux et la colonne X est colinéaire à la colonne J vecteur propre associé à la valeur propre 1. On conclut $\lambda = 1$.

(c) L'étude qui précède a montré qu'un vecteur propre associé à une valeur propre λ vérifiant $|\lambda| = 1$ est colinéaire à J. Les vecteurs propres associés à la valeur propre 1 sont donc tous éléments de $\mathrm{Vect}(J)$ et l'espace propre correspondant est de dimension 1.

4.7.4 Polynôme caractéristique

Exercice 29 *
Soit $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant $\det(A) < 0$. Établir que A possède au moins une valeur propre réelle.

Solution

méthode
‖ Les valeurs propres de A sont les racines de χ_A (Th. 9 p. 123).

Le polynôme caractéristique de A est unitaire de degré n et son coefficient constant est $(-1)^n \det(A)$:

$$\chi_A = X^n - \mathrm{tr}(A)X^{n-1} + \cdots + (-1)^n \det(A).$$

Lorsque $\det(A) < 0$, on a

$$(-1)^n \chi_A(0) = \det(A) < 0 \quad \text{et} \quad \lim_{t \to -\infty} (-1)^n \chi_A(t) = +\infty.$$

La fonction $t \mapsto \chi_A(t)$ est continue sur l'intervalle $]-\infty\,;0]$ et l'on peut appliquer le théorème des valeurs intermédiaires pour affirmer l'existence d'une racine au polynôme χ_A. Ainsi [1], la matrice A possède une valeur propre dans $]-\infty\,;0[$.

Exercice 30 *
Deux matrices de $\mathcal{M}_n(\mathbb{K})$ ayant même polynôme caractéristique et même rang sont-elles nécessairement semblables ?

1. On peut aussi raisonner ainsi : le déterminant d'une matrice est le produit de ses valeurs propres complexes comptées avec multiplicité, si la matrice est réelle, les valeurs propres complexes sont deux à deux conjuguées et, en l'absence de valeurs propres réelles, le déterminant de A est positif car uniquement produit de facteurs $\lambda\overline{\lambda}$.

Solution

méthode

Il suffit que deux matrices triangulaires aient la même diagonale pour posséder le même polynôme caractéristique.

Pour $n \geqslant 2$, considérons les matrices $A = \mathrm{I}_n$ et $B = \mathrm{I}_n + E_{i,j}$ (avec $i \neq j$ dans $[\![1\,;n]\!]$). Elles ont toutes deux $(X-1)^n$ pour polynôme caractéristique et sont inversibles donc de rang n. Cependant, elles ne sont pas semblables car une matrice semblable à I_n s'écrit $P^{-1}\mathrm{I}_n P$ avec P inversible et est donc nécessairement égale à I_n.

Pour $n = 1$, l'égalité du polynôme caractéristique entraîne l'égalité des matrices.

Exercice 31 **

Soit u un endomorphisme d'un espace vectoriel réel E de dimension $n \geqslant 2$.

(a) Exprimer le polynôme caractéristique de u en fonction de $\mathrm{tr}(u)$ lorsque $\mathrm{rg}(u) = 1$.

(b) Même question en utilisant aussi $\mathrm{tr}(u^2)$ lorsque $\mathrm{rg}(u) = 2$.

Solution

(a) **méthode**

Lorsqu'il n'est pas réduit au vecteur nul, le noyau d'un endomorphisme est le sous-espace propre associé à la valeur propre 0.

Par la formule du rang, $\dim \mathrm{Ker}(u) = n - \mathrm{rg}(u) = n-1 \geqslant 1$ et 0 est donc valeur propre de u. De plus, le sous-espace propre associé est de dimension $n-1$ et la multiplicité de la valeur propre 0 est donc au moins égale à $n-1$ (Th. 14 p. 125). Ceci signifie que X^{n-1} divise le polynôme caractéristique de u. Or on sait aussi

$$\chi_u = X^n - \mathrm{tr}(u)X^{n-1} + \cdots + (-1)^n \det(u)$$

et donc

$$\chi_u = X^n - \mathrm{tr}(u)X^{n-1}.$$

En substance, soulignons que $\mathrm{tr}(u)$ est valeur propre[1] de u.

(b) **méthode**

On figure l'endomorphisme u dans une base adaptée à son noyau.

Par la formule du rang, $\dim \mathrm{Ker}(u) = n-2$. Considérons alors une base de E dont les $n-2$ derniers vecteurs constituent une base du noyau de u. Les dernières colonnes de la matrice de u dans cette base sont nulles et celle-ci s'écrit

$$M = \begin{pmatrix} a & b & 0 & \cdots & 0 \\ c & d & 0 & \cdots & 0 \\ * & * & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ * & * & 0 & \cdots & 0 \end{pmatrix} \quad \text{avec} \quad a,b,c,d \in \mathbb{R}.$$

1. Voir aussi le sujet 41 p. 175.

Pour $\lambda \in \mathbb{R}$,

$$\chi_u(\lambda) = \begin{vmatrix} \lambda - a & -b & 0 & \cdots & 0 \\ -c & \lambda - d & 0 & & 0 \\ * & * & \lambda & & (0) \\ \vdots & \vdots & & \ddots & \\ * & * & (0) & & \lambda \end{vmatrix} = \lambda^{n-2}\big(\lambda^2 - \underbrace{(a+d)}_{=\mathrm{tr}(u)}\lambda + ad - bc\big).$$

En calculant M^2 à l'aide d'un produit par blocs, on a aussi $\mathrm{tr}(u^2) = a^2 + 2bc + d^2$ et donc

$$\begin{aligned} ad - bc &= \frac{1}{2}\Big((a+d)^2 - \big(a^2 + 2bc + d^2\big)\Big) \\ &= \frac{1}{2}\Big(\big(\mathrm{tr}(u)\big)^2 - \mathrm{tr}\big(u^2\big)\Big). \end{aligned}$$

On peut alors conclure [1]

$$\chi_u = X^{n-2}\left(X^2 - \mathrm{tr}(u)X + \frac{1}{2}\Big(\big(\mathrm{tr}(u)\big)^2 - \mathrm{tr}\big(u^2\big)\Big)\right).$$

Exercice 32 ** (Matrice compagnon)

Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$ polynôme de $\mathbb{K}[X]$ et

$$A = \begin{pmatrix} 0 & & (0) & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & 0 & \vdots \\ (0) & & 1 & -a_{n-1} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}).$$

Exprimer le polynôme caractéristique de A en fonction de P.

Solution

Soit $\lambda \in \mathbb{K}$. Calculons $\chi_A(\lambda) = \det(\lambda \mathrm{I}_n - A)$.

méthode

‖ On peut commencer par étudier le cas $n = 4$.

On suppose $n = 4$ et l'on calcule le déterminant en faisant apparaître des zéros au

1. Plus généralement, si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres complexes comptées avec multiplicité d'une matrice figurant u, le coefficient de X^{n-2} dans χ_u est la somme des doubles produits $\lambda_i\lambda_j$. Celle-ci se déduit de la somme des λ_i et de celle des λ_i^2 données respectivement par la trace de u et de u^2. On trouve alors que, sans hypothèse sur le rang, le coefficient de X^{n-2} dans χ_u est $\frac{1}{2}\big((\mathrm{tr}\,u)^2 - \mathrm{tr}(u^2)\big)$.

début de la première ligne

$$\chi_A(\lambda) = \begin{vmatrix} \lambda & 0 & 0 & a_0 \\ -1 & \lambda & 0 & a_1 \\ 0 & -1 & \lambda & a_2 \\ 0 & 0 & -1 & \lambda + a_3 \end{vmatrix}$$
$$\underset{L_1 \leftarrow L_1 + \lambda L_2}{=} \begin{vmatrix} 0 & \lambda^2 & 0 & a_0 + a_1\lambda \\ -1 & \lambda & 0 & a_1 \\ 0 & -1 & \lambda & a_2 \\ 0 & 0 & -1 & \lambda + a_3 \end{vmatrix}$$
$$\underset{L_1 \leftarrow L_1 + \lambda^2 L_3}{=} \begin{vmatrix} 0 & 0 & \lambda^3 & a_0 + a_1\lambda + a_2\lambda^2 \\ -1 & \lambda & 0 & a_1 \\ 0 & -1 & \lambda & a_2 \\ 0 & 0 & -1 & \lambda + a_3 \end{vmatrix}$$
$$\underset{L_1 \leftarrow L_1 + \lambda^3 L_4}{=} \begin{vmatrix} 0 & 0 & 0 & P(\lambda) \\ -1 & \lambda & 0 & a_1 \\ 0 & -1 & \lambda & a_2 \\ 0 & 0 & -1 & \lambda + a_3 \end{vmatrix}.$$

On développe ensuite le déterminant selon la première ligne pour terminer le calcul

$$\chi_A(\lambda) = (-1)^{1+4} P(\lambda) \begin{vmatrix} -1 & \lambda & 0 \\ 0 & -1 & \lambda \\ 0 & 0 & -1 \end{vmatrix} = P(\lambda).$$

De façon générale, on calcule le polynôme caractéristique en taille n en commençant par réaliser l'opération [1] $L_1 \leftarrow L_1 + \lambda L_2 + \cdots + \lambda^{n-1} L_n$

$$\chi_A(\lambda) = \begin{vmatrix} \lambda & & (0) & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & \lambda & a_{n-2} \\ (0) & & -1 & \lambda + a_{n-1} \end{vmatrix} = \begin{vmatrix} 0 & \cdots & \cdots & 0 & P(\lambda) \\ -1 & \lambda & & (0) & a_1 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \lambda & a_{n-2} \\ (0) & & & -1 & \lambda + a_{n-1} \end{vmatrix}$$

On développe ensuite selon la première ligne

$$\chi_A(\lambda) = (-1)^{n+1} P(\lambda) \begin{vmatrix} -1 & \lambda & & (0) \\ & \ddots & \ddots & \\ & & \ddots & \lambda \\ (0) & & & -1 \end{vmatrix}_{[n-1]} = (-1)^{n+1} P(\lambda) \times (-1)^{n-1} = P(\lambda).$$

Finalement, le polynôme caractéristique [2] de A est égal à P.

1. On peut aussi développer selon la dernière colonne en étant très attentif aux mineurs introduits.
2. On retient que tout polynôme unitaire de degré n peut se voir comme le polynôme caractéristique d'une matrice carrée de taille n bien choisie.

Exercice 33 **

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. En multipliant à droite et à gauche la matrice

$$M = \begin{pmatrix} \lambda \mathrm{I}_n & A \\ B & \mathrm{I}_n \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{K})$$

par des matrices triangulaires par blocs convenables, établir [1] $\chi_{AB} = \chi_{BA}$.

Solution

méthode

On multiplie M par une matrice triangulaire par blocs afin d'obtenir un produit lui aussi triangulaire par blocs.

D'une part,

$$\begin{pmatrix} \lambda \mathrm{I}_n & A \\ B & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \mathrm{I}_n & \mathrm{O}_n \\ -B & \mathrm{I}_n \end{pmatrix} = \begin{pmatrix} \lambda \mathrm{I}_n - AB & A \\ \mathrm{O}_n & \mathrm{I}_n \end{pmatrix}.$$

D'autre part,

$$\begin{pmatrix} \mathrm{I}_n & \mathrm{O}_n \\ -B & \lambda \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \lambda \mathrm{I}_n & A \\ B & \mathrm{I}_n \end{pmatrix} = \begin{pmatrix} \lambda \mathrm{I}_n & A \\ \mathrm{O}_n & \lambda \mathrm{I}_n - BA \end{pmatrix}.$$

Le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux et donc

$$\det(M) \times 1 = \chi_{AB}(\lambda) \quad \text{et} \quad \lambda^n \det(M) = \lambda^n \chi_{BA}(\lambda).$$

On en déduit $\lambda^n \chi_{AB}(\lambda) = \lambda^n \chi_{BA}(\lambda)$ puis l'égalité des polynômes [2] χ_{AB} et χ_{BA}.

4.7.5 Matrices diagonalisables

Exercice 34 *

Soit $a, b, c \in \mathbb{R}$. La matrice suivante est-elle diagonalisable dans $\mathcal{M}_3(\mathbb{R})$?

$$M = \begin{pmatrix} 0 & -b & c \\ a & 0 & -c \\ -a & b & 0 \end{pmatrix}.$$

1. On trouvera une autre démonstration de cette égalité dans le sujet 30 du chapitre 5 de l'ouvrage *Exercices d'analyse MP*.

2. Pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$, on montre par le même procédé $X^p \chi_{AB} = X^n \chi_{BA}$.

Solution

On calcule le polynôme caractéristique de M en $\lambda \in \mathbb{R}$. Par développement du déterminant selon sa première ligne

$$\chi_M(\lambda) = \begin{vmatrix} \lambda & b & -c \\ -a & \lambda & c \\ a & -b & \lambda \end{vmatrix} = \lambda \begin{vmatrix} \lambda & c \\ -b & \lambda \end{vmatrix} - b \begin{vmatrix} -a & c \\ a & \lambda \end{vmatrix} - c \begin{vmatrix} -a & \lambda \\ a & -b \end{vmatrix} = \lambda^3 + \lambda(ab + bc + ca).$$

Le polynôme caractéristique de M est donc $\chi_M = X\big(X^2 + (ab + bc + ca)\big)$.

méthode

‖ On discute selon le signe du réel $\delta = ab + bc + ca$.

Cas : $\delta < 0$. La matrice M est diagonalisable car elle est de taille 3 et possède 3 valeurs propres distinctes : 0, $\sqrt{-\delta}$ et $-\sqrt{-\delta}$.

Cas : $\delta > 0$. La matrice M n'est pas diagonalisable dans[1] $\mathcal{M}_3(\mathbb{R})$ car son polynôme caractéristique n'est pas scindé sur $\mathbb{R}$.

Cas : $\delta = 0$. La matrice M admet 0 pour seule et unique valeur propre. Si M est diagonalisable, elle est semblable à une matrice diagonale avec sur la diagonale ses valeurs propres, donc des 0. C'est alors la matrice nulle et la réciproque est immédiate. Ainsi, lorsque $\delta = 0$, la matrice M est diagonalisable si, et seulement si, $a = b = c = 0$.

Exercice 35 *

Soit $P = X^n - \big(a_{n-1}X^{n-1} + \cdots + a_1X + a_0\big)$ un polynôme unitaire réel de degré n.

(a) Déterminer les sous-espaces propres de la matrice

$$M = \begin{pmatrix} 0 & 1 & & (0) \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-1} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

(b) À quelle condition relative au polynôme P, la matrice M est-elle diagonalisable ?

Solution

(a) **méthode**

‖ On étudie directement[2] l'équation aux éléments propres $MX = \lambda X$.

Soit $\lambda \in \mathbb{R}$ et $X \in \mathcal{M}_{n,1}(\mathbb{R})$ une colonne de coefficients $x_0, x_1, \ldots, x_{n-1}$. L'étude de

1. Cette matrice est cependant diagonalisable dans $\mathbb{C}$ car y possède 3 valeurs propres distinctes.
2. On peut aussi calculer le polynôme caractéristique de M en s'inspirant du sujet 32 p. 166 mais ce n'est pas nécessaire.

l'équation $MX = \lambda X$ conduit à la résolution du système

$$\begin{cases} x_1 = \lambda x_0 \\ x_2 = \lambda x_1 \\ \vdots \\ x_{n-1} = \lambda x_{n-2} \\ a_0 x_0 + \cdots + a_{n-1} x_{n-1} = \lambda x_{n-1} \end{cases} \quad \text{soit} \quad \begin{cases} x_1 = \lambda x_0 \\ x_2 = \lambda^2 x_0 \\ \vdots \\ x_{n-1} = \lambda^{n-2} x_0 \\ 0 = P(\lambda) x_0. \end{cases}$$

Si λ n'est pas racine de P, la dernière équation donne $x_0 = 0$ et l'on en déduit la nullité de tous les coefficients de X : λ n'est pas valeur propre de P.

Si λ est racine de P, la dernière équation du système se simplifie ce qui permet de définir une solution non nulle à l'équation $MX = \lambda X$: λ est valeur propre et le sous-espace propre associé est

$$E_\lambda(M) = \left\{ \begin{pmatrix} x_0 \\ \lambda x_0 \\ \vdots \\ \lambda^{n-1} x_0 \end{pmatrix} \;\middle|\; x_0 \in \mathbb{R} \right\} = \text{Vect} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}.$$

(b) Tous les sous-espaces propres de M sont de dimension 1. La matrice M de $\mathcal{M}_n(\mathbb{R})$ est alors diagonalisable si, et seulement si, elle possède exactement n valeurs propres distinctes. Cela revient à dire que le polynôme P possède exactement n racines réelles distinctes [1].

4.7.6 Endomorphismes diagonalisables

Exercice 36 *

Soit f un endomorphisme d'un espace vectoriel réel E de dimension $n \geqslant 1$ admettant exactement n valeurs propres distinctes. À quelle condition portant sur un vecteur x de E peut-on affirmer que la famille $\big(x, f(x), \ldots, f^{n-1}(x)\big)$ est une base de E ?

Solution

Notons $\lambda_1, \ldots, \lambda_n$ les n valeurs propres distinctes de f et $e_1, \ldots, e_n$ des vecteurs propres associés. La famille $e = (e_1, \ldots, e_n)$ est une base [2] de E dans laquelle la matrice de f est diagonale.

Soit x un vecteur de E.

méthode

On introduit les coordonnées du vecteur x dans la base e de diagonalisation.

1. Celles-ci sont nécessairement simples car P est de degré n.

2. La famille e est constituée de $n = \dim E$ vecteurs de E et c'est une famille libre car formée de vecteurs propres associés à des valeurs propres deux à deux distinctes : ceci reproduit une démonstration possible du Th. 16 p. 126.

On peut écrire

$$x = x_1 e_1 + \cdots + x_n e_n \quad \text{avec} \quad x_1, \ldots, x_n \in \mathbb{R}.$$

Pour tout $i \in [\![1\,;n]\!]$, on a $f(e_i) = \lambda_i e_i$ car e_i est vecteur propre associé à la valeur propre λ_i. Par une récurrence immédiate, on en déduit $f^k(e_i) = \lambda_i^k e_i$ pour tout $k \in \mathbb{N}$. On a donc

$$f^k(x) = \lambda_1^k x_1 e_1 + \cdots + \lambda_n^k x_n e_n.$$

On peut alors étudier si la famille $\big(x, f(x), \ldots, f^{n-1}(x)\big)$ est une base en calculant son déterminant dans la base e :

$$\begin{aligned}\det_e\big(x, f(x), \ldots, f^{n-1}(x)\big) &= \det \mathrm{Mat}_e\big(x, f(x), \ldots, f^{n-1}(x)\big) \\ &= \begin{vmatrix} x_1 & \lambda_1 x_1 & \lambda_1^2 x_1 & \cdots & \lambda_1^{n-1} x_1 \\ x_2 & \lambda_2 x_2 & \lambda_2^2 x_2 & \cdots & \lambda_2^{n-1} x_2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_n & \lambda_n x_n & \lambda_n^2 x_n & \cdots & \lambda_n^{n-1} x_n \end{vmatrix}.\end{aligned}$$

En factorisant x_i sur chaque ligne, on fait apparaître un déterminant de Vandermonde dont la valeur est connue [1]

$$\det_e\big(x, f(x), \ldots, f^{n-1}(x)\big) = x_1 x_2 \ldots x_n \prod_{1 \leqslant i < j \leqslant n} \underbrace{(\lambda_j - \lambda_i)}_{\neq 0}.$$

La famille $\big(x, f(x), \ldots, f^{n-1}(x)\big)$ est donc une base si, et seulement si, $x_1 x_2 \ldots x_n \neq 0$. Les vecteurs x cherchés sont donc ceux n'ayant pas de coordonnées nulles dans la base de diagonalisation e : x est la somme de vecteurs propres associés à chacune des n valeurs propres.

Exercice 37 **

Soit f un endomorphisme diagonalisable d'un espace vectoriel E de dimension $n \geqslant 1$.

(a) Montrer qu'un endomorphisme g commute avec f si, et seulement si, les sous-espaces propres de f sont stables par g.

(b) On suppose de nouveau f diagonalisable et l'on note $\lambda_1, \ldots, \lambda_m$ ses valeurs propres et $\alpha_1, \ldots, \alpha_m$ leurs multiplicités respectives. Montrer que l'ensemble des endomorphismes qui commutent avec f est un sous-espace vectoriel de $\mathcal{L}(E)$ et donner sa dimension.

Solution

(a) ($\Longleftarrow$) Lorsque deux endomorphismes commutent, on sait que les sous-espaces propres de l'un sont stables [2] pour l'autre.

($\Longrightarrow$) Supposons que tous les sous-espaces propres de f sont stables par g.

1. Voir sujet 33 p. 101.
2. Voir sujet 13 p. 146.

méthode

On vérifie que les applications linéaires $f \circ g$ et $g \circ f$ sont égales sur chaque espace propre de f.

Soit λ une valeur propre de f et x un vecteur de l'espace propre $E_\lambda(f)$. On a $f(x) = \lambda x$ et donc par linéarité

$$(g \circ f)(x) = g\big(f(x)\big) = g(\lambda x) = \lambda g(x).$$

Aussi, $g(x)$ est élément de $E_\lambda(f)$ car l'espace $E_\lambda(f)$ est stable par g et donc

$$(f \circ g)(x) = f\big(g(x)\big) = \lambda g(x).$$

Ainsi, les deux applications linéaires $f \circ g$ et $g \circ f$ sont égales sur chaque espace propre de f. Or f est diagonalisable et l'espace E est donc la somme directe des sous-espaces propres de f. Ceci permet de décomposer n'importe quel vecteur de E en somme de vecteurs sur lesquels les applications $f \circ g$ et $g \circ f$ sont égales. On en déduit que ces applications linéaires sont égales sur E : les endomorphismes f et g commutent.

(b) **méthode**

La stabilité des espaces propres se traduit par une représentation matricielle diagonale par blocs.

L'endomorphisme f étant diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_m$, on peut écrire la décomposition en somme directe

$$E = E_1 \oplus \cdots \oplus E_m \quad \text{avec} \quad E_k = E_{\lambda_k}(f) \text{ pour tout } k \in [\![1\,;m]\!].$$

Les espaces E_k sont de dimension α_k la multiplicité de la valeur propre λ_k.

Considérons ensuite une base e adaptée à la décomposition précédente. Les endomorphismes g qui commutent avec f sont ceux pour lesquels chaque espace E_k est stable. Ils correspondent [1] aux endomorphismes dont la matrice dans e est de la forme

$$\begin{pmatrix} A_1 & & (0) \\ & \ddots & \\ (0) & & A_m \end{pmatrix} \quad \text{avec} \quad A_k \in \mathcal{M}_{\alpha_k}(\mathbb{K}).$$

L'ensemble des matrices de cette forme est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension [2] $\alpha_1^2 + \cdots + \alpha_m^2$. Par l'isomorphisme de représentation matricielle dans la base e, l'ensemble C_f des endomorphismes qui commutent avec f est un sous-espace vectoriel de $\mathcal{L}(E)$ de même dimension.

1. Un endomorphisme g qui commute avec f est aussi entièrement déterminé par les endomorphismes qu'il induit sur les espaces E_k : on peut mettre en correspondance l'ensemble des endomorphismes qui commutent avec f avec l'espace $\mathcal{L}(E_1) \times \cdots \times \mathcal{L}(E_m)$.

2. La dimension de cet espace correspond au nombre de matrices élémentaires $E_{i,j}$ qui lui appartiennent.

Exercice 38 **

Soit f un endomorphisme d'un espace vectoriel réel E de dimension $n \geqslant 1$ possédant exactement n valeurs propres distinctes.

(a) Déterminer les dimensions des sous-espaces propres de f.

(b) Soit g un endomorphisme de E vérifiant $g^2 = f$. Montrer que g et f commutent. En déduire que les vecteurs propres de f sont aussi des vecteurs propres de g.

(c) Combien y a-t-il d'endomorphismes g de E solutions de l'équation $g^2 = f$?

Solution

(a) L'endomorphisme f possède n valeurs propres en dimension n, il est donc diagonalisable et ses sous-espaces propres sont des droites vectorielles (Th. 16 p. 126).

(b) Puisque $g^2 = f$, on peut écrire $g \circ f = g^3 = f \circ g$ et donc les endomorphismes f et g commutent. Les sous-espaces propres de f sont stables [1] par g.

méthode

Les vecteurs propres d'un endomorphisme sont ceux engendrant une droite vectorielle stable (Th. 3 p. 121).

Soit x un vecteur propre de f associé à une valeur propre λ. Le vecteur x est un vecteur non nul de la droite vectorielle $E_\lambda(f)$ et engendre donc celle-ci : $E_\lambda(f) = \text{Vect}(x)$. Or cette droite vectorielle est stable par g et donc x est vecteur propre de g (Th. 3 p. 121).

(c) **méthode**

Une base diagonalisant f diagonalise aussi les solutions de l'équation $g^2 = f$.

L'endomorphisme f étant diagonalisable, on peut introduire $e = (e_1, \ldots, e_n)$ une base de vecteurs propres dans laquelle la matrice de f est

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \ldots, \lambda_n$ ses valeurs propres deux à deux distinctes.

Soit g un endomorphisme de E solution de l'équation $g^2 = f$. L'étude qui précède assure que les vecteurs de la base e sont des vecteurs propres de g. La matrice de g dans cette base est donc diagonale.

Résoudre l'équation $g^2 = f$ revient alors à résoudre l'équation $X^2 = D$ d'inconnue X une matrice diagonale réelle : il y a autant [2] d'endomorphismes g solutions que de matrices X diagonales vérifiant $X^2 = D$. En notant $x_1, \ldots, x_n$ les coefficients diagonaux de

1. Voir sujet 37 p. 171.

2. Le morphisme de représentation matricielle dans la base e est une bijection qui transforme les solutions de l'équation $g^2 = f$ en celles de l'équation $X^2 = D$.

la matrice X, résoudre l'équation $X^2 = D$ équivaut à résoudre le système constitué des équations $x_i^2 = \lambda_i$ pour $i \in [\![1\,;n]\!]$.

S'il existe un λ_i strictement négatif, il n'y a pas de solutions réelles à l'équation $x_i^2 = \lambda_i$ et il n'existe pas de matrices X telles que $X^2 = D$.

Si tous les λ_i sont positifs, chaque équation $x_i^2 = \lambda_i$ admet deux solutions $\sqrt{\lambda_i}$ et $-\sqrt{\lambda_i}$. Les matrices X solutions sont alors les

$$\begin{pmatrix} \pm\sqrt{\lambda_1} & & (0) \\ & \ddots & \\ (0) & & \pm\sqrt{\lambda_n} \end{pmatrix}.$$

Si aucune des valeurs propres n'est nulle, il y a 2^n solutions. Si l'une des valeurs propres est nulle, celle-ci est forcément unique car les valeurs propres sont supposées deux à deux distinctes et il y a 2^{n-1} solutions.

Exercice 39 ***

Soit u un endomorphisme d'un espace vectoriel complexe E de dimension finie non nulle vérifiant [1] :

« Tout sous-espace vectoriel stable par u admet un supplémentaire stable ».

Montrer que l'endomorphisme u est diagonalisable.

Solution

Commençons par exposer l'idée générale. Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie non nulle possède au moins une valeur propre (Th. 11 p. 124). On peut donc introduire λ une valeur propre de u. L'espace $E_\lambda(u)$ est stable par u et possède par hypothèse un supplémentaire stable F. Si l'espace F est réduit au vecteur nul, on peut conclure que u est diagonalisable. Sinon, l'endomorphisme induit par u sur F possède au moins une valeur propre μ. Celle-ci est aussi valeur propre de u et l'espace $E_\lambda(u) \oplus E_\mu(u)$ est stable par u. Il possède donc un supplémentaire stable G ce qui permet de répéter le raisonnement jusqu'à épuisement... Exprimons maintenant une résolution concise.

méthode

‖ On introduit l'espace somme des sous-espaces propres de u.

Le sous-espace vectoriel F déterminé par

$$F = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u)$$

est stable par u car somme de sous-espaces vectoriels stables. Par hypothèse, il admet un supplémentaire stable G. Si l'espace G n'est pas réduit au vecteur nul, l'endomorphisme induit par u sur G possède un vecteur propre qui sera aussi vecteur propre de u donc élément de F. C'est contradictoire car F et G sont en somme directe et seul le vecteur nul leur est commun. On en déduit $G = \{0_E\}$ puis $F = E$: l'endomorphisme u diagonalisable.

1. On dit que l'endomorphisme u est *semi-simple*.

4.7.7 Trigonalisation

Exercice 40 *

Montrer qu'une matrice triangulaire inférieure est trigonalisable et proposer une matrice réalisant cette trigonalisation.

Solution

Soit A une matrice triangulaire inférieure de $\mathcal{M}_n(\mathbb{K})$

$$A = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (*) & & \lambda_n \end{pmatrix}.$$

Son polynôme caractéristique est $\chi_A = (X-\lambda_1)\dots(X-\lambda_n)$. Il s'agit d'un polynôme scindé sur $\mathbb{K}$ et l'on peut donc affirmer que la matrice A est trigonalisable (Th. 22 p. 128). Retrouvons ce résultat en proposant de plus une matrice de trigonalisation.

méthode

On forme une base trigonalisant l'endomorphisme a canoniquement associé à A en renversant l'ordre des vecteurs.

Notons $e = (e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$. La matrice A étant triangulaire inférieure, on a pour tout indice $j \in [\![1\,;n]\!]$

$$a(e_j) \in \text{Vect}(e_j, \dots, e_n). \qquad (*)$$

Considérons alors la base renversée $e' = (e'_1, \dots, e'_n)$ avec $e'_j = e_{n+1-j}$, autrement dit, la base $e' = (e_n, \dots, e_1)$. La propriété $(*)$ donne pour tout indice j

$$a(e'_j) \in \text{Vect}(e'_1, \dots, e'_j).$$

Ainsi, la base e' trigonalise l'endomorphisme a (Th. 20 p. 127) et la matrice de passage P de la base e à la base e' (qui est une matrice de permutation) réalise une trigonalisation de la matrice A

$$A = PTP^{-1} \quad \text{avec} \quad P = \begin{pmatrix} (0) & & 1 \\ & \iddots & \\ 1 & & (0) \end{pmatrix} \text{ et } T = \begin{pmatrix} \lambda_n & & (*) \\ & \ddots & \\ (0) & & \lambda_1 \end{pmatrix}.$$

Exercice 41 *

Montrer que la trace d'une matrice réelle de rang 1 en est valeur propre.

Solution

méthode

La trace d'une matrice réelle est la somme de ses valeurs propres complexes comptées avec multiplicité [1] (Th. 24 p. 128).

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice de rang 1. Par la formule du rang son noyau est un sous-espace vectoriel de dimension $n-1$. Or le noyau d'une matrice est aussi le sous-espace propre associé à la valeur propre 0. Celle-ci est donc de multiplicité au moins égale à $n-1$. Cependant, la matrice A possède exactement n valeurs propres complexes comptées avec multiplicité, il en reste donc encore une à déterminer [2]. La trace de A étant la somme de toutes les valeurs propres complexes comptées avec multiplicité, on peut affirmer que cette trace détermine la dernière valeur propre de A et celle-ci est réelle.

Notons que l'on peut aussi résoudre ce sujet en observant que la matrice A est semblable à

$$\begin{pmatrix} 0 & \cdots & 0 & \alpha_1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \alpha_n \end{pmatrix}.$$

Exercice 42 **

Soit $A, B \in \mathcal{M}_n(\mathbb{C})$ vérifiant $AB = BA$.

(a) Montrer que les matrices A et B ont au moins un vecteur propre en commun.

(b) Établir que les matrices A et B sont simultanément [3] trigonalisables.

Solution

(a) On introduit les endomorphismes a et b de $\mathbb{C}^n$ canoniquement associés aux matrices A et B. Ceux-ci commutent et possèdent [4] donc un vecteur propre en commun.

(b) On raisonne par récurrence sur la taille $n \in \mathbb{N}^*$ des matrices.

Pour $n = 1$, la propriété est entendue car une matrice de taille 1 est triangulaire.

Supposons la propriété vérifiée au rang $n - 1 \geqslant 1$ et considérons $A, B \in \mathcal{M}_n(\mathbb{C})$ vérifiant $AB = BA$.

méthode

Le premier vecteur d'une base de trigonalisation (ou la première colonne d'une matrice de trigonalisation) est un vecteur propre.

Soit e_1 un vecteur propre commun aux endomorphismes a et b canoniquement associés aux matrices A et B : $a(e_1) = \lambda_1 e_1$ et $b(e_1) = \mu_1 e_1$ avec $\lambda_1, \mu_1 \in \mathbb{C}$. Le vecteur e_1

1. Il importe d'être précis : la trace n'est pas seulement « la somme des valeurs propres ».

2. Celle-ci peut d'ailleurs être égale à 0 : une matrice triangulaire supérieure stricte de rang 1 possède 0 pour seule valeur propre.

3. Autrement dit, il existe une même matrice $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $P^{-1}AP$ et $P^{-1}BP$ sont triangulaires supérieures.

4. L'endomorphisme complexe a possède au moins une valeur propre. Le sous-espace propre de a associé est stable par b et l'endomorphisme que b y induit admet un vecteur propre : voir sujet 24 p. 157.

n'est pas nul, on complète celui-ci en une base e de $\mathbb{C}^n$ et l'on forme la matrice de passage $P \in \mathrm{GL}_n(\mathbb{C})$ de la base canonique à la base e. Par la formule changement de base, on obtient

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * \\ 0 & A' \end{pmatrix} \quad \text{et} \quad P^{-1}BP = \begin{pmatrix} \mu_1 & * \\ 0 & B' \end{pmatrix} \quad \text{avec} \quad A', B' \in \mathcal{M}_{n-1}(\mathbb{C}).$$

Puisque les matrices A et B commutent, les matrices $P^{-1}AP$ et $P^{-1}BP$ commutent aussi et un calcul par blocs donne $A'B' = B'A'$. Par hypothèse de récurrence, il existe alors une matrice $Q' \in \mathrm{GL}_{n-1}(\mathbb{C})$ telle que $Q'^{-1}A'Q'$ et $Q'^{-1}B'Q'$ sont toutes deux triangulaires supérieures :

$$Q'^{-1}A'Q' = \begin{pmatrix} \lambda_2 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{et} \quad Q'^{-1}B'Q' = \begin{pmatrix} \mu_2 & & (*) \\ & \ddots & \\ (0) & & \mu_n \end{pmatrix}.$$

Considérons ensuite la matrice inversible

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \quad \text{d'inverse} \quad Q^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & Q'^{-1} \end{pmatrix}.$$

Un calcul par blocs donne alors

$$Q^{-1}(P^{-1}AP)Q = \begin{pmatrix} \lambda_1 & * \\ 0 & Q'^{-1}A'Q' \end{pmatrix} = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

et l'on obtient un résultat analogue pour l'expression de $Q^{-1}(P^{-1}BP)Q$.

Finalement, la matrice $R = PQ \in \mathrm{GL}_n(\mathbb{C})$ trigonalise simultanément A et B. La récurrence est établie [1].

Exercice 43 ***

Soit u et v deux endomorphismes d'un espace vectoriel complexe E de dimension finie $n \geqslant 1$ vérifiant

$$u \circ v - v \circ u = v.$$

(a) Montrer que le noyau de v n'est pas réduit au vecteur nul.

(b) En déduire u possède un vecteur propre dans $\mathrm{Ker}(v)$.

(c) Établir qu'il existe une base de trigonalisation commune à u et v dans laquelle la matrice de v est triangulaire supérieure stricte.

1. Cette résolution est essentiellement basée sur l'existence d'un vecteur propre commun et sur la propagation de l'hypothèse au rang inférieur grâce à un calcul par blocs. Si l'on suppose $AB = \mathrm{O}_n$, on peut établir l'existence d'un vecteur propre commun (voir sujet 24 p. 157) et justifier à nouveau que les matrices sont simultanément trigonalisables.

Solution

(a) Par l'absurde, si $\mathrm{Ker}(v) = \{0_E\}$ on peut affirmer que l'endomorphisme v est inversible et écrire

$$v^{-1} \circ \underbrace{u \circ v}_{=v+v\circ u} = \mathrm{Id}_E + u.$$

En considérant la trace des deux membres, on obtient l'absurdité

$$\mathrm{tr}(u) = \mathrm{tr}(u) + \dim E \quad \text{car} \quad \mathrm{tr}\big(v^{-1} \circ (u \circ v)\big) = \mathrm{tr}\big((u \circ v) \circ v^{-1}\big) = \mathrm{tr}(u).$$

Ainsi, le noyau de v n'est pas réduit au vecteur nul.

(b) **méthode**

‖ Le noyau de v est stable par u.

En effet, si x appartient à $\mathrm{Ker}(v)$, $u(x)$ en est aussi élément car

$$v\big(u(x)\big) = u\big(v(x)\big) - v(x) = u(0_E) - 0_E = 0_E.$$

On peut alors introduire l'endomorphisme induit par u sur l'espace complexe $\mathrm{Ker}(v)$ qui est de dimension finie non nulle. Cet endomorphisme induit admet un vecteur propre qui est vecteur propre de u.

(c) On raisonne par récurrence sur la dimension $n \geqslant 1$ de l'espace E.

Pour $n = 1$, il n'y a rien à démontrer.

Supposons la propriété établie au rang $n - 1 \geqslant 1$.

Soit u et v deux endomorphismes d'un espace complexe E de dimension n vérifiant $u \circ v - v \circ u = v$. L'étude qui précède permet d'introduire un vecteur propre e_1 de u appartenant au noyau de v. On complète ce vecteur non nul en une base $e = (e_1, \ldots, e_n)$ de l'espace E. Les matrices de u et v dans la base e sont de la forme

$$A = \begin{pmatrix} \lambda_1 & * \\ 0 & A' \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & * \\ 0 & B' \end{pmatrix} \quad \text{avec} \quad \lambda_1 \in \mathbb{C}, A', B' \in \mathcal{M}_{n-1}(\mathbb{C}).$$

méthode

‖ On applique l'hypothèse de récurrence aux endomorphismes[1] figurés par les blocs A' et B'.

Notons E' l'espace engendré par les vecteurs de la famille $e' = (e_2, \ldots, e_n)$. Les matrices A' et B' figurent dans la base e' des endomorphismes u' et v'. Ceux-ci ne sont pas des endomorphismes induits par u et v sur l'espace E' car on ignore si ce dernier est

1. On peut aussi raisonner comme dans le sujet précédent et former une matrice de trigonalisation convenable de A et B à partir de l'introduction d'une matrice trigonalisant A' et B'. La démarche exposée ici correspond au point de vue vectoriel de ce raisonnement.

stable. Cependant, si l'on introduit p la projection sur E' parallèlement à $\mathrm{Vect}(e_1)$, on peut écrire

$$u'(x) = p\big(u(x)\big) \quad \text{et} \quad v'(x) = p\big(v(x)\big) \quad \text{pour tout } x \text{ de } E'. \qquad (*)$$

L'égalité $u \circ v - v \circ u = v$ donne $AB - BA = B$. Après calculs par blocs, on en déduit $A'B' - B'A' = B'$ et donc $u' \circ v' - v' \circ u' = u'$. Par hypothèse de récurrence, il existe une base $(e'_2, \dots, e'_n)$ de l'espace E' dans laquelle les matrices des endomorphismes u' et v' sont respectivement triangulaire supérieure et triangulaire supérieure stricte :

$$\mathrm{Mat}_{(e'_2,\dots,e'_n)}(u') = \begin{pmatrix} \lambda_2 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{(e'_2,\dots,e'_n)}(v') = \begin{pmatrix} 0 & & (*) \\ & \ddots & \\ (0) & & 0 \end{pmatrix}.$$

Considérons alors la famille $(e_1, e'_2, \dots, e'_n)$ qui est une base de E et formons la matrice des endomorphismes u et v dans celle-ci. Pour tout $j \in [\![2\,;n]\!]$, les relations $(*)$ donnent

$$u(e'_j) = u'(e'_j) + \alpha_j.e_1 \quad \text{et} \quad v(e'_j) = v'(e'_j) + \beta_j.e_1 \quad \text{avec} \quad \alpha_j, \beta_j \in \mathbb{C}$$

et donc

$$\mathrm{Mat}_{(e_1,e'_2,\dots,e'_n)}(u) = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{(e_1,e'_2,\dots,e'_n)}(v) = \begin{pmatrix} 0 & & (*) \\ & \ddots & \\ (0) & & 0 \end{pmatrix}.$$

La récurrence est établie [1].

4.7.8 Applications de la réduction

Exercice 44 *

On étudie l'équation (E) : $M^2 - M = A$ d'inconnue $M \in \mathcal{M}_2(\mathbb{R})$ avec

$$A = \begin{pmatrix} 4 & 1 \\ 4 & 4 \end{pmatrix}.$$

(a) Diagonaliser la matrice A en précisant une matrice de passage P.

(b) Soit $M \in \mathcal{M}_2(\mathbb{R})$ solution de (E). Justifier que la matrice $P^{-1}MP$ est diagonale.

(c) Déterminer toutes les matrices solutions de l'équation (E).

1. On observe que l'endomorphisme v est nilpotent (voir sujet 20 p. 153).

Solution

(a) Le polynôme caractéristique de A est $\chi_A = X^2 - 8X + 12$ de racines 2 et 6. La matrice A est de taille 2 et possède deux valeurs propres distinctes, elle est diagonalisable. Après résolution, ses sous-espaces propres sont

$$E_2(A) = \text{Vect}\begin{pmatrix}1\\-2\end{pmatrix} \quad \text{et} \quad E_6(A) = \text{Vect}\begin{pmatrix}1\\2\end{pmatrix}.$$

On peut donc diagonaliser la matrice A en écrivant

$$A = PDP^{-1} \quad \text{avec} \quad P = \begin{pmatrix}1 & 1\\-2 & 2\end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix}2 & 0\\0 & 6\end{pmatrix}.$$

(b) **méthode**

‖ On observe que $P^{-1}MP$ et D commutent.

En multipliant l'égalité $M^2 - M = A$ par P à droite et P^{-1} à gauche, on obtient

$$P^{-1}M^2P - P^{-1}MP = D \quad \text{avec} \quad P^{-1}M^2P = \left(P^{-1}MP\right)^2.$$

La matrice $X = P^{-1}MP$ vérifie alors l'équation $X^2 - X = D$ et par conséquent commute avec D car

$$XD = X(X^2 - X) = X^3 - X^2 = (X^2 - X)X = DX.$$

En introduisant les coefficients de X, on a

$$DX = \begin{pmatrix}2a & 2b\\6c & 6d\end{pmatrix}, \quad XD = \begin{pmatrix}2a & 6b\\2c & 6d\end{pmatrix} \quad \text{avec} \quad X = \begin{pmatrix}a & b\\c & d\end{pmatrix}.$$

L'égalité $DX = XD$ entraîne alors $b = c = 0$ et la matrice X est diagonale [1].

(c) L'étude qui précède assure que les solutions de l'équation $M^2 - M = A$ sont à rechercher uniquement parmi les matrices de la forme $M = PXP^{-1}$ avec X matrice diagonale. On a alors

$$\begin{aligned} M^2 - M = A &\iff X^2 - X = D\\ &\iff \begin{cases}x^2 - x = 2\\ y^2 - y = 6\end{cases} \quad \text{avec } x \text{ et } y \text{ les coefficients diagonaux de } X\\ &\iff \begin{cases}x = 2 \text{ ou } x = -1\\ y = 3 \text{ ou } y = -2.\end{cases}\end{aligned}$$

Les matrices M solutions sont les quatre [2] matrices $M_i = PX_iP^{-1}$ pour $i \in [\![1\,;4]\!]$ avec

$$X_1 = \begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}, \; X_3 = \begin{pmatrix}2 & 0\\0 & -2\end{pmatrix}, \; X_3 = \begin{pmatrix}-1 & 0\\0 & 3\end{pmatrix} \text{ et } X_4 = \begin{pmatrix}-1 & 0\\0 & -2\end{pmatrix}.$$

1. Plus généralement, seule les matrices diagonales commutent avec une matrice diagonale à coefficients diagonaux deux à deux distincts.

2. Cette équation de degré 2 en l'inconnue M présente plus de deux solutions car le produit matriciel n'est pas intègre. L'équation $M^2 = I_2$ possède quant à elle une infinité de solutions !

On obtient

$$M_1 = \begin{pmatrix} 5/2 & 1/4 \\ 1 & 5/2 \end{pmatrix},\ M_2 = \begin{pmatrix} 0 & -1 \\ -4 & 0 \end{pmatrix},\ M_3 = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \text{ et } M_4 = \begin{pmatrix} -3/2 & -1/4 \\ -1 & -3/2 \end{pmatrix}.$$

Exercice 45 **

On considère trois suites réelles (x_n), (y_n) et (z_n) vérifiant, pour tout $n \in \mathbb{N}$,

$$(\Sigma): \begin{cases} x_{n+1} = 2x_n + y_n + z_n \\ y_{n+1} = x_n - y_n + z_n \\ z_{n+1} = x_n + y_n - z_n. \end{cases}$$

À quelle condition sur (x_0, y_0, z_0), ces trois suites convergent-elles ?

Solution

Pour $n \in \mathbb{N}$, introduisons la colonne X_n de coefficients x_n, y_n et z_n. Le système (Σ) se traduit par l'équation matricielle $X_{n+1} = AX_n$ avec

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Une récurrence immédiate donne alors $X_n = A^n X_0$ ce qui soulève le problème du calcul de A^n.

méthode

On réduit la matrice A afin de déterminer un changement d'inconnues permettant de découpler les relations de récurrence.

Après calculs, le polynôme caractéristique de A est $\chi_A = (X+2)(X^2 - 2X - 2)$ de racines -2 et $1 \pm \sqrt{3}$. La matrice A possède 3 valeurs propres et est de taille 3, elle est donc diagonalisable et peut s'écrire $A = PDP^{-1}$ avec

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1+\sqrt{3} & 0 \\ 0 & 0 & 1-\sqrt{3} \end{pmatrix}$$

et une matrice P inversible qu'il n'est pas nécessaire de déterminer pour le moment.

La relation $X_{n+1} = AX_n$ peut alors se relire $P^{-1}X_{n+1} = DP^{-1}X_n$ ce qui se relit encore $Y_{n+1} = DY_n$ avec $Y_n = P^{-1}X_n$. En notant u_n, v_n, w_n les coefficients de Y_n, on obtient le système

$$\begin{cases} u_{n+1} = -2u_n \\ v_{n+1} = (1+\sqrt{3})v_n \\ w_{n+1} = (1-\sqrt{3})w_n \end{cases} \text{ et donc } \begin{cases} u_n = (-2)^n u_0 \\ v_n = (1+\sqrt{3})^n v_0 \\ w_n = (1-\sqrt{3})^n w_0. \end{cases}$$

Si les suites (x_n), (y_n) et (z_n) convergent, les suites (u_n), (v_n) et (w_n) convergent aussi car $Y_n = P^{-1}X_n$ et la réciproque est encore vraie puisque $X_n = PY_n$. Or les deux

suites $\left((-2)^n\right)$ et $\left((1+\sqrt{3})^n\right)$ divergent tandis que $\left((1-\sqrt{3})^n\right)$ converge vers 0. Les trois suites (x_n), (y_n) et (z_n) sont donc convergentes si, et seulement si, $u_0 = v_0 = 0$. Cela signifie encore que $X_0 = PY_0$ appartient au sous-espace propre associé à la valeur propre $1-\sqrt{3}$. Après détermination de celui-ci, on aboutit à la condition

$$x_0 = (1-\sqrt{3})y_0 = (1-\sqrt{3})z_0.$$

Exercice 46 **

Soit $n \geqslant 2$. Montrer que les matrices

$$M(a_0, a_1, \ldots, a_{n-1}) = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_1 & \cdots & a_{n-1} & a_0 \end{pmatrix} \quad \text{avec} \quad (a_0, a_1, \ldots, a_{n-1}) \in \mathbb{C}^n$$

sont simultanément diagonalisables[1].

Solution

méthode

On exprime $M(a_0, a_1, \ldots, a_{n-1})$ à l'aide de la matrice $J = M(0, 1, 0, \ldots, 0)$ et de ses puissances.

Considérons la matrice

$$J = M(0,1,0,\ldots,0) = \begin{pmatrix} 0 & 1 & & (0) \\ \vdots & \ddots & \ddots & \\ & & & 1 \\ 0 & (0) & \ddots & \\ 1 & 0 & \cdots & 0 \end{pmatrix}.$$

En posant le produit matriciel, on obtient

$$J^2 = \begin{pmatrix} 0 & 0 & 1 & & (0) \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & & \ddots & \ddots & 1 \\ 1 & (0) & & \ddots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix} = M(0,0,1,0,\ldots,0)$$

et l'on peut aisément anticiper le résultat obtenu pour J^3, J^4, etc. Cependant, pour calculer avec « élégance » les puissances de J, on introduit l'endomorphisme j de $\mathbb{C}^n$

1. On dit que les matrices d'une famille sont *simultanément diagonalisables* lorsqu'il existe une même matrice inversible P permettant de les diagonaliser.

qui lui est canoniquement associé. En notant $e_1, \dots, e_n$ les vecteurs de la base canonique de $\mathbb{C}^n$, la lecture des colonnes de J donne

$$j(e_1) = e_n,\ j(e_2) = e_1,\ \dots,\ j(e_n) = e_{n-1}.$$

En adoptant une notation circulaire consistant à poser $e_k = e_\ell$ lorsque $k \equiv \ell\ [n]$, on peut écrire $j(e_i) = e_{i-1}$ pour tout indice i. On a alors $j^2(e_i) = e_{i-2}$, $j^3(e_i) = e_{i-3}$, etc. Par récurrence sur $k \in [\![0\,;n-1]\!]$, on montre $j^k(e_i) = e_{i-k}$ et l'on obtient

$$J^k = \begin{pmatrix} 0 & \mathrm{I}_{n-k} \\ \mathrm{I}_k & 0 \end{pmatrix} = M(\underbrace{0,\dots,0}_{k \text{ valeurs}},1,0,\dots,0).$$

Ainsi, on peut écrire

$$M(a_0,a_1,\dots,a_{n-1}) = \sum_{k=0}^{n-1} a_k J^k.$$

Il suffit alors de savoir diagonaliser J pour en déduire une diagonalisation de ses puissances puis de $M(a_0,a_1,\dots,a_{n-1})$.

Calculons le polynôme caractéristique de J en $\lambda \in \mathbb{C}$. En développant le déterminant selon la première colonne[1]

$$\chi_J(\lambda) = \begin{vmatrix} \lambda & -1 & & (0) \\ \vdots & \ddots & \ddots & \\ & & & \\ 0 & (0) & \ddots & -1 \\ -1 & 0 & \cdots & \lambda \end{vmatrix}_{[n]} = \lambda \begin{vmatrix} \lambda & & (*) \\ & \ddots & \\ (0) & & \lambda \end{vmatrix}_{[n-1]} + (-1)^{n+2} \times \begin{vmatrix} -1 & & (0) \\ & \ddots & \\ (*) & & -1 \end{vmatrix}_{[n-1]} = \lambda^n - 1.$$

Le polynôme caractéristique de J est donc $\chi_J = X^n - 1$. Ses racines sont les racines[2] n-ièmes de l'unité $\omega_0, \dots, \omega_{n-1}$: il y en a exactement n ce qui assure que la matrice J de $\mathcal{M}_n(\mathbb{C})$ est diagonalisable. On peut donc écrire $J = PDP^{-1}$ avec P inversible et D matrice diagonale de coefficients diagonaux les ω_i pour i allant de 0 à $n-1$. On a alors $J^k = PD^kP^{-1}$ puis

$$PM(a_0,a_1,\dots,a_{n-1})P^{-1} = P\Delta(a_0,a_1,\dots,a_{n-1})P^{-1}$$

avec $\Delta(a_0,a_1,\dots,a_{n-1})$ la matrice diagonale de coefficients diagonaux les $Q(\omega_i)$ pour i allant de 0 à $n-1$ où Q désigne le polynôme[3] donné par

$$Q = a_0 + a_1X + \cdots + a_{n-1}X^{n-1}.$$

Ainsi, la matrice de passage P diagonalise toutes les matrices $M(a_0,a_1,\dots,a_{n-1})$.

1. La transposée de J est une matrice compagnon (sujet 32 p. 166) dont le polynôme caractéristique est connu.

2. Rappelons que ω_k désigne $\mathrm{e}^{\frac{2ik\pi}{n}}$ pour tout $k \in \mathbb{Z}$.

3. Au chapitre suivant, on parle de polynôme en une matrice. En anticipant sur cette notion, on peut écrire $M(a_0,a_1,\dots,a_{n-1}) = Q(J)$ et $\Delta(a_0,a_1,\dots,a_{n-1}) = Q(D)$.

Exercice 47 **

Soit P un polynôme unitaire de degré $n \geqslant 1$ à coefficients entiers :

$$P = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \quad \text{avec} \quad a_0, a_1, \ldots, a_{n-1} \in \mathbb{Z}.$$

On note $\lambda_1, \ldots, \lambda_n$ les racines complexes de P comptées avec multiplicité.

(a) Déterminer une matrice à coefficients entiers dont le polynôme caractéristique est P.

(b) En déduire que, pour $q \in \mathbb{N}^*$, le polynôme unitaire P_q dont les racines sont exactement les $\lambda_1^q, \ldots, \lambda_n^q$ est à coefficients entiers.

(c) Déterminer P_2 lorsque $P = X^3 - 3X + 1$.

Solution

(a) **méthode**

‖ On introduit une matrice compagnon [1].

Considérons la matrice

$$A = \begin{pmatrix} 0 & & (0) & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & 0 & \vdots \\ (0) & & 1 & -a_{n-1} \end{pmatrix}.$$

Celle-ci est à coefficients entiers et son polynôme caractéristique est exactement P.

(b) **méthode**

‖ Les valeurs propres complexes de A^q sont les λ^q avec λ valeur propre de A.

Les valeurs propres de A comptées avec multiplicité sont exactement les racines du polynôme $P = \chi_A$. Dans le cadre complexe, la matrice A est trigonalisable semblable à une matrice triangulaire supérieure T où figurent sur la diagonale ses valeurs propres

$$T = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}.$$

La matrice A^q est alors semblable à

$$T^q = \begin{pmatrix} \lambda_1^q & & (*) \\ & \ddots & \\ (0) & & \lambda_n^q \end{pmatrix}.$$

1. Voir sujet 32 p. 166.

Les valeurs propres de A^q sont donc les $\lambda_1^q, \ldots, \lambda_n^q$. Celles-ci sont les racines comptées avec multiplicité de son polynôme caractéristique et, ce dernier étant unitaire, il s'agit du polynôme P_q. Cependant, la matrice A est à coefficients entiers et, par produit, la matrice A^q l'est aussi. Son polynôme caractéristique P_q est alors lui aussi à coefficients entiers [1].

(c) On introduit la matrice compagnon A associée au polynôme $X^3 - 3X + 1$ puis on calcule le polynôme caractéristique de son carré en développant directement selon une rangée

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 3 & -1 \\ 1 & 0 & 3 \end{pmatrix}$$

puis

$$P_2 = \chi_{A^2} = X^3 - 6X^2 + 9X - 1.$$

4.7.9 Nilpotence

Exercice 48 *

Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = BA$ avec A nilpotente. Calculer $\mathrm{tr}(AB)$.

Solution

méthode

Une matrice nilpotente est semblable à une matrice triangulaire stricte (Th. 26 p. 129) et est donc de trace nulle.

Puisque la matrice A est nilpotente de taille n, on sait $A^n = \mathrm{O}_n$. Or les matrices A et B commutent et l'on a donc aussi

$$(AB)^n = A^n B^n = \mathrm{O}_n.$$

On en déduit que la matrice AB est nilpotente. Elle est donc semblable à une matrice triangulaire supérieure où figurent uniquement des zéros sur la diagonale : on peut directement calculer sa trace

$$\mathrm{tr}(AB) = 0.$$

Exercice 49 *

Soit $n \geqslant 2$ et $A = (a_{i,j})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$ avec [2] $a_{i,j} = \delta_{i,j+1}$ pour tous $i, j \in [\![1\,;n]\!]$. Existe-t-il une matrice $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $B^2 = A$?

1. De manière générale, le déterminant d'une matrice se calcule dans l'anneau de ses coefficients.
2. $\delta_{x,y}$ désigne le symbole de Kronecker, égal à 1 si $x = y$ et 0 sinon.

Solution

méthode

‖ L'indice de nilpotence d'une matrice est inférieur à sa taille.

Visualisons les coefficients de la matrice A

$$A = \begin{pmatrix} 0 & & (0) \\ 1 & \ddots & \\ & \ddots & \ddots \\ (0) & & 1 \quad 0 \end{pmatrix}.$$

La matrice A est triangulaire inférieure stricte donc nilpotente. Pour déterminer son indice de nilpotence, on introduit a l'endomorphisme de $\mathbb{K}^n$ qui lui est canoniquement associé. En notant $e_1, \dots, e_n$ les vecteurs de la base canonique de $\mathbb{K}^n$, on a

$$a(e_1) = e_2,\ a(e_2) = e_3,\ \dots,\ a(e_{n-1}) = e_n \text{ et } a(e_n) = 0_{\mathbb{K}^n}.$$

Par récurrence sur $k \geqslant 1$, on obtient

$$a^k(e_j) = \begin{cases} e_{j+k} & \text{si } j+k \leqslant n \\ 0_{\mathbb{K}^n} & \text{sinon.} \end{cases}$$

On en déduit $a^n = 0$ et $a^{n-1} \neq 0$. Ainsi, l'endomorphisme a est nilpotent d'indice n exactement. Il en est de même pour la matrice A.

Par l'absurde, s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ tel que $B^2 = A$, celle-ci est nilpotente car $B^{2n} = A^n = \mathrm{O}_n$. Son indice de nilpotence étant nécessairement inférieure à sa taille, on a aussi $B^n = \mathrm{O}_n$ et donc $A^{n-1} = B^{2n-2} = \mathrm{O}_n$ car $2n-2 \geqslant n$. C'est absurde.

Il n'existe donc pas [1] de matrice B de carré égal à A.

Exercice 50 **

Soit u et v deux endomorphismes d'un espace réel E de dimension $n \geqslant 1$. On suppose que u et v commutent et que v est nilpotent. Montrer $\det(u+v) = \det(u)$.

Solution

Discutons selon que u est inversible ou non.

Cas : u est inversible. On peut factoriser $\det(u)$ dans le calcul de $\det(u+v)$ et écrire

$$\det(u+v) = \det\Big(u \circ \big(\mathrm{Id}_E + u^{-1} \circ v\big)\Big) = \det(u)\det(\mathrm{Id}_E + w) \quad \text{avec} \quad w = u^{-1} \circ v.$$

Puisque les endomorphismes u et v commutent, il en est de même pour les endomorphismes u^{-1} et v. On en déduit la nilpotence [2] de w car

$$w^n = \big(u^{-1} \circ v\big)^n = \big(u^{-1}\big)^n \circ v^n = 0 \quad \text{car} \quad v^n = 0.$$

1. Une démonstration alternative consiste à étudier les noyaux itérés (sujet 12 p. 82). Si $B^2 = A$ alors $\dim \mathrm{Ker}\big(B^2\big) = 1$ donc $\dim \mathrm{Ker}(B) = 1$ puis $\dim \mathrm{Ker}\big(B^k\big) = 1$ pour tout $k \geqslant 1$ ce qui contredit $A^n = \mathrm{O}_n$.

2. L'espace E étant de dimension n, la nilpotence de v assure $v^n = 0$.

méthode

Tout endomorphisme nilpotent d'un espace de dimension finie peut être figuré par une matrice triangulaire supérieure stricte (Th. 25 p. 129).

Introduisons une base de E dans laquelle la matrice de w est égale à une matrice triangulaire supérieure T. Par représentation matricielle dans cette base, on a

$$\det(\mathrm{Id}_E + w) = \det(\mathrm{I}_n + T) = \begin{vmatrix} 1 & & (*) \\ & \ddots & \\ (0) & & 1 \end{vmatrix} = 1.$$

On en déduit $\det(u+v) = \det(u)$.

Cas : u non inversible.

méthode

Puisque u et v commutent, le noyau de u est stable[1] par v.

L'endomorphisme u n'étant pas inversible, l'espace $\mathrm{Ker}(u)$ n'est pas réduit au vecteur nul. De plus cet espace est stable par v ce qui permet d'introduire l'endomorphisme v' induit par v sur $\mathrm{Ker}(u)$. Tout comme v, l'endomorphisme v' est nilpotent et 0 est son unique valeur propre. Ainsi, il existe dans $\mathrm{Ker}(u)$ un vecteur non nul annulant v. Ce vecteur annule aussi $u+v$ qui n'est donc pas inversible. On conclut[2] à nouveau $\det(u+v) = 0 = \det(u)$.

Exercice 51 ***

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

Montrer que A est nilpotente si, et seulement si, $\mathrm{tr}(A^p) = 0$ pour tout $p \in [\![1\,;n]\!]$.

Solution

($\Longrightarrow$) Supposons la matrice A nilpotente. Elle est semblable à une matrice triangulaire supérieure stricte T et donc $\mathrm{tr}(A) = \mathrm{tr}(T) = 0$. Aussi, pour tout $p \in [\![1\,;n]\!]$, A^p est semblable à T^p et donc $\mathrm{tr}(A^p) = \mathrm{tr}(T^p) = 0$.

($\Longleftarrow$) Supposons $\mathrm{tr}(A^p) = 0$ pour tout $p \in [\![1\,;n]\!]$.

méthode

On montre que 0 est la seule valeur propre complexe de A.

Notons $\lambda_1, ..., \lambda_n$ les valeurs propres complexes comptées avec multiplicité de la matrice A. La matrice A est semblable dans $\mathcal{M}_n(\mathbb{C})$ à une matrice triangulaire T dont les coefficients diagonaux sont les valeurs précédentes. Pour tout $p \in [\![1\,;n]\!]$, la matrice A^p est semblable à T^p et donc

$$\mathrm{tr}(A^p) = \mathrm{tr}(T^p) = \sum_{i=1}^{n} \lambda_i^p = 0.$$

1. Voir sujet 13 p. 146.

2. Dans le cadre complexe, le sujet peut se résoudre en exploitant la commutation de u et v pour réaliser une trigonalisation simultanée : voir sujet 42 p. 176.

On peut alors dire que $(\lambda_1, \dots, \lambda_n) \in \mathbb{C}^n$ est solution du système

$$(\Sigma_n): \begin{cases} \lambda_1 + \lambda_2 + \cdots + \lambda_n = 0 \\ \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 = 0 \\ \qquad\qquad\qquad\qquad \vdots \\ \lambda_1^n + \lambda_2^n + \cdots + \lambda_n^n = 0. \end{cases}$$

L'élément $(\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$ est évidemment solution de (Σ_n). Montrons par récurrence [1] sur $n \in \mathbb{N}^*$ qu'il n'y en a pas d'autres.

Pour $n = 1$, l'affirmation est immédiate.

Supposons le résultat vrai au rang $n - 1 \geqslant 1$. Soit $(\lambda_1, \dots, \lambda_n)$ une solution de Σ_n. On introduit le polynôme $P = (X - \lambda_1) \dots (X - \lambda_n)$. Les racines de P étant les λ_i, la somme $P(\lambda_1) + \cdots + P(\lambda_n)$ est nulle. Aussi, en introduisant les coefficients a_k de P, on peut écrire $P = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n$ et alors

$$0 = \sum_{i=1}^{n} P(\lambda_i) = n a_0 + a_1 \underbrace{\sum_{i=1}^{n} \lambda_i}_{=0} + a_2 \underbrace{\sum_{i=1}^{n} \lambda_i^2}_{=0} + \cdots + a_n \underbrace{\sum_{i=1}^{n} \lambda_i^n}_{=0}.$$

On en déduit que le coefficient constant a_0 est nul et donc P possède une racine nulle. Quitte à redéfinir l'indexation des λ_i, on peut supposer $\lambda_n = 0$ et simplifier les équations du système (Σ_n) pour affirmer que $(\lambda_1, \dots, \lambda_{n-1})$ est solution du système (Σ_{n-1}). Par l'hypothèse de récurrence, on obtient alors $\lambda_1 = \cdots = \lambda_{n-1} = 0$ et l'on peut conclure.

La récurrence est établie.

Finalement, 0 est la seule valeur propre complexe de la matrice A. La matrice A est donc semblable dans $\mathcal{M}_n(\mathbb{C})$ à une matrice triangulaire supérieure stricte, elle est nilpotente.

4.8 Exercices d'approfondissement

Exercice 52 *

L'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ est-il convexe ?

Solution

Lorsque $n = 1$ toutes les matrices de $\mathcal{M}_1(\mathbb{R})$ sont diagonalisables et l'ensemble $\mathcal{M}_1(\mathbb{R})$, qui s'identifie à la droite réelle, est convexe. Montrons que pour $n \geqslant 2$ cette propriété n'est plus vraie en commençant par étudier le cas $n = 2$.

1. Présentons une démarche alternative. On suppose par l'absurde l'existence d'une solution non nulle. On simplifie tous les λ_i nuls du système (Σ_n) et l'on regroupe ensemble ceux qui sont égaux afin de former des équations du type (E_k): $\lambda_1^k \alpha_1 + \cdots + \lambda_m^k \alpha_m = 0$ avec des λ_i non nuls deux à deux distincts et $\alpha_i \in \mathbb{N}^*$. Les équations $(E_1), \dots, (E_m)$ forment un système d'inconnues les $\alpha_1, \dots, \alpha_m$. Or ce système est de Cramer car sa matrice est liée à une matrice de Vandermonde. Sa seule solution est alors la solution évidente $(\alpha_1, \dots, \alpha_m) = (0, \dots, 0)$: c'est absurde !

méthode

On choisit une matrice non diagonalisable M et l'on détermine deux matrices diagonalisables A et B dont M est le milieu.

La matrice

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

n'est pas diagonalisable car 0 est sa seule valeur propre alors que ce n'est pas la matrice nulle. Les matrices

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

sont diagonalisables car possèdent chacune deux valeurs propres distinctes. Enfin,

$$M = \frac{1}{2}(A + B) \in [A\,;B]$$

et donc $[A\,;B]$ n'est pas entièrement inclus dans l'ensemble des matrices diagonalisables.

On peut généraliser cet exemple aux matrices de taille $n \geqslant 3$ en complétant les matrices précédentes de blocs nuls :

$$M' = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(A' + B') \quad \text{avec} \quad A' = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \text{ et } B' = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}.$$

Ainsi, l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ n'est pas une partie convexe lorsque $n \geqslant 2$. En revanche, il s'agit d'une partie connexe par arcs [1].

Exercice 53 *

Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. On suppose qu'il existe M dans $\mathcal{M}_n(\mathbb{K})$ de rang r tel que

$$AM = MB.$$

Montrer que le PGCD de χ_A et χ_B est de degré au moins r.

Solution

méthode

Une matrice de rang r est équivalente à la matrice $J_r = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix}$ de même type.

On peut introduire des matrices inversibles P et Q de taille n telles que $M = QJ_rP^{-1}$. La relation $AM = MB$ donne alors

$$CJ_r = J_rD \quad \text{avec} \quad C = Q^{-1}AQ \text{ et } D = P^{-1}BP.$$

On écrit les matrices C et D par blocs

$$C = \begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} D_{1,1} & D_{1,2} \\ D_{2,1} & D_{2,2} \end{pmatrix} \quad \text{avec} \quad C_{1,1}, D_{1,1} \in \mathcal{M}_r(\mathbb{K}).$$

1. Voir le sujet 19 du chapitre 5 de l'ouvrage *Exercices d'analyse MP*.

En opérant des produits par blocs, l'égalité $CJ_r = J_rD$ donne

$$C_{1,1} = D_{1,1}, \quad C_{2,1} = O_{n-r,r} \quad \text{et} \quad D_{1,2} = O_{r,n-r}.$$

Les matrices C et D sont donc triangulaires avec un bloc d'indice $(1,1)$ identique. Les matrices A et B étant respectivement semblables à C et D, on obtient

$$\chi_A = \chi_C = \chi_{C_{1,1}}\chi_{C_{2,2}} \quad \text{et} \quad \chi_B = \chi_D = \chi_{D_{1,1}}\chi_{D_{2,2}}$$

avec $\chi_{C_{1,1}} = \chi_{D_{1,1}}$ polynôme de degré r.

On peut alors conclure que les polynômes χ_A et χ_B possèdent un facteur commun de degré au moins r et donc que leur PGCD est de degré au moins égal à r.

Exercice 54 **

Une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ est dite *magique* si les sommes de ses coefficients par ligne et par colonne sont toutes égales. On note J la colonne de $\mathcal{M}_{n,1}(\mathbb{R})$ dont tous les coefficients sont égaux à 1.

(a) Montrer que la matrice A est magique si, et seulement si, J est vecteur propre de A et de ${}^tA$.

(b) Vérifier que l'ensemble $\mathcal{MG}_n$ des matrices magiques de taille n est une sous-algèbre de $\mathcal{M}_n(\mathbb{R})$. Que dire de l'inverse d'une matrice magique inversible ?

On introduit le produit scalaire canonique[1] sur l'espace $\mathcal{M}_{n,1}(\mathbb{R})$ et les espaces $D = \text{Vect}(J)$ et $H = D^\perp$.

(c) Montrer qu'une matrice A de $\mathcal{M}_n(\mathbb{R})$ est magique si, et seulement si, elle laisse stable les espaces D et H.

(d) En déduire la dimension de l'algèbre des matrices magiques de $\mathcal{M}_n(\mathbb{R})$.

Solution

(a) Les coefficients de la colonne AJ sont les sommes des coefficients de A sur chaque ligne. Les coefficients de la colonne ${}^tAJ$ sont les sommes des coefficients de A sur chaque colonne. Si la matrice A est magique, ces différentes sommes sont toutes égales à une même constante s et donc

$$AJ = sJ \quad \text{et} \quad {}^tAJ = sJ \quad \text{avec} \quad J \neq 0.$$

On en déduit que J est vecteur propre de A et de ${}^tA$.

Inversement, si J est vecteur propre de A et de ${}^tA$ respectivement associé aux valeurs propres λ et μ, on a $AJ = \lambda J$ et ${}^tAJ = \mu J$. Les valeurs λ et μ sont nécessairement égales car[2]

$${}^tJAJ = {}^tJ(AJ) = \lambda\,{}^tJJ = n\lambda \quad \text{et} \quad {}^tJAJ = {}^t({}^tAJ)J = \mu\,{}^tJJ = n\mu.$$

On en déduit que les sommes des coefficients de A sur chaque ligne et sur chaque colonne sont toutes égales : la matrice A est magique.

1. Le produit scalaire de deux colonnes X, Y de $\mathcal{M}_{n,1}(\mathbb{R})$ est donné par $(X,Y) = {}^tXY$.

2. Le calcul ${}^tJAJ$ détermine la somme de tous les coefficients de A, somme qui peut être organisée par ligne ou par colonne.

(b) L'ensemble $\mathcal{MG}_n$ des matrices magiques de taille n est une partie de l'algèbre des matrices carrées $\mathcal{M}_n(\mathbb{K})$. La matrice unité I_n est magique et, pour tous $\lambda, \mu \in \mathbb{K}$ et tous $A, B \in \mathcal{MG}_n$, on vérifie par un simple produit que la colonne J est vecteur propre des matrices $\lambda A + \mu B$ et AB ainsi que de leurs transposées : $\mathcal{MG}_n$ est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.

Soit A une matrice magique inversible. Montrons que son inverse est aussi une matrice magique.

méthode

On introduit l'application φ qui a $M \in \mathcal{MG}_n$ associe AM et l'on montre qu'il s'agit d'un isomorphisme d'espaces vectoriels [1].

L'application φ proposée est définie au départ de l'espace $\mathcal{MG}_n$ et à valeurs dans lui-même. Cette application est linéaire et aussi injective car, pour tout $M \in \mathcal{MG}_n$,

$$\begin{aligned} AM = \mathrm{O}_n &\implies A^{-1}AM = \mathrm{O}_n \\ &\implies M = \mathrm{O}_n. \end{aligned}$$

Puisque l'espace $\mathcal{MG}_n$ est de dimension finie, l'application φ est un automorphisme de $\mathcal{MG}_n$. Par surjectivité, il existe une matrice $B \in \mathcal{MG}_n$ telle que $AB = \mathrm{I}_n$. Cette matrice B est nécessairement l'inverse de A et donc $A^{-1} \in \mathcal{MG}_n$.

(c) La droite vectorielle D est stable par A si, et seulement si, J est vecteur propre de A. L'hyperplan H est stable par A si, et seulement si [2] sa droite normale $H^\perp = D$ est stable par ${}^tA$, c'est-à-dire si, et seulement si, J est vecteur propre de ${}^tA$. La résolution de la première question assure alors que la matrice A est magique si, et seulement si, les deux espaces D et H sont stables par A.

(d) **méthode**

La stabilité de deux espaces supplémentaires se lit matriciellement par une représentation diagonale par blocs dans une base adaptée (Th. 2 p. 120).

Les espaces D et H sont supplémentaires dans $\mathcal{M}_{n,1}(\mathbb{R})$. Si l'on introduit une base adaptée à l'écriture $\mathcal{M}_{n,1}(\mathbb{R}) = D \oplus H$, l'étude au-dessus assure qu'une matrice A est magique si, et seulement si, l'endomorphisme canoniquement associé à A est représenté dans cette base par une matrice de la forme

$$\begin{pmatrix} \alpha & 0 \\ 0 & M \end{pmatrix} \quad \text{avec} \quad \alpha \in \mathbb{R} \text{ et } M \in \mathcal{M}_{n-1}(\mathbb{R}).$$

Un tel endomorphisme est alors entièrement déterminé par ses restrictions aux espaces D et H qui sont des endomorphismes. L'espace des matrices magiques est donc isomorphe à $\mathcal{L}(D) \times \mathcal{L}(H)$ ce qui détermine sa dimension :

$$\dim \mathcal{MG}_n = (\dim D)^2 + (\dim H)^2 = 1 + (n-1)^2.$$

1. On peut aussi remarquer que $AJ = sJ$ (avec s nécessairement non nul) entraîne $\frac{1}{s}J = A^{-1}J$ et l'on a une égalité analogue pour la transposée.

2. Voir sujet 16 p. 150.

Exercice 55 ***

Soit $n \geqslant 2$. Déterminer les valeurs propres de la comatrice de $A \in \mathcal{M}_n(\mathbb{C})$ en fonction de celles de A.

Solution

La matrice A est complexe, elle est admet donc n valeurs propres comptées avec multiplicité que nous notons $\lambda_1, \ldots, \lambda_n$.

La comatrice de A est la matrice des cofacteurs de A et celle-ci vérifie :

$$ {}^t\big(\mathrm{Com}(A)\big)A = \det(A)\mathrm{I}_n. \qquad (*)$$

méthode

‖ On commence par étudier le cas où la matrice A est inversible.

Supposons la matrice A inversible. On peut exprimer la comatrice de A en fonction de son inverse

$$\mathrm{Com}(A) = \det(A)\,{}^t\big(A^{-1}\big).$$

Les valeurs propres de A sont nécessairement non nulles et sont les inverses des valeurs propres de A. En effet [1], A est trigonalisable et l'on peut écrire avec P inversible

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{et} \quad P^{-1}A^{-1}P = \big(P^{-1}AP\big)^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & & (*) \\ & \ddots & \\ (0) & & \frac{1}{\lambda_n} \end{pmatrix}.$$

De plus, les valeurs propres d'une matrice sont aussi celles de sa transposée. On en déduit que les valeurs propres de $\mathrm{Com}(A)$ comptées avec multiplicité sont les

$$\frac{\det(A)}{\lambda_1}, \ldots, \frac{\det(A)}{\lambda_n} \quad \text{avec} \quad \det(A) = \lambda_1 \ldots \lambda_n.$$

Après simplification de ces fractions, on obtient que les valeurs propres de $\mathrm{Com}(A)$ sont tous les produits possibles de $n-1$ facteurs choisis parmi les n valeurs propres de A.

Supposons maintenant la matrice A non inversible. La matrice A est donc de rang inférieur à $n-1$. Si $\mathrm{rg}(A) \leqslant n-2$, tous les déterminants des matrices de taille $n-1$ extraites de A sont nuls [2]. Les cofacteurs de A sont alors tous nuls et la comatrice de A est nulle : 0 est son unique valeur propre et celle-ci est de multiplicité n. Il reste à étudier le cas $\mathrm{rg}(A) = n-1$.

Lorsque $\mathrm{rg}(A) = n-1$, la matrice A possède $n-1$ valeurs propres non nulles et une valeur propre nulle. Quitte à redéfinir l'indexation de celles-ci, on peut supposer $\lambda_n = 0$.

Le déterminant de A étant nul, la relation $(*)$ devient

$$ {}^t\big(\mathrm{Com}(A)\big)A = \mathrm{O}_n.$$

1. Une résolution par les éléments propres est aussi possible : $AX = \lambda X \iff A^{-1}Y = \frac{1}{\lambda}Y$ avec $Y = AX$ et $Y \neq 0$ si, et seulement si, $X \neq 0$.

2. Rappelons que si l'on peut extraire d'une matrice A une matrice carrée inversible de taille p alors $\mathrm{rg}(A) \geqslant p$.

L'image de A est donc incluse dans le noyau de ${}^t\big(\mathrm{Com}(A)\big)$. On en déduit

$$\dim \mathrm{Ker}\big(\mathrm{Com}(A)\big) = \dim \mathrm{Ker}\Big({}^t\big(\mathrm{Com}(A)\big)\Big) \geqslant \mathrm{rg}(A) = n-1.$$

Lorsqu'il n'est pas réduit au vecteur nul, le noyau correspond au sous-espace propre associé à la valeur propre 0. On en déduit que 0 est valeur propre de $\mathrm{Com}(A)$ de multiplicité au moins $n-1$. Puisque la trace d'une matrice complexe est la somme de ses valeurs propres comptées avec multiplicité, $\mu = \mathrm{tr}\big(\mathrm{Com}(A)\big)$ détermine la dernière valeur propre [1] restant à calculer.

méthode

‖ On introduit la comatrice de $A_\varepsilon = A + \varepsilon \mathrm{I}_n$ avec ε qui tend vers 0^+.

Soit $\varepsilon > 0$. Les valeurs propres de la matrice A_ε sont les $\lambda_1 + \varepsilon, \ldots, \lambda_{n-1} + \varepsilon$ et ε car, en reprenant les notations qui précèdent,

$$P^{-1}A_\varepsilon P = P^{-1}AP + \varepsilon \mathrm{I}_n = \begin{pmatrix} \lambda_1 + \varepsilon & & (*) \\ & \ddots & \\ & & \lambda_{n-1} + \varepsilon \\ (0) & & & \varepsilon \end{pmatrix}.$$

Pour $\varepsilon > 0$ petit [2], aucune de ces valeurs propres n'est nulle et la matrice A_ε est inversible. On en déduit que les valeurs propres de $\mathrm{Com}(A_\varepsilon)$ sont

$$\frac{\det(A_\varepsilon)}{\lambda_1 + \varepsilon}, \ldots, \frac{\det(A_\varepsilon)}{\lambda_{n-1} + \varepsilon}, \frac{\det(A_\varepsilon)}{\varepsilon}$$

avec

$$\det(A_\varepsilon) = (\lambda_1 + \varepsilon) \ldots (\lambda_{n-1} + \varepsilon)\varepsilon.$$

On peut alors calculer la trace de la comatrice de A_ε

$$\mathrm{tr}\big(\mathrm{Com}(A_\varepsilon)\big) = \frac{\det(A_\varepsilon)}{\lambda_1 + \varepsilon} + \cdots + \frac{\det(A_\varepsilon)}{\lambda_{n-1} + \varepsilon} + \frac{\det(A_\varepsilon)}{\varepsilon}.$$

Enfin, la trace de la comatrice de A_ε est une fonction continue de ε car polynomiale [3]. On a donc par passage à la limite

$$\mu = \mathrm{tr}\big(\mathrm{Com}(A)\big) = \lim_{\varepsilon \to 0^+} \Bigg(\underbrace{\frac{\det(A_\varepsilon)}{\lambda_1 + \varepsilon}}_{\to 0} + \cdots + \underbrace{\frac{\det(A_\varepsilon)}{\lambda_{n-1} + \varepsilon}}_{\to 0} + \underbrace{\frac{\det(A_\varepsilon)}{\varepsilon}}_{\to \lambda_1 \ldots \lambda_{n-1}} \Bigg) = \lambda_1 \ldots \lambda_{n-1}.$$

En résumé, dans tous les cas, les n valeurs propres de $\mathrm{Com}(A)$ sont tous les produits possibles [4] de $n-1$ facteurs choisis parmi les n valeurs propres de A.

1. Voir le sujet 41 p. 175 pour une démonstration détaillée dans un cadre analogue.

2. Il suffit de choisir ε inférieur à la valeur absolue des valeurs propres réelles négatives de A, s'il y en a.

3. La trace est la somme des coefficients diagonaux et chacun est une fonction polynomiale en ε car déterminant d'une matrice extraite de $A + \varepsilon \mathrm{I}_n$.

4. Une démonstration moins directe est aussi possible et plus rapide. Lorsque deux matrices sont

Exercice 56 * (Théorème de Perron)**

Soit $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ deux éléments de $\mathbb{R}^n$. On écrit $x \leqslant y$ pour signifier $x_i \leqslant y_i$ pour tout $i \in [\![1\,;n]\!]$.

On étudie une matrice A carrée de taille n dont tous les coefficients sont strictement positifs et l'on introduit le compact $K = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n \mid x_1 + \cdots + x_n = 1\}$.

(a) Pour $x \in K$, justifier que l'on peut introduire

$$\theta(x) = \max\{\alpha \in \mathbb{R}_+ \mid \alpha x \leqslant Ax\}.$$

(b) Montrer qu'il existe $y \in K$ tel que $\theta(x) \leqslant \theta(y)$ pour tout $x \in K$.

(c) Montrer que y est vecteur propre de A associé à la valeur propre $\mu = \theta(y)$.

(d) Établir que les valeurs propres complexes λ de A vérifient $|\lambda| < \mu$.

(e) Vérifier que les éléments de y sont tous strictement positifs puis que l'espace propre complexe associé à la valeur propre μ est de dimension 1.

Solution

(a) Soit $x = (x_1, \dots, x_n)$ un élément de K. On a $\alpha x \leqslant Ax$ si, et seulement si, pour tout $i \in [\![1\,;n]\!]$,

$$\alpha x_i = [\alpha x]_i \leqslant [Ax]_i = \sum_{j=1}^{n} a_{i,j} x_j.$$

Lorsque x_i est nul, cette inégalité est assurément vérifiée car les $a_{i,j}$ et les x_j sont positifs. Lorsque x_i n'est pas nul, cette inégalité est vérifiée si, et seulement si,

$$\alpha \leqslant \frac{[Ax]_i}{x_i}.$$

Sachant que x n'est pas le vecteur nul, au moins l'un des x_i est non nul et donc

$$\theta(x) = \max\{\alpha \in \mathbb{R}_+ \mid \alpha x \leqslant Ax\} = \min_{\substack{1 \leqslant i \leqslant n \\ x_i \neq 0}} \left(\frac{[Ax]_i}{x_i} \right).$$

(b) **méthode**

> On montre l'existence de la borne supérieure des $\theta(x)$ pour x parcourant K avant d'établir que celle-ci est un max.

L'ensemble $\Theta = \{\theta(x) \mid x \in K\}$ est une partie non vide de $\mathbb{R}$. De plus, pour tout $x = (x_1, \dots, x_n) \in K$, si x_{i_0} désigne le plus grand élément parmi $x_1, \dots, x_n$, on a

$$\theta(x) \leqslant \frac{[Ax]_{i_0}}{x_{i_0}} = \frac{1}{x_{i_0}} \sum_{j=1}^{n} a_{i_0,j} \underbrace{x_j}_{\leqslant x_{i_0}} \leqslant \sum_{j=1}^{n} a_{i_0,j} \leqslant M$$

semblables, on peut montrer que leurs comatrices le sont aussi (voir sujet 40 p. 110). De plus, la comatrice d'une matrice triangulaire supérieure est triangulaire inférieure. En trigonalisant la matrice A, les valeurs propres de sa comatrice sont les cofacteurs diagonaux de la matrice triangulaire formée.

avec

$$M = \max_{1\leqslant i\leqslant n}\left(\sum_{j=1}^{n}|a_{i,j}|\right).$$

L'ensemble Θ est donc une partie non vide et majorée de $\mathbb{R}$ ce qui autorise à introduire sa borne supérieure que nous notons μ. Vérifions que celle-ci est un max. Par la réalisation séquentielle [1] d'une borne supérieure, on peut introduire une suite (x_k) d'éléments de K telle que $\theta(x_k)$ converge vers μ. La partie K étant compacte, on peut extraire de la suite (x_k) une suite convergeant dans K. Quitte à ne conserver de la suite (x_k) que les termes déterminant une suite convergente, on peut supposer que (x_k) converge vers un élément y de K.

Pour tout $k \in \mathbb{N}$, on a $\theta(x_k)x_k \leqslant Ax_k$. Par passage à la limite des inégalités larges coordonnées par coordonnées, on obtient $\mu y \leqslant Ay$. On en déduit $\mu \leqslant \theta(y)$ puis l'égalité $\mu = \theta(y)$ car μ est la borne supérieure des valeurs prises par θ.

(c) Posons $z = Ay - \mu y$. Par construction, les éléments du vecteur z sont tous positifs.

méthode

‖ Par l'absurde, on suppose z non nul et l'on étudie $y' \in K$ colinéaire à Ay.

Si z n'est pas nul, tous les éléments du vecteur Az sont strictement positifs car les coefficients de la matrice A sont strictement positifs et les éléments du vecteur z sont positifs avec au moins l'un d'eux non nul. On peut alors introduire $\varepsilon > 0$ assez petit tel que $\varepsilon Ay \leqslant Az$ ce qui se relit

$$(\mu + \varepsilon)Ay \leqslant A(Ay).$$

Enfin, on choisit $\lambda \in \mathbb{R}^+$ tel que $y' = \lambda Ay$ soit élément de K. Ceci est possible car Ay est un vecteur non nul dont tous les éléments sont positifs ce qui permet de prendre λ égal à l'inverse de la somme de ceux-ci. On a alors $(\mu + \varepsilon).y' \leqslant Ay'$ et donc $\theta(y') \geqslant \mu + \varepsilon > \mu$. C'est absurde car contredit la définition de μ comme borne supérieure des $\theta(x)$ pour x parcourant K.

On en déduit que z est le vecteur nul et donc $Ay = \mu y$ avec $y \neq 0_{\mathbb{R}^n}$.

(d) Soit λ une valeur propre complexe de A et $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ un vecteur propre associé. Sans perte de généralité, on peut supposer $|x_1| + \cdots + |x_n| = 1$ quitte à considérer un vecteur colinéaire à x. Introduisons aussi $x^+ = \big(|x_1|, \ldots, |x_n|\big)$ qui est élément de K. Pour tout $i \in [\![1\,;n]\!]$,

$$|\lambda|\,|x_i| = |\lambda x_i| = \left|\sum_{j=1}^{n} a_{i,j}x_j\right| \leqslant \sum_{j=1}^{n} a_{i,j}\,|x_j|\,. \qquad (*)$$

Ceci donne $|\lambda|\,x^+ \leqslant Ax^+$ et donc $|\lambda| \leqslant \theta\big(x^+\big) \leqslant \mu$.

De plus, si $|\lambda|$ est égal à μ, il y a égalité [2] dans chacune des inégalités précédemment écrites. En particulier, $\theta\big(x^+\big) = \mu$ ce qui entraîne que x^+ est vecteur propre associé à

1. Voir Th 15 du chapitre 6 de l'ouvrage *Exercices d'analyse MPSI*.
2. Voir sujet 32 du chapitre 3 de l'ouvrage *Exercices d'analyse MPSI*.

la valeur propre μ comme on l'a démontré pour le vecteur y précédemment. De plus, il y a aussi égalité dans l'inégalité triangulaire $(*)$ ce qui signifie que les complexes $a_{i,j}x_j$ figurent sur une même demi-droite issue de l'origine. Les $a_{i,j}$ étant positifs, ceci revient à dire qu'il existe $\varphi \in \mathbb{R}$ tel que $x_j = |x_j|\,\mathrm{e}^{\mathrm{i}\varphi}$ pour tout $j \in [\![1\,;n]\!]$. Le vecteur x est alors colinéaire à x^+ et est donc associé à la même valeur propre μ.

En résumé, les valeurs propres complexes de A vérifient $|\lambda| \leqslant \mu$ et seule la valeur $\lambda = \mu$ satisfait l'égalité.

(e) Le vecteur y est non nul et ses éléments $y_1, \ldots, y_n$ sont tous positifs. De plus, la matrice A est à coefficients strictement positifs et donc les éléments de Ay sont strictement positifs. Ceux de y le sont alors aussi.

Soit $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ un vecteur propre associé à la valeur propre μ.

méthode

On montre que x et y sont colinéaires en formant un vecteur combinaison linéaire de x et y dont tous les coefficients sont des réels positifs sauf un qui est nul.

Par l'étude du cas d'égalité dans la question précédente, on peut affirmer que x est colinéaire au vecteur $x^+ = \left(|x_1|, \ldots, |x_n|\right)$. Sans perte de généralité, on peut donc supposer les éléments x_i du vecteur x tous réels positifs. Considérons ensuite le vecteur

$$z = x_i y - y_i x$$

avec $i \in [\![1\,;n]\!]$ déterminé de sorte que

$$\frac{x_i}{y_i} = \max_{1 \leqslant j \leqslant n} \left(\frac{x_j}{y_j}\right).$$

Les éléments du vecteur z sont positifs et celui d'indice i est nul par construction. Or, par combinaison linéaire de vecteurs propres, $Az = \mu z$. Si z n'est pas le vecteur nul, tous les éléments de Az sont strictement positifs, notamment celui d'indice i ce qui est absurde. On en déduit que le vecteur z est nul et donc x et y sont colinéaires.

Finalement, l'espace propre associé à la valeur propre μ est de dimension [1] 1.

1. Cette étude généralise celle déjà menée dans le sujet 28 p. 162.

CHAPITRE 5

Réduction Algébrique

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$, u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension quelconque et n désigne un entier naturel non nul.

5.1 Polynômes d'un endomorphisme

5.1.1 Valeur d'un polynôme en un endomorphisme

Soit P un polynôme de $\mathbb{K}[X]$:

$$P = \sum_{k=0}^{N} a_k X^k = a_0 + a_1 X + \cdots + a_N X^N \quad \text{avec} \quad N \geqslant \deg(P).$$

Définition

On appelle *valeur du polynôme* P en u l'endomorphisme[1] de E donné par

$$P(u) \overset{\text{déf}}{=} \sum_{k=0}^{N} a_k u^k = a_0 \mathrm{Id}_E + a_1 u + \cdots + a_N u^N.$$

La valeur de $P = X^3 - 2X + 1$ en u est $P(u) = u^3 - 2u + \mathrm{Id}_E$.
La valeur de l'endomorphisme $P(u)$ en un vecteur x de E se note[2] $P(u)(x)$.

1. Les puissances de u se comprennent comme des itérés de composition : $u^k = u \circ \cdots \circ u$ (k facteurs).
2. Il ne faut pas écrire $P\big(u(x)\big)$: les puissances de $u(x)$ n'ont a *priori* aucun sens.

Théorème 1
L'application

$$\Phi_u : \begin{cases} \mathbb{K}[X] \to \mathcal{L}(E) \\ P \mapsto P(u) \end{cases}$$

est un morphisme de $\mathbb{K}$-algèbres.

Par le morphisme Φ_u, une factorisation

$$X^3 - 2X + 1 = (X-1)(X^2 + X - 1)$$

donne l'identité

$$u^3 - 2u + \mathrm{Id}_E = (u - \mathrm{Id}_E) \circ (u^2 + u - \mathrm{Id}_E).$$

Plus généralement, toute identité polynomiale se transpose aux endomorphismes.

5.1.2 Polynôme en un endomorphisme

Définition
On dit qu'un endomorphisme v de E est un *polynôme en u* lorsqu'il existe $P \in \mathbb{K}[X]$ tel que $v = P(u)$.

Pour $\lambda \in \mathbb{K}$ et $\alpha \in \mathbb{N}$, $(u - \lambda \mathrm{Id}_E)^\alpha$ est un polynôme en l'endomorphisme u.

Définition
On appelle *algèbre engendrée par l'endomorphisme u* l'ensemble $\mathbb{K}[u]$ des polynômes en u :

$$\mathbb{K}[u] \overset{\text{déf}}{=} \{P(u) \mid P \in \mathbb{K}[X]\} = \mathrm{Vect}\{u^n \mid n \in \mathbb{N}\}.$$

Théorème 2
$\mathbb{K}[u]$ est une sous-algèbre commutative de $\mathcal{L}(E)$. Celle-ci contient l'endomorphisme u et est incluse dans toute sous-algèbre de $\mathcal{L}(E)$ contenant u.

En particulier, les polynômes en u commutent avec u.

5.1.3 Polynômes annulateurs

Définition
On appelle *polynôme annulateur* de u tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(u) = 0$.

Le polynôme nul annule tout endomorphisme.
Le polynôme $X^2 - X$ est annulateur des projections vectorielles.

Théorème 3
L'ensemble des polynômes annulateurs de l'endomorphisme u est un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$.

Si un polynôme P annule u, les polynômes multiples de P annulent aussi u.

Si λ est valeur propre de u alors, pour tout polynôme P de $\mathbb{K}[X]$, le scalaire $P(\lambda)$ est valeur propre [1] de $P(u)$ associé aux mêmes vecteurs propres. On en déduit le résultat qui suit :

Théorème 4
Les valeurs propres de u figurent parmi les racines de ses polynômes annulateurs.

Cependant, un polynôme annulateur de u peut admettre des racines qui ne sont pas valeurs propres de u.

5.1.4 Polynômes annulateurs en dimension finie

On suppose l'espace E de dimension finie.

Théorème 5 (Théorème de Cayley-Hamilton)
Le polynôme caractéristique χ_u est annulateur de u.

L'ensemble des polynômes annulateurs de l'endomorphisme u est alors un idéal de $\mathbb{K}[X]$ non réduit au polynôme nul. Il existe donc un unique polynôme unitaire Π_u tel que les polynômes annulateurs de u sont exactement les multiples de Π_u.

Définition
Ce polynôme Π_u est appelé *polynôme minimal* de l'endomorphisme u.

Le polynôme Π_u est annulateur de u et diviseur du polynôme caractéristique χ_u. C'est par conséquent un polynôme de degré d inférieur à n dont les racines dans $\mathbb{K}$ sont exactement les valeurs propres de u.

Si P est un polynôme de $\mathbb{K}[X]$ et si R est le reste de la division euclidienne de P par le polynôme minimal Π_u, on vérifie $P(u) = R(u)$ avec $\deg(R) < \deg(\Pi_u)$. On en déduit :

Théorème 6
Si d est le degré du polynôme minimal Π_u, la famille $\left(\mathrm{Id}_E, u, \ldots, u^{d-1}\right)$ est une base l'algèbre $\mathbb{K}[u]$ des polynômes en u.

5.1.5 Polynômes d'une matrice carrée

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$ et P un polynôme de $\mathbb{K}[X]$:

$$P = \sum_{k=0}^{N} a_k X^k = a_0 + a_1 X + \cdots + a_N X^N \quad \text{avec} \quad N \geqslant \deg(P).$$

1. Voir sujet 1 p. 202.

Définition

On appelle *valeur du polynôme* P en A la matrice de $\mathcal{M}_n(\mathbb{K})$ donnée par

$$P(A) \overset{\text{déf}}{=} \sum_{k=0}^{N} a_k A^k = a_0 \mathrm{I}_n + a_1 A + \cdots + a_N A^N.$$

Si la matrice A est triangulaire, les coefficients diagonaux de $P(A)$ sont remarquables :

$$\text{Pour } A = \begin{pmatrix} \lambda_1 & & (*) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}, \quad P(A) = \begin{pmatrix} P(\lambda_1) & & (*) \\ & \ddots & \\ (0) & & P(\lambda_n) \end{pmatrix}.$$

Si A est la matrice dans une base e d'un endomorphisme u d'un espace E de dimension finie, $P(A)$ figure l'endomorphisme $P(u)$ dans la même base.

Les résultats qui précèdent se transposent aux matrices :

Théorème 7

L'application

$$\Phi_A : \begin{cases} \mathbb{K}[X] \to \mathcal{M}_n(\mathbb{K}) \\ P \mapsto P(A) \end{cases}$$

est un morphisme de $\mathbb{K}$-algèbres.

L'image $\mathbb{K}[A]$ du morphisme Φ_A est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$. Ses éléments sont les *polynômes en* A. Celle-ci est incluse dans toute sous-algèbre de $\mathcal{M}_n(\mathbb{K})$ contenant A.

Le noyau du morphisme Φ_A est un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$. Celui-ci regroupe les polynômes vérifiant $P(A) = \mathrm{O}_n$. Ces derniers sont les *polynômes annulateurs* de A.

Théorème 8

Les valeurs propres de A figurent parmi les racines de ses polynômes annulateurs.

Encore une fois, il est possible qu'une racine d'un polynôme annulateur ne soit pas valeur propre de la matrice qu'il annule.

Théorème 9 (Théorème de Cayley-Hamilton)

Le polynôme caractéristique χ_A est annulateur de A.

Enfin, on peut introduire le *polynôme minimal* de A : celui-ci est unitaire, annulateur et tout polynôme annulateur de A en est multiple, notamment le polynôme caractéristique. Comme pour le polynôme caractéristique, ses racines sont exactement les valeurs propres de A.

5.2 Réduction et polynômes annulateurs

5.2.1 Lemme de décomposition des noyaux

Théorème 10 (Lemme de décomposition des noyaux)
Si $P_1, \dots, P_m$ sont des éléments de $\mathbb{K}[X]$ deux à deux premiers entre eux de produit égal à P,

$$\operatorname{Ker}\big(P(u)\big) = \bigoplus_{k=1}^{m} \operatorname{Ker}\big(P_k(u)\big).$$

En particulier, si $\lambda_1, \dots, \lambda_m$ sont des valeurs propres deux à deux distinctes de u, on retrouve que les sous-espaces propres associés sont en somme directe puisque les polynômes $X - \lambda_k$ sont deux à deux premiers entre eux.

5.2.2 Diagonalisabilité

Dans la suite, l'espace E est supposé de dimension finie.

Théorème 11
On a équivalence entre :
(i) u est diagonalisable ;
(ii) u annule un polynôme scindé sur $\mathbb{K}$ à racines simples ;
(iii) le polynôme minimal de u est scindé sur $\mathbb{K}$ à racines simples.

Le polynôme minimal de u est alors le produit des $(X - \lambda)$ avec λ parcourant les valeurs propres de u.

Ce résultat se transpose aux matrices carrées :

Théorème 12
Une matrice A de $\mathcal{M}_n(\mathbb{K})$ est diagonalisable si, et seulement si, elle annule un polynôme scindé sur $\mathbb{K}$ à racines simples.

5.2.3 Trigonalisabilité

Théorème 13
On a équivalence entre :
(i) u est trigonalisable ;
(ii) u annule un polynôme scindé sur $\mathbb{K}$;
(iii) le polynôme minimal de u est scindé sur $\mathbb{K}$.
De plus, l'espace E est alors la somme directe de sous-espaces stables par u sur chacun desquels u induit la somme d'une homothétie et d'un endomorphisme nilpotent.

Ce résultat se transpose aux matrices carrées. En particulier, lorsque $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable, celle-ci est semblable à une matrice diagonale par blocs où chaque bloc diagonal est de la forme $\lambda \mathrm{I}_\alpha + N$ avec λ valeur propre de A, α sa multiplicité et N une matrice nilpotente de taille α.

5.2.4 Réduction d'un endomorphisme induit

Si F est un sous-espace vectoriel stable par l'endomorphisme u, on peut introduire l'endomorphisme induit u_F. Le polynôme minimal de celui-ci est alors un diviseur du polynôme minimal de u.

Théorème 14

Soit F un sous-espace vectoriel stable par un endomorphisme u.

a) Si u est diagonalisable, l'endomorphisme induit u_F l'est aussi ;

b) Si u est trigonalisable, l'endomorphisme induit u_F l'est aussi.

Lorsqu'un endomorphisme est diagonalisable, ses sous-espaces stables sont exactement ceux admettant une base de vecteurs propres [1].

5.3 Exercices d'apprentissage

5.3.1 Polynômes d'un endomorphisme, d'une matrice carrée

Exercice 1

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E et P un polynôme de $\mathbb{K}[X]$.

Montrer que si $\lambda \in \mathbb{K}$ est une valeur propre de u, $P(\lambda)$ est une valeur propre de $P(u)$.

Solution

Soit $\lambda \in \mathbb{K}$ une valeur propre de u et $x \neq 0_E$ un vecteur propre associé.

méthode

‖ On montre $P(u)(x) = P(\lambda)x$ en commençant par le cas $P = X^k$.

Vérifions l'égalité [2] $u^k(x) = \lambda^k.x$ par récurrence sur $k \in \mathbb{N}$.

Pour $k = 0$, l'application u^0 est l'endomorphisme Id_E puisqu'il s'agit du neutre pour le produit de composition. Sachant de plus $\lambda^0 = 1$, l'égalité $u^0(x) = \lambda^0 x$ est vraie.

Supposons la propriété établie au rang $k \geqslant 0$. Au rang suivant, on obtient par linéarité

$$u^{k+1}(x) = \big(u \circ u^k\big)(x) = u\big(u^k(x)\big) = u\big(\lambda^k x\big) = \lambda^k u(x) = \lambda^{k+1} x.$$

La récurrence est établie.

1. Voir sujet 8 p. 207.

2. Il importe d'interpréter correctement $P(u)(x)$ lorsque $P = X^k$. Ce n'est pas $P\big(u(x)\big) = \big(u(x)\big)^k$ ce qui signifierait une puissance pour une multiplication qui n'est pas introduite. Il faut plutôt comprendre $u^k(x) = (u \circ \cdots \circ u)(x)$ (composition à k facteurs).

Poursuivons en décrivant le polynôme P par l'introduction de ses coefficients

$$P = a_0 + a_1 X + \cdots + a_N X^N \quad \text{avec} \quad a_0, \ldots, a_N \in \mathbb{K} \text{ et } N \geqslant \deg(P).$$

L'endomorphisme $P(u)$ est alors $P = a_0 \mathrm{Id}_E + a_1 u + \cdots + a_N u^N$ et donc

$$P(u)(x) = \sum_{k=0}^{N} \big(a_k u^k(x)\big) = \sum_{k=0}^{N} \big(a_k \lambda^k x\big) = \left(\sum_{k=0}^{N} a_k \lambda^k\right) x = P(\lambda)x.$$

Sachant le vecteur x non nul, on peut affirmer que $P(\lambda)$ est valeur propre de $P(u)$ associée au vecteur propre x.

Finalement, lorsque λ est valeur propre de u, $P(\lambda)$ est valeur propre de $P(u)$ associée aux mêmes vecteurs propres.

Exercice 2

Montrer qu'un endomorphisme nilpotent d'un espace non réduit au vecteur nul admet une et une seule valeur propre qui est 0.

Solution

Ce résultat est déjà connu en dimension finie (Th. 25 p. 129). Il s'agit ici de le retrouver sans hypothèse de dimension. Soit u un endomorphisme nilpotent d'un espace E non réduit au vecteur nul. Introduisons $p \in \mathbb{N}^*$ tel que $u^p = 0$.

méthode

Les valeurs propres sont racines des polynômes annulateurs (Th. 4 p. 199).

L'égalité $u^p = 0$ signifie que le polynôme X^p est annulateur de l'endomorphisme u. Les valeurs propres de u figurent donc parmi ses racines. On en déduit que seul 0 peut être valeur propre de u. Inversement, vérifions que 0 est valeur propre de u.

méthode

0 est valeur propre d'un endomorphisme si, et seulement si, celui-ci n'est pas injectif.

Par l'absurde, si l'endomorphisme u est injectif, les compositions $u \circ u$, $u \circ u \circ u$, etc. sont toutes injectives car la composition de deux injections définit une injection. Cependant, l'application u^p n'est pas injective car c'est l'application nulle. C'est absurde.

On en déduit que u n'est pas injectif, c'est-à-dire que 0 est valeur propre de u.

Exercice 3

Soit u un endomorphisme bijectif d'un espace de dimension finie $n \geqslant 1$. Montrer que son inverse u^{-1} est un polynôme en u.

Solution

méthode

Le théorème de Cayley-Hamilton (Th. 5 p. 199) détermine[1] un polynôme annulateur de u permettant d'exprimer u^{-1}.

Le polynôme caractéristique de u peut s'écrire

$$\chi_u = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \quad \text{avec} \quad a_0 = (-1)^n \det(u) \neq 0.$$

L'égalité $\chi_u(u) = 0$ donne alors

$$u^n + a_{n-1}u^{n-1} + \cdots + a_1 u + a_0 \mathrm{Id}_E = 0.$$

En réorganisant les membres, on obtient

$$u^n + a_{n-1}u^{n-1} + \cdots + a_1 u = -a_0 \mathrm{Id}_E.$$

En composant par l'application u^{-1} à gauche (ou à droite), il vient

$$u^{n-1} + a_{n-1}u^{n-2} + \cdots + a_1 \mathrm{Id}_E = -a_0 u^{-1}$$

puis

$$u^{-1} = -\frac{1}{a_0}\left(u^{n-1} + a_{n-1}u^{n-2} + \cdots + a_1 \mathrm{Id}_E\right) \in \mathbb{K}[u].$$

Exercice 4

(a) Déterminer le polynôme minimal de chacune des matrices réelles suivantes.

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

(b) Exploiter ces polynômes minimaux pour exprimer A^n, B^n et C^n pour $n \in \mathbb{N}$.

Solution

(a) **méthode**

Le polynôme minimal d'une matrice (resp. d'un endomorphisme) est un diviseur du polynôme caractéristique qui a les mêmes racines.

Le polynôme caractéristique de A est $\chi_A = X^2 - 5X + 6 = (X-2)(X-3)$. Le polynôme minimal Π_A le divise et ses racines sont 2 et 3 : on en déduit $\Pi_A = (X-2)(X-3)$.

Le polynôme caractéristique de B est $\chi_B = X^2 - 4X + 4 = (X-2)^2$. Le polynôme minimal Π_B le divise et 2 est sa seule racine. Le polynôme Π_B peut donc être égal à $X - 2$ ou $(X-2)^2$. Cependant, le polynôme $X - 2$ n'annule pas B car $B \neq 2.\mathrm{I}_2$. On en déduit $\Pi_B = (X-2)^2$.

1. On peut aussi introduire le polynôme minimal de u sachant que 0 n'en est pas racine.

Après quelques calculs, le polynôme caractéristique de C est $\chi_C = (X-2)^2(X-3)$. Le polynôme minimal Π_C le divise et ses racines sont 2 et 3. Le polynôme Π_C peut donc être égal à $(X-2)(X-3)$ ou $(X-2)^2(X-3)$. En évaluant $(X-2)(X-3)$ en C, on constate que ce polynôme est annulateur et donc $\Pi_C = (X-2)(X-3)$.

(b) Soit $n \in \mathbb{N}$.

méthode

On détermine le reste de la division euclidienne de X^n par le polynôme minimal.

Pour calculer A^n, on pose la division euclidienne de X^n par le polynôme Π_A. Celle-ci s'écrit

$$X^n = (X-2)(X-3)Q_n + R_n \quad \text{avec} \quad Q_n, R_n \in \mathbb{R}[X] \text{ et } \deg(R_n) < 2.$$

En notant a_n et b_n les coefficients du polynôme R_n, l'égalité ci-dessus devient

$$X^n = (X-2)(X-3)Q_n + a_nX + b_n. \qquad (*)$$

Pour déterminer a_n et b_n, on évalue $(*)$ en les racines 2 et 3 afin d'éliminer le facteur Q_n. Ceci produit le système

$$\begin{cases} 2a_n + b_n = 2^n \\ 3a_n + b_n = 3^n. \end{cases}$$

Après résolution, $(*)$ devient

$$X^n = (X-2)(X-3)Q_n + \left(3^n - 2^n\right)X + \left(3.2^n - 2.3^n\right). \qquad (**)$$

On évalue alors $(**)$ en la matrice A pour obtenir

$$\begin{aligned} A^n &= (A - 2\mathrm{I}_2)(A - 3\mathrm{I}_3)Q_n(A) + \left(3^n - 2^n\right)A + \left(3.2^n - 2.3^n\right)\mathrm{I}_2 \\ &= 3^n(A - 2\mathrm{I}_2) + 2^n(3\mathrm{I}_2 - A) \quad \text{car} \quad (A - 2\mathrm{I}_2)(A - 3\mathrm{I}_3) = \Pi_A(A) = \mathrm{O}_n. \end{aligned}$$

Pour calculer B^n, on opère de façon semblable en réalisant la division euclidienne de X^n par $(X-2)^2$

$$X^n = (X-2)^2Q_n + a_nX + b_n \quad \text{avec} \quad a_n, b_n \in \mathbb{R} \text{ et } Q_n \in \mathbb{R}[X]. \qquad (\triangle)$$

Pour déterminer a_n et b_n, on évalue $(\triangle)$ en la racine 2 et l'on dérive $(\triangle)$ avant d'évaluer à nouveau en 2. On forme ainsi le système

$$\begin{cases} 2a_n + b_n = 2^n \\ a_n \qquad\quad = n2^{n-1}. \end{cases}$$

Après résolution, $(\triangle)$ devient

$$X^n = (X-2)^2Q_n + n2^{n-1}X - (n-1)2^n. \qquad (\triangle\triangle)$$

Pour terminer, on évalue $(\triangle\triangle)$ en la matrice B pour obtenir

$$B^n = n2^{n-1}B - (n-1)2^n \mathrm{I}_2.$$

Enfin, le polynôme minimal de C étant celui de A, le calcul de C^n se résout en évaluant $(**)$ en C :

$$C^n = 3^n(C - 2\mathrm{I}_2) + 2^n(3\mathrm{I}_2 - C).$$

5.3.2 Réduction et polynômes annulateurs

Exercice 5

Soit u un endomorphisme d'un espace réel E vérifiant $u^3 = \mathrm{Id}_E$. Justifier que les espaces $\mathrm{Ker}(u - \mathrm{Id}_E)$ et $\mathrm{Ker}\big(u^2 + u + \mathrm{Id}_E\big)$ sont supplémentaires.

Solution

méthode

L'application du lemme des noyaux (Th. 10 p. 201) avec un polynôme annulateur permet de réaliser une décomposition en somme directe d'un espace.

Le polynôme $X^3 - 1 = (X-1)\big(X^2 + X + 1\big)$ est annulateur de u et les facteurs $X - 1$ et $X^2 + X + 1$ sont premiers entre eux[1]. On peut appliquer le lemme des noyaux et conclure

$$E = \mathrm{Ker}\big(u^3 - \mathrm{Id}_E\big) = \mathrm{Ker}(u - \mathrm{Id}_E) \oplus \mathrm{Ker}\big(u^2 + u + \mathrm{Id}_E\big).$$

Exercice 6

Soit φ une forme linéaire non nulle sur un espace E de dimension finie, a un vecteur de E et f l'endomorphisme de E déterminé par

$$f(x) = \varphi(a)x - \varphi(x)a \quad \text{pour tout } x \in E.$$

Calculer $f \circ f$ et former une condition nécessaire et suffisante portant sur a et φ pour que f soit diagonalisable.

Solution

Pour $x \in E$, on obtient

$$\begin{aligned}(f \circ f)(x) &= f\big(f(x)\big) = f\big(\varphi(a)x - \varphi(x)a\big) \\ &= \varphi(a)f(x) - \varphi(x)\underbrace{f(a)}_{=0_E} = \varphi(a)f(x).\end{aligned}$$

Par conséquent, $f \circ f = \varphi(a)f$.

1. Pour que deux polynômes de $\mathbb{K}[X]$ soient premiers entre eux, il faut et il suffit qu'ils n'aient pas de racines complexes en commun.

méthode

Un endomorphisme est diagonalisable si, et seulement si, il annule un polynôme scindé à racines simples (Th. 11 p. 201).

Le polynôme $P = X^2 - \varphi(a)X = X\big(X - \varphi(a)\big)$ est annulateur de f.

Cas : $\varphi(a) \neq 0$. Le polynôme P est scindé sur $\mathbb{K}$ et à racines simples : l'endomorphisme f est diagonalisable [1].

Cas : $\varphi(a) = 0$. Le polynôme P est égal à X^2 ce qui ne permet plus de conclure aussi immédiatement. Cependant, 0 est la seule racine de ce polynôme et donc la seule valeur propre possible (Th. 4 p. 199) de f. On en déduit que f est diagonalisable si, et seulement si [2], f est l'endomorphisme nul. Sachant que l'expression de f se simplifie en $f(x) = \varphi(x)a$ lorsque $\varphi(a) = 0$ et sachant que la forme linéaire φ est non nulle, l'endomorphisme f est diagonalisable si, et seulement si, a est le vecteur nul.

En résumé, f est diagonalisable si, et seulement si, $\varphi(a) \neq 0$ ou $a = 0_E$.

Exercice 7

Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ vérifiant $A^3 + A^2 + A = \mathrm{O}_n$. Montrer que le rang de A est pair et que sa trace est un entier.

Solution

Le polynôme $X^3 + X^2 + X = X\big(X^2 + X + 1\big)$ est annulateur de A. Il n'est pas scindé sur $\mathbb{R}$ mais peut s'écrire $X(X - j)\big(X - j^2\big)$ dans $\mathbb{C}[X]$ avec $j = \mathrm{e}^{2\mathrm{i}\pi/3}$.

méthode

La matrice A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ car annule un polynôme scindé sur $\mathbb{C}$ à racines simples (Th. 12 p. 201).

Dans $\mathcal{M}_n(\mathbb{C})$, la matrice A est semblable à une matrice diagonale D où, sur la diagonale, figurent ses valeurs propres comptées avec multiplicité. Les valeurs propres de A sont racines du polynôme annulateur $X(X - j)\big(X - j^2\big)$, ce ne peut donc qu'être $0, j$ et $j^2 = \overline{j}$. Notons p, q et r les multiplicités de chacune, quitte à ce que celles-ci soient nulles si les valeurs associées ne sont pas valeurs propres. La matrice A étant réelle, les multiplicités des valeurs propres conjuguées sont égales et donc $q = r$. Le rang et la trace de la matrice A étant égaux au rang et à la trace de D, on conclut

$$\mathrm{rg}(A) = 2q \in 2\mathbb{N} \quad \text{et} \quad \mathrm{tr}(A) = p \times 0 + qj + qj^2 = -q \in \mathbb{Z}.$$

Exercice 8

Soit u un endomorphisme diagonalisable d'un espace E de dimension finie $n \geqslant 1$. Montrer qu'un sous-espace vectoriel de E est stable par u si, et seulement si, il possède une base formée de vecteurs propres de u.

1. L'étude des éléments propres de f montre que $\varphi(a)$ est valeur propre d'espace propre $\mathrm{Ker}(\varphi)$ tandis que 0 est valeur propre d'espace propre $\mathrm{Vect}(a)$.

2. Un endomorphisme diagonalisable dont λ est la seule valeur propre est $\lambda \mathrm{Id}_E$ (voir sujet 5 p. 134).

Solution

($\Longleftarrow$) Soit F un sous-espace vectoriel possédant une base $(e_1, \ldots, e_p)$ formée de vecteurs propres de u. Pour tout $j \in [\![1\,;p]\!]$, on peut écrire $u(e_j) = \lambda_j e_j \in F$ en introduisant λ_j la valeur propre associée au vecteur propre e_j. Par combinaison linéaire, on a alors $u(x) \in F$ pour tout $x \in F$. Ainsi, un sous-espace vectoriel possédant une base de vecteurs propres est stable.

($\Longrightarrow$) Soit F un sous-espace vectoriel stable par u. On peut introduire u' l'endomorphisme induit par u sur F.

méthode

Les endomorphismes induits par un endomorphisme diagonalisable sont diagonalisables (Th. 14 p. 202).

Puisque u est diagonalisable, l'endomorphisme induit u' l'est aussi et il existe une base de F formée de vecteurs propres de u'. Or les vecteurs propres de u' sont des vecteurs propres de u et l'on peut conclure.

5.4 Exercices d'entraînement

5.4.1 Polynômes d'un endomorphisme, d'une matrice carrée

Exercice 9 *

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension $n \geqslant 1$.
On suppose qu'il existe un vecteur $x_0 \in E$ tel que la famille $\big(x_0, u(x_0), \ldots, u^{n-1}(x_0)\big)$ est libre. Montrer que les polynômes en u sont les seuls endomorphismes qui commutent avec u.

Solution

On sait déjà que les polynômes en u commutent avec u. Inversement, considérons v un endomorphisme commutant avec u. La famille $\big(x_0, u(x_0), \ldots, u^{n-1}(x_0)\big)$ étant libre et formée de $n = \dim E$ vecteurs de E, c'est une base de E. On peut alors écrire

$$v(x_0) = a_0 x_0 + a_1 u(x_0) + \cdots + a_{n-1} u^{n-1}(x_0) \quad \text{avec} \quad a_0, a_1, \ldots, a_{n-1} \in \mathbb{K}.$$

Ceci invite à introduire l'endomorphisme $w = a_0 \mathrm{Id}_E + a_1 u + \cdots + a_{n-1} u^{n-1}$ qui est un polynôme en u et à vérifier $v = w$.

méthode

On peut montrer que deux applications linéaires sont égales en observant qu'elles prennent les mêmes valeurs sur les vecteurs d'une base.

Par construction, les applications v et w sont égales en le vecteur x_0. Plus généralement, pour $k \in [\![0\,;n-1]\!]$, on peut écrire en exploitant la commutation de v et w avec u^k

$$v\big(u^k(x_0)\big) = u^k\big(v(x_0)\big) = u^k\big(w(x_0)\big) = w\big(u^k(x_0)\big).$$

Les applications linéaires v et w sont donc égales sur chacun des vecteurs de la base $(x_0, u(x_0), \dots, u^{n-1}(x_0))$, elles sont donc égales sur E. Finalement [1], $v \in \mathbb{K}[u]$.

Exercice 10 *

Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. On suppose qu'il existe un polynôme $P \in \mathbb{K}[X]$ vérifiant

$$AB = P(A) \quad \text{et} \quad P(0) \neq 0.$$

Montrer que A est inversible et que A et B commutent.

Solution

méthode

On transforme l'égalité $AB = P(A)$ afin de déterminer une matrice C telle que $AC = \mathrm{I}_n$.

Le polynôme P peut s'écrire

$$P = P(0) + a_1 X + \cdots + a_N X^N \quad \text{avec} \quad a_1, \dots, a_N \in \mathbb{K}.$$

L'égalité $AB = P(A)$ donne alors

$$AB = P(0)\mathrm{I}_n + a_1 A + \cdots + a_N A^N$$

et donc

$$A\Big(B - \big(a_1 \mathrm{I}_n + a_2 A + \cdots + a_N A^{N-1}\big)\Big) = \underbrace{P(0)}_{\neq 0}\mathrm{I}_n.$$

On en déduit que A est inversible et son inverse est

$$A^{-1} = \frac{1}{P(0)}\Big(B - \big(a_1 \mathrm{I}_n + a_2 A + \cdots + a_N A^{N-1}\big)\Big).$$

En reprenant les calculs en sens inverse, l'égalité $A^{-1}A = \mathrm{I}_n$ donne $BA = P(A)$ et l'on peut conclure que A et B commutent.

5.4.2 Polynômes annulateurs

Exercice 11 *

Soit u un endomorphisme d'un espace vectoriel réel. On suppose qu'il existe un polynôme P annulateur de u dont 0 est racine simple. Montrer $\mathrm{Ker}(u) = \mathrm{Ker}(u^2)$.

1. La liberté de la famille $(x_0, u(x_0), \dots, u^{n-1}(x_0))$ entraîne que le polynôme minimal de u est de degré au moins égal à n et donc égal à n. L'espace $\mathbb{K}[u]$ des endomorphismes qui commutent avec u est donc de dimension n (voir sujet 19 du chapitre 8 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*).

Solution

L'inclusion $\mathrm{Ker}(u) \subset \mathrm{Ker}(u^2)$ est toujours vraie. Il s'agit d'établir l'inclusion réciproque. Soit $x \in \mathrm{Ker}(u^2)$.

méthode

Le polynôme annulateur P s'écrit $P = aX + X^2Q$ avec $a \neq 0$ et $Q \in \mathbb{K}[X]$.

L'égalité $P(u)(x) = 0_E$ donne

$$au(x) + \big(u^2 \circ Q(u)\big)(x) = 0_E. \qquad (*)$$

Les endomorphismes u^2 et $Q(u)$ commutent et donc

$$\big(u^2 \circ Q(u)\big)(x) = \big(Q(u) \circ u^2\big)(x) = Q(u)\big(u^2(x)\big) = Q(u)(0_E) = 0_E.$$

L'égalité $(*)$ devient $au(x) = 0_E$ et donc $u(x) = 0_E$ car $a \neq 0$. Ainsi, $x \in \mathrm{Ker}(u)$ et l'on peut conclure $\mathrm{Ker}(u) = \mathrm{Ker}(u^2)$ par double inclusion.

Exercice 12 **

Soit u un endomorphisme d'un espace vectoriel E. On suppose qu'il existe deux polynômes $P, Q \in \mathbb{K}[X]$ premiers entre eux vérifiant $(PQ)(u) = 0$.

(a) On suppose l'espace E de dimension finie. Montrer

$$\mathrm{Ker}\big(P(u)\big) \oplus \mathrm{Im}\big(P(u)\big) = E.$$

(b) On ne suppose plus l'espace E de dimension fini. Le résultat précédent est-il encore vrai ?

Solution

(a) Les polynômes P et Q étant premiers entre eux, le lemme de décomposition des noyaux (Th. 10 p. 201) assure que les espaces $\mathrm{Ker}\big(P(u)\big)$ et $\mathrm{Ker}\big(Q(u)\big)$ sont en somme directe. Or l'égalité $Q(u) \circ P(u) = (PQ)(u) = 0$ donne $\mathrm{Im}\big(P(u)\big) \subset \mathrm{Ker}\big(Q(u)\big)$ et les espaces $\mathrm{Ker}\big(P(u)\big)$ et $\mathrm{Im}\big(P(u)\big)$ sont donc aussi en somme directe. De plus, la formule du rang appliquée à l'endomorphisme $P(u)$ donne

$$\dim \mathrm{Ker}\big(P(u)\big) + \dim \mathrm{Im}\big(P(u)\big) = \dim E$$

et donc

$$\mathrm{Ker}\big(P(u)\big) \oplus \mathrm{Im}\big(P(u)\big) = E.$$

(b) Vérifions que le résultat précédent demeure en dimension quelconque. Comme ci-dessus, on sait déjà que les espaces $\mathrm{Ker}\big(P(u)\big)$ et $\mathrm{Im}\big(P(u)\big)$ sont en somme directe. Il reste à établir que leur somme est égale à E. Soit x un élément de E.

méthode

Les polynômes P et Q étant premiers entre eux, on peut introduire des polynômes V et W vérifiant $PV + QW = 1$ et produire une démonstration proche de celle du lemme de décomposition des noyaux (Th. 10 p. 201).

En évaluant la relation $PV + QW = 1$ en u, on obtient l'identité

$$\mathrm{Id}_E = P(u) \circ V(u) + Q(u) \circ W(u).$$

En calculant en le vecteur x, il vient alors

$$x = a + b \quad \text{avec} \quad a = P(u) \circ V(u)(x) \quad \text{et} \quad b = Q(u) \circ W(u)(x).$$

On a immédiatement $a \in \mathrm{Im}\big(P(u)\big)$ car $a = P(u)(y)$ avec $y = V(u)(x)$. On a aussi b élément de $\mathrm{Ker}\big(P(u)\big)$ puisque

$$P(u)(b) = (PQ)(u) \circ W(u)(x) = 0_E \quad \text{car} \quad PQ(u) = 0.$$

Ainsi, tout vecteur de E s'écrit comme la somme d'un vecteur de $\mathrm{Im}\big(P(u)\big)$ et d'un vecteur de $\mathrm{Ker}\big(P(u)\big)$ et l'on peut conclure à l'égalité

$$\mathrm{Ker}\big(P(u)\big) \oplus \mathrm{Im}\big(P(u)\big) = E.$$

Exercice 13 **

Soit $\lambda, \mu \in \mathbb{R}$ et p, q, f trois endomorphismes d'un espace vectoriel réel E vérifiant

$$\begin{cases} f = \lambda p + \mu q \\ f^2 = \lambda^2 p + \mu^2 q \\ f^3 = \lambda^3 p + \mu^3 q. \end{cases}$$

Exprimer f^n en fonction de λ, μ, p et q pour tout $n \in \mathbb{N}^*$.

Solution

méthode

On détermine un polynôme annulateur de f à partir duquel on calcule f^n par une division euclidienne.

Par élimination de p, à l'aide de la première et de la deuxième équation d'une part, et de la deuxième et troisième équation d'autre part, on a

$$f^2 - \lambda f = \mu(\mu - \lambda)q \quad \text{et} \quad f^3 - \lambda f^2 = \mu^2(\mu - \lambda)q.$$

En combinant ces deux équations, on élimine q et l'on parvient à

$$\big(f^3 - \lambda f^2\big) - \mu\big(f^2 - \lambda f\big) = f^3 - (\lambda + \mu)f^2 + \lambda\mu f = 0.$$

Ainsi, le polynôme $P = X^3 - (\lambda + \mu)X^2 + \lambda\mu X = X(X - \lambda)(X - \mu)$ est annulateur de f.

Soit $n \in \mathbb{N}^*$. On réalise la division euclidienne de X^{n-1} par $(X - \lambda)(X - \mu)$:

$$X^{n-1} = (X - \lambda)(X - \mu)Q_n + R_n \quad \text{avec} \quad R_n = a_n X + b_n \text{ et } a_n, b_n \in \mathbb{R}.$$

En multipliant cette relation par X, on obtient

$$X^n = X(X - \lambda)(X - \mu)Q_n + a_n X^2 + b_n X. \qquad (*)$$

En évaluant en f, il vient

$$f^n = a_n f^2 + b_n f = (a_n\lambda^2 + b_n\lambda)p + (a_n\mu^2 + b_n\mu)q.$$

Or en évaluant $(*)$ en λ et μ, on obtient aussi

$$a_n\lambda^2 + b_n\lambda = \lambda^n \quad \text{et} \quad a_n\mu^2 + b_n\mu = \mu^n.$$

On peut donc conclure

$$f^n = \lambda^n p + \mu^n q \quad \text{pour tout } n \geqslant 1.$$

5.4.3 Réduction et polynômes annulateurs

Exercice 14 *

Soit f un endomorphisme d'un espace vectoriel réel E de dimension 3.
On suppose que 1 et -1 sont valeurs propres de f et que $f^4 = f^2$. Montrer que f est diagonalisable.

Solution

méthode

‖ La relation $f^4 = f^2$ détermine un polynôme annulateur de f.

Le polynôme $X^4 - X^2 = X^2(X-1)(X+1)$ est annulateur de f. Ce polynôme est scindé mais n'est pas à racines simples... On étudie[1] les valeurs propres de f. Celles-ci sont racines du polynôme annulateur $X^4 - X^2$ et donc $\mathrm{Sp}(f) \subset \{0, 1, -1\}$. Aussi, 1 et -1 sont par hypothèse valeurs propres. Deux cas sont alors possibles.

Cas : $\mathrm{Sp}(f) = \{0, 1, -1\}$. L'endomorphisme f est diagonalisable car possède 3 valeurs propres en dimension 3.

Cas : $\mathrm{Sp}(f) = \{1, -1\}$. L'endomorphisme f est inversible car 0 n'en est pas valeur propre. On peut alors simplifier la relation $f^4 = f^2$ en composant par f^{-1} et affirmer $f^2 = \mathrm{Id}_E$. Le polynôme $X^2 - 1 = (X-1)(X+1)$ est donc annulateur de f, il est aussi scindé à racines simples et, par conséquent, f est diagonalisable (Th. 11 p. 201).

Exercice 15 *

Soit $A \in \mathcal{M}_3(\mathbb{R})$ vérifiant $A^3 = \mathrm{I}_3$ et $A \neq \mathrm{I}_3$.

(a) Déterminer les valeurs propres réelles de A.

(b) Déterminer les valeurs propres complexes de A.

1. On peut aussi étudier le polynôme minimal de f : celui-ci est de degré inférieur à 3, divise $X^4 - X^2$ et $1, -1$ en sont racines, il est nécessairement scindé à racines simples.

Solution

(a) **méthode**

‖ Les valeurs propres sont racines des polynômes annulateurs (Th. 8 p. 200).

Le polynôme $X^3 - 1$ annule A et 1 est sa seule racine réelle donc la seule valeur propre réelle possible de A. Cependant, la matrice A est réelle et de taille impaire, elle possède [1] donc au moins une valeur propre. On en déduit que 1 est l'unique valeur propre réelle de A.

(b) Dans $\mathbb{C}$, le polynôme $X^3 - 1$ possède trois racines complexes qui sont les racines troisième [2] de l'unité 1, j et j^2. Le spectre complexe de A est donc inclus dans $\{1, j, j^2\}$. De plus, comme on l'a vu au-dessus, 1 est valeur propre de A. Aussi, la matrice A étant réelle, ses valeurs propres sont deux à deux conjuguées et donc

$$j \in \mathrm{Sp}_{\mathbb{C}}(A) \iff j^2 \in \mathrm{Sp}_{\mathbb{C}}(A).$$

Ainsi, le spectre complexe de A est égal à $\{1\}$ ou à $\{1, j, j^2\}$. Cependant, la matrice A est diagonalisable dans $\mathcal{M}_3(\mathbb{C})$ car annule $X^3 - 1 = (X-1)(X-j)(X-j^2)$ qui est un polynôme complexe scindé à racines simples (Th. 12 p. 201). Par l'absurde, si 1 est la seule valeur propre de A, la matrice A est semblable à la matrice I_3 et donc égale à I_3. Le sujet exclut cette possibilité et l'on peut conclure

$$\mathrm{Sp}_{\mathbb{C}}(A) = \{1, j, j^2\}.$$

Exercice 16 *

Soit $M \in \mathcal{M}_n(\mathbb{R})$ vérifiant $M^2 - {}^tM = \mathrm{I}_n$. Montrer que M est diagonalisable.

Solution

méthode

‖ On détermine un polynôme annulateur scindé à racines simples (Th. 12 p. 201).

On réorganise les membres afin de séparer M et ${}^tM$

$${}^tM = M^2 - \mathrm{I}_n.$$

On élève au carré chaque membre

$$\left({}^tM\right)^2 = \left(M^2 - \mathrm{I}_n\right)^2 = M^4 - 2M^2 + \mathrm{I}_n.$$

Or

$$\left({}^tM\right)^2 = {}^t\left(M^2\right) = {}^t\left({}^tM + I_n\right) = M + \mathrm{I}_n$$

et donc

$$M + \mathrm{I}_n = M^4 - 2M^2 + \mathrm{I}_n.$$

1. Le polynôme caractéristique de A est réel et de degré impair, il possède au moins une racine réelle.
2. Rappelons $j = \mathrm{e}^{2\mathrm{i}\pi/3}$ et $\overline{j} = j^2$.

On en déduit que le polynôme $X^4 - 2X^2 - X$ est annulateur de M. Or

$$X^4 - 2X^2 - X = X(X+1)\underbrace{(X-\alpha)(X-\beta)}_{=X^2-X-1} \quad \text{avec} \quad \alpha = \frac{1+\sqrt{5}}{2} \text{ et } \beta = \frac{1-\sqrt{5}}{2}.$$

La matrice M annule un polynôme réel scindé à racines simples, elle est donc diagonalisable dans $\mathcal{M}_n(\mathbb{R})$.

Exercice 17 **

(a) Déterminer toutes les matrices de $\mathcal{M}_n(\mathbb{C})$ vérifiant

$$M^n = \mathrm{I}_n \quad \text{et} \quad \mathrm{tr}(M) = n.$$

(b) Déterminer toutes les matrices de $\mathcal{M}_n(\mathbb{C})$ vérifiant

$$M(M-\mathrm{I}_n)^3 = \mathrm{O}_n \quad \text{et} \quad \mathrm{tr}(M) = 0.$$

Solution

(a) Soit M une matrice solution.

méthode

La trace de la matrice complexe M est la somme de ses valeurs propres comptées avec multiplicité.

La matrice M annule le polynôme $X^n - 1$. Les valeurs propres de M figurent donc parmi les racines de ce polynôme : ce sont des racines n-ièmes de l'unité. Or toutes les racines n-ièmes ont une partie réelle inférieure à 1 et seule la racine 1 est de partie réelle égale à 1. Par l'absurde, si l'une des valeurs propres de M est différente de 1, la partie réelle de $\mathrm{tr}(M)$ est strictement inférieure à n ce qui contredit l'hypothèse $\mathrm{tr}(M) = n$. On en déduit que seul 1 est valeur propre de M et celle-ci est de multiplicité n.

Enfin, M est diagonalisable car annule le polynôme $X^n - 1$ qui est scindé sur $\mathbb{C}$ et à racines simples. On peut alors conclure $M = \mathrm{I}_n$ car 1 est son unique[1] valeur propre.

La réciproque est immédiate.

(b) Soit M une matrice solution. Celle-ci annule le polynôme $X(X-1)^3$. Les valeurs propres de M ne peuvent donc qu'être 0 ou 1. Or la trace de M est nulle et celle-ci est la somme des valeurs propres de M comptées avec multiplicité. On en déduit que seule 0 est valeur propre de M et celle-ci est de multiplicité n. Cependant, on ne peut pas conclure aussi rapidement que dans l'étude précédente car on ignore si la matrice M est diagonalisable.

méthode

Si λ n'est pas valeur propre de M, la matrice $M - \lambda \mathrm{I}_n$ est inversible.

1. Voir sujet 5 p. 134.

Puisque 1 n'est pas valeur propre de M, la matrice $M-\mathrm{I}_n$ est inversible. En multipliant par son inverse, on peut simplifier la relation $M(M-\mathrm{I}_n)^3=\mathrm{O}_n$ pour obtenir $M=\mathrm{O}_n$.

La réciproque est immédiate.

Exercice 18 **

Soit $A, B \in \mathcal{M}_n(\mathbb{R})$ vérifiant $AB=BA$ et M la matrice de $\mathcal{M}_{2n}(\mathbb{R})$ donnée par

$$M=\begin{pmatrix} A & B \\ 0 & A \end{pmatrix}.$$

(a) Montrer que pour tout polynôme P de $\mathbb{R}[X]$

$$P(M)=\begin{pmatrix} P(A) & P'(A)B \\ 0 & P(A) \end{pmatrix}.$$

(b) Énoncer une condition nécessaire et suffisante portant sur A et B pour que M soit diagonalisable.

Solution

(a) Montrons par récurrence[1] l'égalité

$$M^k=\begin{pmatrix} A^k & kA^{k-1}B \\ 0 & A^k \end{pmatrix}$$

pour tout[2] $k \in \mathbb{N}^*$.

Pour $k=1$, l'égalité est vérifiée. Supposons celle-ci vraie au rang $k \geqslant 1$.

$$M^{k+1}=MM^k=\begin{pmatrix} A & B \\ 0 & A \end{pmatrix}\begin{pmatrix} A^k & kA^{k-1}B \\ 0 & A^k \end{pmatrix}=\begin{pmatrix} A^{k+1} & kA^kB+BA^k \\ 0 & A^{k+1} \end{pmatrix}.$$

Or A et B commutent et donc $BA^k=A^kB$ puis

$$M^{k+1}=\begin{pmatrix} A^{k+1} & (k+1)A^kB \\ 0 & A^{k+1} \end{pmatrix}.$$

La récurrence est établie.

Considérons ensuite $P=a_0+a_1X+\cdots+a_NX^N$ un polynôme réel. On a

$$\begin{aligned} P(M) &= a_0\mathrm{I}_{2n}+a_1M+\cdots+a_NM^N \\ &= \begin{pmatrix} \sum_{k=0}^{N} a_kA^k & \sum_{k=1}^{N} ka_kA^{k-1}B \\ 0 & \sum_{k=0}^{N} a_kA^k \end{pmatrix}=\begin{pmatrix} P(A) & P'(A)B \\ 0 & P(A) \end{pmatrix}. \end{aligned}$$

1. On peut aussi appliquer la formule du binôme en écrivant $M=\left(\begin{smallmatrix} A & 0 \\ 0 & A \end{smallmatrix}\right)+\left(\begin{smallmatrix} 0 & B \\ 0 & 0 \end{smallmatrix}\right)$ où les deux matrices introduites commutent et la seconde est nilpotente d'indice 2.

2. On considère séparément $k=0$ afin de ne pas écrire A^{-1} puisque l'on ignore si A est inversible.

(b) **méthode**

On analyse la diagonalisabilité de M à l'aide d'un polynôme annulateur.

Supposons que M est diagonalisable. La matrice M annule un polynôme réel P scindé à racines simples. Les calculs qui précédent montrent que A annule alors ce polynôme et donc A est nécessairement diagonalisable. Au surplus, on a $P'(A)B = O_n$.

méthode

On montre que la matrice B est nulle en vérifiant que la matrice $P'(A)$ est inversible.

En notant $\lambda_1, \dots, \lambda_n$ les valeurs propres de A comptées avec multiplicité, les matrices A et $P'(A)$ sont respectivement semblables à

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{et} \quad P'(D) = \begin{pmatrix} P'(\lambda_1) & & (0) \\ & \ddots & \\ (0) & & P'(\lambda_n) \end{pmatrix}.$$

Or les valeurs propres de A sont racines de P et ce polynôme n'admet que des racines simples. On en déduit que les coefficients $P'(\lambda_1), \dots, P'(\lambda_n)$ sont tous non nuls et la matrice $P'(A)$ est inversible. L'identité $P'(A)B = O_n$ donne[1] alors $B = O_n$.

En résumé, si M est diagonalisable, A est diagonalisable et B est nulle.

Inversement, si A est diagonalisable et B nulle la matrice M est diagonalisable car un polynôme annulateur de A scindé sur $\mathbb{R}$ à racines simples annule aussi[2] M.

Exercice 19 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$. À quelle condition la matrice M de $\mathcal{M}_{2n}(\mathbb{R})$ suivante est-elle diagonalisable ?

$$M = \begin{pmatrix} 0 & A \\ -2A & 3A \end{pmatrix}.$$

Solution

méthode

On exploite une diagonalisation de la matrice $\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$ afin de diagonaliser M par blocs.

Après études des éléments propres, on peut écrire

$$\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} = P \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} P^{-1} \quad \text{avec} \quad P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

1. Une variante élégante est la suivante. Les polynômes P et P' sont premiers entre eux car sans racines complexes en commun. On peut écrire une relation de Bézout $1 = UP + VP'$. En évaluant celle-ci en A et en multipliant par B, on conclut $B = U(A)P(A)B + V(A)P'(A)B = O_n$. En fait, $P'(A)$ est inversible d'inverse $V(A)$.

2. Aussi, si P est une matrice inversible diagonalisant A, la matrice $Q = \left(\begin{smallmatrix} P & 0 \\ 0 & P \end{smallmatrix}\right)$ diagonalise M.

Réalisons une extension par blocs de la matrice P en considérant

$$Q = \begin{pmatrix} \mathrm{I}_n & \mathrm{I}_n \\ \mathrm{I}_n & 2\mathrm{I}_n \end{pmatrix}.$$

Par le calcul d'un produit par blocs, on vérifie que Q est inversible d'inverse

$$Q^{-1} = \begin{pmatrix} 2\mathrm{I}_n & -\mathrm{I}_n \\ -\mathrm{I}_n & \mathrm{I}_n \end{pmatrix}.$$

Toujours par produit par blocs, on obtient

$$M' = Q^{-1}MQ = \begin{pmatrix} A & 0 \\ 0 & 2A \end{pmatrix}.$$

Soit Π un polynôme réel. On vérifie [1]

$$\Pi(M') = \begin{pmatrix} \Pi(A) & 0 \\ 0 & \Pi(2A) \end{pmatrix}.$$

Les polynômes annulateurs de M (qui sont aussi annulateurs de M') annulent la matrice A : si M est diagonalisable, M annule un polynôme scindé sur $\mathbb{R}$ à racines simples et la matrice A l'annule aussi, elle est diagonalisable [2].

Inversement, si A est diagonalisable, on peut introduire une matrice $R \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = RDR^{-1}$ avec D matrice diagonale de taille n. En considérant la matrice inversible

$$S = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} \quad \text{d'inverse} \quad S^{-1} = \begin{pmatrix} R^{-1} & 0 \\ 0 & R^{-1} \end{pmatrix}$$

on a

$$(QS)^{-1}MQS = \begin{pmatrix} D & 0 \\ 0 & 2D \end{pmatrix}$$

et donc M est diagonalisable.

Finalement, la matrice M est diagonalisable si, et seulement si, la matrice A l'est.

Exercice 20 ***

Soit f un endomorphisme d'un espace vectoriel complexe E de dimension finie $n \geqslant 1$.

(a) On suppose que f est diagonalisable. Montrer que f^2 est diagonalisable et que les noyaux de f et f^2 sont égaux.

On étudie désormais la propriété inverse.

(b) Par un exemple, montrer que si f^2 est diagonalisable, f n'est pas nécessairement diagonalisable.

(c) On suppose f^2 diagonalisable et f inversible. Montrer que f est diagonalisable.

(d) On suppose f^2 diagonalisable et $\mathrm{Ker}(f) = \mathrm{Ker}(f^2)$. Montrer à nouveau que f est diagonalisable.

1. La démarche est semblable à celle déjà détaillée dans le sujet 18 p. 215 : on vérifie par récurrence l'identité lorsque le polynôme est X^n avant de généraliser à tout polynôme par combinaison linéaire.

2. La matrice A est la matrice d'un endomorphisme induit par un endomorphisme figuré par M, celui-ci est donc diagonalisable lorsque M l'est.

Solution

(a) Supposons f diagonalisable et introduisons $e = (e_1, \ldots, e_n)$ une base de E formée de vecteurs propres de f. La matrice de f dans cette base est

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \ldots, \lambda_n$ les valeurs propres de f respectivement associées aux vecteurs $e_1, \ldots, e_n$. La matrice de f^2 dans la base e est alors

$$D^2 = \begin{pmatrix} \lambda_1^2 & & (0) \\ & \ddots & \\ (0) & & \lambda_n^2 \end{pmatrix}$$

et donc f^2 est diagonalisable [1]. De plus, les matrices D et D^2 ont le même rang car les éventuels zéros figurant sur la diagonale d'une matrice figurent aux mêmes places sur la diagonale de l'autre. Ainsi, $\mathrm{rg}(f) = \mathrm{rg}(f^2)$ et, puisque l'inclusion de $\mathrm{Ker}(f)$ dans $\mathrm{Ker}(f^2)$ est connue, on peut conclure [2] $\mathrm{Ker}(f) = \mathrm{Ker}(f^2)$.

(b) L'endomorphisme de $\mathbb{C}^2$ canoniquement associé à la matrice suivante convient [3]

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Celui-ci n'est pas diagonalisable car 0 est sa seule valeur propre alors que ce n'est pas l'endomorphisme nul. Cependant, son carré est l'endomorphisme nul qui est diagonalisable.

(c) Supposons f^2 diagonalisable.

méthode

‖ On forme un polynôme annulateur de f à partir du polynôme minimal de f^2.

Soit P le polynôme minimal de f^2. Celui-ci est scindé sur $\mathbb{C}$ et à racines simples car f^2 est diagonalisable. De plus, 0 n'est pas racine de P car 0 n'est pas valeur propre de f^2 puisque f, et donc aussi f^2, est inversible. On peut alors écrire

$$P = \prod_{k=1}^{m} (X - \lambda_k)$$

1. Plus généralement, une base diagonalisant un endomorphisme f diagonalise aussi les polynômes en f.

2. Cette étude reprend dans le cadre vectoriel celle du sujet 7 p. 136.

3. La matrice A est choisie triangulaire non diagonale avec des coefficients diagonaux identiques afin qu'elle ne soit pas diagonalisable. Au surplus, cette matrice est choisie non inversible car les questions de la suite du sujet soulignent que si f est inversible, la diagonalisabilité de f^2 entraîne celle de f.

avec $\lambda_1, \dots, \lambda_m$ des complexes deux à deux distincts et non nuls. L'égalité $P(f^2) = 0$ donne

$$\prod_{k=1}^{m} (f^2 - \lambda_k \mathrm{Id}_E) = 0$$

et donc

$$Q = \prod_{k=1}^{m} (X^2 - \lambda_k)$$

est un polynôme annulateur de f. Montrons que celui-ci est à racines simples.

Soit $k \in [\![1\,;m]\!]$. Le complexe non nul λ_k possèdent [1] deux racines carrées distinctes δ_k et $-\delta_k$ permettant d'écrire la factorisation

$$X^2 - \lambda_k = (X - \delta_k)(X + \delta_k) \quad \text{avec} \quad \delta_k \neq 0.$$

On a alors

$$Q = \prod_{k=1}^{m} (X - \delta_k)(X + \delta_k).$$

Le polynôme Q est à racines simples [2] car les $\delta_1, \dots, \delta_m$ et les $-\delta_1, \dots, -\delta_m$ sont deux à deux distincts puisque, s'ils ne sont pas opposés, ils sont de carrés distincts.

Finalement, f est diagonalisable.

(d) Supposons f^2 diagonalisable et $\mathrm{Ker}(f) = \mathrm{Ker}(f^2)$. Introduisons de nouveau P le polynôme minimal de f^2. Celui-ci est toujours à racines simples mais 0 peut en être racine. Si ce n'est pas le cas, on retrouve l'étude précédente où f est inversible. Sinon, on écrit

$$P = X \prod_{k=1}^{m} (X - \lambda_k)$$

avec $\lambda_1, \dots, \lambda_m$ les valeurs propres non nulles de f^2. L'égalité $P(f^2) = 0$ donne alors

$$f^2 \circ \prod_{k=1}^{m} (f^2 - \lambda_k \mathrm{Id}_E) = f^2 \circ Q(f) = 0 \quad \text{avec} \quad Q = \prod_{k=1}^{m} (X^2 - \lambda_k).$$

Ainsi, on a $\mathrm{Im}(Q(f)) \subset \mathrm{Ker}(f^2)$. Or, par hypothèse, les noyaux de f et f^2 sont égaux et donc $\mathrm{Im}(Q(f)) \subset \mathrm{Ker}(f)$. Ceci donne la relation simplifiée $f \circ Q(f) = 0$.

Enfin, comme au-dessus, on introduit un complexe δ_k de carré λ_k et l'on peut affirmer que f est diagonalisable car annule le polynôme scindé sur $\mathbb{C}$ à racines simples suivant :

$$R = X \prod_{k=1}^{m} (X - \delta_k)(X + \delta_k).$$

1. Si l'on écrit $\lambda = |\lambda| e^{i\theta}$, les racines carrées complexes de λ sont $\delta = \sqrt{|\lambda|} e^{i\theta/2}$ et $-\delta$.

2. L'argumentation suivante est aussi possible. Par dérivation d'un produit Q' est la somme des $2X \prod_{k \neq j}^{m} (X^2 - \lambda_k)$ pour j allant de 1 à m. Une racine de Q annule l'un des facteurs $X^2 - \lambda_k$ et donc tous les termes de la somme donnant Q' sauf un : les polynômes Q et Q' sont sans racines communes et les racines de Q sont donc simples.

5.4.4 Théorème de Cayley-Hamilton

Exercice 21 * (Endomorphisme unipotent)

Soit f un endomorphisme d'un espace vectoriel complexe E de dimension $n \geqslant 1$. On suppose que f possède 1 pour seule valeur propre.

Justifier que l'endomorphisme $f - \mathrm{Id}_E$ est nilpotent.

Solution

méthode

On calcule [1] le polynôme caractéristique de f que l'on sait annulateur (Th. 5 p. 199).

Puisque 1 est la seule valeur propre de l'endomorphisme f, c'est aussi la seule racine de son polynôme caractéristique dans $\mathbb{C}$. Or celui-ci est unitaire de degré n et est scindé sur $\mathbb{C}$, il s'écrit $\chi_f = (X-1)^n$. Le théorème de Cayley-Hamilton assure que ce polynôme annule f et donc $(f - \mathrm{Id}_E)^n = 0$. L'endomorphisme $f - \mathrm{Id}_E$ est nilpotent.

Exercice 22 **

Soit A, B, M trois matrices de $\mathcal{M}_n(\mathbb{C})$ telles que $AM = MB$ avec $M \neq \mathrm{O}_n$.

(a) Montrer que pour tout $P \in \mathbb{C}[X]$, on a $P(A)M = MP(B)$.

(b) Montrer que A et B ont au moins une valeur propre en commun.

Solution

(a) **méthode**

On vérifie l'identité pour $P = X^k$ avec $k \in \mathbb{N}$ avant de généraliser à tout polynôme.

Montrons par récurrence $A^k M = MB^k$ pour tout $k \in \mathbb{N}$.

Lorsque $k = 0$, la relation $A^0 M = MB^0$ se relit simplement $M = M$. Supposons la propriété vraie au rang $k \geqslant 0$. Au rang suivant, on a

$$A^{k+1}M = A\big(A^k M\big) = A\big(MB^k\big) = (AM)B^k = (MB)B^k = MB^{k+1}.$$

La récurrence est établie.

Considérons ensuite P un polynôme de $\mathbb{C}[X]$. En introduisant ses coefficients on peut écrire

$$P = a_0 + a_1 X + \cdots + a_N X^N = \sum_{k=0}^{N} a_k X^k$$

et alors

$$P(A)M = \sum_{k=0}^{N} a_k A^k M = \sum_{k=0}^{N} a_k MB^k = MP(B).$$

1. On peut aussi trigonaliser l'endomorphisme f et observer que $f - \mathrm{Id}_E$ est alors figuré par une matrice triangulaire supérieure stricte.

(b) **méthode**

‖ On considère le polynôme $P = \chi_B$ qui est annulateur de B (Th. 9 p. 200).

Pour $P = \chi_B$, la relation $P(A)M = MP(B)$ entraîne $P(A)M = \mathrm{O}_n$. La matrice $P(A)$ ne peut être inversible car sinon $M = \big(P(A)\big)^{-1}P(A)M = \mathrm{O}_n$ ce que le sujet exclut. On en déduit la nullité du déterminant de $P(A)$. Or l'étude est menée dans le cadre des nombres complexes et le polynôme caractéristique de B peut donc s'écrire

$$P = \prod_{i=1}^{n}(X - \lambda_i)$$

avec λ_i les valeurs propres de B comptées avec multiplicité. L'égalité $\det\big(P(A)\big) = 0$ donne alors

$$\det\left(\prod_{i=1}^{n}(A - \lambda_i \mathrm{I}_n)\right) = \prod_{i=1}^{n}\det(A - \lambda_i \mathrm{I}_n) = 0.$$

Par conséquent, il existe $i \in [\![1\,;n]\!]$ tel que $\det(A - \lambda_i \mathrm{I}_n) = 0$. Le scalaire λ_i est alors valeur propre de A : les matrices A et B ont une valeur propre commune [1].

Exercice 23 ***

Soit f un endomorphisme d'un espace vectoriel complexe E de dimension finie $n \geqslant 1$ vérifiant

$$\mathrm{rg}(f - \lambda \mathrm{Id}_E) = \mathrm{rg}(f - \lambda \mathrm{Id}_E)^2 \quad \text{pour tout } \lambda \in \mathrm{Sp}(f).$$

Montrer que f diagonalisable.

Solution

Soit $\lambda_1, \ldots, \lambda_m$ les valeurs propres de f de multiplicités respectives $\alpha_1, \ldots, \alpha_m$. On peut écrire

$$\chi_f = \prod_{k=1}^{m}(X - \lambda_k)^{\alpha_k}$$

car f est un endomorphisme d'un espace complexe et son polynôme caractéristique est donc scindé [2] sur $\mathbb{C}$. Ce polynôme est annulateur de f et les facteurs $(X - \lambda_k)^{\alpha_k}$ sont deux à deux premiers entre eux. Le lemme de décomposition des noyaux donne alors

$$E = \mathrm{Ker}\big(\chi_f(f)\big) = \bigoplus_{k=1}^{m}\mathrm{Ker}(f - \lambda_k \mathrm{Id}_E)^{\alpha_k}. \qquad (*)$$

méthode

‖ En dimension finie, un endomorphisme u tel que $\mathrm{rg}(u) = \mathrm{rg}\big(u^2\big)$ vérifie [3] aussi $\mathrm{Ker}(u) = \mathrm{Ker}\big(u^2\big)$ et donc $\mathrm{Ker}(u) = \mathrm{Ker}\big(u^2\big) = \mathrm{Ker}\big(u^3\big) = \cdots$.

1. Le résultat du sujet 53 p. 189 permet de proposer une alternative à la démarche suivie ici.
2. Le polynôme χ_f est unitaire, scindé et ses racines sont les $\lambda_1, \ldots, \lambda_m$ de multiplicités $\alpha_1, \ldots, \alpha_m$.
3. Voir sujet 9 p. 79 et sujet 12 p. 82.

Soit $k \in [\![1\,;m]\!]$ et $u = f - \lambda_k \mathrm{Id}_E$. L'hypothèse d'étude donne $\mathrm{rg}(u) = \mathrm{rg}(u^2)$ ce qui permet d'affirmer $\mathrm{Ker}(u) = \mathrm{Ker}(u^2) = \mathrm{Ker}(u^{\alpha_k})$. L'égalité $(*)$ devient alors

$$E = \bigoplus_{k=1}^{m} \mathrm{Ker}(f - \lambda_k \mathrm{Id}_E)$$

qui signifie que l'espace E est la somme directe des espaces propres de f : l'endomorphisme f est diagonalisable.

5.4.5 Polynôme minimal

Exercice 24 *

Soit a un réel et L l'endomorphisme de $\mathcal{M}_n(\mathbb{R})$ (avec $n \geqslant 2$) défini par

$$L(M) = aM + \mathrm{tr}(M)\mathrm{I}_n.$$

(a) Déterminer les éléments propres de L.

(b) En déduire le polynôme minimal de L.

Solution

(a) Si la matrice M est de trace nulle, on a $L(M) = aM$ et inversement. Le réel a est donc valeur propre de L et le sous-espace propre associé est l'hyperplan [1] des matrices de trace nulle.

Aussi, $L(\mathrm{I}_n) = (a+n)\mathrm{I}_n$ et donc $a+n$ est valeur propre de L associée au vecteur propre I_n. Puisque la somme des dimensions des espaces propres est inférieure à la dimension de l'espace, il ne peut y avoir d'autres valeurs propres et l'espace propre associé à la valeur propre $a+n$ est la droite vectorielle engendrée par I_n.

(b) **méthode**

> Lorsqu'un endomorphisme est diagonalisable, son polynôme minimal est le produit des $X - \lambda$ avec λ parcourant ses valeurs propres.

L'endomorphisme L est diagonalisable car la somme des dimensions de ses sous-espaces propres est égale à la dimension de $\mathcal{M}_n(\mathbb{R})$. Le polynôme minimal de L est donc

$$\Pi_L = (X - a)\big(X - (a+n)\big).$$

Exercice 25 **

Montrer que le polynôme minimal et le polynôme caractéristique d'une matrice réelle ont les mêmes facteurs irréductibles.

1. L'ensemble des matrices de trace nulle est un hyperplan car il s'agit du noyau de la trace qui est une forme linéaire non nulle.

Solution

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Puisque le polynôme Π_A divise le polynôme caractéristique χ_A, les facteurs irréductibles de Π_A se retrouvent dans χ_A. Inversement, considérons P un facteur irréductible de χ_A.

méthode

Les polynômes irréductibles réels sont les polynômes de degré 1 et ceux de degré 2 sans racines réelles.

Quitte à considérer un polynôme associé, on peut supposer le polynôme P unitaire.

Si le polynôme P est de degré 1, il s'écrit $X - \lambda$ auquel cas λ est une racine de χ_A donc une valeur propre de A. Celle-ci est alors racine du polynôme minimal car celui-ci est annulateur de A. Ainsi, $P = X - \lambda$ est un facteur irréductible de Π_A.

Si le polynôme P est de degré 2 sans racines réelles, il s'écrit $(X - \lambda)(X - \overline{\lambda})$ avec λ un nombre complexe non réel. Le nombre λ est alors une valeur propre complexe de A et c'est donc une racine complexe du polynôme annulateur Π_A. Dans la décomposition en facteurs irréductibles réels de Π_A, il existe un facteur irréductible $Q \in \mathbb{R}[X]$ dont λ est une racine complexe non réelle. Celui est nécessairement de degré 2 et $\overline{\lambda}$ en est aussi racine. Ce facteur Q est donc associé à P et l'on peut affirmer que P divise Π_A.

Exercice 26 **

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie $n \geqslant 1$.

Montrer que la multiplicité de $\lambda \in \mathbb{K}$ en tant que racine du polynôme minimal Π_u est le plus petit entier naturel p vérifiant

$$\mathrm{Ker}(u - \lambda \mathrm{Id}_E)^p = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)^{p+1}.$$

Solution

Notons α la multiplicité de λ en tant que racine du polynôme minimal Π_u. On peut écrire $\Pi_u = (X - \lambda)^\alpha Q$ avec Q un polynôme de $\mathbb{K}[X]$ vérifiant $Q(\lambda) \neq 0$.

Montrons par double inclusion l'égalité $\mathrm{Ker}(u - \lambda \mathrm{Id}_E)^\alpha = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)^{\alpha+1}$, c'est-à-dire $\mathrm{Ker}\left(v^\alpha\right) = \mathrm{Ker}\left(v^{\alpha+1}\right)$ en introduisant l'endomorphisme $v = u - \lambda \mathrm{Id}_E$.

L'inclusion $\mathrm{Ker}\left(v^\alpha\right) \subset \mathrm{Ker}\left(v^{\alpha+1}\right)$ est bien connue. Étudions l'inclusion réciproque.

méthode

On introduit [1] une relation de Bézout exprimant que les polynômes $X - \lambda$ et Q sont premiers entre eux.

Les polynômes $X - \lambda$ et Q sont premiers entre eux car λ n'est pas racine de Q. On peut donc introduire deux polynômes V et W de $\mathbb{K}[X]$ vérifiant

$$1 = V(X - \lambda) + WQ.$$

On multiplie cette relation par $(X - \lambda)^\alpha$ pour écrire

$$(X - \lambda)^\alpha = V(X - \lambda)^{\alpha+1} + W\Pi_u.$$

1. Il est aussi possible d'écrire $Q = Q(\lambda) + a_1(X - \lambda) + a_2(X - \lambda)^2 + \cdots$ et montrer l'inclusion en adaptant l'étude du sujet 11 p. 209.

En évaluant cette identité en l'endomorphisme u, il vient

$$v^\alpha = (u-\lambda \mathrm{Id}_E)^\alpha = V(u)\circ(u-\lambda\mathrm{Id}_E)^{\alpha+1} + W(u)\circ \underbrace{\Pi_u(u)}_{=0} = V(u)\circ v^{\alpha+1}.$$

Par conséquent [1], $\mathrm{Ker}\big(v^{\alpha+1}\big) \subset \mathrm{Ker}\big(v^\alpha\big)$.

En résumé, le plus petit entier p qui vérifie $\mathrm{Ker}(u-\lambda\mathrm{Id}_E)^p = \mathrm{Ker}(u-\lambda\mathrm{Id}_E)^{p+1}$ est inférieur à α car α vérifie ce type d'identité. Inversement, considérons le plus petit entier p tel que $\mathrm{Ker}(u-\lambda\mathrm{Id}_E)^p = \mathrm{Ker}(u-\lambda\mathrm{Id}_E)^{p+1}$, autrement dit, tel que $\mathrm{Ker}\big(v^p\big) = \mathrm{Ker}\big(v^{p+1}\big)$.

méthode

Lorsque $\mathrm{Ker}\big(v^{p+1}\big) = \mathrm{Ker}\big(v^p\big)$, on a aussi [2] $\mathrm{Ker}\big(v^{p+k}\big) = \mathrm{Ker}\big(v^p\big)$ pour tout $k \in \mathbb{N}$.

En particulier, $\mathrm{Ker}\big(v^p\big) = \mathrm{Ker}\big(v^\alpha\big)$. Montrons alors que $(X-\lambda)^pQ$ est annulateur de u. L'égalité

$$\Pi_u(u) = (u-\lambda\mathrm{Id}_E)^\alpha \circ Q(u) = 0_E$$

donne $\mathrm{Im}\big(Q(u)\big) \subset \mathrm{Ker}(u-\lambda\mathrm{Id}_E)^\alpha$ et donc $\mathrm{Im}\big(Q(u)\big) \subset \mathrm{Ker}(u-\lambda\mathrm{Id}_E)^p$. On en déduit

$$(u-\lambda\mathrm{Id}_E)^p \circ Q(u) = 0_E.$$

Le polynôme $(X-\lambda)^pQ$ est donc annulateur de u et le polynôme minimal Π_u le divise. On conclut $p \geqslant \alpha$ puis $p = \alpha$.

Exercice 27 **

Soit u un endomorphisme d'un espace vectoriel complexe E de dimension finie $n \geqslant 1$. On note $\lambda_1, \ldots, \lambda_m$ les valeurs propres sans répétitions de u et $\alpha_1, \ldots, \alpha_m$ leurs multiplicités respectives. Montrer que, pour tout $k \in [\![1\,;m]\!]$,

$$\dim \mathrm{Ker}(u-\lambda_k\mathrm{Id}_E)^{\alpha_k} = \alpha_k.$$

Solution

Avec les notations introduites, le polynôme caractéristique de u s'écrit

$$\chi_u = \prod_{k=1}^{m}(X-\lambda_k)^{\alpha_k}.$$

Par le théorème de Cayley-Hamilton, le polynôme caractéristique de u est annulateur de u. Les facteurs $(X-\lambda_k)^{\alpha_k}$ étant deux à deux premiers entre eux, on peut appliquer le lemme des noyaux et écrire [3]

$$E = \mathrm{Ker}\big(\chi_u(u)\big) = \bigoplus_{k=1}^{m} \mathrm{Ker}(u-\lambda_k\mathrm{Id}_E)^{\alpha_k}.$$

1. Avec des notations entendues, si $f = g\circ h$, on a $\mathrm{Ker}(h) \subset \mathrm{Ker}(f)$.
2. Voir sujet 12 p. 82.
3. Les espaces $\mathrm{Ker}(u-\lambda_k\mathrm{Id}_E)^{\alpha_k}$ se nomment les *sous-espaces caractéristiques* de u.

méthode

Le polynôme caractéristique de u peut être calculé par la matrice diagonale par blocs que l'on obtient lorsque l'on figure u dans une base adaptée à la décomposition en somme directe précédente.

Chaque espace $\mathrm{Ker}(u-\lambda_k \mathrm{Id}_E)^{\alpha_k}$ est stable par u car u et $(u-\lambda_k \mathrm{Id}_E)^{\alpha_k}$ commutent. Dans une base adaptée à la décomposition en somme directe précédente, la matrice de u est diagonale par blocs et ses blocs diagonaux figurent les endomorphismes u_k induits par u sur les espaces $\mathrm{Ker}(u-\lambda_k \mathrm{Id}_E)^{\alpha_k}$. Autrement dit, la matrice de u dans cette base adaptée est de la forme

$$M = \begin{pmatrix} A_1 & & (0) \\ & \ddots & \\ (0) & & A_m \end{pmatrix}$$

avec A_k matrice représentant u_k. On peut alors calculer le polynôme caractéristique de u à l'aide des polynômes caractéristiques des endomorphismes induits u_k

$$\chi_u = \chi_M = \prod_{k=1}^{m} \chi_{A_k} = \prod_{k=1}^{m} \chi_{u_k}.$$

Fixons ensuite un indice $k \in [\![1\,;m]\!]$. Le polynôme $(X-\lambda_k)^{\alpha_k}$ annule l'endomorphisme induit u_k et donc λ_k est sa seule valeur propre. Le polynôme caractéristique de l'endomorphisme u_k est scindé sur $\mathbb{C}$ et s'écrit donc $(X-\lambda_k)^{\beta_k}$ avec β_k la dimension de l'espace $\mathrm{Ker}(u-\lambda_k \mathrm{Id}_E)^{\alpha_k}$.

Finalement,

$$\chi_u = \prod_{k=1}^{m} (X-\lambda_k)^{\beta_k}.$$

Par unicité de la décomposition d'un polynôme complexe en facteurs irréductibles, on conclut, pour tout $k \in [\![1\,;m]\!]$,

$$\dim \mathrm{Ker}(u-\lambda_k \mathrm{Id}_E)^{\alpha_k} = \beta_k = \alpha_k.$$

Exercice 28 ***

Soit u un endomorphisme d'un espace vectoriel complexe E de dimension finie non nulle de polynôme minimal Π_u.

(a) Soit $x \in E$. Justifier l'existence d'un unique polynôme unitaire P_x vérifiant, pour tout $P \in \mathbb{C}[X]$,

$$P(u)(x) = 0_E \iff P_x \mid P.$$

(b) Soit x et y deux vecteurs de E. On suppose que P_x et P_y sont premiers entre eux. Déterminer P_{x+y}.

(c) Soit λ une valeur propre de u et α sa multiplicité dans le polynôme minimal Π_u. Montrer l'existence d'un vecteur $x \in E$ tel que $P_x = (X-\lambda)^{\alpha}$.

(d) Conclure qu'il existe un vecteur x de E tel que $P_x = \Pi_u$.

Solution

(a) Commençons par établir l'unicité du polynôme P_x. Supposons P_x et Q_x deux polynômes solutions. Puisque P_x se divise lui-même, on a $P_x(u)(x) = 0_E$ et donc Q_x divise P_x. Un raisonnement symétrique donne que P_x divise Q_x et donc $P_x = Q_x$ car ces deux polynômes sont unitaires. Montrons maintenant l'existence du polynôme P_x.

méthode

On vérifie que l'ensemble des polynômes P tels que $P(u)(x) = 0_E$ est un idéal de $\mathbb{C}[X]$ non réduit au polynôme nul.

Posons

$$I_x = \{P \in \mathbb{C}[X] \mid P(u)(x) = 0_E\}.$$

La partie I_x est non vide et non réduite à l'élément nul car le polynôme minimal Π_u lui appartient. Si P et Q sont deux éléments de I_x, le polynôme $P + Q$ appartient à I_x car

$$(P + Q)(u)(x) = P(u)(x) + Q(u)(x) = 0_E.$$

De plus, si R est un polynôme quelconque de $\mathbb{C}[X]$, le polynôme PR appartient aussi à I_x car

$$(PR)(u) = \big(R(u) \circ P(u)\big)(x) = R(u)\big(P(u)(x)\big) = R(u)(0_E) = 0_E.$$

Ainsi, I_x est un idéal de $\mathbb{C}[X]$ et il existe un polynôme $P_x \in \mathbb{C}[X]$ tel que $I_x = P_x\mathbb{C}[X]$ (Th. 7 p. 39). Ce polynôme P_x n'est pas nul car I_x n'est pas réduit au polynôme nul et, quitte à le multiplier par une constante bien choisie, on peut supposer P_x unitaire.

(b) Montrons par double divisibilité que P_{x+y} est le produit P_xP_y.

D'une part, par linéarité

$$\begin{aligned}(P_xP_y)(u)(x+y) &= (P_xP_y)(u)(x) + (P_xP_y)(u)(y)\\ &= \big(P_y(u) \circ P_x(u)\big)(x) + \big(P_x(u) \circ P_y(u)\big)(y)\\ &= P_y(u)\big(\underbrace{P_x(u)(x)}_{=0_E}\big) + P_x(u)\big(\underbrace{P_y(u)(y)}_{=0_E}\big) = 0_E\end{aligned}$$

et donc P_{x+y} divise P_xP_y.

D'autre part, en écrivant $x = x + y - y$, on a P_x diviseur de $P_{x+y}P_{-y}$ avec $P_{-y} = P_y$. Or P_x est premier avec P_y et le lemme de Gauss assure que P_x divise P_{x+y}. Un raisonnement symétrique donne que P_y divise aussi P_{x+y}. Enfin, les polynômes P_x et P_y sont premiers entre eux et l'on peut affirmer que leur produit P_xP_y divise encore P_{x+y}.

On peut alors conclure $P_{x+y} = P_xP_y$ car ces deux polynômes sont unitaires et se divisent mutuellement.

(c) Par définition de α, le polynôme minimal Π_u s'écrit $(X - \lambda)^\alpha Q$ avec $Q(\lambda) \neq 0$.

Un vecteur tel que voulu est à chercher dans $\mathrm{Ker}(u - \lambda\mathrm{Id}_E)^\alpha$. Soit x un vecteur de $\mathrm{Ker}(u - \lambda\mathrm{Id}_E)^\alpha$. Le polynôme $P = (X - \lambda)^\alpha$ vérifie $P(u)(x) = 0_E$ et donc P_x le

divise. Ainsi, le polynôme P_x s'écrit $(X-\lambda)^{\beta_x}$ avec $\beta_x \leqslant \alpha$ (l'entier β_x dépendant du choix de x).

méthode

Par l'absurde, s'il n'existe pas de vecteurs x tels que $\beta_x = \alpha$, on contredit la minimalité [1] du polynôme Π_u.

Supposons que pour tout vecteur x de $\mathrm{Ker}(u-\lambda \mathrm{Id}_E)^\alpha$ la valeur de β_x est strictement inférieure à α. On a donc $\beta_x \leqslant \alpha - 1$ et par conséquent $(u-\lambda \mathrm{Id}_E)^{\alpha-1}(x) = 0_E$. On dispose alors de l'inclusion [2]

$$\mathrm{Ker}(u-\lambda \mathrm{Id}_E)^\alpha \subset \mathrm{Ker}(u-\lambda \mathrm{Id}_E)^{\alpha-1}.$$

Or $\Pi_u(u) = 0$ donne $\mathrm{Im}\big(Q(u)\big) \subset \mathrm{Ker}(u-\lambda \mathrm{Id}_E)^\alpha$ et donc $\mathrm{Im}\big(Q(u)\big) \subset \mathrm{Ker}(u-\lambda \mathrm{Id}_E)^{\alpha-1}$. Ainsi, $(X-\lambda)^{\alpha-1}Q$ est annulateur de u. C'est absurde car le polynôme minimal Π_u ne le divise pas !

Finalement, il existe un vecteur x dans E tel que $P_x = (X-\lambda)^\alpha$.

(d) Le polynôme minimal de u s'écrit dans $\mathbb{C}[X]$

$$\Pi_u = \prod_{k=1}^{m} (X-\lambda_k)^{\alpha_k}$$

avec $\lambda_1, \ldots, \lambda_m$ les racines sans répétitions de Π_u (ce sont aussi les valeurs propres de u) et $\alpha_1, \ldots, \alpha_m$ leur multiplicités respectives. L'étude qui précède assure l'existence pour tout $k \in [\![1\,;m]\!]$ d'un vecteur x_k pour lequel $P_{x_k} = (X-\lambda_k)^{\alpha_k}$. Ces polynômes étant deux à deux premiers entre eux, l'application répétée du résultat de la question (b) assure que le vecteur $x = x_1 + \cdots + x_m$ convient.

5.4.6 Applications

Exercice 29 *

Soit A la matrice donnée par

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

(a) Déterminer le polynôme minimal de la matrice A.

On étudie l'équation $M^2 - M = A$ d'inconnue $M \in \mathcal{M}_2(\mathbb{R})$.

(b) Justifier que les solutions de cette équation sont diagonalisables et déterminer les valeurs propres possibles de celles-ci.

(c) Déterminer les matrices M solutions par l'introduction d'un polynôme annulateur.

1. Cette étude entre en résonance avec celle du sujet 26 p. 223.
2. On a même l'égalité car l'inclusion en sens inverse est toujours vraie.

Solution

(a) Le polynôme minimal Π_A divise le polynôme caractéristique $\chi_A = X(X-2)$, il possède les mêmes racines et est unitaire : $\Pi_A = X^2 - 2X$.

(b) Soit $M \in \mathcal{M}_2(\mathbb{R})$ une matrice solution. L'égalité $A^2 - 2A = A(A - 2\mathrm{I}_2) = \mathrm{O}_2$ donne $(M^2 - M)(M^2 - M - 2\mathrm{I}_2) = \mathrm{O}_2$ et donc

$$P = (X^2 - X)(X^2 - X - 2) = X(X-1)(X+1)(X-2)$$

est annulateur de M. Celui-ci est scindé sur $\mathbb{R}$ et à racines simples et la matrice M est donc diagonalisable. De plus, les valeurs propres possibles de M sont les racines de P, à savoir les nombres $0, 1, -1$ et 2.

(c) Soit $M \in \mathcal{M}_2(\mathbb{R})$ une matrice solution. Celle-ci est diagonalisable et ses deux valeurs propres ne peuvent être égales. En effet, si la matrice diagonalisable M ne possède qu'une seule valeur propre λ, celle-ci est semblable à $\lambda\mathrm{I}_2$, donc égale à $\lambda\mathrm{I}_2$, or cette dernière n'est pas solution de l'équation.

méthode

> Si λ et μ sont les deux valeurs propres de M alors $(M - \lambda\mathrm{I}_2)(M - \mu\mathrm{I}_2) = \mathrm{O}_2$ car M est diagonalisable.

Cas : $\mathrm{Sp}(M) = \{0, 2\}$. On a conjointement $M^2 - 2M = \mathrm{O}_2$ et l'équation $M^2 - M = A$. On en déduit $M = A$ en faisant la différence de ces deux relations.

Cas : $\mathrm{Sp}(M) = \{0, -1\}$. On a $M^2 + M = \mathrm{O}_2$ et $M^2 - M = A$ donc $M = -\frac{1}{2}A$.

Cas : $\mathrm{Sp}(M) = \{0, 1\}$. On a $M^2 - M = \mathrm{O}_2$ et $M^2 - M = A$ ce qui est incompatible[1].

Les cas $\mathrm{Sp}(M) = \{1, -1\}$, $\mathrm{Sp}(M) = \{1, 2\}$ et $\mathrm{Sp}(M) = \{-1, 2\}$ sont analogues et conduisent respectivement à $M = \mathrm{I}_2 - A$, $M = \frac{1}{2}A + \mathrm{I}_2$ et une incompatibilité.

Inversement, les matrices proposées sont solutions. On peut le vérifier par le calcul ou en reprenant le raisonnement en sens inverse : on détermine pour chaque cas les valeurs propres de M en fonction de celles de A ce qui propose un polynôme annulateur de M à partir duquel on retrouve l'équation $M^2 - M = A$.

Exercice 30 **

Soit (u_n) une suite réelle vérifiant, pour tout $n \in \mathbb{N}$,

$$u_{n+3} + 4u_{n+2} + 5u_{n+1} + 2u_n = 0.$$

Pour tout $n \in \mathbb{N}$, on pose $X_n \in \mathcal{M}_{3,1}(\mathbb{R})$ la colonne de coefficients u_n, u_{n+1}, u_{n+2}.

(a) Déterminer une matrice $A \in \mathcal{M}_3(\mathbb{R})$ telle que $X_{n+1} = AX_n$.

(b) Exprimer u_n en fonction de u_0, u_1, u_2 et $n \in \mathbb{N}$.

1. Si $\mathrm{Sp}(M) = \{0, 1\}$, les valeurs propres de $M^2 - M$ sont toutes deux égales à 0 : on ne retrouve pas les deux valeurs propres de A. Ceci explique pourquoi cette situation conduit à une incompatibilité.

Solution

(a) Pour tout $n \in \mathbb{N}$

$$X_{n+1} = \begin{pmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix} = AX_n \quad \text{avec} \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{pmatrix}.$$

(b) Par récurrence, il vient $X_n = A^n X_0$ pour tout $n \in \mathbb{N}$. Ceci invite au calcul de A^n.

méthode

On calcule A^n à partir d'un polynôme annulateur et d'une division euclidienne.

Après quelques calculs[1], on obtient $\chi_A = (X+1)^2(X+2)$ et l'on sait par le théorème de Cayley-Hamilton que ce polynôme est annulateur[2] de A. Soit $n \in \mathbb{N}$. La division euclidienne de X^n par χ_A s'écrit

$$X^n = \chi_A Q + R \quad \text{avec} \quad R = aX^2 + bX + c \text{ et } a, b, c \in \mathbb{R}.$$

Pour rendre les calculs qui suivent plus simples, on exprime plutôt le reste R dans la base de Taylor formée des polynômes $1, (X+1)$ et $(X+1)^2$

$$R = \alpha(X+1)^2 + \beta(X+1) + \gamma \quad \text{avec} \quad \alpha, \beta, \gamma \in \mathbb{R}.$$

On détermine ensuite α, β et γ en évaluant l'identité

$$X^n = \chi_A Q + \alpha(X+1)^2 + \beta(X+1) + \gamma \qquad (*)$$

en -2, en -1 et en dérivant $(*)$ avant d'évaluer à nouveau en -1. On forme ainsi le système

$$\begin{cases} \alpha - \beta + \gamma = (-2)^n \\ \gamma = (-1)^n \\ \beta = n(-1)^{n-1}. \end{cases}$$

On en déduit

$$R = (-1)^n\big(2^n - n - 1\big)(X+1)^2 + (-1)^{n-1} n(X+1) + (-1)^n.$$

Enfin, en évaluant $(*)$, en la matrice A

$$A^n = (-1)^n\big(2^n - n - 1\big)(A + \mathrm{I}_3)^2 + (-1)^{n-1} n(A + \mathrm{I}_3) + (-1)^n \mathrm{I}_3.$$

Il suffit alors de calculer la première ligne de $A + \mathrm{I}_3$ et de $(A + \mathrm{I}_3)^2$ pour pouvoir exprimer u_n en fonction de u_0, u_1 et u_2 :

$$u_n = (-1)^n\big(\big(2^n - 2n\big)u_0 + \big(2^{n+1} - 3n - 2\big)u_1 + \big(2^n - n - 1\big)u_2\big).$$

1. On pourra initier le calcul du polynôme caractéristique par la transformation $C_1 \leftarrow C_1 - C_2 + C_3$.
2. Il s'agit même de son polynôme minimal car la matrice A n'est pas diagonalisable (voir sujet 35 p. 169).

Exercice 31 **

Soit u et v deux endomorphismes diagonalisables d'un espace vectoriel E de dimension finie non nulle. Montrer que u et v commutent si, et seulement si, u et v sont simultanément [1] diagonalisables.

Solution

($\Longleftarrow$) Supposons qu'il existe une base e de l'espace E dans laquelle les endomorphismes u et v sont figurés par des matrices diagonales D_u et D_v. Les matrices diagonales commutent entre elles et donc $D_u D_v = D_v D_u$ puis $u \circ v = v \circ u$: les endomorphismes u et v commutent.

($\Longrightarrow$) Supposons que les endomorphismes u et v commutent.

méthode

Lorsque deux endomorphismes commutent, les sous-espaces propres de l'un sont stables [2] pour l'autre.

Notons $\lambda_1, \ldots, \lambda_m$ les valeurs propres deux à deux distinctes de l'endomorphisme u. Celui-ci étant diagonalisable, on sait

$$E = E_{\lambda_1}(u) \oplus \cdots \oplus E_{\lambda_m}(u).$$

Pour tout $k \in [\![1\,;m]\!]$, l'espace propre $E_{\lambda_k}(u)$ est stable par v. L'endomorphisme v étant diagonalisable, l'endomorphisme qu'il induit sur l'espace $E_{\lambda_k}(u)$ est aussi diagonalisable (Th. 14 p. 202). On peut donc former une base $e^k = \left(e_1^k, \ldots, e_{\alpha_k}^k\right)$ de l'espace $E_{\lambda_k}(u)$ (avec α_k la dimension de $E_{\lambda_k}(u)$) constituée de vecteurs propres de v. Celle-ci est aussi constituée de vecteurs propres de u puisque les vecteurs non nuls de $E_{\lambda_k}(u)$ sont vecteurs propres associés à la valeur propre λ_k. Enfin, en considérant la famille obtenue en accolant les différentes bases $e^1, \ldots, e^m$, on constitue une base de E car E est la somme directe des espaces $E_{\lambda_k}(u)$. Cette base est formée de vecteurs propres communs à u et v, c'est une base de diagonalisation commune [3].

Exercice 32 **

Soit u un endomorphisme diagonalisable d'un espace vectoriel complexe E de dimension finie $n \geqslant 1$. On étudie l'équation $v^2 = u$ d'inconnue $v \in \mathcal{L}(E)$.

(a) Montrer qu'il existe au moins un endomorphisme v de E solution.

(b) Justifier que l'on peut déterminer une solution v_0 qui soit un polynôme en u.

1. Ceci signifie l'existence d'une base de diagonalisation commune aux endomorphismes u et v.
2. Voir sujet 13 p. 146.
3. Une conséquence de ce résultat est que les endomorphismes combinaisons linéaires de u et v sont diagonalisables.

Solution

(a) Soit $e = (e_1, \dots, e_n)$ une base de vecteurs propres de u. La matrice de u dans cette base est de la forme

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \dots, \lambda_n$ les valeurs propres de u comptées avec multiplicité.

méthode

On détermine une solution en introduisant des racines carrées des nombres complexes $\lambda_1, \dots, \lambda_n$.

Soit $j \in [\![1\,;n]\!]$. Le complexe λ_j peut s'écrire $\lambda_j = |\lambda_j|\,\mathrm{e}^{\mathrm{i}\theta_j}$ avec un argument $\theta_j \in \mathbb{R}$. Posons alors $\mu_j = \sqrt{|\lambda_j|}\mathrm{e}^{\mathrm{i}\theta_j/2}$ de sorte que $\mu_j^2 = \lambda_j$. Enfin, considérons l'endomorphisme v figuré dans la base e par la matrice suivante

$$\Delta = \begin{pmatrix} \mu_1 & & (0) \\ & \ddots & \\ (0) & & \mu_n \end{pmatrix}.$$

Par construction, Δ vérifie $\Delta^2 = D$ et l'on a donc $v^2 = u$. L'endomorphisme v est solution de l'équation étudiée.

(b) **méthode**

On définit un polynôme interpolateur qui envoie les valeurs propres de u sur des racines carrées complexes de celles-ci.

Introduisons de nouveau $\lambda_1, \dots, \lambda_m$ les valeurs propres de u, mais cette fois-ci sans répétitions, quitte à introduire leurs multiplicités $\alpha_1, \dots, \alpha_m$. Comme au-dessus, en considérant les modules et arguments des λ_k, on peut définir des μ_k complexes vérifiant $\mu_k^2 = \lambda_k$. Par interpolation [1] de Lagrange, on peut introduire un polynôme P tel que $P(\lambda_k) = \mu_k$ pour tout $k \in [\![1\,;m]\!]$. Considérons enfin l'endomorphisme $v_0 = P(u)$. Dans une base adaptée à la supplémentarité de ses sous-espaces propres, l'endomorphisme u est figuré par la matrice D suivante

$$D = \begin{pmatrix} \lambda_1 \mathrm{I}_{\alpha_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_m \mathrm{I}_{\alpha_m} \end{pmatrix}.$$

1. Pour réaliser cette interpolation, il importe que les points $\lambda_1, \dots, \lambda_m$ où l'on impose les valeurs du polynôme soient deux à deux distincts : c'est pour cette raison que l'on ne considère plus les valeurs propres comptées avec multiplicité de peur que, pour deux valeurs propres identiques, on impose des racines carrées complexes différentes à cause du choix arbitraire de l'argument.

L'endomorphisme v_0 est alors représenté dans cette base par la matrice

$$\Delta = P(D) = \begin{pmatrix} P(\lambda_1)\mathrm{I}_{\alpha_1} & & (0) \\ & \ddots & \\ (0) & & P(\lambda_m)\mathrm{I}_{\alpha_m} \end{pmatrix} = \begin{pmatrix} \mu_1\mathrm{I}_{\alpha_1} & & (0) \\ & \ddots & \\ (0) & & \mu_m\mathrm{I}_{\alpha_m} \end{pmatrix}.$$

Comme au-dessus, on a par construction $\Delta^2 = D$ et donc $v_0^2 = u$. Ainsi, v_0 est une solution de l'équation qui a la particularité d'être un polynôme en u.

Exercice 33 ***

Soit u un endomorphisme d'un espace vectoriel réel E de dimension finie non nulle. Montrer qu'il existe une droite vectorielle ou un plan vectoriel stable par u.

Solution

méthode

On détermine un polynôme irréductible réel P tel que $\mathrm{Ker}\big(P(u)\big)$ n'est pas réduit au vecteur nul.

Le polynôme caractéristique de u peut s'écrire $\chi_u = P_1P_2\dots P_m$ avec $P_1, P_2, \dots, P_m$ une liste pouvant comporter des répétitions de polynômes irréductibles unitaires de $\mathbb{R}[X]$. Puisque $\chi_u(u) = P_1(u)\circ\cdots\circ P_m(u) = 0$, les endomorphismes $P_k(u)$ ne peuvent être tous injectifs. Il existe donc au moins un indice $k \in [\![1\,;m]\!]$ tel que $\mathrm{Ker}\big(P_k(u)\big)$ n'est pas réduit au vecteur nul. Introduisons alors x un vecteur non nul de ce noyau et discutons selon le degré du polynôme irréductible P_k.

Si le polynôme P_k est de degré 1, il s'écrit $P_k = X - \lambda$. Dans ce cas, x est vecteur propre de u associé à la valeur propre λ et la droite $D = \mathrm{Vect}(x)$ est stable par u.

Si le polynôme P_k est de degré 2, il s'écrit $P_k = X^2 + \mu X + \nu$ avec $\mu^2 - 4\nu < 0$. Le vecteur x vérifie alors $u^2(x) + \mu u(x) + \nu x = 0_E$ et donc $u^2(x)$ appartient au sous-espace vectoriel $P = \mathrm{Vect}\big(x, u(x)\big)$. Ce dernier est alors stable par u car l'image par u d'une combinaison linéaire de x et $u(x)$ est une combinaison linéaire de $u(x)$ et $u^2(x)$, c'est donc un élément de P. Enfin, la dimension de P est inférieure[1] à 2 et cela suffit à affirmer l'existence d'une droite ou d'un plan stable par u.

5.5 Exercices d'approfondissement

Exercice 34 *

Soit A une matrice carrée de taille 2 à coefficients entiers. On suppose que $A^n = \mathrm{I}_2$ pour une certaine valeur de $n \in \mathbb{N}^*$. Montrer que $A^{12} = \mathrm{I}_2$.

1. En fait, P est un plan car x n'est pas vecteur propre de u puisque P est un polynôme sans racines réelles annulateur de l'endomorphisme induit par u sur $\mathrm{Ker}\big(P(u)\big)$.

Solution

méthode

Les valeurs propres de A sont des racines de l'unité dont la somme est un entier.

La matrice A est diagonalisable dans $\mathcal{M}_2(\mathbb{C})$ car annule le polynôme $X^n - 1$ qui est scindé sur $\mathbb{C}$ à racines simples. De plus, les valeurs propres de A sont racines de ce polynôme, ce sont des racines de l'unité.

Si les deux valeurs propres de A sont réelles, ce ne peuvent être que 1 et -1. Les spectres de A possibles sont alors

$$\{1\}, \quad \{-1\} \quad \text{et} \quad \{1,-1\}. \tag{$*$}$$

Si A possède une valeur propre λ complexe non réelle, $\overline{\lambda}$ est aussi valeur propre de A car la matrice A est à coefficients réels. La trace de A vaut alors $\lambda + \overline{\lambda} = 2\operatorname{Re}\lambda$ et c'est un entier car A est à coefficients entiers. Puisque λ est une racine de l'unité non réelle, on a $|\operatorname{Re}(\lambda)| < 1$ et donc $\operatorname{tr}(A) \in [\![-1\,;1]\!]$. Aussi, le déterminant de A vaut $\lambda\overline{\lambda} = |\lambda|^2 = 1$ et, selon la valeur de la trace de A, les polynômes caractéristiques[1] de A possibles sont

$$X^2 - X + 1, \quad X^2 + 1, \quad \text{et} \quad X^2 + X + 1.$$

Les spectres de A respectifs sont alors

$$\{-j, -j^2\}, \quad \{\mathrm{i}, -\mathrm{i}\}, \quad \text{et} \quad \{j, j^2\}. \tag{$**$}$$

Dans tous les cas listés dans $(*)$ et $(**)$, les valeurs propres λ vérifient[2] $\lambda^{12} = 1$. Il suffit alors de diagonaliser la matrice A dans $\mathcal{M}_2(\mathbb{C})$ pour vérifier $A^{12} = \mathrm{I}_2$.

Exercice 35 *

Soit E un espace vectoriel complexe de dimension finie $n \geqslant 1$.

(a) Montrer qu'il existe un polynôme réel P_n vérifiant $\sqrt{1+x} = P_n(x) + \mathrm{O}(x^n)$ quand le réel x tend vers 0.

(b) Établir que X^n divise alors le polynôme $P_n^2 - X - 1$.

(c) Soit f un endomorphisme de E nilpotent. Montrer qu'il existe un endomorphisme g de E vérifiant $g^2 = \mathrm{Id}_E + f$.

(d) Soit maintenant f un endomorphisme de E ne possédant qu'une seule valeur propre λ non nulle[3]. Montrer qu'il existe un endomorphisme g de E vérifiant $g^2 = f$.

1. Le polynôme caractéristique d'une matrice carrée A de taille 2 est $X^2 - \operatorname{tr}(A)X + \det(A)$.
2. La valeur 12 correspond au PPCM des ordres dans $(\mathbb{U}, \times)$ des valeurs propres possibles.
3. Lorsque $\lambda = 0$, l'équation étudiée peut ne pas avoir de solutions, voir sujet 49 p. 185.

Solution

(a) **méthode**

‖ On écrit un développement limité de $\sqrt{1+x}$ quand x tend vers 0.

La fonction $x \mapsto \sqrt{1+x}$ est de classe $\mathcal{C}^\infty$ sur l'intervalle $]-1\,;+\infty[$ contenant 0. Elle admet donc un développement limité à tout ordre en 0. En particulier, un développement limité à l'ordre n donne l'écriture [1]

$$\sqrt{1+x} \underset{x\to 0}{=} \underbrace{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}}_{=P_n(x)} + \underbrace{a_n x^n + o\big(x^n\big)}_{=O(x^n)}.$$

Ceci détermine un polynôme réel P_n de degré inférieur à $n-1$ convenable.

(b) Par élévation au carrée

$$P_n^2(x) \underset{x\to 0}{=} \big(\sqrt{1+x} + O\big(x^n\big)\big)^2 = 1 + x + \underbrace{2\sqrt{1+x} \times O\big(x^n\big) + O\big(x^n\big)^2}_{=O(x^n)}$$

et donc

$$P_n^2(x) - x - 1 \underset{x\to 0}{=} O\big(x^n\big).$$

méthode

‖ On introduit α la multiplicité de 0 en tant que racine de $P_n^2 - X - 1$.

On peut écrire $P_n^2 - X - 1 = X^\alpha Q$ avec $Q(0) \neq 0$ et donc

$$x^\alpha Q(x) \underset{x\to 0}{=} O\big(x^n\big) \quad \text{puis} \quad x^{\alpha-n}Q(x) \underset{\substack{x\to 0\\ x\neq 0}}{=} O(1).$$

Nécessairement $\alpha - n \geqslant 0$ car sinon la fonction en premier membre n'est pas bornée au voisinage de 0. On peut alors affirmer que 0 est racine de multiplicité au moins n du polynôme $P_n^2 - X - 1$ et donc que X^n divise celui-ci.

(c) Puisque f est un endomorphisme nilpotent d'un espace de dimension n, on sait que $f^n = 0$. Le polynôme X^n annule alors f et donc annule aussi $P_n^2 - X - 1$ qui est un multiple de X^n. L'endomorphisme $g = P_n(f)$ vérifie alors

$$g^2 = P_n^2(f) = f + \mathrm{Id}_E.$$

(d) Puisque E est un espace vectoriel complexe, on peut affirmer que le polynôme caractéristique de f est scindé. Or λ est sa seule racine et donc $\chi_f = (X-\lambda)^n$. En vertu du théorème de Cayley-Hamilton, on a alors

$$(f - \lambda \mathrm{Id}_E)^n = 0.$$

1. Pour la suite de l'étude, il n'est pas nécessaire d'exprimer exactement les coefficients a_k.

Introduisons ensuite un complexe [1] $\mu \in \mathbb{C}^*$ vérifiant $\mu^2 = \lambda$ et l'endomorphisme

$$g = \mu P_n(h) \quad \text{avec} \quad h = \frac{1}{\mu^2}(f - \lambda \mathrm{Id}_E) = \frac{1}{\mu^2} f - \mathrm{Id}_E.$$

Puisque l'endomorphisme h est nilpotent, les calculs de la question précédente permettent de conclure

$$g^2 = \mu^2 P_n^2(h) = \mu^2(h + \mathrm{Id}_E) = f.$$

Exercice 36 **

Soit G un sous-groupe de $\big(\mathrm{GL}_n(\mathbb{R}), \times\big)$ tel que $M^2 = \mathrm{I}_n$ pour tout $M \in G$.

(a) Montrer que le groupe G est commutatif.

(b) Établir que les éléments de G sont simultanément [2] diagonalisables.

(c) En déduire $\mathrm{Card}(G) \leqslant 2^n$.

(d) Application : Montrer que $\big(\mathrm{GL}_n(\mathbb{R}), \times\big)$ et $\big(\mathrm{GL}_m(\mathbb{R}), \times\big)$ sont isomorphes si, et seulement si, $n = m$.

Solution

(a) Pour tout élément A de G, on a $A^{-1} = A$ car $A^2 = \mathrm{I}_n$. Pour tous A et B dans G, on a $AB \in G$ donc

$$AB = (AB)^{-1} = B^{-1}A^{-1} = BA.$$

Le groupe G est commutatif.

(b) Les éléments de G annulent le polynôme $X^2 - 1 = (X - 1)(X + 1)$ qui est scindé à racines simples. Les matrices appartenant à G sont donc diagonalisables et seules 1 et -1 peuvent en être valeurs propres.

Montrons par récurrence forte sur la taille $n \geqslant 1$ des matrices l'existence d'une même matrice P inversible de taille n les diagonalisant toutes.

Pour $n = 1$, il n'y a rien à démontrer.

Supposons le résultat vrai jusqu'au rang $n - 1 \geqslant 1$. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$ dont tous les éléments M vérifient $M^2 = \mathrm{I}_n$. S'il n'existe pas d'autres éléments dans G que I_n et $-\mathrm{I}_n$, la propriété est acquise. Sinon, il existe un élément $A \in G$ différent de I_n et de $-\mathrm{I}_n$. Pour cette matrice 1 et -1 sont valeurs propres.

méthode

On traduit matriciellement que les sous-espaces propres associés aux valeurs propres 1 et -1 de A sont stables par tous les éléments de G.

1. En écrivant $\lambda = |\lambda| e^{i\theta}$, le nombre $\mu = \sqrt{|\lambda|} e^{i\theta/2}$ convient.

2. Autrement dit, il existe une matrice $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $P^{-1}MP$ est diagonale pour toute matrice M de G.

Puisque $A^2 = \mathrm{I}_n$, la matrice A est semblable à une matrice diagonale où figurent sur la diagonale ses valeurs propres 1 et -1. Il existe donc une matrice inversible P vérifiant

$$P^{-1}AP = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & -\mathrm{I}_{n-r} \end{pmatrix} \quad \text{avec} \quad r \in [\![1\,;n-1]\!].$$

Soit M un élément de G. On peut écrire par blocs

$$P^{-1}MP = \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}$$

avec M_1 et M_4 des matrices carrées de tailles respectives r et $n-r$.

Puisque A et M commutent, les matrices $P^{-1}AP$ et $P^{-1}MP$ commutent ce qui entraîne la nullité des blocs M_2 et M_3. De plus, sachant $M^2 = \mathrm{I}_n$, on a aussi $M_1^2 = \mathrm{I}_r$ et $M_4^2 = \mathrm{I}_{n-r}$. Les ensembles G' et G'' constitués respectivement des matrices M_1 et M_4 obtenues lorsque M parcourt G sont alors des sous-groupes [1] de respectivement $\mathrm{GL}_r(\mathbb{R})$ et $\mathrm{GL}_{n-r}(\mathbb{R})$, dont les éléments sont de carrés égaux à la matrice identité. Par hypothèse de récurrence, il existe Q' matrice inversible de taille r telle que $Q'^{-1}M'Q'$ est diagonale pour toute matrice M' dans G'. Il existe aussi Q'' matrice inversible de taille $n-r$ telle que l'on dispose d'une propriété analogue pour les matrices M'' de G''. Posons alors

$$Q = \begin{pmatrix} Q' & 0 \\ 0 & Q'' \end{pmatrix} \quad \text{inversible et d'inverse} \quad Q^{-1} = \begin{pmatrix} Q'^{-1} & 0 \\ 0 & Q''^{-1} \end{pmatrix}.$$

Pour tout $M \in G$, un calcul par blocs assure que $(PQ)^{-1}M(PQ)$ est diagonale : les matrices de G sont toutes diagonalisables par la matrice de passage PQ.

La récurrence est établie.

(c) Par une même matrice de passage, les matrices appartenant à G sont semblables à des matrices diagonales dont les coefficients diagonaux ne peuvent être que 1 et -1. Il n'existe que 2^n matrices de ce type dans $\mathcal{M}_n(\mathbb{R})$, on en déduit $\mathrm{Card}(G) \leqslant 2^n$.

(d) Supposons qu'il existe un isomorphisme de $\big(\mathrm{GL}_n(\mathbb{R}), \times\big)$ vers $\big(\mathrm{GL}_m(\mathbb{R}), \times\big)$. Considérons l'ensemble G formé des matrices diagonales M de $\mathcal{M}_n(\mathbb{R})$ vérifiant $M^2 = \mathrm{I}_n$. L'ensemble G est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$ de cardinal exactement 2^n. Par l'isomorphisme φ, l'ensemble $G' = \varphi(G)$ est un sous-groupe de $\mathrm{GL}_m(\mathbb{R})$ dont tous les éléments sont de carrés égaux à l'identité car, pour tout $M \in G$,

$$\varphi(M)^2 = \varphi\big(M^2\big) = \varphi(\mathrm{I}_n) = \mathrm{I}_m.$$

Par l'étude qui précède, on peut affirmer $2^n = \mathrm{Card}(G') \leqslant 2^m$. On en déduit $n \leqslant m$. Un raisonnement symétrique donne l'inégalité complémentaire et donc $n = m$.

Réciproquement, si $n = m$, les deux groupes sont évidemment isomorphes.

1. Un calcul par blocs assure que l'application $M \mapsto M_1$ est un morphisme du groupe $(G, \times)$ vers le groupe $\big(\mathrm{GL}_r(\mathbb{K}), \times\big)$ et G' est l'image de ce morphisme donc un sous-groupe. Il en est de même pour G''.

Exercice 37 ***

Soit u un endomorphisme d'un espace vectoriel complexe E de dimension $n \geqslant 1$. On note $\lambda_1, \ldots, \lambda_m$ les valeurs propres sans répétitions de u et $\alpha_1, \ldots, \alpha_m$ leurs multiplicités respectives. On suppose que les sous-espaces propres de u sont tous de dimension 1.

(a) Soit $k \in [\![1\,;m]\!]$. Montrer que, pour tout $p \in [\![1\,;\alpha_k]\!]$, le noyau de $(u - \lambda_k \mathrm{Id}_E)^p$ est de dimension p.

(b) Soit F un sous-espace vectoriel de E stable par u. Montrer qu'il existe un polynôme unitaire Q de $\mathbb{C}[X]$ tel que $F = \mathrm{Ker}\big(Q(u)\big)$.

(c) Combien l'endomorphisme u possède-t-il de sous-espaces vectoriels stables ?

Solution

(a) **méthode**

Si u et v sont deux endomorphismes de E, on sait [1]

$$\dim \mathrm{Ker}(u \circ v) \leqslant \dim \mathrm{Ker}(u) + \dim \mathrm{Ker}(v).$$

Par la propriété ci-dessus, on montre pour tout $p \in \mathbb{N}$

$$\dim \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)^{p+1} \leqslant \dim \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)^p + \underbrace{\dim \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)}_{=1}. \qquad (*)$$

Une récurrence facile donne alors

$$\dim \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)^p \leqslant p.$$

Le polynôme caractéristique de u est scindé sur $\mathbb{C}$ et s'écrit

$$\chi_u = \prod_{k=1}^{m} (X - \lambda_k)^{\alpha_k}.$$

Par le théorème de Cayley-Hamilton, ce polynôme est annulateur de u. Les facteurs $(X - \lambda_k)^{\alpha_k}$ étant deux à deux premiers entre eux, on peut appliquer le lemme des noyaux et écrire

$$E = \bigoplus_{k=1}^{m} \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k}.$$

On a donc

$$\dim E = \sum_{k=1}^{m} \underbrace{\dim \mathrm{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k}}_{\leqslant \alpha_k}$$

mais aussi

$$\dim E = \deg(\chi_u) = \sum_{k=1}^{m} \alpha_k.$$

1. Voir sujet 15 p. 85.

On en déduit [1] que, pour tout $k \in [\![1\,;m]\!]$,

$$\dim \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k} = \alpha_k.$$

De plus, la comparaison $(*)$ entraîne alors

$$\dim \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{p} = p \quad \text{pour tout } p \in [\![1\,;\alpha_k]\!].$$

(b) Soit F un sous-espace vectoriel stable par u. On peut introduire u' l'endomorphisme induit par u sur F. Le polynôme caractéristique Q de u' divise χ_u et peut donc s'écrire

$$Q = \prod_{k=1}^{m} (X - \lambda_k)^{\beta_k} \quad \text{avec} \quad \beta_k \in [\![0\,;\alpha_k]\!] \text{ pour tout } k \in [\![1\,;m]\!].$$

De plus, le théorème de Cayley-Hamilton assure que le polynôme Q annule u' et donc

$$F \subset \operatorname{Ker}\big(Q(u)\big).$$

Or le lemme de décomposition des noyaux donne

$$\operatorname{Ker}\big(Q(u)\big) = \bigoplus_{k=1}^{m} \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\beta_k}$$

et l'étude de la question précédente assure alors

$$\dim \operatorname{Ker}\big(Q(u)\big) = \sum_{k=1}^{m} \beta_k = \deg(Q) = \dim F.$$

Par inclusion et égalité des dimensions, on conclut $\operatorname{Ker}\big(Q(u)\big) = F$.

(c) L'étude qui précède assure que les sous-espaces stables par u sont de la forme

$$F = \bigoplus_{k=1}^{m} \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\beta_k}$$

avec, pour tout $k \in [\![1\,;m]\!]$, $\beta_k \in [\![0\,;\alpha_k]\!]$ qui s'identifie à la multiplicité de λ_k pour l'endomorphisme induit par u sur F. Inversement, un tel espace est stable par u et l'on peut définir une correspondance bijective entre les sous-espaces vectoriels F stables par u et les éléments

$$(\beta_1, \ldots, \beta_m) \in [\![0\,;\alpha_1]\!] \times \cdots \times [\![0\,;\alpha_m]\!].$$

Il y a donc exactement $(\alpha_1 + 1) \ldots (\alpha_m + 1)$ sous-espaces vectoriels stables par u.

1. Cette propriété est aussi établie de façon générale dans le sujet 27 p. 224.

Exercice 38 * (Décomposition de Dunford)**

Soit u un endomorphisme d'un espace vectoriel E de dimension finie non nulle dont le polynôme caractéristique est scindé. On souhaite établir l'existence et l'unicité d'un couple (d,n) d'endomorphismes de E avec d diagonalisable et n nilpotent vérifiant

$$u = d + n \quad \text{et} \quad d \circ n = n \circ d.$$

On note $\lambda_1, \ldots, \lambda_m$ les valeurs propres sans répétitions de u et $\alpha_1, \ldots, \alpha_m$ leurs multiplicités respectives.

(a) Justifier

$$E = \bigoplus_{k=1}^{m} \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k}.$$

Établir que les projecteurs[1] p_k associés à cette écriture sont des polynômes en u.

(b) On pose $d = \lambda_1 p_1 + \cdots + \lambda_m p_m$ et $n = u - d$. Vérifier que le couple (d,n) est solution du problème posé.

(c) Montrer que c'est le seul couple possible.

Solution

(a) Le polynôme caractéristique de u s'écrit

$$\chi_u = \prod_{k=1}^{m} (X - \lambda_k)^{\alpha_k}.$$

Par le théorème de Cayley-Hamilton, χ_u est annulateur de u et les facteurs $(X - \lambda_k)^{\alpha_k}$ étant deux à deux premiers entre eux, on peut appliquer le lemme des noyaux pour écrire la décomposition en somme directe

$$E = \bigoplus_{k=1}^{m} F_k \quad \text{avec} \quad F_k = \operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k}.$$

Étudions la projection p_k sur l'espace[2] F_k parallèlement à l'espace G_k égal à la somme directe des F_j pour $j \neq k$.

$$E = F_k \oplus G_k \quad \text{avec} \quad G_k = \bigoplus_{\substack{1 \leqslant j \leqslant m \\ j \neq k}} \operatorname{Ker}(u - \lambda_j \mathrm{Id}_E)^{\alpha_j}.$$

Par le lemme de décomposition des noyaux, on a aussi

$$G_k = \operatorname{Ker}\big(Q(u)\big) \quad \text{avec} \quad Q = \prod_{\substack{1 \leqslant j \leqslant m \\ j \neq k}} (X - \lambda_j)^{\alpha_j}.$$

1. L'endomorphisme p_k est la projection sur $\operatorname{Ker}(u - \lambda_k \mathrm{Id}_E)^{\alpha_k}$ parallèlement à la somme directe des autres noyaux.

2. Les espaces F_k correspondent aux sous-espaces caractéristiques de u.

méthode

On exprime que Q et $(X-\lambda_k)^{\alpha_k}$ sont premiers entre eux par une relation de Bézout.

Puisque λ_k n'est pas racine de Q, les polynômes Q et $(X-\lambda_k)^{\alpha_k}$ sont premiers entre eux. On peut alors introduire deux polynômes V et W tels que

$$1 = VQ + W(X-\lambda_k)^{\alpha_k}.$$

En évaluant cette identité en u, on obtient

$$\mathrm{Id}_E = \underbrace{V(u)\circ Q(u)}_{=p} + \underbrace{W(u)\circ(u-\lambda_k\mathrm{Id}_E)^{\alpha_k}}_{=q}.$$

Pour $x \in E$, posons $a = p(x)$ et $b = q(x)$ de sorte que $x = a+b$. Les vecteurs a et b sont respectivement éléments de F_k et de G_k car

$$(u-\lambda_k.\mathrm{Id}_E)^{\alpha_k}(a) = \big(V(u)\circ\underbrace{\chi_u(u)}_{=0}\big)(x) = 0_E \quad \text{et} \quad Q(u)(b) = \big(W(u)\circ\underbrace{\chi_u(u)}_{=0}\big)(x) = 0_E.$$

Ainsi, p est la projection sur F_k parallèlement à G_k, c'est donc l'endomorphisme p_k. Enfin, par construction, celui-ci se révèle être un polynôme en u.

(b) L'endomorphisme d est diagonalisable car toute base de E adaptée à la décomposition $E = F_1 \oplus \cdots \oplus F_m$ est formée de vecteurs propres [1] de d. L'endomorphisme d est un polynôme en u car combinaison linéaire de polynômes en u. Par conséquent, d commute avec u et aussi avec $n = u-d$ qui encore un polynôme en u. On a évidemment $u = n+d$ et il ne reste plus qu'à vérifier que l'endomorphisme n est nilpotent.

Soit $k \in [\![1\,;m]\!]$. L'espace F_k est stable par n car c'est le noyau d'un endomorphisme qui commute avec n. Or pour tout $x \in F_k$

$$n(x) = u(x) - d(x) = u(x) - \lambda_k x = (u-\lambda_k\mathrm{Id}_E)(x)$$

et donc

$$\begin{aligned} n^{\alpha_k}(x) &= n^{\alpha_k-1}\big(\underbrace{(u-\lambda_k\mathrm{Id}_E)(x)}_{\in F_k}\big) = n^{\alpha_k-2}\big(\underbrace{(u-\lambda_k\mathrm{Id}_E)^2(x)}_{\in F_k}\big) \\ &= \cdots = (u-\lambda_k\mathrm{Id}_E)^{\alpha_k}(x) = 0_E. \end{aligned}$$

En posant $\alpha = \max(\alpha_1,\ldots,\alpha_m)$, on a alors $n^\alpha(x) = 0_E$ pour tous les vecteurs x des espaces F_k et donc, par linéarité, pour tout $x \in E$: l'endomorphisme n est nilpotent.

(c) Soit (d', n') un autre couple solution du problème posé.

méthode

On montre que l'endomorphisme $d-d'$ est diagonalisable et nilpotent.

1. Les espaces F_k correspondent aux sous-espaces propres de d.

L'endomorphisme d' commute avec n' donc avec $u = d' + n'$ et encore avec d qui est un polynôme en u. Or les endomorphismes d et d' sont diagonalisables, ils sont donc simultanément diagonalisables [1]. En particulier, $d - d'$ est diagonalisable.

Parallèlement, l'endomorphisme n' commute avec u et donc aussi avec n qui est un polynôme en u. Or ces deux endomorphismes sont nilpotents. En posant α et α' les indices de nilpotence de n et n', la formule du binôme donne

$$\begin{aligned}(n'-n)^{\alpha+\alpha'} &= \sum_{k=0}^{\alpha+\alpha'} (-1)^k \binom{\alpha+\alpha'}{k} n'^{\alpha+\alpha'-k} n^k \\ &= \sum_{k=0}^{\alpha} (-1)^k \binom{\alpha+\alpha'}{k} \underbrace{n'^{\alpha+\alpha'-k}}_{=0} n^k + \sum_{k=\alpha+1}^{\alpha+\alpha'} (-1)^k \binom{\alpha+\alpha'}{k} n'^{\alpha+\alpha'-k} \underbrace{n^k}_{=0} \\ &= 0.\end{aligned}$$

Ainsi, $n' - n$ est nilpotent.

Finalement, l'endomorphisme $d - d'$, qui est aussi $n' - n$, est à la fois diagonalisable et nilpotent : c'est l'endomorphisme nul. On conclut l'unicité $(d', n') = (d, n)$.

1. Voir sujet 31 p. 230.

CHAPITRE 6

Compléments sur les espaces préhilbertiens

E désigne un espace vectoriel réel quelconque et n un entier naturel non nul.

6.1 Quelques rappels

Un *produit scalaire*[1] sur un espace vectoriel réel E est une forme bilinéaire symétrique définie positive communément notée $(\cdot \mid \cdot)$ ou $\langle \cdot, \cdot \rangle$.

Avec des notations entendues, le produit scalaire canonique sur $\mathbb{R}^n$ est défini par

$$(x \mid y) = x_1 y_1 + \cdots + x_n y_n$$

et le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ par la relation

$$\langle X, Y \rangle = {}^tXY = x_1 y_1 + \cdots + x_n y_n.$$

Plus généralement, on définit le produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathbb{R})$ en posant[2]

$$\langle A, B \rangle = \operatorname{tr}({}^tAB) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{i,j} b_{i,j}$$

1. On renvoie le lecteur à l'ouvrage *Exercices d'algèbre et de probabilités MPSI* pour le détails des notions qui suivent.

2. Voir sujet 2 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI.*

Lorsque l'on convient qu'un espace E est muni d'un produit scalaire $(\cdot \mid \cdot)$, on dit qu'il s'agit d'un *espace préhilbertien* réel. Celui-ci est alors normé par la *norme euclidienne* donnée par

$$\|x\| = \sqrt{(x \mid x)} \quad \text{pour tout } x \in E.$$

Rappelons l'inégalité de Cauchy-Schwarz :

$$\left|(a \mid b)\right| \leqslant \|a\| \, \|b\| \quad \text{pour tous } a \text{ et } b \text{ vecteurs de } E$$

avec égalité si, et seulement si, a et b sont colinéaires.

Des vecteurs x et y d'un espace préhilbertien réel sont dits *orthogonaux* lorsque $(x \mid y) = 0$. Une famille de vecteurs est qualifiée d'*orthogonale* si elle est constituée de vecteurs deux à deux orthogonaux. Elle est dite *orthonormale* si les vecteurs sont de plus unitaires. Une telle famille est assurément libre [1].

Le théorème de Pythagore affirme que si $e = (e_1, \ldots, e_n)$ est une famille orthogonale de vecteurs de E,

$$\left\| \sum_{i=1}^{n} e_i \right\|^2 = \sum_{i=1}^{n} \|e_i\|^2 .$$

Si A désigne une partie d'un espace préhilbertien E, l'espace $A^\perp$ réunit les vecteurs de E qui sont orthogonaux à tous les vecteurs de A :

$$A^\perp = \big\{ x \in E \,\big|\, \forall a \in A,\ (a \mid x) = 0 \big\}.$$

Un *espace euclidien* est un espace préhilbertien de dimension finie. Un tel espace possède une base orthonormale. Les coordonnées $x_1, \ldots, x_n$ d'un vecteur x dans une base orthonormale $(e_1, \ldots, e_n)$ sont données par

$$x_k = (e_k \mid x) \quad \text{pour tout } k \in [\![1\,;n]\!].$$

Si $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ sont les coordonnées de vecteurs x et y dans une base orthonormale, on a

$$(x \mid y) = x_1 y_1 + \cdots + x_n y_n \quad \text{et} \quad \|x\|^2 = x_1^2 + \cdots + x_n^2.$$

Si X et Y sont les colonnes des coordonnées des vecteurs x et y, on a aussi

$$(x \mid y) = {}^t XY \quad \text{et} \quad \|x\|^2 = {}^t XX.$$

6.2 Compléments

6.2.1 Représentation d'une forme linéaire

Soit E un espace euclidien. Pour tout vecteur a de E, l'application $\varphi_a : E \to \mathbb{R}$ définie par $\varphi_a(x) = (a \mid x)$ est une forme linéaire sur E. Le résultat qui suit assure que les formes linéaires sur E sont toutes de ce type :

1. Plus généralement, une famille orthogonale ne comportant pas le vecteur nul est libre.

Théorème 1
Si φ est une forme linéaire sur un espace euclidien E, il existe un unique vecteur a de E vérifiant $\varphi = \varphi_a$, c'est-à-dire $\varphi(x) = (a|x)$ pour tout $x \in E$.

Si $e = (e_1, \ldots, e_n)$ désigne une base orthonormale de E, le vecteur a est déterminé par

$$a = \sum_{k=1}^{n} (e_k \,|\, a) e_k = \sum_{k=1}^{n} \varphi(e_k) e_k.$$

Lorsque φ n'est pas nulle, a est vecteur normal à l'hyperplan $H = \mathrm{Ker}(\varphi)$.

6.2.2 Somme directe orthogonale

Soit E un espace préhilbertien.

Définition
On dit que deux sous-espaces F et G de E sont *orthogonaux* lorsqu'ils sont formés de vecteurs deux à deux orthogonaux, c'est-à-dire si $(x \,|\, y) = 0$ pour tout $x \in F$ et tout $x \in G$.

L'orthogonalité des sous-espaces vectoriels F et G signifie l'inclusion[1] $F \subset G^{\perp}$.

Théorème 2
Si $F_1, \ldots, F_m$ sont des sous-espaces vectoriels de E deux à deux orthogonaux, ils sont en somme directe.

Définition
Lorsque les sous-espaces vectoriels $F_1, \ldots, F_m$ sont deux à deux orthogonaux, on dit que leur somme $F_1 + \cdots + F_m$ est *directe orthogonale*. Celle-ci peut être notée

$$F_1 \oplus^{\perp} \cdots \oplus^{\perp} F_m.$$

Si F est un sous-espace vectoriel d'un espace euclidien E, $F^{\perp}$ est un supplémentaire de F dans E. En particulier,

$$\dim F^{\perp} = \dim E - \dim F.$$

Plus généralement :

Théorème 3
Si F est un sous-espace vectoriel de dimension finie d'un espace préhilbertien E, l'espace $F^{\perp}$ est un supplémentaire de F dans E.

On dit alors que $F^{\perp}$ est *le supplémentaire orthogonal* de F.

1. Ou, et c'est équivalent, $G \subset F^{\perp}$.

6.2.3 Projection orthogonale sur un sous-espace de dimension finie

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E de dimension quelconque [1]. On peut écrire $F \oplus^{\perp} F^{\perp} = E$.

Définition

On appelle *projection orthogonale* sur F la projection p_F sur F parallèlement à $F^{\perp}$. Si x est un vecteur de E, $p_F(x)$ se nomme le *projeté orthogonal* de x sur F.

Théorème 4

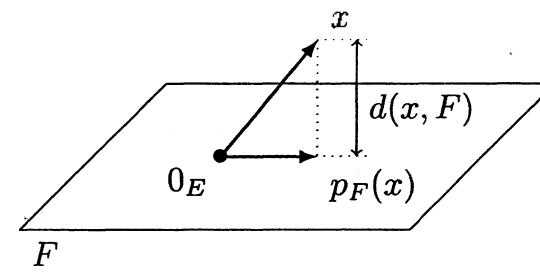
Soit x un vecteur de E. Pour tout vecteur y de F

$$\|x - y\| \geqslant \|x - p_F(x)\|$$

avec égalité si, et seulement si, $y = p_F(x)$.

En particulier,

$$d(x, F) \overset{\text{déf}}{=} \inf_{y \in F} \|x - y\| = \|x - p_F(x)\|.$$



La détermination du projeté orthogonal d'un vecteur permet de calculer sa distance au sous-espace vectoriel.

Théorème 5

Si $(e_1, \ldots, e_r)$ est une base orthonormale de F,

$$p_F(x) = \sum_{k=1}^{r} (e_k \,|\, x) e_k \quad \text{pour tout } x \in E.$$

Le théorème de Pythagore donne alors :

Théorème 6 (Inégalité de Bessel)

Si $(e_1, \ldots, e_r)$ est une famille orthonormale de vecteurs de E,

$$\sum_{k=1}^{r} (e_k \,|\, x)^2 \leqslant \|x\|^2 \quad \text{pour tout } x \in E.$$

Si $(e_n)_{n \in \mathbb{N}}$ est une famille orthonormale de vecteurs d'un espace préhilbertien E de dimension infinie alors, pour tout $x \in E$, la série numérique $\sum (e_n \,|\, x)^2$ converge et

$$\sum_{n=0}^{+\infty} (e_n \,|\, x)^2 \leqslant \|x\|^2 .$$

1. La notion de projection orthogonale a déjà été présentée en première année dans les espaces euclidiens. On l'étend ici aux espaces préhilbertiens de dimensions infinies.

La suite $\big((e_n \,|\, x)\big)_{n\in\mathbb{N}}$ est donc de carré sommable.

6.2.4 Suite orthonormale totale de vecteurs

Soit E un espace préhilbertien de dimension infinie.

Définition

On dit qu'une suite $(e_n)_{n\in\mathbb{N}}$ de vecteurs de E est *totale* si l'espace vectoriel qu'elle engendre est une partie dense de E, autrement dit, si

$$\overline{\mathrm{Vect}\big\{e_n \,\big|\, n\in\mathbb{N}\big\}} = E.$$

Lorsque l'on introduit les espaces $F_n = \mathrm{Vect}(e_0,\dots,e_n)$, on a alors, pour tout x de E,

$$d(x,F_n) \xrightarrow[n\to+\infty]{} 0.$$

Théorème 7

Si $(e_n)_{n\in\mathbb{N}}$ une suite totale d'éléments de E alors, pour tout $x\in E$,

$$p_n(x) \xrightarrow[n\to+\infty]{} x \quad \text{c'est-à-dire} \quad \big\|p_n(x)-x\big\| \xrightarrow[n\to+\infty]{} 0$$

avec p_n la projection orthogonale sur l'espace $F_n = \mathrm{Vect}(e_0,\dots,e_n)$.

Si $(e_n)_{n\in\mathbb{N}}$ est une suite orthonormale totale d'éléments de E alors [1], pour tout $x\in E$,

$$x = \sum_{n=0}^{+\infty} (e_n \,|\, x)e_n.$$

6.3 Exercices d'apprentissage

Exercice 1

On note $E = \mathbb{R}[X]$ et l'on considère l'application $\varphi\colon E\times E\to\mathbb{R}$ donnée par

$$\varphi(P,Q) = \int_0^{+\infty} P(t)Q(t)\mathrm{e}^{-t}\,\mathrm{d}t.$$

(a) Montrer que φ définit un produit scalaire sur E.

(b) Pour $p,q\in\mathbb{N}$, calculer $\varphi\big(X^p,X^q\big)$.

(c) Orthonormaliser par le procédé de Schmidt la famille $\big(1,X,X^2\big)$.

(d) Calculer

$$\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^{+\infty} \big(t^3-(at^2+bt+c)\big)^2\,\mathrm{d}t.$$

1. La convergence de la série a lieu pour la norme euclidienne.

Solution

(a) **méthode**

On vérifie que l'application φ est bien définie en constatant la convergence de l'intégrale donnant $\varphi(P,Q)$.

Soit $P,Q \in E$. La fonction $f\colon t \mapsto P(t)Q(t)\mathrm{e}^{-t}$ est définie et continue par morceaux sur $[0\,;+\infty[$. Elle est négligeable devant $1/t^2$ quand t tend vers $+\infty$ car

$$t^2 f(t) = t^2 P(t)Q(t)\mathrm{e}^{-t} \xrightarrow[t\to+\infty]{} 0$$

puisqu'une fonction polynomiale est négligeable devant $t \mapsto \mathrm{e}^{-t}$ en $+\infty$. La fonction f est donc intégrable sur $[0\,;+\infty[$ et l'intégrale définissant $\varphi(P,Q)$ est convergente. L'application φ est donc bien définie de $E \times E$ vers $\mathbb{R}$.

méthode

On vérifie que la forme φ est bilinéaire, symétrique et définie positive.

Soit $\lambda,\mu \in \mathbb{R}$ et $P,Q,R \in E$.

On vérifie sans peine la symétrie $\varphi(P,Q) = \varphi(Q,P)$. Avec convergence de chacune des intégrales écrites, on a aussi

$$\int_0^{+\infty} P(t)\big(\lambda Q(t) + \mu R(t)\big)\,\mathrm{d}t = \lambda \int_0^{+\infty} P(t)Q(t)\,\mathrm{d}t + \mu \int_0^{+\infty} P(t)R(t)\,\mathrm{d}t$$

et donc $\varphi(P,\lambda Q + \mu R) = \lambda\varphi(P,Q) + \mu\varphi(P,R)$. On en déduit que φ est linéaire en sa deuxième variable et donc bilinéaire par symétrie.

Il reste à montrer qu'elle est définie positive. Par positivité de l'intégrale, on a

$$\varphi(P,P) = \int_0^{+\infty} \underbrace{P(t)^2\mathrm{e}^{-t}}_{\geqslant 0}\,\mathrm{d}t \geqslant 0.$$

De plus, si $\varphi(P,P) = 0$, on obtient une intégrale nulle d'une fonction continue et positive qui est donc la fonction nulle : $P(t)^2\mathrm{e}^{-t} = 0$ pour tout $t \in [0\,;+\infty[$. On en déduit que le polynôme P admet une infinité de racines, c'est le polynôme nul.

Finalement, φ est un produit scalaire sur E.

(b) Pour $n \in \mathbb{N}$, posons

$$I_n = \int_0^{+\infty} t^n\mathrm{e}^{-t}\,\mathrm{d}t$$

de sorte que $\varphi\big(X^p,X^q\big) = I_{p+q}$. Par une intégration par parties généralisée[1], on a pour tout $n \geqslant 1$

$$\int_0^{+\infty} t^n\mathrm{e}^{-t}\,\mathrm{d}t = \Big[-t^n\mathrm{e}^{-t}\Big]_0^{+\infty} + n\int_0^{+\infty} t^{n-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

1. Les calculs sont détaillés dans le sujet 17 du chapitre 2 de l'ouvrage *Exercices d'analyse MP*.

Ainsi, $I_n = nI_{n-1}$ et, sachant $I_0 = 1$, on conclut $I_n = n!$ puis

$$\varphi(X^p, X^q) = (p+q)!$$

(c) La famille $(1, X, X^2)$ est libre, il est donc licite de l'orthonormaliser par le procédé de Schmidt [1].

On pose $P_0 = 1$ puis $P_1 = X + \lambda P_0$ avec λ tel que $(P_0 \mid P_1) = 0$. On résout l'équation $1 + \lambda = 0$ et donc $P_1 = X - 1$. On pose ensuite $P_2 = X^2 + \lambda P_0 + \mu P_1$ avec λ et μ tels que $(P_0 \mid P_2) = 0$ et $(P_1 \mid P_2) = 0$. On résout alors les deux équations $2 + \lambda = 0$ et $4 + \mu = 0$ ce qui donne $P_2 = X^2 - 4X + 2$. Enfin, on divise chaque polynôme par sa norme [2] pour former la famille orthonormale (Q_0, Q_1, Q_2) cherchée :

$$Q_0 = 1, \quad Q_1 = X - 1 \quad \text{et} \quad Q_2 = \frac{1}{2}(X^2 - 4X + 2).$$

(d) **méthode**

> La borne inférieure cherchée est liée à la distance de X^3 à l'espace $\mathbb{R}_2[X]$ pour la norme euclidienne associée au produit scalaire φ.

Pour $(a, b, c) \in \mathbb{R}^3$, on remarque

$$\int_0^{+\infty} \left(t^3 - (at^2 + bt + c)\right)^2 \mathrm{d}t = \left\| X^3 - (aX^2 + bX + c) \right\|^2$$

et donc [3]

$$\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^{+\infty} \left(t^3 - (at^2 + bt + c)\right)^2 \mathrm{d}t = \inf_{P\in\mathbb{R}_2[X]} \left\| X^3 - P \right\|^2 = d(X^3, \mathbb{R}_2[X])^2.$$

Cette distance se calcule à partir du projeté orthogonal de X^3 sur $F = \mathbb{R}_2[X]$.

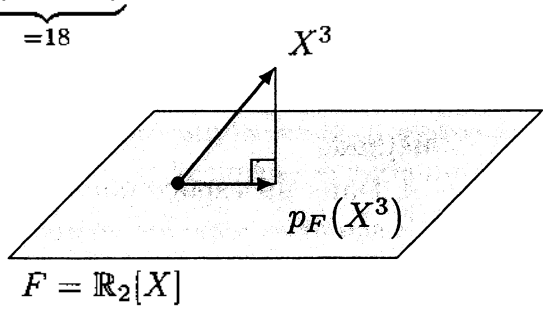
$$d(X^3, F) = \left\| X^3 - p_F(X^3) \right\|.$$

On peut calculer ce projeté à l'aide de la base orthonormale (Q_0, Q_1, Q_2) de F (Th. 5 p. 246) :

$$p_F(X^3) = \underbrace{\varphi(Q_0, X^3)}_{=6} Q_0 + \underbrace{\varphi(Q_1, X^3)}_{=18} Q_1 + \underbrace{\varphi(Q_2, X^3)}_{=18} Q_2 = 9X^2 - 18X + 6.$$

Il est inutilement fastidieux de calculer directement la norme de $X^3 - p_F(X^3)$... Par le théorème de Pythagore, on peut écrire :

$$\left\| X^3 \right\|^2 = d(X^3, F)^2 + \left\| p_F(X^3) \right\|^2.$$



1. La démarche est présentée dans le sujet 3 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

2. Pour calculer la norme de Q_2, il est intéressant de remarquer $(Q_2 \mid Q_2) = (Q_2 \mid X^2)$ ce qui allège considérablement les calculs.

3. L'élévation au carrée est une bijection croissante sur $\mathbb{R}_+$: dans la deuxième égalité, la borne inférieure du carré est le carré de la borne inférieure car les quantités engagées sont positives.

avec

$$\left\|p_F\left(X^3\right)\right\|^2 = 6^2 + 18^2 + 18^2 = 684$$

car 6, 18 et 18 sont les coordonnées du projeté dans une base orthonormale.

Finalement,

$$\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^{+\infty} \left(t^3 - \left(at^2 + bt + c\right)\right)^2 \mathrm{d}t = \left\|X^3\right\|^2 - 684 = 36.$$

Exercice 2

Soit f un endomorphisme d'un espace euclidien E muni d'une base orthonormale $e = (e_1, \ldots, e_n)$. Montrer

$$\mathrm{tr}(f) = \sum_{k=1}^{n} (e_k \,|\, f(e_k)).$$

Solution

méthode

Les coordonnées $x_1, \ldots, x_n$ d'un vecteur x de E dans unc base orthonormale $e = (e_1, \ldots, e_n)$ sont données par $x_k = (e_k \,|\, x)$.

La trace de f est égale à la trace de sa matrice représentative $A = (a_{i,j})$ dans la base e. La j-ème colonne de la matrice A est constituée des coordonnées dans e du vecteur $f(e_j)$. Le coefficient $a_{i,j}$ correspond donc à la i-ème coordonnée de $f(e_j)$. On l'obtient par le calcul $a_{i,j} = (e_i \,|\, f(e_j))$. On en déduit

$$\mathrm{tr}(f) = \mathrm{tr}(A) = \sum_{k=1}^{n} a_{k,k} = \sum_{k=1}^{n} (e_k \,|\, f(e_k)).$$

Exercice 3

Soit φ une forme linéaire sur $E = \mathcal{M}_n(\mathbb{R})$. Montrer qu'il existe une matrice A dans $\mathcal{M}_n(\mathbb{R})$ vérifiant $\varphi(M) = \mathrm{tr}(AM)$ pour tout $M \in \mathcal{M}_n(\mathbb{R})$.

Solution

méthode

Dans un espace euclidien E, les formes linéaires correspondent aux produits scalaires avec les vecteurs de E (Th. 1 p. 245).

On introduit le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ donné par

$$\langle A, B\rangle = \mathrm{tr}\left({}^tAB\right).$$

Puisque φ est une forme linéaire sur l'espace euclidien $\mathcal{M}_n(\mathbb{R})$, il existe $A' \in \mathcal{M}_n(\mathbb{R})$ telle que

$$\varphi(M) = \langle A', M\rangle \quad \text{pour tout } M \in \mathcal{M}_n(\mathbb{R}).$$

En posant $A = {}^tA'$, on obtient [1]

$$\varphi(M) = \operatorname{tr}(AM) \quad \text{pour tout } M \in \mathcal{M}_n(\mathbb{R}).$$

Exercice 4

On munit l'espace $E = \mathcal{C}\big([a\,;b],\mathbb{R}\big)$ du produit scalaire

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,\mathrm{d}t.$$

Pour $n \in \mathbb{N}$, on note f_n la fonction de E définie par $f_n(t) = t^n$ et $\mathcal{P}$ l'ensemble des fonctions polynomiales sur $[a\,;b]$.

(a) Justifier que la famille $(p_n)_{n\in\mathbb{N}}$ est totale.

(b) Déterminer l'orthogonal de $\mathcal{P}$.

Solution

(a) L'espace vectoriel engendré par les fonctions f_n est $\mathcal{P}$. Il s'agit donc d'établir que l'espace $\mathcal{P} = \operatorname{Vect}(p_n)_{n\in\mathbb{N}}$ est dense dans E pour la norme euclidienne.

méthode

Par le théorème de Weierstrass [2], toute fonction continue sur un segment peut être uniformément approchée par une fonction polynôme.

Soit f une fonction élément de E et $\varepsilon > 0$. Il existe une fonction polynomiale $\varphi \in \mathcal{P}$ vérifiant $\big|f(t) - \varphi(t)\big| \leqslant \varepsilon$ pour tout $t \in [a\,;b]$. On a alors

$$\|f-\varphi\| = \left(\int_a^b \underbrace{\big(f(t)-\varphi(t)\big)^2}_{\leqslant \varepsilon^2}\,\mathrm{d}t\right)^{1/2} \leqslant \sqrt{b-a}\,\varepsilon.$$

Par conséquent, la partie $\mathcal{P}$ est dense [3] dans E et l'on peut affirmer que la famille $(f_n)_{n\in\mathbb{N}}$ est totale.

(b) Soit f une fonction de l'orthogonal de $\mathcal{P}$. Par la densité qui précède, il existe une suite (φ_n) de fonctions polynomiales qui converge vers f pour la norme euclidienne. On a alors

$$\|f\|^2 = (f\,|\,f) = (f\,|\,f-\varphi_n) + \underbrace{(f\,|\,\varphi_n)}_{=0} \quad \text{car} \quad f \in \mathcal{P}^{\perp}.$$

1. On peut aussi justifier l'existence de la matrice A de façon plus élémentaire comme dans le sujet 32 du chapitre 9 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

2. Voir Th 7 du chapitre 7 de l'ouvrage *Exercices d'analyse MP*.

3. La norme euclidienne est dominée par la norme uniforme sur $[a\,;b]$ et le théorème de Weierstrass assure que la partie $\mathcal{P}$ est dense dans E pour la norme uniforme : elle l'est donc aussi pour la norme euclidienne.

Par l'inégalité de Cauchy-Schwarz

$$\|f\|^2 \leqslant \|f\| \underbrace{\|f-\varphi_n\|}_{\to 0} \xrightarrow[n\to+\infty]{} 0 \quad \text{donc} \quad f=0.$$

Ainsi [1], seule la fonction nulle appartient à $\mathcal{P}^\perp$ et l'on peut conclure [2] que $\mathcal{P}^\perp = \{0\}$.

6.4 Exercices d'entraînement

6.4.1 Généralités sur les espaces préhilbertiens

Exercice 5 *

Soit E un espace préhilbertien réel et $f, g\colon E \to E$ deux applications vérifiant

$$\langle f(x), y\rangle = \langle x, g(y)\rangle \quad \text{pour tout } (x,y) \in E^2.$$

Montrer que les applications f et g sont linéaires.

Solution

méthode

On peut montrer que deux vecteurs x et y de E sont égaux en observant [3]

$$\langle x, z\rangle = \langle y, z\rangle \quad \text{pour tout } z \in E.$$

Soit $\lambda, \mu \in \mathbb{R}$ et $x, y \in E$. Pour tout vecteur $z \in E$, l'hypothèse d'étude donne

$$\langle f(\lambda x + \mu y), z\rangle = \langle \lambda x + \mu y, g(z)\rangle.$$

On développe le second membre par linéarité du produit scalaire en la première variable

$$\langle f(\lambda x + \mu y), z\rangle = \lambda\langle x, g(z)\rangle + \mu\langle y, g(z)\rangle.$$

On poursuit le calcul en exploitant à nouveau l'hypothèse

$$\langle f(\lambda x + \mu y), z\rangle = \lambda\langle f(x), z\rangle + \mu\langle f(y), z\rangle = \langle \lambda f(x) + \mu f(y), z\rangle.$$

Par différence de membres, on obtient

$$\big\langle f(\lambda x + \mu y) - \big(\lambda f(x) + \mu f(y)\big), z\big\rangle = 0.$$

Le vecteur $f(\lambda x + \mu y) - \big(\lambda f(x) + \mu f(y)\big)$ est orthogonal à tout vecteur de E, il est donc nul. Ainsi, on obtient que l'application f est linéaire. Le raisonnement est identique pour g.

1. On peut comparer cette étude à celle menée avec la norme uniforme dans le sujet 12 du chapitre 7 de l'ouvrage *Exercices d'analyse MP*.

2. En particulier, $(\mathcal{P}^\perp)^\perp \neq \mathcal{P}$. Pour F sous-espace vectoriel, la formule $(F^\perp)^\perp = F$ est valable dans un espace euclidien, ou plus généralement lorsque F est de dimension finie. Elle n'est pas vraie en général.

3. Il ne s'agit pas de simplifier directement par z mais d'exploiter que $x - y$ est orthogonal à tout vecteur de E.

Exercice 6 *

Soit A une partie d'un espace préhilbertien E. Montrer

$$A^\perp = \left((A^\perp)^\perp\right)^\perp.$$

Solution

méthode

Pour A et B deux parties de E, on sait

$$A \subset (A^\perp)^\perp \quad \text{et} \quad A \subset B \implies B^\perp \subset A^\perp.$$

Par la première propriété utilisée avec $A^\perp$ au lieu de A, on obtient une première inclusion $A^\perp \subset \left((A^\perp)^\perp\right)^\perp$.

Par la deuxième propriété utilisée avec $(A^\perp)^\perp$ au lieu de B, on obtient l'inclusion réciproque $\left((A^\perp)^\perp\right)^\perp \subset A^\perp$.

Par double inclusion, on peut affirmer l'égalité.

Exercice 7 *

Soit A une partie d'un espace préhilbertien E.

(a) Montrer que l'orthogonal de A est une partie fermée.

(b) Montrer que A et $\overline{A}$ ont le même orthogonal.

Solution

(a) **méthode**

On vérifie que $A^\perp$ contient les limites de ses suites convergentes.

Considérons (x_n) une suite d'éléments de $A^\perp$ convergeant vers un vecteur x de E. Montrons que x appartient à $A^\perp$. Soit a un élément de A. Pour tout $n \in \mathbb{N}$, on sait $(a\,|\,x_n) = 0$. Vérifions que ceci entraîne $(a\,|\,x) = 0$. Par l'inégalité de Cauchy-Schwarz [1], on a

$$\left|(a\,|\,x) - (a\,|\,x_n)\right| = \left|(a\,|\,x - x_n)\right| \leqslant \|a\| \underbrace{\|x - x_n\|}_{\to 0} \xrightarrow[n\to+\infty]{} 0$$

et donc

$$(a\,|\,x) = \lim_{n\to+\infty} (a\,|\,x_n) = 0.$$

Ainsi, x est orthogonal à tout élément de A et donc $x \in A^\perp$. La partie $A^\perp$ contient les limites de ses suites convergentes, c'est une partie fermée.

(b) On sait $A \subset \overline{A}$ et l'on a donc une première inclusion $\overline{A}^\perp \subset A^\perp$.

1. On peut montrer que le produit scalaire est une application continue lorsque E est muni de la norme euclidienne car $|(x\,|\,y) - (x_0\,|\,y_0)| \leqslant \|x - x_0\|\,\|y\| + \|x_0\|\,\|y - y_0\|$.

Inversement, soit a un élément de $A^\perp$ et x un élément de $\overline{A}$. Il existe une suite (x_n) d'éléments de A convergeant vers x. Pour tout $n \in \mathbb{N}$, on a $(a \mid x_n) = 0$ et, comme au-dessus, on obtient à la limite $(a \mid x) = 0$. Ainsi, a est orthogonal à tout élément de $\overline{A}$ et l'on peut affirmer la seconde inclusion $A^\perp \subset \overline{A}^\perp$.

Finalement [1], $A^\perp = \overline{A}^\perp$.

Exercice 8 **

On note $E = \ell^2(\mathbb{N}, \mathbb{R})$ l'ensemble des suites réelles (u_n) telles que la série $\sum u_n^2$ converge.

(a) Montrer que E est un espace vectoriel réel.

Pour $u, v \in E$, on pose

$$\langle u, v \rangle = \sum_{n=0}^{+\infty} u_n v_n.$$

(b) Montrer que $\langle \cdot , \cdot \rangle$ définit un produit scalaire sur E.

On note F le sous-espace vectoriel de E constitué des suites nulles à partir d'un certain rang et v une suite élément de E qui n'appartient pas à F.

(c) Déterminer $F^\perp$. Les espaces F et $F^\perp$ sont-ils supplémentaires ?

(d) On pose $G = \text{Vect}(v)$. Comparer $F^\perp + G^\perp$ et $(F \cap G)^\perp$.

Solution

(a) **méthode**

On vérifie que $E = \ell^2(\mathbb{N}, \mathbb{R})$ est un sous-espace vectoriel de l'espace $\mathbb{R}^{\mathbb{N}}$ des suites réelles.

L'ensemble E contient la suite nulle, c'est donc une partie non vide de l'espace $\mathbb{R}^{\mathbb{N}}$. Pour $\lambda \in \mathbb{R}$ et $u \in E$, on a immédiatement la convergence de $\sum(\lambda u_n)^2$ et donc $\lambda u \in E$. Soit $u, v \in E$. Étudions la suite $u+v$ qui est de terme général $u_n + v_n$. Pour tout naturel n

$$(u_n + v_n)^2 = u_n^2 + 2u_n v_n + v_n^2.$$

méthode

On exploite l'inégalité $2ab \leqslant a^2 + b^2$ valable pour tous a et b réels.

On en déduit

$$(u_n + v_n)^2 \leqslant 2(u_n^2 + v_n^2).$$

Par comparaison de séries à termes positifs, on obtient la convergence de $\sum(u_n + v_n)^2$ et l'on peut affirmer que $u + v$ appartient à E.

Finalement, $E = \ell^2(\mathbb{N}, \mathbb{R})$ est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$, c'est donc un espace vectoriel réel.

1. Par ce résultat, on retrouve rapidement celui de la deuxième question du sujet 4 p. 251 : sachant que $\mathcal{P}$ est dense dans E, $\mathcal{P}^\perp$ est réduit à la fonction nulle.

(b) Commençons par vérifier que l'application $\langle \cdot , \cdot \rangle$ est bien définie en étudiant la convergence de la série définissant $\langle u, v \rangle$. Soit $u, v \in \ell^2(\mathbb{N}, \mathbb{R})$. En exploitant de nouveau l'inégalité $2ab \leqslant a^2 + b^2$, on obtient

$$|u_n v_n| = |u_n|\,|v_n| \leqslant \frac{1}{2}\big(|u_n|^2 + |v_n|^2\big) = \frac{1}{2}\big(u_n^2 + v_n^2\big).$$

Par comparaison de séries à termes positifs, on peut affirmer que la série $\sum u_n v_n$ est absolument convergente et donc convergente. Ainsi, l'application $\langle \cdot , \cdot \rangle$ est bien définie de $E \times E$ vers $\mathbb{R}$.

Vérifions ensuite que $\langle \cdot , \cdot \rangle$ est une forme bilinéaire symétrique définie positive. On introduit $u, v, w \in E$ et $\lambda, \mu \in \mathbb{R}$. Par commutativité de la multiplication réelle, on obtient la propriété de symétrie

$$\langle v, u \rangle = \sum_{n=0}^{+\infty} v_n u_n = \sum_{n=0}^{+\infty} u_n v_n = \langle u, v \rangle.$$

Aussi, on peut écrire avec convergence des séries introduites

$$\langle u, \lambda v + \mu w \rangle = \sum_{n=0}^{+\infty} u_n(\lambda v_n + \mu w_n) = \lambda \sum_{n=0}^{+\infty} u_n v_n + \mu \sum_{n=0}^{+\infty} u_n w_n = \lambda \langle u, v \rangle + \mu \langle u, w \rangle.$$

L'application $\langle \cdot , \cdot \rangle$ est donc linéaire en sa deuxième variable et par conséquent bilinéaire. Enfin, par sommation de termes positifs, on obtient

$$\langle u, u \rangle = \sum_{n=0}^{+\infty} u_n^2 \geqslant 0$$

et, si $\langle u, u \rangle = 0$, on conclut que la suite u est nulle par nullité d'une somme de termes tous positifs.

Finalement, $\langle \cdot , \cdot \rangle$ définit un produit scalaire sur $E = \ell^2(\mathbb{N}, \mathbb{R})$.

(c) Soit $u \in F^{\perp}$.

méthode

La suite u est orthogonale aux suites élémentaires e^p (avec $p \in \mathbb{N}$) déterminées par

$$e^p(n) = \delta_{n,p} = \begin{cases} 1 & \text{si } n = p \\ 0 & \text{sinon} \end{cases} \quad \text{pour tout } n \in \mathbb{N}.$$

Pour tout $p \in \mathbb{N}$, la suite e^p est élément de F et donc

$$\langle u, e^p \rangle = \sum_{n=0}^{+\infty} u_n \delta_{n,p} = u_p = 0.$$

On en déduit que la suite u est nulle ce qui donne $F^\perp \subset \{0\}$. Inversement, la suite nulle appartient à l'orthogonal de F et donc $F^\perp = \{0\}$.

Les sous-espaces vectoriels F et $F^\perp$ ne sont pas supplémentaires [1] car

$$F \oplus F^\perp = F \oplus \{0\} = F \neq E.$$

(d) D'une part, $F^\perp + G^\perp = \{0\} + G^\perp = G^\perp$. D'autre part, $(F \cap G)^\perp = \{0\}^\perp = E$. On a donc l'inclusion [2] $F^\perp + G^\perp \subset (F \cap G)^\perp$ et celle-ci est stricte car $G^\perp \neq E$ puisque v est une suite non nulle : elle appartient à $G \subset E$ sans appartenir à $G^\perp$.

Exercice 9 **

Soit S l'ensemble des vecteurs de norme 1 d'un espace préhilbertien réel E.
Montrer que si x et y sont deux éléments distincts de S alors, pour tout $\lambda \in \mathbb{R}$,

$$\lambda \neq 0 \text{ et } \lambda \neq 1 \implies (1-\lambda)x + \lambda y \notin S.$$

Solution

méthode

On observe que la fonction $f : \lambda \in \mathbb{R} \mapsto \left\|(1-\lambda)x + \lambda.y\right\|^2$ est une fonction polynomiale de degré 2.

Par l'identité remarquable

$$\|a+b\|^2 = \|a\|^2 + 2\langle a, b\rangle + \|b\|^2$$

on développe la norme exprimant $f(\lambda)$

$$f(\lambda) = (1-\lambda)^2 \underbrace{\|x\|^2}_{=1} + 2\lambda(1-\lambda)\langle x, y\rangle + \lambda^2 \underbrace{\|y\|^2}_{=1}.$$

L'expression $f(\lambda)$ est donc polynomiale de degré inférieur à 2. Le coefficient de λ^2 dans celle-ci est $2(1 - \langle x, y\rangle)$. Par l'absurde, si $\langle x, y\rangle = 1$, on a égalité dans l'inégalité de Cauchy-Schwarz

$$\langle x, y\rangle = \left|\langle x, y\rangle\right| = \|x\| \, \|y\|.$$

Les vecteurs x et y sont donc colinéaires. Or ils sont aussi de même norme et distincts.

1. Aussi l'inclusion de F dans $(F^\perp)^\perp$ est stricte.

2. Lorsque F et G sont des sous-espaces vectoriels de E, on a les propriétés $(F+G)^\perp = F^\perp \cap G^\perp$ et $F^\perp + G^\perp \subset (F \cap G)^\perp$. Cette dernière inclusion devient une égalité dans un espace euclidien (voir sujet 6 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*).

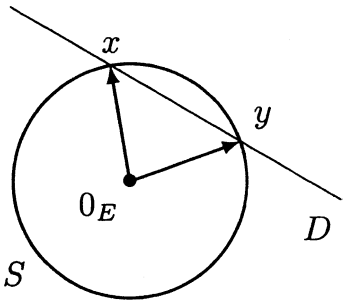
Ils sont donc opposés mais alors $\langle x, y\rangle = -1$. C'est absurde.

Ainsi, $f(\lambda)$ est une expression polynomiale du second degré exactement. Puisque celle-ci prend la valeur 1 en $\lambda = 0$ et en $\lambda = 1$, elle ne peut reprendre la valeur 1 pour aucun autre λ.

Géométriquement, ce résultat signifie que la droite affine

$$D = \left\{(1-\lambda)x + \lambda y \;\middle|\; \lambda \in \mathbb{R}\right\}$$

passant par x et y ne recoupe pas la sphère unité S. Ceci est évident si l'on figure la situation dans un plan contenant x et y.



6.4.2 Espaces euclidiens

Exercice 10 *

Montrer que la norme euclidienne associée au produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ vérifie :

$$\|AB\| \leqslant \|A\|\,\|B\| \quad \text{pour tous } A \text{ et } B \text{ de } \mathcal{M}_n(\mathbb{R}).$$

Solution

Soit $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices de $\mathcal{M}_n(\mathbb{R})$. Le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ est donné par $\langle A, B\rangle = \operatorname{tr}({}^tAB)$ et la norme euclidienne par

$$\|A\| = \sqrt{\operatorname{tr}({}^tAA)} = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}^2\right)^{1/2}.$$

Étudions la matrice $C = AB = (c_{i,j})$. Pour tous i et j de $[\![1\,;n]\!]$, le coefficient d'indice (i,j) de la matrice C est

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} \quad \text{donc} \quad \|C\|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_{i,k}b_{k,j}\right)^2.$$

méthode

Par l'inégalité de Cauchy-Schwarz, on sait

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leqslant \left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right)$$

pour tous $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ réels.

Par cette inégalité

$$\|C\|^2 \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left(\sum_{k=1}^{n} a_{i,k}^2\right)\left(\sum_{k=1}^{n} b_{k,j}^2\right)\right).$$

On réorganise ensuite le calcul des sommes

$$\|C\|^2 \leqslant \sum_{i=1}^{n}\left(\left(\sum_{k=1}^{n} a_{i,k}^2\right)\underbrace{\sum_{j=1}^{n}\left(\sum_{k=1}^{n} b_{k,j}^2\right)}_{=\|B\|^2}\right) = \underbrace{\sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{i,k}^2\right)}_{=\|A\|^2}\|B\|^2 = \|A\|^2\|B\|^2.$$

Par croissance de la fonction racine carrée, on conclut $\|AB\| = \|C\| \leqslant \|A\|\,\|B\|$.

Exercice 11 **

Soit a et b deux vecteurs unitaires d'un espace euclidien E. On étudie l'endomorphisme f de E donné par

$$f(x) = x - (a\,|\,x)b.$$

(a) À quelle condition l'endomorphisme f est-il bijectif ?

(b) À quelle condition l'endomorphisme f est-il diagonalisable ?

Solution

(a) L'application f est un endomorphisme d'un espace de dimension finie, il suffit d'étudier son noyau pour savoir s'il est bijectif. Soit x élément de E. On a

$$x \in \mathrm{Ker}(f) \iff x = (a\,|\,x)b.$$

Dans cette équation, l'inconnue x apparaît sous deux écritures : x et $(a\,|\,x)$.

méthode

On étudie le produit scalaire avec a des deux membres de l'équation afin de déterminer $(a\,|\,x)$.

Si x est solution de l'équation $x = (a\,|\,x)b$, on a $(a\,|\,x) = (a\,|\,x)(a\,|\,b)$. Ceci conduit à discuter selon que $(a\,|\,b) = 1$ ou non.

Cas : $(a\,|\,b) \neq 1$. On obtient $(a\,|\,x) = 0$ puis $x = 0.b = 0_E$. Le noyau de f se réduit au vecteur nul, l'endomorphisme est bijectif.

Cas : $(a\,|\,b) = 1$. L'équation précédente n'apporte rien mais on observe $f(b) = 0_E$ et le noyau de f n'est donc pas réduit au vecteur nul : l'endomorphisme f n'est pas bijectif.

En résumé, f est bijectif si, et seulement si, $(a\,|\,b) \neq 1$. Cette condition peut cependant être exprimée plus simplement. Les vecteurs a et b étant unitaires, lorsque $(a\,|\,b) = 1$ on a égalité dans l'inégalité de Cauchy-Schwarz

$$\left|(a\,|\,b)\right| = \|a\|\,\|b\|.$$

On en déduit que a et b sont colinéaires et même égaux car unitaires et de produit scalaire positif. La réciproque étant immédiate, on peut affirmer que $(a\,|\,b) = 1$ si, et seulement si, les vecteurs a et b sont égaux.

Finalement, l'endomorphisme est bijectif[1] si, et seulement si, $a \neq b$.

1. Lorsque $a = b$, l'endomorphisme f est la projection orthogonale sur l'hyperplan de vecteur normal a.

(b) **méthode**

On peut résoudre cette question en étudiant l'équation aux éléments propres $f(x) = \lambda x$ ce qui conduit à des calculs semblables aux précédents. On peut aussi rechercher un polynôme annulateur [1] de f.

Nous privilégions cette dernière méthode. Pour $x \in E$

$$f\big(f(x)\big) = f(x) - \big(a \,|\, f(x)\big)b = f(x) - \big(1 - (a\,|\,b)\big)\underbrace{(a\,|\,x)b}_{=x-f(x)}$$
$$= \big(2 - (a\,|\,b)\big)f(x) + \big((a\,|\,b) - 1\big)x.$$

Le polynôme $X^2 + \big((a \mid b) - 2\big)X + \big(1 - (a \mid b)\big)$ est annulateur de f et possède deux racines : 1 et $1 - (a\,|\,b)$. On poursuit en discutant selon que ces racines sont distinctes ou non.

Cas : $(a\,|\,b) \neq 0$. Le polynôme annulateur est scindé sur $\mathbb{R}$ et à racines simples, l'endomorphisme f est diagonalisable.

Cas : $(a\,|\,b) = 0$. Le polynôme annulateur possède une seule racine 1 et c'est donc la seule valeur propre possible pour f. L'endomorphisme est alors diagonalisable si, et seulement si, $f = \mathrm{Id}_E$. Or ceci n'est pas le cas car $f(a) = a - b \neq a$.

En résumé, l'endomorphisme f est diagonalisable [2] si, et seulement si, $(a\,|\,b) \neq 0$.

Exercice 12 **

Soit $E = \mathbb{R}_n[X]$ avec $n \in \mathbb{N}$.

(a) Montrer l'existence et l'unicité d'un polynôme A de E tel que

$$P(0) = \int_0^1 A(t)P(t)\,\mathrm{d}t \quad \text{pour tout } P \in E.$$

(b) Établir que le polynôme A est de degré n exactement.

Solution

(a) **méthode**

L'application $P \mapsto P(0)$ est une forme linéaire sur l'espace euclidien $\mathbb{R}_n[X]$, elle peut donc être représentée par un produit scalaire (Th. 1 p. 245).

On définit un produit scalaire [3] sur $\mathbb{R}[X]$, et donc *a fortiori* sur $E = \mathbb{R}_n[X]$, en posant

$$\langle P, Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t.$$

1. L'endomorphisme $g\colon x \mapsto (a \mid x)b$ est de rang 1 et peut donc être annulé par un polynôme de degré 2, il en est alors de même de $f = \mathrm{Id}_E - g$.

2. En étudiant les éléments propres, on obtient que l'espace propre associé à la valeur propre 1 est l'hyperplan $\{a\}^\perp$ tandis que l'espace propre associé à la valeur propre $1 - (a\,|\,b)$ est la droite $\mathrm{Vect}(b)$.

3. Voir sujet 1 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI.*

Puisque l'application $P \mapsto P(0)$ est une forme linéaire sur $\mathbb{R}[X]$, il existe un unique polynôme A dans $\mathbb{R}_n[X]$ tel que cette forme linéaire corresponde au produit scalaire avec A. Autrement dit, il existe un unique polynôme $A \in E$ vérifiant

$$P(0) = \langle A, P \rangle = \int_0^1 A(t)P(t)\,\mathrm{d}t \quad \text{pour tout } P \in E.$$

(b) Par l'absurde, si le degré de A est strictement inférieur à n, le polynôme $P = XA$ est élément de $\mathbb{R}_n[X]$ et donc

$$\int_0^1 tA(t)^2\,\mathrm{d}t = \langle A, P \rangle = P(0) = 0.$$

Cependant, la fonction $t \mapsto tA(t)^2$ est continue et positive sur $[0\,;1]$, la nullité de l'intégrale entraîne alors que $tA(t)^2 = 0$ pour tout $t \in [0\,;1]$. On en déduit que le polynôme A est nul puisqu'il possède une infinité de racines. Ceci est absurde [1].

Exercice 13 ** (Inégalité d'Hadamard)

Soit E un espace euclidien orienté de dimension $n \geqslant 1$.

(a) Montrer que, pour toute famille $(x_1, \ldots, x_n)$ de vecteurs [2] de E,

$$\left|[x_1, \ldots, x_n]\right| \leqslant \|x_1\| \ldots \|x_n\|.$$

(b) Dans quels cas y a-t-il égalité ?

Solution

(a) **méthode**

‖ On orthonormalise la famille $(x_1, \ldots, x_n)$ par le procédé de Schmidt.

Si la famille $(x_1, \ldots, x_n)$ est liée, son produit mixte est nul et l'inégalité est vraie.

Si la famille $(x_1, \ldots, x_n)$ est libre, on peut l'orthonormaliser par le procédé de Schmidt ce qui forme une base orthonormale $e = (e_1, \ldots, e_n)$ de l'espace euclidien E qui est de dimension n.

Si la base e est directe, le produit mixte correspond au déterminant dans cette base. Sinon, le produit mixte est opposé à ce déterminant. Dans les deux cas

$$\left|[x_1, \ldots, x_n]\right| = \left|\det_e(x_1, \ldots, x_n)\right|$$

avec $\det_e(x_1, \ldots, x_n)$ qui est le déterminant de la matrice $A = (a_{i,j})$ figurant la famille $(x_1, \ldots, x_n)$ dans la base orthonormale e. Pour tout indice (i, j), le coefficient $a_{i,j}$

1. Le même raisonnement peut être repris pour établir que la propriété de la première question est fausse lorsque $E = \mathbb{R}[X]$. En substance, on peut souligner que le théorème de représentation des formes linéaires peut ne pas être valable en dimension infinie.

2. $[x_1, \ldots, x_n]$ désigne le produit mixte de la famille $(x_1, \ldots, x_n)$, c'est-à-dire le déterminant de cette famille dans n'importe quelle base orthonormale directe de E.

est la i-ème coordonnée dans e du vecteur x_j et donc

$$a_{i,j} = \langle e_i, x_j \rangle.$$

Par le procédé de Schmidt, $x_j \in \text{Vect}(e_1, \dots, e_j)$ et la matrice A est triangulaire supérieure de la forme

$$\begin{pmatrix} \langle e_1, x_1 \rangle & \langle e_1, x_2 \rangle & \cdots & \langle e_1, x_n \rangle \\ & \langle e_2, x_2 \rangle & \ddots & \vdots \\ & & \ddots & \langle e_{n-1}, x_n \rangle \\ (0) & & & \langle e_n, x_n \rangle \end{pmatrix}.$$

On a donc

$$\left|[x_1, \dots, x_n]\right| = \left|\langle e_1, x_1 \rangle \dots \langle e_n, x_n \rangle\right|.$$

Enfin, la base e étant constituée de vecteurs unitaires, l'inégalité de Cauchy-Schwarz donne

$$\left|\langle e_j, x_j \rangle\right| \leqslant \|e_j\| \, \|x_j\| = \|x_j\|$$

ce qui permet de conclure

$$\left|[x_1, \dots, x_n]\right| \leqslant \|x_1\| \dots \|x_n\|.$$

(b) Si la famille est liée, il y a égalité si, et seulement si, l'un des vecteurs x_j est nul.

Si la famille est libre, on reprend les notations qui précèdent et l'on peut affirmer qu'il y a égalité si, et seulement si,

$$\|x_j\| = \left|\langle e_j, x_j \rangle\right| \quad \text{pour tout } j \in [\![1\,;n]\!].$$

Par égalité dans l'inégalité de Cauchy-Schwarz, cela revient à dire que x_j est colinéaire à e_j pour tout $j \in [\![1\,;n]\!]$. La famille $(x_1, \dots, x_n)$ est alors orthogonale. Inversement, si la famille $(x_1, \dots, x_n)$ est orthogonale, il y a égalité [1].

Exercice 14 **

Soit a, b deux vecteurs unitaires d'un espace euclidien E de dimension $n \geqslant 2$.
Déterminer le maximum et le minimum de la fonction $f : x \mapsto (a|x)(b|x)$ définie sur la sphère unité fermée de E.

Solution

La sphère unité fermée S est une partie compacte de E et l'application f y est continue car produit de deux formes linéaires [2] : la fonction f admet donc un minimum et un maximum sur la boule unité fermée.

1. Le produit mixte mesure le volume algébrique d'un parallélépipède en dimension n, il y a égalité dans l'inégalité lorsque celui-ci est « droit » ou « plat ».

2. Rappelons qu'en dimension finie les parties compactes sont les parties fermées et bornées tandis que les applications linéaires au départ d'un espace de dimension finie sont assurément continues.

méthode

On exprime le vecteur x par ses coordonnées dans une base orthonormale adaptée à la situation.

On commence par isoler les cas où $\mathrm{Vect}(a,b)$ ne serait pas un plan.

Cas : $a=b$. On a $f(x)=(a\,|\,x)^2$. Le maximum cherché vaut 1, il est atteint en $x=a$. Le minimum vaut 0 et est atteint en tout vecteur unitaire orthogonal à a.

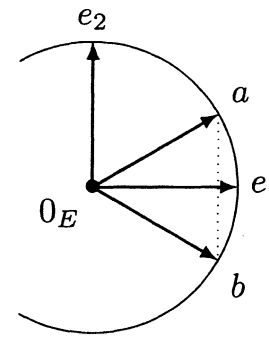
Cas : $a=-b$. On a $f(x)=-(a\,|\,x)^2$. Le maximum vaut 0 et le minimum -1.

Dans les cas restants, les vecteurs $a+b$ et $a-b$ constituent une famille orthogonale car

$$(a+b\,|\,a-b)=\|a\|^2-\|b\|^2=0.$$

Posons alors

$$e_1=\frac{1}{\|a+b\|}(a+b)\quad\text{et}\quad e_2=\frac{1}{\|a-b\|}(a-b).$$



Les vecteurs e_1 et e_2 forment une famille orthonormale que l'on peut compléter en une base orthonormale $e=(e_1,e_2,\dots,e_n)$. Soit x un vecteur de la sphère unité fermée S et $x_1,\dots,x_n$ ses coordonnées dans la base e

$$x=x_1e_1+x_2e_2+\cdots+x_ne_n\quad\text{avec}\quad x_1^2+x_2^2+\cdots+x_n^2=1.$$

On a

$$(a\,|\,x)=x_1\frac{1+(a\,|\,b)}{\|a+b\|}+x_2\frac{1-(a\,|\,b)}{\|a-b\|}$$

et une écriture analogue pour $(b\,|\,x)$ permettant d'exprimer $f(x)$:

$$f(x)=x_1^2\left(\frac{1+(a\,|\,b)}{\|a+b\|}\right)^2-x_2^2\left(\frac{1-(a\,|\,b)}{\|a-b\|}\right)^2=\underbrace{\frac{1+(a\,|\,b)}{2}}_{\geqslant 0}x_1^2-\underbrace{\frac{1-(a\,|\,b)}{2}}_{\geqslant 0}x_2^2.$$

Sachant $x_1^2+x_2^2+\cdots+x_n^2=1$, cette expression est maximale lorsque x_1 vaut 1 et les autres valeurs x_i sont nulles. Elle est minimale pour $x_2=1$ et les autres valeurs nulles.

Finalement,

$$\max_{x\in S}f(x)=\frac{1+(a\,|\,b)}{2}\quad\text{et}\quad\min_{x\in S}f(x)=\frac{(a\,|\,b)-1}{2}.$$

On notera que ces formules conviennent aussi pour les cas initialement isolés.

Exercice 15 * (Famille obtusangle[1])**

Soit $x_1,x_2,\dots,x_{n+2}$ des vecteurs d'un espace euclidien E de dimension $n\geqslant 1$. Montrer qu'il est impossible que $(x_i\,|\,x_j)<0$ pour tous les indices i et j distincts compris entre 1 et $n+2$.

1. Selon que $(x\,|\,y)$ est positif, nul ou négatif, on dit que les deux vecteurs x et y forment un angle *aigu*, *droit* ou *obtus*.

Solution

méthode

On commence par étudier les cas $n = 1$ et $n = 2$.

Cas : $n = 1$. Par l'absurde, supposons disposer d'une famille (x_1, x_2, x_3) de vecteurs d'une droite E telle que

$$(x_1 \,|\, x_2) < 0, \quad (x_2 \,|\, x_3) < 0 \quad \text{et} \quad (x_3 \,|\, x_1) < 0.$$

L'espace E étant de dimension 1, les vecteurs x_1, x_2 et x_3 sont deux à deux colinéaires. En particulier, on peut écrire $x_1 = \lambda_1 x_3$ et $x_2 = \lambda_2 x_3$ car x_3 est non nul. On a alors

$$(x_1 \,|\, x_3) = \lambda_1 \|x_3\|^2 \quad \text{et} \quad (x_2 \,|\, x_3) = \lambda_2 \|x_3\|^2$$

avec $\lambda_1 < 0$ et $\lambda_2 < 0$. Ceci entraîne

$$(x_1 \,|\, x_2) = \underbrace{\lambda_1 \lambda_2}_{>0} \|x_3\|^2 > 0$$

ce qui contredit l'hypothèse $(x_1 \,|\, x_2) < 0$.

Cas : $n = 2$. Par l'absurde, supposons disposer d'une famille de vecteurs (x_1, x_2, x_3, x_4) telle que

$$\forall (i,j) \in [\![1\,;4]\!]^2, \quad i \neq j \implies (x_i \,|\, x_j) < 0.$$

En décomposant les vecteurs x_1, x_2, x_3 selon les droites $\mathrm{Vect}(x_4)$ et $\{x_4\}^\perp$, on peut écrire

$$x_1 = y_1 + \lambda_1 x_4, \quad x_2 = y_2 + \lambda_2 x_4 \quad \text{et} \quad x_3 = y_3 + \lambda_3 x_4$$

avec $y_1, y_2, y_3 \in \{x_4\}^\perp$. L'inégalité $(x_i \,|\, x_4) < 0$ donne $\lambda_i < 0$ pour tout $i \in [\![1\,;3]\!]$.

On a alors, pour $i \neq j$ dans $[\![1\,;3]\!]$,

$$(x_i \,|\, x_j) = \lambda_i \lambda_j \|x_4\|^2 + (y_i \,|\, y_j) \quad \text{et} \quad (x_i \,|\, x_j) < 0.$$

Or $\lambda_i \lambda_j > 0$ et donc $(y_i \,|\, y_j) < 0$. Les vecteurs y_1, y_2, y_3 appartenant à la droite $\{x_4\}^\perp$, l'étude du cas $n = 1$ permet de conclure à une absurdité.

Cas général :

méthode

On raisonne par récurrence en projetant la famille de vecteurs sur l'hyperplan de vecteur normal égal au dernier vecteur de la famille.

Montrons par récurrence sur $n \geqslant 1$ qu'il ne peut exister dans un espace euclidien de dimension $n \geqslant 1$ de famille $(x_1, \ldots, x_{n+2})$ de vecteurs vérifiant $(x_i \,|\, x_j) < 0$ pour tous les indices $i \neq j$.

Les cas $n = 1$ et $n = 2$ ont été résolus ci-dessus.

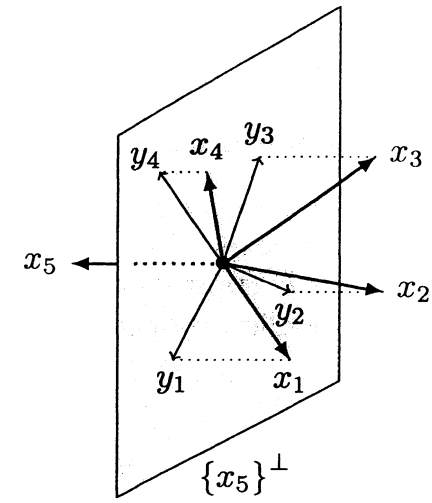
Supposons la propriété établie au rang $n - 1 \geqslant 1$ et supposons par l'absurde que $(x_1, \ldots, x_{n+2})$ soit une famille de vecteurs d'un espace euclidien de dimension n vérifiant l'hypothèse $(x_i \,|\, x_j) < 0$ pour $i \neq j$. On introduit les vecteurs $y_1, \ldots, y_{n+1}$ obtenus par projection orthogonale de $x_1, \ldots, x_{n+1}$ sur l'hyperplan $H = \{x_{n+2}\}^\perp$.

Pour tout $i \in [\![1\,;n+1]\!]$, on peut écrire

$$x_i = y_i + \lambda_i x_{n+2}$$

avec $\lambda_i < 0$ car $(x_i \,|\, x_{n+2}) = \lambda_i \|x_{n+2}\|^2 < 0$. On a alors, pour $i \neq j$ dans $[\![1\,;n+1]\!]$,

$$(x_i \,|\, x_j) = \underbrace{\lambda_i \lambda_j}_{>0} \|x_{n+2}\|^2 + (y_i \,|\, y_j) < 0.$$



Les vecteurs $y_1, \ldots, y_{n+1}$ constituent alors une famille d'un espace euclidien de dimension $n-1$ vérifiant [1] $(y_i \,|\, y_j) < 0$ pour tous les indices $i \neq j$. L'hypothèse de récurrence assure que ceci est absurde.

La récurrence est établie.

6.4.3 Projection orthogonale et distance

Exercice 16 *

On munit $\mathcal{M}_n(\mathbb{R})$ de son produit scalaire canonique $\langle A, B\rangle = \mathrm{tr}({}^tAB)$.

(a) Montrer que les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ des matrices symétriques et antisymétriques sont supplémentaires et orthogonaux.

(b) Calculer la distance à $\mathcal{S}_3(\mathbb{R})$ de la matrice

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 1 & 4 & 3 \end{pmatrix}.$$

Solution

(a) **méthode**

> En montrant que des espaces sont orthogonaux, on peut immédiatement affirmer qu'ils sont en somme directe (Th. 2 p. 245).

Soit $S \in \mathcal{S}_n(\mathbb{R})$ et $A \in \mathcal{A}_n(\mathbb{R})$. On a ${}^tS = S$ et ${}^tA = -A$. On en déduit

$$\langle S, A\rangle = \mathrm{tr}\big({}^tSA\big) = \mathrm{tr}(SA) \quad \text{et} \quad \langle S, A\rangle = \langle A, S\rangle = \mathrm{tr}\big({}^tAS\big) = \mathrm{tr}(-AS) = -\,\mathrm{tr}(AS).$$

Or $\mathrm{tr}(AS) = \mathrm{tr}(SA)$ et donc $\langle A, S\rangle = 0$. Ainsi, les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogo-

1. Les vecteurs y_i et y_j se situant d'un même côté de l'hyperplan sur lequel on projette, la propriété de former un angle obtus est conservée par la projection.

naux et donc en somme directe. Aussi, pour tout $M \in \mathcal{M}_n(\mathbb{R})$, on peut écrire [1]

$$M = \frac{1}{2}(M + {}^tM) + \frac{1}{2}(M - {}^tM) \qquad (*)$$

avec

$$\frac{1}{2}(M + {}^tM) \in \mathcal{S}_n(\mathbb{R}) \quad \text{et} \quad \frac{1}{2}(M - {}^tM) \in \mathcal{A}_n(\mathbb{R}).$$

La somme des espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ est donc égale à $\mathcal{M}_n(\mathbb{R})$ et l'on peut affirmer que ceux-ci sont supplémentaires et orthogonaux [2].

(b) **méthode**

La distance de M à l'espace $\mathcal{S}_3(\mathbb{R})$ est égale à la distance de M à son projeté orthogonal sur $\mathcal{S}_3(\mathbb{R})$.

Pour calculer le projeté orthogonal de M sur $\mathcal{S}_3(\mathbb{R})$, il suffit de décomposer la matrice M en la somme d'une matrice symétrique et d'une matrice antisymétrique. La formule $(*)$ réalise cette décomposition :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 2 & 4 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

La distance de M à $\mathcal{S}_3(\mathbb{R})$ correspond alors à la norme de la matrice antisymétrique de l'écriture ci-dessus.

méthode

Le produit scalaire $\langle A, B\rangle = \operatorname{tr}({}^tAB)$ correspond simplement au calcul

$$\langle A, B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}b_{i,j}.$$

La distance de M à $\mathcal{S}_3(\mathbb{R})$ est donc $d\big(M, \mathcal{S}_3(\mathbb{R})\big) = \sqrt{1^2 + 1^2 + (-1)^2 + (-1)^2} = 2$.

Exercice 17 **

Soit E un espace euclidien de dimension n muni d'une base orthonormale e et p la projection orthogonale sur un sous-espace vectoriel F muni d'une base orthonormale $(x_1, \ldots, x_m)$. Montrer que la matrice A de p dans la base e est

$$A = \sum_{k=1}^{m} X_k {}^tX_k$$

avec $X_1, \ldots, X_m$ les colonnes des coordonnées des vecteurs $x_1, \ldots, x_m$ dans e.

1. En peut aussi affirmer la supplémentarité par un argument de dimension car $\dim \mathcal{S}_n(\mathbb{R}) = \frac{n(n+1)}{2}$ et $\dim \mathcal{A}_n(\mathbb{R}) = \frac{n(n-1)}{2}$ de somme égale à $n^2 = \dim \mathcal{M}_n(\mathbb{R})$.

2. Chacun est alors l'orthogonal de l'autre.

Solution

méthode

On sait exprimer le projeté orthogonal d'un vecteur dans une base orthonormale de l'espace sur lequel on projette (Th. 5 p. 246).

Pour tout $x \in E$, on a

$$p(x) = \sum_{k=1}^{m} (x_k \,|\, x) x_k.$$

En notant X la colonne des coordonnées du vecteur x dans la base orthonormale e, on peut exprimer le produit scalaire de x_k et x par le calcul matriciel suivant :

$$(x_k \,|\, x) = {}^t X_k X.$$

La colonne des coordonnées du vecteur image $p(x)$ est alors

$$AX = \sum_{k=1}^{m} ({}^t X_k X) X_k.$$

Or ${}^t X_k X$ est un réel et donc $({}^t X_k X) X_k = X_k ({}^t X_k X)$ ce qui permet d'écrire

$$AX = \sum_{k=1}^{p} X_k {}^t X_k X = \left(\sum_{k=1}^{p} X_k {}^t X_k \right) X.$$

On identifie [1] alors la matrice A ce qui valide la formule proposée.

Exercice 18 **

Soit p un projecteur d'un espace euclidien E vérifiant $\langle p(x), x \rangle \geqslant 0$ pour tout x de E. Montrer que p est un projecteur orthogonal.

Solution

méthode

Un projecteur p projette sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$. Il est orthogonal si, et seulement si, $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont des sous-espaces orthogonaux.

Soit $x \in \mathrm{Im}(p)$ et $y \in \mathrm{Ker}(p)$. Considérons [2] $z = x + \lambda y$ avec $\lambda \in \mathbb{R}$. On a par hypothèse

$$\langle p(x + \lambda y), x + \lambda . y \rangle \geqslant 0.$$

Sachant $p(x) = x$ et $p(y) = 0_E$, ceci donne

$$\|x\|^2 + \lambda \langle x, y \rangle \geqslant 0 \quad \text{pour tout } \lambda \in \mathbb{R}.$$

1. La matrice A d'une application linéaire u est l'unique matrice qui caractérise le calcul vectoriel $y = u(x)$ par le produit matriciel $Y = AX$ avec X et Y les colonnes des coordonnées de x et y dans des bases préalablement introduites. On peut aussi considérer X une colonne élémentaire et employer l'égalité pour vérifier que les matrices sont identiques colonne par colonne.

2. On peut aussi introduire $z' = \lambda x + y$ ce qui conduit à $\lambda^2 \|x\|^2 + \lambda \langle x, y \rangle \geqslant 0$. Si $\langle x, y \rangle \neq 0$ on obtient une absurdité en considérant un équivalent quand λ tend vers 0.

Par l'absurde, si $\langle x, y\rangle \neq 0$, la fonction affine $\lambda \mapsto \|x\|^2 + \lambda\langle x, y\rangle$ change de signe ce qui contredit la propriété au-dessus. On en déduit $\langle x, y\rangle = 0$ et les espaces $\mathrm{Im}(p)$ et $\mathrm{Ker}(p)$ sont orthogonaux.

Exercice 19 ***

On munit l'espace $E = \mathcal{C}^1\big([-1\,;1],\mathbb{R}\big)$ du produit scalaire $\langle\,\cdot\,,\,\cdot\,\rangle$ donné par

$$\langle u, v\rangle = \int_{-1}^{1} \big(u(t)v(t) + u'(t)v'(t)\big)\,\mathrm{d}t$$

et l'on introduit les sous-espaces vectoriels

$$F = \big\{f \in E \,\big|\, f(-1) = f(1) = 0\big\} \quad \text{et} \quad G = \big\{g \in E \,\big|\, g \text{ est de classe } \mathcal{C}^2 \text{ et } g'' = g\big\}.$$

(a) Montrer que les espaces F et G sont supplémentaires et orthogonaux.

Soit a et b deux réels et

$$F_{a,b} = \big\{u \in E \,\big|\, u(-1) = a \text{ et } u(1) = b\big\}.$$

(b) Calculer

$$\inf_{u\in F_{a,b}} \int_{-1}^{1} \big(u(t)^2 + u'(t)^2\big)\,\mathrm{d}t.$$

Solution

(a) Soit $f \in F$ et $g \in G$. En écrivant $g''(t)$ au lieu de $g(t)$ dans le calcul intégral, on a

$$\langle f, g\rangle = \int_{-1}^{1} \big(f(t)g''(t) + f'(t)g'(t)\big)\,\mathrm{d}t = \Big[f(t)g'(t)\Big]_{-1}^{1} = 0.$$

Les espaces F et G sont donc orthogonaux et *a fortiori* en somme directe. Montrons que leur somme est égale à E.

Soit u une fonction de E. Déterminons des fonctions f dans F et g dans G telles que $u = f + g$.

Analyse : Supposons que (f, g) désigne un couple de fonctions convenables. La fonction g étant solution de l'équation différentielle linéaire du second ordre à coefficients constants $y'' - y = 0$, il existe deux réels α et β tels que [1]

$$g(t) = \alpha\,\mathrm{ch}\,t + \beta\,\mathrm{sh}\,t \quad \text{pour tout } t \in [-1\,;1].$$

Les conditions $f(-1) = f(1) = 0$ définissent alors un système permettant de déterminer les réels α et β

$$\begin{cases} \alpha\,\mathrm{ch}(1) - \beta\,\mathrm{sh}(1) = u(-1) \\ \alpha\,\mathrm{ch}(1) + \beta\,\mathrm{sh}(1) = u(1) \end{cases} \iff \begin{cases} \alpha = \frac{1}{2\,\mathrm{ch}(1)}\big(u(1) + u(-1)\big) \\ \beta = \frac{1}{2\,\mathrm{sh}(1)}\big(u(1) - u(-1)\big). \end{cases}$$

1. La solution générale de l'équation $y'' - y = 0$ s'exprime indifféremment $y(t) = \lambda \mathrm{e}^t + \mu \mathrm{e}^{-t}$ avec $\lambda, \mu \in \mathbb{R}$ ou $y(t) = \alpha\,\mathrm{ch}\,t + \beta\,\mathrm{sh}\,t$ avec $\alpha, \beta \in \mathbb{R}$: on privilégie cette dernière écriture afin d'exploiter la symétrie de l'intervalle d'étude.

On en déduit la fonction g puis la fonction $f = u - g$.

Synthèse : Posons f et g les fonctions déterminées sur $[-1\,;1]$ par

$$g(t) = \frac{u(1)+u(-1)}{2\,\mathrm{ch}(1)}\,\mathrm{ch}\,t + \frac{u(1)-u(-1)}{2\,\mathrm{sh}(1)}\,\mathrm{sh}\,t \quad \text{et} \quad f(t) = u(t) - g(t).$$

La fonction g appartient à G et $u = f + g$. On vérifie par le calcul que la fonction f satisfait $f(1) = f(-1) = 0$ et appartient donc à F.

Finalement, les espaces F et G sont supplémentaires et orthogonaux dans E.

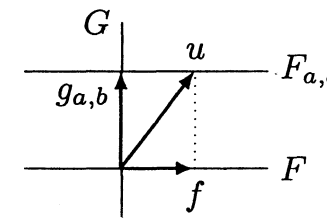
(b) La borne inférieure étudiée revient à chercher la distance du vecteur nul à $F_{a,b}$.

méthode

$F_{a,b}$ est un espace affine obtenu par translation du sous-espace vectoriel F : la distance du vecteur nul à celui se déduit de l'intersection de $F_{a,b}$ et de G.

À l'aide des calculs qui précèdent, on peut déterminer l'unique fonction $g_{a,b}$ appartenant à $F_{a,b} \cap G$

$$g_{a,b}(t) = \frac{a+b}{2\,\mathrm{ch}(1)}\,\mathrm{ch}\,t + \frac{a-b}{2\,\mathrm{sh}(1)}\,\mathrm{sh}\,t.$$



Les fonctions de $F_{a,b}$ sont de la forme $u = g_{a,b} + f$ avec f parcourant F. Par orthogonalité de f et $g_{a,b}$

$$\int_{-1}^{1} \left(u(t)^2 + u'(t)^2\right) \mathrm{d}t = \|u\|^2 = \|f\|^2 + \|g_{a,b}\|^2.$$

Cette quantité est minimale lorsque f est la fonction nulle. On en déduit

$$\inf_{u \in F_{a,b}} \int_{-1}^{1} \left(u(t)^2 + u'(t)^2\right) \mathrm{d}t = \|g_{a,b}\|^2 = \frac{\left(a^2+b^2\right)\mathrm{ch}(2) - 2ab}{\mathrm{sh}(2)}.$$

Les derniers calculs sont résolus en observant que les fonctions ch et sh sont orthogonales et

$$\|\mathrm{ch}\|^2 = \|\mathrm{sh}\|^2 = \int_{-1}^{1} \underbrace{\mathrm{ch}^2(t) + \mathrm{sh}^2(t)}_{=\mathrm{ch}(2t)}\,\mathrm{d}t = \left[\frac{1}{2}\,\mathrm{sh}(2t)\right]_{-1}^{1} = \mathrm{sh}\,2.$$

6.4.4 Produit scalaire et transposition matricielle

L'espace $\mathcal{M}_{n,1}(\mathbb{R})$ des colonnes de taille n est muni du produit scalaire

$$\langle X, Y\rangle = {}^tXY$$

et de la norme euclidienne associée.

Exercice 20 *

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Vérifier $\mathrm{Im}\left({}^tA\right) = \left(\mathrm{Ker}(A)\right)^{\perp}$.

Solution

méthode

‖ On vérifie que les espaces $\mathrm{Ker}(A)$ et $\mathrm{Im}({}^tA)$ sont orthogonaux.

Soit $X \in \mathrm{Ker}(A)$ et $Y \in \mathrm{Im}({}^tA)$. On a $AX = 0$ et l'on peut écrire $Y = {}^tAX'$ avec X' une colonne. On a alors

$$\langle X, Y\rangle = {}^tXY = {}^tX\,{}^tAX' = {}^t(AX)X' = 0 \quad \text{car} \quad AX = 0.$$

Les espaces $\mathrm{Ker}(A)$ et $\mathrm{Im}({}^tA)$ sont donc orthogonaux ce qui permet d'écrire[1]

$$\mathrm{Im}({}^tA) \subset \big(\mathrm{Ker}(A)\big)^{\perp}.$$

De plus, le rang d'une matrice est le rang de sa transposée et la formule du rang donne

$$\mathrm{rg}({}^tA) = \mathrm{rg}(A) = n - \dim \mathrm{Ker}(A) = \dim\big(\mathrm{Ker}(A)\big)^{\perp}.$$

Par inclusion et égalité des dimensions, on peut conclure[2] $\mathrm{Im}({}^tA) = \big(\mathrm{Ker}(A)\big)^{\perp}$.

Exercice 21 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

(a) Comparer les espaces $\mathrm{Ker}(A)$ et $\mathrm{Ker}({}^tAA)$.

(b) Comparer les espaces $\mathrm{Im}(A)$ et $\mathrm{Im}(A\,{}^tA)$.

Solution

(a) On a immédiatement $\mathrm{Ker}(A) \subset \mathrm{Ker}({}^tAA)$ car, si $X \in \mathrm{Ker}(A)$, on a $AX = 0$ et donc aussi ${}^tAAX = 0$.

méthode

‖ On établit l'inclusion réciproque en étudiant $\|AX\|^2$.

Soit $X \in \mathrm{Ker}({}^tAA)$. On a ${}^tAAX = 0$. En considérant la norme euclidienne sur l'espace des colonnes,

$$\|AX\|^2 = {}^t(AX)AX = {}^tX({}^tAAX) = 0.$$

Ainsi, $AX = 0$ ce qui établit l'inclusion réciproque $\mathrm{Ker}({}^tAA) \subset \mathrm{Ker}(A)$ et, finalement, $\mathrm{Ker}({}^tAA) = \mathrm{Ker}(A)$.

(b) L'inclusion $\mathrm{Im}(A\,{}^tA) \subset \mathrm{Im}(A)$ est immédiate car, si une colonne Y s'écrit $A\,{}^tAX$, on peut aussi l'écrire $Y = AX'$ avec $X' = {}^tAX$.

1. Ou, et c'est équivalent, l'inclusion $\mathrm{Ker}(A) \subset \big(\mathrm{Im}({}^tA)\big)^{\perp}$. L'orthogonalité des espaces ne suffit cependant pas à affirmer leur égalité.

2. Lorsque la matrice A est symétrique (ou antisymétrique), on obtient $\mathrm{Im}(A) = \big(\mathrm{Ker}(A)\big)^{\perp}$.

méthode

On montre l'égalité par un argument de dimension et de conservation du rang par transposition.

Par l'étude de la question précédente et la formule du rang, on peut écrire

$$\operatorname{rg}({}^tAA) = n - \dim \operatorname{Ker}({}^tAA) = n - \dim \operatorname{Ker}(A) = \operatorname{rg}(A).$$

En appliquant cette formule à la matrice ${}^tA$ au lieu de A, il vient

$$\operatorname{rg}(A\,{}^tA) = \operatorname{rg}({}^tA) \quad \text{donc} \quad \operatorname{rg}(A\,{}^tA) = \operatorname{rg}(A).$$

Finalement, par inclusion et égalité des dimensions, on conclut $\operatorname{Im}(A\,{}^tA) = \operatorname{Im}(A)$.

Exercice 22 **

Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ vérifiant $A^2 = \mathrm{O}_n$.

(a) Établir $\operatorname{Ker}({}^tA + A) = \operatorname{Ker}(A) \cap \operatorname{Ker}({}^tA)$.

(b) En déduire

$${}^tA + A \in \mathrm{GL}_n(\mathbb{R}) \iff \operatorname{Im}(A) = \operatorname{Ker}(A).$$

Solution

(a) On a immédiatement

$$\operatorname{Ker}(A) \cap \operatorname{Ker}({}^tA) \subset \operatorname{Ker}({}^tA + A)$$

car une colonne X annulant A et ${}^tA$ annule aussi ${}^tA + A$.

Inversement, soit $X \in \operatorname{Ker}({}^tA + A)$. On sait ${}^tAX + AX = 0$. Afin d'exploiter l'hypothèse $A^2 = \mathrm{O}_n$, on multiplie à gauche par A ce qui donne

$$A\,{}^tAX + A^2X = A\,{}^tAX = 0.$$

méthode

On fait apparaître une norme euclidienne en multipliant à gauche par ${}^tX$.

On obtient

$${}^tX \underbrace{A\,{}^tAX}_{=0} = {}^t\left({}^tAX\right){}^tAX = \left\|{}^tAX\right\|^2.$$

On en déduit ${}^tAX = 0$. La relation initiale ${}^tAX + AX = 0$ donne alors $AX = 0$. Ainsi, la colonne X appartient à $\operatorname{Ker}(A)$ et $\operatorname{Ker}({}^tA)$.

Finalement, on a obtenu l'égalité demandée par double inclusion.

(b) ($\Longrightarrow$) On suppose la matrice ${}^tA + A$ inversible. On a alors

$$\operatorname{Ker}(A) \cap \operatorname{Ker}({}^tA) = \operatorname{Ker}({}^tA + A) = \{0\}.$$

Les espaces $\mathrm{Ker}(A)$ et $\mathrm{Ker}({}^tA)$ sont donc en somme directe ce qui entraîne

$$\dim \mathrm{Ker}(A) + \dim \mathrm{Ker}({}^tA) \leqslant n.$$

Par la formule du rang, il vient alors

$$\dim \mathrm{Ker}(A) \leqslant n - \dim \mathrm{Ker}({}^tA) = \mathrm{rg}({}^tA) = \mathrm{rg}(A).$$

Cependant, l'hypothèse $A^2 = O_n$ entraîne $\mathrm{Im}(A) \subset \mathrm{Ker}(A)$. Les espaces $\mathrm{Im}(A)$ et $\mathrm{Ker}(A)$ sont donc égaux.

($\Longleftarrow$) Supposons $\mathrm{Im}(A) = \mathrm{Ker}(A)$. Étudions le noyau de ${}^tA + A$. Soit X une colonne élément de ce noyau. La colonne X appartient alors à $\mathrm{Ker}(A) \cap \mathrm{Ker}({}^tA)$. En particulier, X appartient à $\mathrm{Ker}(A)$ donc à $\mathrm{Im}(A)$. On peut alors introduire un colonne X' telle que $X = AX'$. De plus, X appartient à $\mathrm{Ker}({}^tA)$ et alors

$${}^tAX = {}^tAAX' = 0.$$

En multipliant à gauche par la ligne ${}^tX'$, on obtient

$$\|X\|^2 = \|AX'\|^2 = {}^tX' \underbrace{{}^tAAX'}_{=0} = 0$$

et donc X est la colonne nulle. Ainsi, le noyau de ${}^tA + A$ est réduit à l'élément nul et l'on peut affirmer que la matrice ${}^tA + A$ est inversible.

Exercice 23 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant $\|AX\| \leqslant \|X\|$ pour toute colonne X de $\mathcal{M}_{n,1}(\mathbb{R})$.

(a) Montrer $\|{}^tAX\| \leqslant \|X\|$ pour toute colonne X de $\mathcal{M}_{n,1}(\mathbb{R})$.

(b) Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$ vérifiant $AX = X$. Montrer ${}^tAX = X$.

Solution

(a) **méthode**

‖ On emploie l'inégalité de Cauchy-Schwarz.

Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. On peut écrire

$$\|{}^tAX\|^2 = {}^t({}^tAX)AX = {}^tXA\,{}^tAX = \langle X, A\,{}^tAX\rangle.$$

Par l'inégalité de Cauchy-Schwarz, on obtient

$$\|{}^tAX\|^2 = \langle X, A\,{}^tAX\rangle \leqslant \|X\|\|A\,{}^tAX\|.$$

L'hypothèse vérifiée par la matrice A permet d'écrire $\|A\,{}^tAX\| \leqslant \|{}^tAX\|$ et donc

$$\|{}^tAX\|^2 \leqslant \|X\|\|{}^tAX\|.$$

On peut alors affirmer $\|{}^tAX\| \leqslant \|X\|$ que la colonne ${}^tAX$ soit nulle ou non.

(b) **méthode**

‖ On étudie $\left\|{}^tAX - X\right\|^2$.

On développe le calcul de la norme euclidienne par identité remarquable

$$\left\|{}^tAX - X\right\|^2 = \left\|{}^tAX\right\|^2 - 2\langle {}^tAX, X\rangle + \|X\|^2.$$

Par la définition du produit scalaire, on remarque

$$\langle {}^tAX, X\rangle = {}^t\left({}^tAX\right)X = {}^tXAX = {}^tXX = \|X\|^2$$

et donc

$$\left\|{}^tAX - X\right\|^2 = \left\|{}^tAX\right\|^2 - \|X\|^2 \leqslant \|X\|^2 - \|X\|^2 = 0.$$

On peut alors conclure ${}^tAX = X$.

6.4.5 Polynômes orthogonaux

Exercice 24 ** (Polynômes orthogonaux de Legendre)

Dans ce sujet, on identifie polynôme et fonction polynomiale associée sur $[-1\,;1]$.
On munit l'espace $E = \mathcal{C}\left([-1\,;1], \mathbb{R}\right)$ du produit scalaire

$$(f\,|\,g) = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t.$$

Pour $n \in \mathbb{N}$, on introduit le polynôme $P_n = U_n^{(n)}$ avec $U_n = (X-1)^n(X+1)^n$.

(a) Montrer que P_n est un polynôme de degré n orthogonal à tout polynôme de degré inférieur à $n-1$.

(b) Établir que, pour toute fonction f de l'espace préhilbertien E,

$$\left\| \sum_{k=0}^{n} \frac{(P_k\,|\,f)}{\|P_k\|^2} P_k - f \right\| \xrightarrow[n\to+\infty]{} 0.$$

Solution

(a) Le polynôme U_n est de degré $2n$ et donc, par dérivation à l'ordre n, le degré de P_n vaut $\deg(U_n) - n = n$.

Soit Q un polynôme de degré inférieur à $n-1$. Calculons $(P_n\,|\,Q)$.

méthode

‖ On procède par intégration par parties où l'on dérive le polynôme Q.

On réalise une première intégration par parties où l'on intègre $P_n = U_n^{(n)}$ en $U_n^{(n-1)}$:

$$(P_n\,|\,Q) = \int_{-1}^{1} P_n(t)Q(t)\,\mathrm{d}t = \left[U_n^{(n-1)}(t)Q(t)\right]_{-1}^{1} - \int_{-1}^{1} U_n^{(n-1)}(t)Q'(t)\,\mathrm{d}t.$$

Les valeurs 1 et -1 sont racines de multiplicité n du polynôme U_n, elles sont donc aussi racines des polynômes $U_n', \dots, U_n^{(n-1)}$. L'égalité précédente devient alors

$$(P_n \,|\, Q) = -\int_{-1}^{1} U_n^{(n-1)}(t) Q'(t)\,\mathrm{d}t.$$

On répète ces intégrations par parties jusqu'à disparition par dérivation du polynôme Q

$$(P_n \,|\, Q) = \int_{-1}^{1} U_n^{(n-2)}(t) Q''(t)\,\mathrm{d}t = \cdots = (-1)^n \int_{-1}^{1} U_n(t) Q^{(n)}(t)\,\mathrm{d}t = 0 \quad \text{car} \quad Q^{(n)} = 0.$$

Le polynôme P_n est donc orthogonal à tout polynôme de Q de degré inférieur à $n-1$.

(b) **méthode**

‖ On introduit une famille orthonormale totale.

La famille $(P_n)_{n\in\mathbb{N}}$ est une famille orthogonale car, pour tous m et n entiers naturels vérifiant $m < n$, on a $(P_n \mid P_m) = 0$ puisque P_m est de degré inférieur à $n-1$. Aucun polynôme P_n n'est nul et l'on peut donc diviser chacun par sa norme afin de former une famille $(Q_n)_{n\in\mathbb{N}}$ orthonormale. Enfin, cette famille $(Q_n)_{n\in\mathbb{N}}$ est totale dans l'espace E. En effet, celle-ci est constituée de polynômes de degrés étagés [1] et engendre donc l'espace $\mathcal{P}$ des fonctions polynomiales sur $[-1\,;1]$. Par le théorème de Weierstrass, $\mathcal{P}$ est une partie dense [2] de E pour la norme uniforme et donc pour la norme euclidienne qu'elle domine. On peut alors écrire (Th. 7 p. 247)

$$\underbrace{\sum_{k=0}^{n} (Q_k \,|\, f) Q_k}_{\text{projeté de } f \text{ sur } \mathbb{R}_n[X]} = \sum_{k=0}^{n} \frac{(P_k \,|\, f)}{\|P_k\|^2} P_k \xrightarrow[n\to+\infty]{} f$$

la convergence étant à comprendre au sens de la norme euclidienne.

Exercice 25 ** (Polynômes orthogonaux de Tchebychev)

Dans ce sujet, on identifie polynôme et fonction polynomiale associée sur $[-1\,;1]$.

On note E l'espace des fonctions continues de $[-1\,;1]$ vers $\mathbb{R}$ et, pour $f, g \in E$, on pose

$$\langle f, g \rangle = \int_{-1}^{1} \frac{f(t)g(t)}{\sqrt{1-t^2}}\,\mathrm{d}t.$$

(a) Montrer que $\langle \cdot\,,\cdot \rangle$ définit un produit scalaire sur E.

On considère la suite de polynômes $(T_n)_{n\in\mathbb{N}}$ déterminée par

$$T_0 = 1,\ T_1 = X \quad \text{et} \quad T_{n+1} = 2XT_n - T_{n-1} \quad \text{pour tout } n \geqslant 1.$$

(b) Soit $n \in \mathbb{N}$. Montrer $T_n(\cos\theta) = \cos(n\theta)$ pour tout réel θ.

(c) Établir que $(T_n)_{n\in\mathbb{N}}$ est une famille orthogonale totale de E.

1. Voir sujet 24 du chapitre 7 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.
2. Cette étude est déjà détaillée dans le sujet 4 p. 251.

Solution

(a) L'intégrale définissant $\langle f, g\rangle$ est généralisée en 1 et en -1. Commençons par justifier sa convergence.

méthode

La fonction $t \mapsto f(t)g(t)$ est bornée sur $[-1\,;1]$ et $t \mapsto \frac{1}{\sqrt{1-t^2}}$ est intégrable sur $]-1\,;1[$.

Soit f et g deux fonctions éléments de E. La fonction fg est continue sur le segment $[-1\,;1]$ donc bornée par un certain réel M. On a alors, pour tout $t \in]-1\,;1[$,

$$\left|\frac{f(t)g(t)}{\sqrt{1-t^2}}\right| = \frac{|f(t)g(t)|}{\sqrt{1-t^2}} \leqslant \frac{M}{\sqrt{1-t^2}}.$$

Or la fonction $t \mapsto \frac{1}{\sqrt{1-t^2}}$ est intégrable sur $]-1\,;1[$ car positive et de primitive $t \mapsto \arcsin t$ qui admet des limites finies en 1 et en -1. Par domination, on peut affirmer l'intégrabilité de $t \mapsto \frac{f(t)g(t)}{\sqrt{1-t^2}}$ sur $]-1\,;1[$ et donc la convergence [1] de l'intégrale définissant $\langle f, g\rangle$.

Ainsi, l'application $\langle \cdot\,, \cdot \rangle$ est bien définie de $E \times E$ vers $\mathbb{R}$. Vérifions ensuite qu'il s'agit d'une forme bilinéaire symétrique définie positive.

Pour $f, g, h \in E$ et $\lambda, \mu \in \mathbb{R}$, on vérifie aisément

$$\langle g, f\rangle = \langle f, g\rangle \quad \text{et} \quad \langle f, \lambda g + \mu h\rangle = \lambda\langle f, g\rangle + \mu\langle f, h\rangle.$$

Aussi, pour $f \in E$

$$\langle f, f\rangle = \int_{-1}^{1} \frac{f(t)^2}{\sqrt{1-t^2}}\,\mathrm{d}t \geqslant 0.$$

De plus, si $\langle f, f\rangle = 0$, on obtient une intégrale nulle d'une fonction continue et positive qui est donc la fonction nulle. On en déduit que f est nulle sur $]-1\,;1[$ puis sur $[-1\,;1]$ par continuité en 1 et en -1.

Finalement, $\langle \cdot\,, \cdot \rangle$ est un produit scalaire sur E.

(b) Soit $\theta \in \mathbb{R}$. On vérifie la propriété $T_n(\cos\theta) = \cos(n\theta)$ par récurrence double sur $n \in \mathbb{N}$.

Pour $n = 0$ et $n = 1$, la vérification est immédiate.

Supposons la propriété établie aux rangs $n-1$ et n (avec $n \geqslant 1$). On a

$$T_{n+1}(\cos\theta) = 2\cos(\theta)T_n(\cos\theta) - T_{n-1}(\cos\theta).$$

Par hypothèses de récurrence, on poursuit le calcul

$$T_{n+1}(\cos\theta) = 2\cos(\theta)\cos(n\theta) - \cos\big((n-1)\theta\big).$$

1. On peut aussi réaliser le changement de variable $t = \cos\theta$ qui transforme l'intégrale étudiée en une intégrale faussement généralisée.

On développe le terme $\cos\big((n-1)\theta\big)$ en $\cos(n\theta)\cos(\theta)+\sin(n\theta)\sin(\theta)$ et l'on obtient

$$T_{n+1}(\cos\theta)=\cos(n\theta)\cos(\theta)-\sin(n\theta)\sin(\theta)=\cos\big((n+1)\theta\big).$$

La récurrence est établie.

(c) Commençons par vérifier que la famille est orthogonale.

méthode

‖ On réalise le changement de variable $t=\cos\theta$.

Soit m et $n\in\mathbb{N}$ distincts. Par le changement de variable proposé avec $\theta\in[0\,;\pi]$

$$\langle T_n,T_m\rangle=\int_0^{\pi}\cos(n\theta)\cos(m\theta)\,\mathrm{d}\theta.$$

On linéarise l'expression trigonométrique par la formule

$$\cos(a)\cos(b)=\frac{1}{2}\big(\cos(a+b)+\cos(a-b)\big)$$

et l'on poursuit le calcul avec des divisions par $m-n$ et $m+n$ qui sont possibles car ces entiers sont non nuls :

$$\begin{aligned}\langle T_n,T_m\rangle&=\frac{1}{2}\int_0^{\pi}\cos\big((n+m)\theta\big)\,\mathrm{d}\theta+\frac{1}{2}\int_0^{\pi}\cos\big((n-m)\theta\big)\,\mathrm{d}\theta\\&=\frac{1}{2}\left[\frac{\sin\big((n+m)\theta\big)}{n+m}\right]_0^{\pi}+\frac{1}{2}\left[\frac{\sin\big((n-m)\theta\big)}{n-m}\right]_0^{\pi}=0.\end{aligned}$$

Montrons maintenant que la famille est totale, c'est-à-dire que l'espace qu'elle engendre est dense dans E. Par récurrence double [1], on vérifie que $\deg(T_n)=n$. La famille $(T_n)_{n\in\mathbb{N}}$ est donc une famille de polynômes de degrés étagés : elle engendre l'espace des fonctions polynomiales sur $[-1\,;1]$. Par le théorème de Weierstrass, ce dernier est dense dans E pour la norme $\|\cdot\|_\infty$ et l'est aussi pour la norme associée au produit scalaire $\langle\cdot\,,\cdot\rangle$ car celle-ci est dominée par $\|\cdot\|_\infty$. En effet, pour toute fonction f de E,

$$\|f\|=\sqrt{\langle f,f\rangle}=\left(\int_{-1}^{1}\frac{f(t)^2}{\sqrt{1-t^2}}\,\mathrm{d}t\right)^{1/2}\leqslant\left(\int_{-1}^{1}\frac{\|f\|_\infty^2}{\sqrt{1-t^2}}\,\mathrm{d}t\right)^{1/2}=\sqrt{\pi}\,\|f\|_\infty.$$

La famille $(T_n)_{n\in\mathbb{N}}$ est donc totale.

1. On retrouvera celle-ci détaillée dans le sujet 28 du chapitre 5 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

Exercice 26 **

Soit I un intervalle non vide de $\mathbb{R}$ et $\omega\colon I \to \mathbb{R}$ une fonction continue à valeurs strictement positives telle que $t \mapsto t^n\omega(t)$ est intégrable sur I pour tout $n \in \mathbb{N}$.
On munit $\mathbb{R}[X]$ du produit scalaire [1]

$$\langle P, Q\rangle = \int_I P(t)Q(t)\omega(t)\,\mathrm{d}t.$$

(a) Établir l'existence et l'unicité d'une suite $(P_n)_{n\in\mathbb{N}}$ formée de polynômes deux à deux orthogonaux et où chaque polynôme P_n est de degré n et de coefficient dominant 1.

Soit $n \geqslant 1$.

(b) Montrer que le polynôme $P_{n+1} - XP_n$ est orthogonal à tout polynôme de degré inférieur à $n-2$.

(c) En déduire l'existence de réels a_n et b_n tels que $P_{n+1} = (X - a_n)P_n - b_nP_{n-1}$.

(d) Vérifier

$$a_n = \frac{\langle XP_n, P_n\rangle}{\|P_n\|^2} \quad \text{et} \quad b_n = \frac{\langle XP_n, P_{n-1}\rangle}{\|P_{n-1}\|^2}.$$

Solution

(a) **méthode**

> Par analyse-synthèse, on montre l'existence et l'unicité de la famille $(P_n)_{n\in\mathbb{N}}$ en observant que P_n est choisi dans l'orthogonal de $\mathbb{R}_{n-1}[X]$.

Analyse : Supposons la famille $(P_n)_{n\in\mathbb{N}}$ convenable. Le polynôme P_0 est de degré 0 et de coefficient dominant 1, ce ne peut être que le polynôme constant égal à 1. Pour $n \geqslant 1$, le polynôme P_n est de degré n et orthogonal aux polynômes $P_0, \ldots, P_{n-1}$. Or ces derniers constituent une famille de polynômes de degrés étagés qui est une base de $\mathbb{R}_{n-1}[X]$. Le polynôme P_n appartient donc à la droite normale de l'hyperplan $\mathbb{R}_{n-1}[X]$ dans $\mathbb{R}_n[X]$. Sur cette droite il figure un seul polynôme de coefficient dominant 1 ce qui détermine P_n de façon unique [2].

Synthèse : Considérons la famille $(P_n)_{n\in\mathbb{N}}$ déterminée par les conditions : $P_0 = 1$ et, pour tout $n \geqslant 1$, P_n est le polynôme de coefficient dominant 1 figurant sur la droite normale de l'hyperplan $\mathbb{R}_{n-1}[X]$ dans $\mathbb{R}_n[X]$. Ce polynôme P_n est de degré n car appartient à $\mathbb{R}_n[X]$ sans appartenir à $\mathbb{R}_{n-1}[X]$ puisqu'il est non nul. Au surplus, il est orthogonal aux polynômes $P_0, \ldots, P_{n-1}$ car ceux-ci appartiennent tous à $\mathbb{R}_{n-1}[X]$.

Finalement, la suite $(P_n)_{n\in\mathbb{N}}$ ainsi définie est solution.

1. On vérifie aisément que cette application définit un produit scalaire notamment parce que la fonction ω est à valeurs strictement positives sur I ce qui entraîne $\langle P, P\rangle > 0$ pour $P \neq 0$.
2. On retrouve ici une mise en place de l'algorithme de Schmidt.

(b) **méthode**

Pour P et Q deux polynômes réels, on remarque

$$\langle XP, Q\rangle = \int_I tP(t)Q(t)\omega(t)\,\mathrm{d}t = \langle P, XQ\rangle.$$

Pour $n = 1$, un polynôme de degré inférieur à $n-2$ est le polynôme nul et la propriété est entendue. On suppose dans la suite $n \geqslant 2$. Soit Q un polynôme de $\mathbb{R}_{n-2}[X]$. On écrit

$$\langle P_{n+1} - XP_n, Q\rangle = \langle P_{n+1}, Q\rangle - \langle XP_n, Q\rangle = \langle P_{n+1}, Q\rangle - \langle P_n, XQ\rangle.$$

Lors de l'analyse de la question précédente on a vu que P_n est orthogonal à tout polynôme de $\mathbb{R}_{n-1}[X]$. Ici, XQ appartient à $\mathbb{R}_{n-1}[X]$ et donc $\langle P_n, XQ\rangle = 0$. Un argument semblable donne aussi $\langle P_{n+1}, Q\rangle = 0$ et donc $\langle P_{n+1} - XP_n, Q\rangle = 0$.

(c) Les polynômes P_{n+1} et XP_n sont tous deux de degré $n+1$ et de coefficient dominant 1. Il y a donc simplification des termes de plus haut degré dans le calcul de $P_{n+1} - XP_n$ ce qui permet d'affirmer

$$P_{n+1} - XP_n \in \mathbb{R}_n[X].$$

Le polynôme $P_{n+1} - XP_n$ est donc combinaison linéaire des polynômes $P_0, \ldots, P_n$ qui constituent une base orthogonale de $\mathbb{R}_n[X]$. Cependant, $P_{n+1} - XP_n$ est aussi orthogonal aux polynômes $P_0, \ldots, P_{n-2}$ et est donc seulement combinaison linéaire des polynômes P_{n-1} et P_n. On peut donc écrire

$$P_{n+1} - XP_n = -a_nP_n - b_nP_{n-1} \quad \text{avec} \quad a_n, b_n \in \mathbb{R}$$

ce qui donne, après réorganisation des membres, la relation voulue.

(d) **méthode**

Le polynôme P_{n+1} est orthogonal aux polynômes P_n et P_{n-1}.

D'une part, $\langle P_{n+1}, P_n\rangle = 0$ donne par linéarité

$$\langle XP_n, P_n\rangle - a_n \underbrace{\langle P_n, P_n\rangle}_{=\|P_n\|^2} - b_n \underbrace{\langle P_{n-1}, P_n\rangle}_{=0} = 0 \quad \text{donc} \quad a_n = \frac{\langle XP_n, P_n\rangle}{\|P_n\|^2}.$$

D'autre part, $\langle P_{n+1}, P_{n-1}\rangle = 0$ fournit

$$\langle XP_n, P_{n-1}\rangle - a_n \underbrace{\langle P_n, P_{n-1}\rangle}_{=0} - b_n \underbrace{\langle P_{n-1}, P_{n-1}\rangle}_{=\|P_{n-1}\|^2} \quad \text{donc} \quad b_n = \frac{\langle XP_n, P_{n-1}\rangle}{\|P_{n-1}\|^2}.$$

On peut aussi remarquer $\langle XP_n, P_{n-1}\rangle = \langle P_n, XP_{n-1}\rangle = \|P_n\|^2$.

6.5 Exercices d'approfondissement

Exercice 27 *

Soit E un espace vectoriel euclidien.

Montrer que l'ensemble $\{(x,y) \in E^2 \mid (x,y) \text{ libre}\}$ est un ouvert de E^2.

Solution

méthode

Par l'étude du cas d'égalité dans l'inégalité de Cauchy-Schwarz on sait :

$$(x,y) \text{ est libre} \iff |(x\,|\,y)| < \|x\|\,\|y\|$$

Considérons l'application $f\colon E^2 \to \mathbb{R}$ définie par

$$f(x,y) = \|x\|\,\|y\| - (x\,|\,y).$$

L'application f est continue sur E^2 car la norme $\|\cdot\|$ est continue et le produit scalaire est continue[1] car bilinéaire au départ d'un espace de dimension finie. L'ensemble $\{(x,y) \in E^2 \mid (x,y) \text{ libre}\}$ est l'image réciproque de l'ouvert $]0\,;+\infty[$ par cette application continue, c'est donc un ouvert relatif à E^2, c'est-à-dire un ouvert de E^2.

Exercice 28 ** (Produit vectoriel)

Soit $\vec{u}$ et $\vec{v}$ deux vecteurs d'un espace euclidien orienté E de dimension 3.

(a) Montrer qu'il existe un unique vecteur noté $\vec{u} \wedge \vec{v}$ dans E vérifiant[2]

$$[\vec{u},\vec{v},\vec{x}] = (\vec{u} \wedge \vec{v}\,|\,\vec{x}) \quad \text{pour tout } \vec{x} \in E.$$

Le vecteur $\vec{u} \wedge \vec{v}$ est appelé *produit vectoriel* de $\vec{u}$ par $\vec{v}$.

(b) Vérifier que $\vec{u} \wedge \vec{v}$ est un vecteur orthogonal à $\vec{u}$ et $\vec{v}$.

(c) Montrer que la famille $(\vec{u},\vec{v})$ est libre si, et seulement si, $\vec{u} \wedge \vec{v}$ est non nul. Observer que la famille $(\vec{u},\vec{v},\vec{u} \wedge \vec{v})$ est alors une base directe.

(d) On introduit une base orthonormale directe $\mathcal{B} = (\vec{i},\vec{j},\vec{k})$ telle que $\vec{u} \in \text{Vect}(\vec{i})$ et $\vec{v} \in \text{Vect}(\vec{i},\vec{j})$. Exprimer le vecteur $\vec{u} \wedge \vec{v}$ et vérifier la formule[3]

$$(\vec{u}\,|\,\vec{v})^2 + \|\vec{u} \wedge \vec{v}\|^2 = \|\vec{u}\|^2\,\|\vec{v}\|^2.$$

1. Plus généralement, le produit scalaire est continue pour la norme euclidienne associée.

2. $[\vec{u},\vec{v},\vec{w}]$ désigne le produit mixte des vecteurs $\vec{u},\vec{v}$ et $\vec{w}$, c'est-à-dire le déterminant de la famille $(\vec{u},\vec{v},\vec{w})$ dans une base orthonormale directe de E.

3. D'autres propriétés classiques sur le produit vectoriel peuvent être établies comme sa bilinéarité ou la formule du double produit vectoriel $\vec{u} \wedge (\vec{v} \wedge \vec{w}) = (\vec{u}\,|\,\vec{w})\vec{v} - (\vec{u}\,|\,\vec{v})\vec{w}$. Notons que le produit vectoriel n'est pas associatif et qu'il est anticommutatif $\vec{v} \wedge \vec{u} = -(\vec{u} \wedge \vec{v})$.

Solution

(a) **méthode**

> On introduit $\vec{u}\wedge\vec{v}$ par le théorème de représentation des formes linéaires (Th. 1 p. 245).

Le produit mixte est une forme multilinéaire sur E, l'application $\varphi\colon \vec{x}\mapsto\left[\vec{u},\vec{v},\vec{x}\right]$ est donc une forme linéaire sur l'espace euclidien E. Il existe alors un unique vecteur $\vec{u}\wedge\vec{v}$ vérifiant $\varphi(\vec{x})=(\vec{u}\wedge\vec{v}\,|\,\vec{x}\,)$ pour tout $\vec{x}\in E$.

(b) **méthode**

> Le produit mixte d'une famille de vecteurs est nul si, et seulement si, cette famille est liée.

On en déduit

$$(\vec{u}\wedge\vec{v}\,|\,\vec{u}\,)=\left[\vec{u},\vec{v},\vec{u}\right]=0\quad\text{et}\quad(\vec{u}\wedge\vec{v}\,|\,\vec{v})=\left[\vec{u},\vec{v},\vec{v}\right]=0.$$

(c) Raisonnons par double implication.

($\Longrightarrow$) Si la famille $(\vec{u},\vec{v}\,)$ est liée, $\varphi(\vec{x})=\left[\vec{u},\vec{v},\vec{x}\right]=0$ pour tout $\vec{x}\in E$ et l'unique vecteur représentant la forme linéaire nulle est le vecteur nul : $\vec{u}\wedge\vec{v}=\vec{0}$.

($\Longleftarrow$) Supposons la famille $(\vec{u},\vec{v}\,)$ libre. On peut introduire un vecteur $\vec{w}$ complétant celle-ci en une base et alors $\left[\vec{u},\vec{v},\vec{w}\right]\neq 0$ ce qui entraîne $\vec{u}\wedge\vec{v}\neq\vec{0}$.

Au surplus, si tel est le cas

$$\left[\vec{u},\vec{v},\vec{u}\wedge\vec{v}\right]=(\vec{u}\wedge\vec{v}\,|\,\vec{u}\wedge\vec{v}\,)=\|\vec{u}\wedge\vec{v}\,\|^2>0$$

et la famille $(\vec{u},\vec{v},\vec{u}\wedge\vec{v}\,)$ est une base directe.

(d) On introduit les coordonnées des vecteurs $\vec{u}$ et $\vec{v}$ dans $\mathcal{B}$

$$\vec{u}=u_1\vec{\imath}\quad\text{et}\quad\vec{v}=v_1\vec{\imath}+v_2\vec{\jmath}\quad\text{avec}\quad u_1,v_1,v_2\in\mathbb{R}.$$

Pour tout $\vec{x}\in E$ de coordonnées x_1,x_2,x_3 dans e, on a alors

$$\left[\vec{u},\vec{v},\vec{x}\right]=\begin{vmatrix}u_1 & v_1 & x_1\\ 0 & v_2 & x_2\\ 0 & 0 & x_3\end{vmatrix}=u_1v_2x_3=\langle u_1v_2\vec{k},\vec{x}\rangle.$$

Par conséquent, $\vec{u}\wedge\vec{v}=u_1v_2\vec{k}$ et donc

$$(\vec{u}\,|\,\vec{v}\,)^2+\|\vec{u}\wedge\vec{v}\,\|^2=(u_1v_1)^2+(u_1v_2)^2=u_1^2\left(v_1^2+v_2^2\right)=\|\vec{u}\,\|^2\,\|\vec{v}\,\|^2.$$

Les différentes propriétés obtenues permettent de positionner le vecteur $\vec{u} \wedge \vec{v}$.

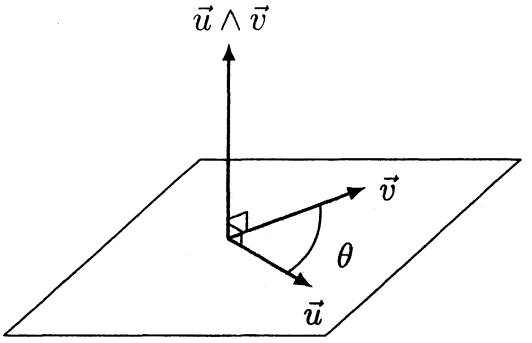
Lorsqu'il n'est pas nul :

— il est orthogonal à $\vec{u}$ et $\vec{v}$;

— la famille $(\vec{u}, \vec{v}, \vec{u} \wedge \vec{v})$ est directe ;

— $\|\vec{u} \wedge \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin\theta$

avec θ l'angle géométrique[1] formé par $\vec{u}$ et $\vec{v}$.



Exercice 29 **

Soit E un espace préhilbertien réel.

(a) Établir l'inclusion $\overline{F} \subset (F^\perp)^\perp$ pour tout sous-espace vectoriel F de E.

On se propose d'établir par un exemple que cette inclusion peut être stricte. On introduit pour cela l'espace $E = \mathbb{R}[X]$ muni du produit scalaire donné par

$$(P \mid Q) = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t.$$

(b) Montrer que

$$H = \left\{ P \in \mathbb{R}[X] \mid \int_{-1}^{1} |t| \, P(t)\,\mathrm{d}t = 0 \right\}$$

est un hyperplan fermé de E.

(c) Soit $Q \in H^\perp$. Établir que, pour tout $P \in \mathbb{R}[X]$,

$$\int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t = \left(\int_{-1}^{1} |t| \, P(t)\,\mathrm{d}t \right) \left(\int_{-1}^{1} Q(t)\,\mathrm{d}t \right).$$

(d) Vérifier que $H^\perp = \{0\}$ et conclure.

Solution

(a) On sait $F \subset (F^\perp)^\perp$ et l'on sait que $(F^\perp)^\perp$ est fermée car l'orthogonal d'une partie est un fermé[2]. Les limites des suites convergentes d'éléments de F appartiennent donc à $(F^\perp)^\perp$, c'est-à-dire $\overline{F} \subset (F^\perp)^\perp$.

(b) **méthode**

‖ On montre que H est le noyau d'une forme linéaire continue.

1. L'*angle géométrique* formé par deux vecteurs $\vec{u}$ et $\vec{v}$ est l'unique angle θ de $[0\,;\pi]$ pour lequel on vérifie $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \, \|\vec{v}\| \cos\theta$. Ce n'est pas un angle orienté mais seulement une mesure de l'écart angulaire formé par deux vecteurs : un angle orienté ne peut être défini qu'à l'intérieur d'un plan orienté et correspond à une mesure modulo 2π.

2. Voir sujet 7 p. 253.

Par définition, H est le noyau de la forme linéaire non nulle

$$\varphi\colon P \mapsto \int_{-1}^{1} |t|\, P(t)\, \mathrm{d}t.$$

Par l'inégalité de Cauchy-Schwarz, on a pour tout P de $\mathbb{R}[X]$

$$\left|\varphi(P)\right| \leqslant \left(\int_{-1}^{1} t^2\, \mathrm{d}t\right)^{1/2} \left(\int_{-1}^{1} P(t)^2\, \mathrm{d}t\right)^{1/2} = \sqrt{\frac{2}{3}}\, \|P\|\,.$$

Ceci détermine un réel k pour lequel $\left|\varphi(P)\right| \leqslant k\,\|P\|$ ce qui assure que l'application linéaire φ est continue. Son noyau, qui est l'image réciproque du fermé $\{0\}$, est donc un hyperplan qui est un fermé relatif à E donc un fermé de E.

(c) **méthode**

> On interprète la différence des deux membres comme le produit scalaire de Q avec un polynôme de H.

Soit $P \in \mathbb{R}[X]$. On étudie la différence des deux membres où l'on considère l'intégrale de $t \mapsto |t|\, P(t)$ comme une constante réelle λ

$$\int_{-1}^{1} P(t)Q(t)\, \mathrm{d}t - \underbrace{\left(\int_{-1}^{1} |t|\, P(t)\, \mathrm{d}t\right)}_{=\lambda\in\mathbb{R}} \left(\int_{-1}^{1} Q(t)\, \mathrm{d}t\right) = \int_{-1}^{1} \underbrace{\big(P(t) - \lambda\big)}_{=R(t)} Q(t)\, \mathrm{d}t.$$

Le polynôme R introduit appartient à H car

$$\int_{-1}^{1} |t|\, R(t)\, \mathrm{d}t = \int_{-1}^{1} |t|\, P(t)\, \mathrm{d}t - \int_{-1}^{1} \lambda\, |t|\, \mathrm{d}t = \lambda - \lambda = 0.$$

Le polynôme Q appartenant à l'orthogonal de H, on a $(R\,|\,Q) = 0$ ce qui produit l'égalité voulue.

(d) En considérant cette fois l'intégrale de Q comme une constante réelle μ, on obtient pour tout $P \in \mathbb{R}[X]$,

$$\int_{-1}^{1} P(t)Q(t)\, \mathrm{d}t - \left(\int_{-1}^{1} |t|\, P(t)\, \mathrm{d}t\right) \underbrace{\left(\int_{-1}^{1} Q(t)\, \mathrm{d}t\right)}_{=\mu\in\mathbb{R}} = \int_{-1}^{1} P(t)\big(Q(t) - \mu\, |t|\big)\, \mathrm{d}t = 0.$$

Ceci entraîne [1] que la fonction $t \mapsto Q(t) - \mu\, |t|$ est nulle sur $[-1\,;1]$. Cependant, la fonction valeur absolue n'est pas polynomiale sur $[-1\,;1]$ et l'on a donc nécessairement $\mu = 0$ puis $Q = 0$. Ainsi, $H^{\perp} = \{0\}$.

Finalement, on a $(H^{\perp})^{\perp} = E$ alors que $\overline{H} = H \neq E$: l'inclusion $\overline{H} \subset \big(H^{\perp}\big)^{\perp}$ est stricte.

1. Voir sujet 4 p. 251.

Exercice 30 ** (Quadrature par la méthode de Gauss)

Soit $a < b$ deux réels. Dans ce sujet, on identifie polynôme et fonction polynomiale associée sur $[a\,;b]$. On munit l'espace $E = \mathcal{C}\big([a\,;b],\mathbb{R}\big)$ du produit scalaire

$$(f\,|\,g) = \int_a^b f(t)g(t)\,\mathrm{d}t.$$

Soit n un entier naturel.

(a) Montrer qu'il existe un polynôme G de degré $n+1$ orthogonal à tout polynôme de $\mathbb{R}_n[X]$.

(b) Montrer que le polynôme G admet exactement $n+1$ racines distinctes toutes dans l'intervalle $]a\,;b[$.

On note $x_0,\ldots,x_n$ les racines de G, $\omega_0,\ldots,\omega_n$ des réels et l'on pose, pour toute fonction f de E,

$$\mathcal{E}(f) = \int_a^b f(t)\,\mathrm{d}t - \sum_{k=0}^{n}\omega_k f(x_k).$$

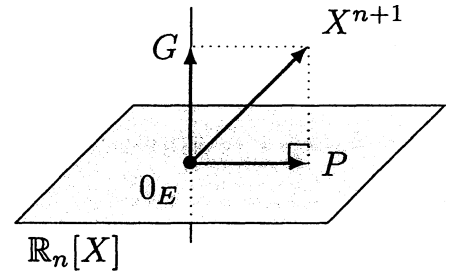
(c) Montrer qu'il est possible de choisir les réels $\omega_0,\ldots,\omega_n$ de sorte que $\mathcal{E}(P) = 0$ pour tout polynôme P de $\mathbb{R}_n[X]$.

(d) Vérifier alors que $\mathcal{E}(P)$ est aussi nul pour tout polynôme P de degré inférieur à $2n+1$.

(e) Justifier que les ω_k sont tous positifs.

Solution

(a) $\mathbb{R}_n[X]$ est un sous-espace vectoriel de dimension finie de E, on peut donc introduire la projection orthogonale sur celui-ci (Th. 3 p. 245). On considère alors le projeté orthogonal P de X^{n+1} et $G = X^{n+1} - P$. Le polynôme G convient car de degré $n+1$ et orthogonal[1] à tout polynôme de $\mathbb{R}_n[X]$.



(b) **méthode**

‖ On introduit un polynôme ayant le même signe que G sur $[a\,;b]$.

Notons $x_1,\ldots,x_p$ les racines de multiplicités impaires de G appartenant à $]a\,;b[$ et considérons le polynôme

$$P = (X - x_1)\ldots(X - x_p).$$

Le produit GP détermine un polynôme de signe constant sur $[a\,;b]$ car ses racines sur $]a\,;b[$ sont de multiplicités paires. La fonction $t \mapsto G(t)P(t)$ est alors continue sur $[a\,;b]$, de signe constant, sans être la fonction nulle : son intégrale sur $[a\,;b]$ ne peut être nulle.

1. On peut aussi employer que $\mathbb{R}_n[X]$ est un hyperplan de $\mathbb{R}_{n+1}[X]$: tout élément non nul de sa droite normale convient.

Le polynôme P est donc de degré au moins égal à $n+1$ et G possède au moins $n+1$ racines de multiplicités impaires dans $]a\,;b[$. Or le polynôme G est de degré $n+1$ et l'on peut donc affirmer que G possède exactement $n+1$ racines, toutes simples et dans l'intervalle $]a\,;b[$.

(c) **méthode**

‖ On introduit les polynômes interpolateurs de Lagrange en les $x_0,\dots,x_n$.

Pour $k\in[\![0\,;n]\!]$, notons L_k le polynôme de degré n prenant la valeur 1 en x_k et la valeur 0 en x_j pour tout $j\neq k$. Pour tout P polynôme de $\mathbb{R}_n[X]$, on peut écrire

$$P=\sum_{k=0}^{n}P(x_k)L_k$$

car[1] les polynômes dans les deux membres sont de degrés inférieurs à n et prennent les mêmes valeurs en les $n+1$ points $x_0,\dots,x_n$. On a alors par linéarité de l'intégrale

$$\int_a^b P(t)\,\mathrm{d}t=\sum_{k=0}^{n}\left(P(x_k)\int_a^b L_k(t)\,\mathrm{d}t\right)=\sum_{k=0}^{n}\omega_k P(x_k)\quad\text{avec}\quad\omega_k=\int_a^b L_k(t)\,\mathrm{d}t.$$

Ainsi, pour ces valeurs ω_k indépendantes de P, on a $\mathcal{E}(P)=0$.

(d) Soit P un polynôme de degré inférieur à $2n+1$.

méthode

‖ On réalise la division euclidienne de P par G.

La division euclidienne de P par G s'écrit

$$P=GQ+R\quad\text{avec}\quad\deg(R)\leqslant n.$$

Puisque le polynôme P est de degré inférieur à $2n+1$, le polynôme quotient Q est de degré inférieur à n et est donc orthogonal à G. Par conséquent,

$$\int_a^b P(t)\,\mathrm{d}t=\underbrace{(G\,|\,Q)}_{=0}+\int_0^1 R(t)\,\mathrm{d}t.$$

Or le polynôme R est de degré inférieur à n et prend les mêmes valeurs que P en les x_k qui sont les racines de G

$$\int_a^b R(t)\,\mathrm{d}t=\sum_{k=0}^{n}\omega_k R(x_k)=\sum_{k=0}^{n}\omega_k P(x_k).$$

On en déduit

$$\int_a^b P(t)\,\mathrm{d}t=\sum_{k=0}^{n}\omega_k P(x_k)\quad\text{puis}\quad\mathcal{E}(P)=0.$$

1. La famille des polynômes de Lagrange en $n+1$ points est une base de $\mathbb{R}_n[X]$.

(e) Soit $j \in [\![0\,;n]\!]$. Le polynôme $P = L_j^2$ est de degré $2n$ et donc

$$\omega_j = \sum_{k=1}^{n} \omega_k L_j^2(\omega_k) = \int_a^b L_j(t)^2 \,\mathrm{d}t \geqslant 0.$$

Exercice 31 ***

Soit E un espace vectoriel réel et $\|\cdot\|$ une norme sur E.

Montrer que la norme $\|\cdot\|$ est euclidienne[1] si, et seulement si,

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big) \quad \text{pour tout } (x,y) \in E^2.$$

Solution

Raisonnons par double implication.

($\Longrightarrow$) Supposons la norme $\|\cdot\|$ euclidienne et notons $(\cdot\,|\,\cdot)$ le produit scalaire associé. Par les identités remarquables

$$\|x+y\|^2 = \|x\|^2 + 2(x\,|\,y) + \|y\|^2 \tag{1}$$
$$\|x-y\|^2 = \|x\|^2 - 2(x\,|\,y) + \|y\|^2 \tag{2}$$

on obtient directement l'identité voulue[2].

($\Longleftarrow$) Supposons l'identité du sujet vérifiée. Une petite analyse est nécessaire pour proposer un produit scalaire dont la norme serait issue. En exploitant l'identité remarquable (1) on pourrait proposer

$$(x\,|\,y) = \frac{1}{2}\Big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\Big)$$

mais cette expression semble mal se prêter à l'hypothèse en cours. En considérant la différence des relations (1) et (2), on propose plutôt

$$(x\,|\,y) = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2\Big).$$

Considérons donc l'application $\varphi\colon E \times E \to \mathbb{R}$ définie par la formule

$$\varphi(x,y) = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2\Big).$$

L'application φ est symétrique et vérifie $\varphi(x,x) = \|x\|^2 > 0$ pour tout $x \in E$ non nul. La difficulté est d'établir qu'elle est bilinéaire. Compte tenu de la symétrie, il suffit d'étudier la linéarité en la deuxième variable.

Soit x, y, z trois vecteurs de E.

1. C'est-à-dire que la norme $\|\cdot\|$ est la norme euclidienne associée à un produit scalaire sur E.
2. Celle-ci est l'identité du parallélogramme : la somme des carrés des diagonales d'un parallélogramme est la somme des carrés de ses quatre côtés.

méthode

On vérifie pour commencer $\varphi(2x, y+z) = 2\varphi(x,y) + 2\varphi(x,z)$.

On écrit

$$\varphi(2x, y+z) = \frac{1}{4}\Big(\big\|(x+y)+(x+z)\big\|^2 - \big\|(x-y)+(x-z)\big\|^2\Big).$$

Grâce à la propriété vérifiée par $\|\cdot\|$, on a

$$\big\|(x+y)+(x+z)\big\|^2 = 2\|x+y\|^2 + 2\|x+z\|^2 - \|y-z\|^2 \text{ et}$$
$$\big\|(x-y)+(x-z)\big\|^2 = 2\|x-y\|^2 + 2\|x-z\|^2 - \|z-y\|^2.$$

Après simplification et organisation des termes

$$\varphi(2x, y+z) = \frac{1}{2}\Big(\|x+y\|^2 - \|x-y\|^2 + \|x+z\|^2 - \|x-z\|^2\Big) = 2\varphi(x,y) + 2\varphi(x,z).$$

En particularisant cette relation à $z = 0_E$, il vient $\varphi(2x, y) = 2\varphi(x,y)$ car $\varphi(x, 0_E) = 0$. On en déduit que, pour tous x, y et z dans E,

$$\varphi(x, y+z) = \varphi(x,y) + \varphi(x,z).$$

Il reste à justifier l'identité $\varphi(x, \lambda y) = \lambda\varphi(x,y)$ pour tout λ réel.

méthode

On introduit la fonction $f : \lambda \mapsto \varphi(x, \lambda y)$ définie sur $\mathbb{R}$ et l'on vérifie que celle-ci est continue et additive.

Soit λ et μ deux réels. Par l'étude qui précède

$$f(\lambda+\mu) = \varphi\big(x, (\lambda+\mu)y\big) = \varphi(x, \lambda y + \mu y) = \varphi(x, \lambda y) + \varphi(x, \mu y) = f(\lambda) + f(\mu).$$

L'application f est donc additive. Elle est aussi continue par opérations sur les fonctions car la norme $\|\cdot\|$ est continue. On sait [1] alors que f est une fonction linéaire. On peut donc écrire

$$\varphi(x, \lambda y) = f(\lambda) = \lambda f(1) = \lambda\varphi(x,y) \quad \text{pour tout } \lambda \in \mathbb{R}.$$

Finalement, φ est bien un produit scalaire sur E et la norme introduite est la norme euclidienne associée.

1. Voir par exemple le sujet 22 du chapitre 7 de l'ouvrage *Exercices d'analyse MPSI*.

CHAPITRE 7

Endomorphismes des espaces euclidiens

E désigne un espace euclidien de produit scalaire [1] $(\cdot \mid \cdot)$ et n un entier naturel non nul.

7.1 Isométries vectorielles

7.1.1 Définition

Définition

On appelle *isométrie vectorielle* [2] de E tout endomorphisme u de E conservant la norme euclidienne, c'est-à-dire vérifiant $\|u(x)\| = \|x\|$ pour tout $x \in E$.

Les isométries vectorielles conservent aussi le produit scalaire et par conséquent l'orthogonalité. Elles transforment une base orthonormale en une base orthonormale et cette propriété caractérise les isométries parmi les endomorphismes.

L'ensemble $\mathrm{O}(E)$ des isométries de E est un groupe pour la composition des applications, on l'appelle *groupe orthogonal* de E.

Théorème 1

Si F est un sous-espace vectoriel stable par une isométrie vectorielle u, le supplémentaire orthogonal $F^{\perp}$ est aussi stable par u.

1. On utilisera aussi indifféremment la notation $\langle \cdot , \cdot \rangle$.
2. On parle aussi d'*automorphisme orthogonal*.

7.1.2 Réduction d'une isométrie en base orthonormale

Les isométries vectorielles correspondent aux endomorphismes représentés par une matrice orthogonale [1] en base orthonormale. En choisissant correctement cette base, on peut proposer une représentation simplifiée :

Théorème 2

Si u est une isométrie vectorielle de E, il existe une base orthonormale de E dans laquelle la matrice de u est diagonale par blocs avec des blocs diagonaux de la forme

$$(1), \quad (-1) \quad \text{ou} \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \not\equiv 0 \ [\pi].$$

Autrement dit, E est la somme directe orthogonale des espaces $E_1(u)$, $E_{-1}(u)$ et de plans sur lesquels l'endomorphisme induit par u est une rotation non triviale.

Toute matrice orthogonale est donc semblable [2] par une matrice de passage orthogonale à une matrice diagonale par blocs avec des blocs diagonaux de la forme

$$(1), \quad (-1) \quad \text{ou} \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \not\equiv 0 \ [\pi].$$

7.1.3 Isométries vectorielles positives en dimension 3

Soit E un espace euclidien orienté de dimension 3. À cause de la nature géométrique de ce qui suit, on adopte une notation fléchée des vecteurs.

Orientation induite

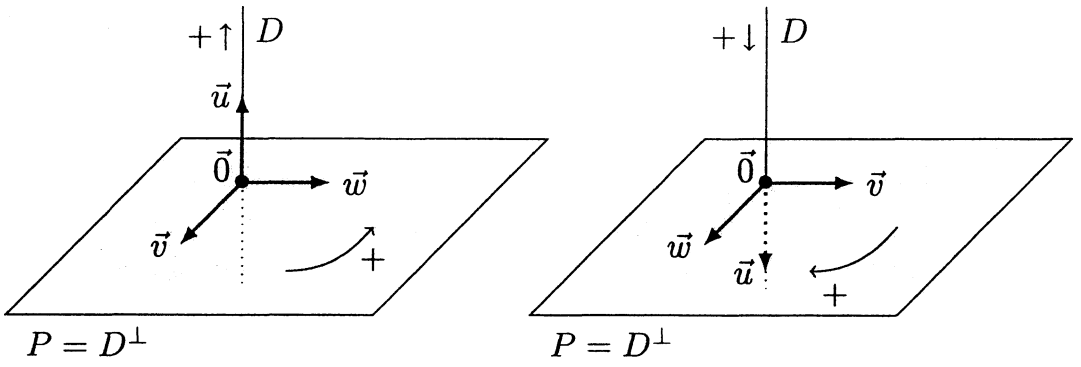
Soit P un plan de l'espace E et $D = P^\perp$ sa droite normale. Il n'existe pas a *priori* d'orientation préférentielle ni sur P, ni sur D.

On choisit arbitrairement une orientation sur D par l'introduction d'un vecteur unitaire $\vec{u}$ déterminant le sens positif : on dit alors que la droite D est un *axe*.

On complète $\vec{u}$ en une base orthonormale directe $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ de l'espace E. La famille $(\vec{v}, \vec{w})$ est alors une base orthonormale du plan P. En choisissant celle-ci pour base orientée de référence de P, on dit que l'on munit le plan P de l'*orientation induite* par celle de D.

1. Une matrice orthogonale de taille n est une matrice A de $\mathcal{M}_n(\mathbb{R})$ vérifiant ${}^tAA = I_n$. L'ensemble $O_n(\mathbb{R})$ de ces matrices est un groupe multiplicatif que l'on appelle *groupe orthogonal* d'ordre n.

2. On dit aussi *orthogonalement semblable* pour signifier que la matrice de passage peut être choisie orthogonale.



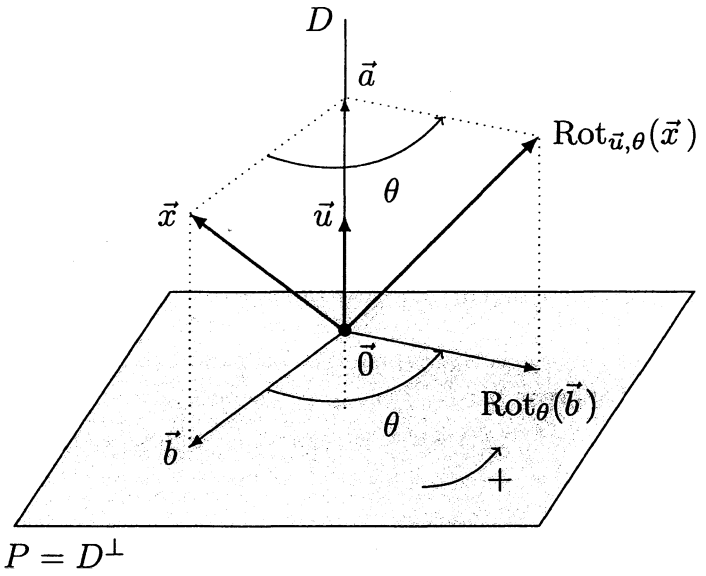
Si l'on renverse l'orientation de D, l'orientation induite sur P est elle aussi inversée.

Rotation de l'espace

Une isométrie positive[1] f de E peut être représentée dans une base orthonormale directe $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ par la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

On introduit alors la droite D dirigée et orientée par le vecteur $\vec{u}$ et le plan $P = \text{Vect}(\vec{v}, \vec{w})$ muni de l'orientation induite. L'isométrie f agit comme l'identité sur l'axe $D = \text{Vect}(\vec{u})$ et comme la rotation d'angle θ sur le plan $P = D^\perp$.



Définition

On dit que f est la *rotation d'axe dirigé et orienté par $\vec{u}$ et d'angle θ*. On la note $\text{Rot}_{\vec{u},\theta}$.

Les isométries positives de E se limitent aux rotations qui viennent d'être décrites, on dit que $\text{SO}(E)$ est le *groupe des rotations* de E.

Il est facile de composer deux rotations de même axe et celles-ci commutent :

$$\forall (\theta, \theta') \in \mathbb{R}^2, \quad \text{Rot}_{\vec{u},\theta} \circ \text{Rot}_{\vec{u},\theta'} = \text{Rot}_{\vec{u},\theta+\theta'} = \text{Rot}_{\vec{u},\theta'} \circ \text{Rot}_{\vec{u},\theta}.$$

L'inverse d'une rotation est une rotation de même axe et d'angle opposé.
La composée de deux rotations d'axes différents est une rotation mais celle-ci n'est pas immédiate à caractériser.

1. Les isométries de E se partitionnent en les isométries positives de déterminant 1 et les négatives de déterminant -1. Les isométries positives conservent l'orientation et l'on parle parfois d'*isométries directes*.

Classification

Soit f une isométrie de l'espace E autre que l'identité. Par le théorème Th. 2 p. 288, on peut affirmer :

- si l'espace $\mathrm{Ker}(f - \mathrm{Id}_E)$ est une droite vectorielle, l'endomorphisme f est de déterminant 1, c'est une rotation autour de cette droite [1] ;
- si l'espace $\mathrm{Ker}(f - \mathrm{Id}_E)$ est un plan vectoriel, f est la réflexion par rapport à ce plan.

Le cas où l'espace $\mathrm{Ker}(f - \mathrm{Id}_E)$ est réduit au vecteur nul sort du cadre du programme. On peut cependant établir que f est alors la composée d'une réflexion et d'une rotation autour de la droite normale au plan de réflexion.

7.2 Endomorphismes symétriques

7.2.1 Définition

Définition

Un endomorphisme u de E est dit *symétrique* lorsque $\big(u(x)\,|\,y\big) = \big(x\,|\,u(y)\big)$ pour tous les vecteurs x et y de E.

Les projections orthogonales sont des endomorphismes symétriques et, inversement, une projection qui est un endomorphisme symétrique est orthogonale.

Théorème 3

Soit $u \in \mathcal{L}(E)$ et $e = (e_1, \ldots, e_n)$ une base orthonormale de E. On a équivalence entre :

(i) u est symétrique ;
(ii) la matrice de u dans e est symétrique.

L'ensemble $\mathcal{S}(E)$ des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathcal{L}(E)$ isomorphe à l'espace $\mathcal{S}_n(\mathbb{R})$ des matrices réelles de taille n. L'espace $\mathcal{S}(E)$ est donc de dimension

$$\frac{n(n+1)}{2}.$$

7.2.2 Réduction des endomorphismes symétriques

Théorème 4

Si F est un sous-espace vectoriel de E stable par un endomorphisme symétrique u, l'espace $F^\perp$ est aussi stable par u.

Les endomorphismes induits par u sur F et $F^\perp$ sont encore symétriques.

1. On verra dans le sujet 2 p. 292 comment en déterminer l'angle.

Théorème 5 (Théorème spectral)
Si u est un endomorphisme symétrique d'un espace euclidien E alors E est la somme directe orthogonale des sous-espaces propres de u.

En conséquence, tout endomorphisme symétrique est diagonalisable dans une base orthonormale [1].

Théorème 6 (Théorème spectral matriciel)
Toute matrice symétrique réelle est diagonalisable par une matrice de passage orthogonale [2].

Ainsi, toute matrice A de $\mathcal{S}_n(\mathbb{R})$ peut s'écrire $A = PDP^{-1}$, ou encore $A = PD\,{}^tP$, avec $P \in \mathrm{O}_n(\mathbb{R})$ et D diagonale.

Une matrice symétrique complexe peut ne pas être diagonalisable.

7.3 Exercices d'apprentissage

7.3.1 Isométries vectorielles

Exercice 1
Soit u une isométrie vectorielle d'un espace euclidien E de produit scalaire $\langle \cdot\,,\,\cdot \rangle$.

(a) Quelles sont les valeurs propres réelles possibles de u ?

(b) Vérifier que les espaces propres associés sont orthogonaux.

(c) On suppose que F et G sont deux sous-espaces vectoriels orthogonaux de E. Que dire des espaces images $u(F)$ et $u(G)$?

Solution

méthode
Une isométrie est un endomorphisme qui conserve le produit scalaire, la norme et l'orthogonalité.

(a) Soit $\lambda \in \mathbb{R}$ une valeur propre de u et x un vecteur propre associé : x est un vecteur non nul vérifiant $u(x) = \lambda x$. En considérant la norme des deux membres de cette égalité on obtient

$$\big\|u(x)\big\| = \|\lambda x\| = |\lambda|\,\|x\|\,.$$

Or $\big\|u(x)\big\| = \|x\|$ car u conserve la norme et donc $|\lambda|\,\|x\| = \|x\|$. En simplifiant par $\|x\|$ qui est non nul, on obtient $|\lambda| = 1$. Ainsi, les seules valeurs propres réelles possibles d'une isométrie vectorielle sont 1 et -1.

1. On dit que les endomorphismes symétriques sont *orthogonalement diagonalisables.*
2. On dit que les matrices symétriques réelles sont *orthogonalement diagonalisables.*

(b) Soit x et y des vecteurs [1] appartenant aux espaces $E_1(u)$ et $E_{-1}(u)$. On a $u(x) = x$ et $u(y) = -y$ donc

$$\langle u(x), u(y)\rangle = \langle x, -y\rangle = -\langle x, y\rangle.$$

Cependant, $\langle u(x), u(y)\rangle = \langle x, y\rangle$ car l'isométrie conserve le produit scalaire. On a donc $\langle x, y\rangle = 0$ et l'on peut affirmer que les espaces $E_1(u)$ et $E_{-1}(u)$ sont orthogonaux.

Soulignons que ces deux résultats sont des conséquences directes du théorème de réduction des isométries en base orthonormale (Th. 2 p. 288).

(c) Si F et G sont orthogonaux, on a $F \subset G^\perp$ et donc $u(F) \subset u(G^\perp)$. Cependant, on sait aussi $u(G^\perp) = \big(u(G)\big)^\perp$ (Th. 1 p. 287) et donc $u(F) \subset \big(u(G)\big)^\perp$: les espaces $u(F)$ et $u(G)$ sont orthogonaux.

Exercice 2

Soit E un espace vectoriel euclidien orienté muni d'une base orthonormale directe $\mathcal{B} = (\vec{\imath}, \vec{\jmath}, \vec{k})$. Décrire les endomorphismes de E figurés dans la base $\mathcal{B}$ par chacune des matrices suivantes :

(a) $A = \dfrac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$ (b) $B = \dfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}$

(c) $C = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ (d) $D = \dfrac{1}{9}\begin{pmatrix} -8 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & -8 \end{pmatrix}$.

Solution

(a) **méthode**

On observe que la matrice A est orthogonale en vérifiant que ses colonnes (ou ses lignes) forment une famille orthonormale pour le produit scalaire usuel.

Notons C_1, C_2 et C_3 les colonnes de A. Le produit scalaire de deux colonnes de coefficients (x, y, z) et (x', y', z') se calcule par la formule $xx' + yy' + zz'$. On obtient alors

$$(C_1 \,|\, C_2) = \frac{1}{9}(2 + 2 - 4) = 0 \qquad \|C_1\|^2 = \frac{1}{9}(1 + 4 + 4) = 1$$

$$(C_2 \,|\, C_3) = \frac{1}{9}(4 - 2 - 2) = 0 \qquad \|C_2\|^2 = \frac{1}{9}(4 + 1 + 4) = 1$$

$$(C_3 \,|\, C_1) = \frac{1}{9}(2 - 4 + 2) = 0 \qquad \|C_3\|^2 = \frac{1}{9}(4 + 4 + 1) = 1.$$

La matrice A est donc orthogonale. L'endomorphisme a figuré dans la base orthonormale $\mathcal{B}$ par la matrice A est donc une isométrie.

1. Nous ne disons pas que x et y sont des vecteurs propres car ceux-ci peuvent être nuls, les espaces $E_1(u)$ et $E_{-1}(u)$ peuvent d'ailleurs être réduits au vecteur nul.

méthode

On détermine la nature de f en étudiant l'ensemble de ses vecteurs invariants : lorsqu'il s'agit d'un plan c'est un réflexion, lorsqu'il s'agit d'une droite c'est une rotation.

Soit $\vec{u} = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$ un vecteur de E. On a $a(\vec{u}) = \vec{u}$ si, et seulement si, $AX = X$ avec X la colonne de coefficients x, y, z. Ceci conduit à la résolution du système

$$\begin{cases} x + 2y + 2z = 3x \\ 2x + y - 2z = 3y \\ 2x - 2y + z = 3z \end{cases} \quad \text{soit} \quad \begin{cases} -2x + 2y + 2z = 0 \\ 2x - 2y - 2z = 0 \\ 2x - 2y - 2z = 0. \end{cases}$$

L'ensemble solution est le plan P d'équation $x - y - z = 0$. On en déduit que a est la réflexion par rapport à ce plan.

(b) On vérifie que les colonnes de B sont unitaires et deux à deux orthogonales. L'endomorphisme b figuré par la matrice B dans la base orthonormale $\mathcal{B}$ est donc une isométrie. Par la résolution de l'équation $b(\vec{u}) = \vec{u}$ d'inconnue $\vec{u} \in E$, on obtient que l'espace des vecteurs invariants par b est la droite $D = \text{Vect}(\vec{\imath} + \vec{k})$: l'isométrie b est une rotation [1] autour de la droite D.

méthode

Pour introduire l'angle de la rotation b, on oriente la droite D.

Orientons [2] la droite D par le vecteur $\vec{u} = \vec{\imath} + \vec{k}$ et notons θ l'angle de la rotation b autour de l'axe D.

méthode

Dans une base adaptée, la rotation b est figurée par la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$.
On a donc $\text{tr}(b) = 1 + 2\cos\theta$.

La trace de l'endomorphisme b peut aussi être calculée à partir de la matrice B ce qui révèle la valeur de $\cos\theta$:

$$1 + 2\cos\theta = \text{tr}(B) = \frac{1}{2} + 0 + \frac{1}{2} = 1 \quad \text{donc} \quad \cos\theta = 0.$$

Il suffit ensuite de connaître le signe de $\sin\theta$ pour déterminer θ à 2π près.

méthode

Si $\vec{v}$ est un vecteur n'appartenant pas à l'axe D, la famille $(\vec{u}, \vec{v}, b(\vec{v}))$ est directe si $\sin\theta > 0$ et indirecte si $\sin\theta < 0$.

Le signe du produit mixte $[\vec{u}, \vec{\imath}, b(\vec{\imath})]$ détermine donc le signe de $\sin\theta$. Ce produit mixte peut être calculé à l'aide des coordonnées dans la base $\mathcal{B}$ car celle-ci est une base

1. On peut aussi calculer le déterminant de B : observer $\det(B) = 1$ assure que b est une rotation.
2. Ce choix est arbitraire, si l'on oriente la droite D par un vecteur opposé au précédent, la mesure angulaire finale est opposée.

orthonormale directe :

$$\left[\vec{u},\vec{\imath},b(\vec{\imath})\right]=\begin{vmatrix}1 & 1 & 1/2\\ 0 & 0 & \sqrt{2}/2\\ 1 & 0 & 1/2\end{vmatrix}=\frac{\sqrt{2}}{2}>0.$$

On a donc $\cos\theta=0$ avec $\sin\theta>0$ donc $\theta\equiv\pi/2\ [2\pi]$.

On peut conclure que b est la rotation d'axe dirigé et orienté par $\vec{\imath}+\vec{k}$ et d'angle $\pi/2$ (il s'agit d'un *quart de tour* direct autour de l'axe dirigé et orienté par $\vec{\imath}+\vec{k}$).

(c) Les calculs sont semblables aux précédents. La matrice C est orthogonale car ses colonnes sont unitaires et deux à deux orthogonales. L'endomorphisme c figuré par C dans la base orthonormale $\mathcal{B}$ est une isométrie. L'ensemble des vecteurs invariants est la droite D dirigée par $\vec{u}=\vec{\imath}+\vec{\jmath}+\vec{k}$: c est une rotation autour de cette droite. Orientons la droite D par ce vecteur $\vec{u}$ et déterminons l'angle θ de la rotation c. On a

$$2\cos\theta+1=\operatorname{tr}(C)=0\quad\text{et}\quad\left[\vec{u},\vec{\imath},c(\vec{\imath})\right]=\begin{vmatrix}1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 0\end{vmatrix}=1>0.$$

On conclut que c est la rotation[1] d'axe dirigé et orienté par $\vec{\imath}+\vec{\jmath}+\vec{k}$ et d'angle $2\pi/3$.

(d) La matrice D est orthogonale et l'endomorphisme d figuré par D dans la base orthonormale $\mathcal{B}$ est une isométrie. L'ensemble des vecteurs invariants est la droite Δ dirigée par $\vec{u}=\vec{\imath}+4\vec{\jmath}+\vec{k}$. Orientons celle-ci par le vecteur $\vec{u}$. L'angle θ de la rotation d vérifie $2\cos\theta+1=\operatorname{tr}(D)=-1$ et donc $\cos\theta=-1$. On en déduit directement $\theta\equiv\pi\ [2\pi]$: la rotation d est une symétrie orthogonale[2] par rapport à la droite Δ. On dit encore que d est un *retournement* d'axe Δ.

7.3.2 Endomorphismes symétriques

Exercice 3

Soit u un endomorphisme symétrique d'un espace euclidien E. Montrer

$$\operatorname{Ker}(u)\oplus^{\perp}\operatorname{Im}(u)=E.$$

Solution

méthode

On établit que les espaces $\operatorname{Ker}(u)$ et $\operatorname{Im}(u)$ sont orthogonaux par la propriété de symétrie $\langle u(x),y\rangle=\langle x,u(y)\rangle$ pour tous x,y de E.

1. Cette affirmation est conforme à l'action de c qui envoie $\vec{\imath}$ sur $\vec{\jmath}$, $\vec{\jmath}$ sur $\vec{k}$ et $\vec{k}$ sur $\vec{\imath}$: la matrice C est une matrice de permutation.

2. On peut anticiper que d est une symétrie orthogonale en observant que la matrice D est orthogonale et symétrique : voir sujet 4 p. 295.

Soit $x \in \mathrm{Ker}(u)$ et $y \in \mathrm{Im}(u)$. On peut introduire $a \in E$ tel que $y = u(a)$ et écrire par la propriété de symétrie

$$\langle x, y\rangle = \langle x, u(a)\rangle = \langle u(x), a\rangle = \langle 0_E, a\rangle = 0.$$

Ainsi, les espaces $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont orthogonaux et donc en somme directe. De plus, la formule du rang donne

$$\dim E = \dim \mathrm{Ker}(u) + \dim \mathrm{Im}(u)$$

donc les espaces $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont supplémentaires et orthogonaux[1].

En particulier, on peut écrire $\mathrm{Ker}(u) = \big(\mathrm{Im}(u)\big)^{\perp}$ et $\mathrm{Im}(u) = \big(\mathrm{Ker}(u)\big)^{\perp}$.

Exercice 4

Déterminer les isométries d'un espace euclidien qui sont aussi des endomorphismes symétriques.

Solution

Soit u une isométrie qui est aussi un endomorphisme symétrique. Pour x et y vecteurs de E, la conservation du produit scalaire donne

$$\langle u(x), u(y)\rangle = \langle x, y\rangle$$

tandis que la propriété de symétrie donne

$$\langle u(x), u(y)\rangle = \langle u\big(u(x)\big), y\rangle = \langle u^2(x), y\rangle.$$

méthode

On peut montrer qu'un vecteur est nul en constatant qu'il est orthogonal à tout vecteur de l'espace.

Par différence, les calculs précédents donnent

$$\langle u^2(x) - x, y\rangle = 0 \quad \text{pour tous } x, y \in E.$$

On en déduit que, pour tout x de E, le vecteur $u^2(x) - x$ est orthogonal à tout vecteur de E et c'est donc le vecteur nul. Ainsi, $u^2 = \mathrm{Id}_E$: l'endomorphisme u est une symétrie[2]. Au surplus, celle-ci est une symétrie orthogonale car ses sous-espaces propres $\mathrm{Ker}(u - \mathrm{Id}_E)$ et $\mathrm{Ker}(u + \mathrm{Id}_E)$ sont orthogonaux (Th. 5 p. 291).

Inversement, une symétrie orthogonale s est une isométrie mais aussi un endomorphisme symétrique car, pour tous x et y de E,

$$\langle s(x), y\rangle \underset{s^2=\mathrm{Id}_E}{=} \big\langle s(x), s\big(s(y)\big)\big\rangle \underset{s\in \mathrm{O}(E)}{=} \langle x, s(y)\rangle.$$

1. Par le théorème spectral, on sait que E est la somme directe orthogonale des sous-espaces propres de u. Ici, $\mathrm{Im}(u)$ correspond à la somme des sous-espaces propres associés aux valeurs propres non nulles.

2. On peut aussi obtenir le résultat matriciellement, si M figure l'endomorphisme u dans une base orthonormale, celle-ci vérifie ${}^tMM = \mathrm{I}_n$ et ${}^tM = M$ donc $M^2 = \mathrm{I}_n$.

Exercice 5

Soit u un endomorphisme symétrique d'un espace euclidien E non réduit au vecteur nul. Montrer [1]

$$\sup_{\|x\|=1} \|u(x)\| = \max_{\lambda \in \mathrm{Sp}(u)} |\lambda|.$$

Solution

méthode

Par le théorème spectral (Th. 5 p. 291), l'endomorphisme symétrique u est diagonalisable dans une base orthonormale.

Soit $e = (e_1, \ldots, e_n)$ une base orthonormale de E formée de vecteurs propres de u. Notons $\lambda_1, \ldots, \lambda_n$ les valeurs propres associées aux vecteurs propres $e_1, \ldots, e_n$ et posons ρ la valeur maximale parmi $|\lambda_1|, \ldots, |\lambda_n|$.

Soit x un vecteur unitaire de E. On peut écrire $x = x_1 e_1 + \cdots + x_n e_n$ avec $x_1, \ldots, x_n$ des réels tels que $x_1^2 + \cdots + x_n^2 = 1$. On a alors $u(x) = (\lambda_1 x_1)e_1 + \cdots + (\lambda_n x_n)e_n$ puis

$$\|u(x)\|^2 = \sum_{i=1}^{n} \lambda_i^2 x_i^2 \leqslant \sum_{|\lambda_i| \leqslant \rho} \sum_{i=1}^{n} \rho^2 x_i^2 = \rho^2.$$

Ainsi, on peut affirmer l'inégalité qui suit avec existence [2] de la borne supérieure introduite

$$\sup_{\|x\|=1} \|u(x)\| \leqslant \rho.$$

Enfin, si i_0 désigne un indice pour lequel $|\lambda_{i_0}| = \rho$, on a $\|u(e_{i_0})\| = \rho$ et $\|e_{i_0}\| = 1$ donc

$$\sup_{\|x\|=1} \|u(x)\| = \max_{\|x\|=1} \|u(x)\| = \rho.$$

Exercice 6

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

(a) Justifier que la matrice ${}^tAA$ est diagonalisable.

(b) Montrer que les valeurs propres de ${}^tAA$ sont toutes positives.

Solution

(a) **méthode**

Les matrices symétriques réelles sont diagonalisables (Th. 6 p. 291).

La transposée d'un produit est le produit des transposées en ordre inverse et donc

$${}^t({}^tAA) = {}^tA\,{}^t({}^tA) = {}^tAA.$$

La matrice ${}^tAA$ est donc diagonalisable car symétrique et réelle.

1. Précisément, la borne supérieure porte sur les vecteurs de E de norme 1.

2. L'existence de la borne supérieure peut aussi être acquise directement en appliquant le théorème des bornes atteintes à la fonction continue $x \mapsto \|u(x)\|$ définie sur le compact non vide formé des vecteurs de norme 1. Au surplus, on peut anticiper que cette borne supérieure est un max.

(b) Soit λ une valeur propre de ${}^tAA$ et X un vecteur propre associé. On a

$$ {}^tAAX = \lambda X \quad \text{avec} \quad X \neq 0. $$

méthode

On multiplie à gauche par ${}^tX$ afin de faire apparaître la norme euclidienne sur les colonnes.

D'une part,

$$ {}^tX\,{}^tAAX = {}^tX({}^tAAX) = \lambda\,{}^tXX = \lambda\langle X, X\rangle = \lambda\,\|X\|^2. $$

D'autre part,

$$ {}^tX\,{}^tAAX = {}^t(AX)AX = \langle AX, AX\rangle = \|AX\|^2. $$

Puisque la colonne X est non nulle, on conclut

$$ \lambda = \frac{\|AX\|^2}{\|X\|^2} \geqslant 0. $$

7.4 Exercices d'entraînement

7.4.1 Isométries et matrices orthogonales

Exercice 7 *

(a) Trouver toutes les matrices de $\mathrm{O}_n(\mathbb{R})$ diagonalisables sur $\mathbb{R}$.

(b) Montrer que toutes les matrices de $\mathrm{O}_n(\mathbb{R})$ sont diagonalisables sur $\mathbb{C}$.

Solution

(a) **méthode**

Les valeurs propres réelles possibles d'une matrices orthogonales sont 1 et -1 (Th. 2 p. 288).

Analyse : Soit $A \in \mathrm{O}_n(\mathbb{R})$ diagonalisable. Les valeurs propres de A ne pouvant être que 1 et -1, la matrice A est semblable à une matrice diagonale D où ne figurent sur la diagonale que des 1 et/ou des -1. Le polynôme $X^2 - 1 = (X-1)(X+1)$ annule D donc aussi A et par conséquent $A^2 = \mathrm{I}_n$.

Synthèse : Soit $A \in \mathrm{O}_n(\mathbb{R})$ vérifiant $A^2 = \mathrm{I}_n$. La matrice A est diagonalisable car annule le polynôme $X^2 - 1$ qui est scindé sur $\mathbb{R}$ à racines simples.

En résumé, les matrices A de $\mathrm{O}_n(\mathbb{R})$ diagonalisables sur $\mathbb{R}$ sont celles vérifiant [1] $A^2 = \mathrm{I}_n$ (ou, et c'est équivalent ${}^tA = A$).

1. Ces matrices figurent les symétries orthogonales en base orthonormale.

(b) Soit $A \in \mathrm{O}_n(\mathbb{R})$.

méthode

La matrice A est semblable à une matrice diagonale par blocs de blocs diagonaux (Th. 2 p. 288)

$$(1), \quad (-1) \quad \text{ou} \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \not\equiv 0 \ [\pi].$$

Soit $\theta \not\equiv 0 \ [\pi]$. La matrice

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

a pour polynôme caractéristique $X^2 - 2\cos(\theta)X + 1$ de racines complexes distinctes $\mathrm{e}^{\mathrm{i}\theta}$ et $\mathrm{e}^{-\mathrm{i}\theta}$. La matrice R_θ est donc diagonalisable [1] sur $\mathbb{C}$ semblable à

$$\begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\theta} \end{pmatrix}.$$

En raisonnant par blocs, on peut affirmer que la matrice A est aussi diagonalisable sur $\mathbb{C}$. Notons que les valeurs propres complexes de A sont toutes de module 1.

Exercice 8 **

Soit $n \in \mathbb{N}$ avec $n \geqslant 2$. Déterminer les matrices de $\mathrm{O}_n(\mathbb{R})$ dont tous les coefficients sont positifs ou nuls.

Solution

Soit $A = (a_{i,j}) \in \mathrm{O}_n(\mathbb{R})$ dont tous les coefficients sont positifs ou nuls.

méthode

On montre que les rangées de A comportent un et un seul coefficient non nul.

La matrice A étant orthogonale, ses colonnes forment une famille orthonormale pour le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ donné par

$$\langle X, Y \rangle = {}^tXY = \sum_{i=1}^{n} x_i y_i.$$

Par l'absurde, supposons que la j-ème colonne de A possède au moins deux coefficients non nuls situés en k-ième et en ℓ-ième ligne. Puisque les colonnes de A sont orthogonales, on a pour tout indice j' différent de j

$$\sum_{i=1}^{n} a_{i,j} a_{i,j'} = 0.$$

1. Voir aussi sujet 4 p. 134.

Sachant que les coefficients sont tous positifs, on obtient $a_{i,j}a_{i,j'} = 0$ pour tout $i \in [\![1\,;n]\!]$. En particulier, $a_{k,j'} = a_{\ell,j'} = 0$ car $a_{k,j}$ et $a_{\ell,j}$ sont non nuls. Ainsi, les $n-1$ colonnes correspondant aux indices autres que j appartiennent à l'espace formé des colonnes dont les k-ième et ℓ-ième coefficients sont nuls. Or cet espace est de dimension $n-2$ car c'est l'intersection de deux hyperplans distincts tandis que les $n-1$ colonnes considérées sont linéairement indépendantes : c'est absurde.

Ainsi, il figure au plus un coefficient non nul sur chaque colonne de A. Celui-ci est forcément égal à 1 car les colonnes de A sont unitaires. Les colonnes de A sont donc élémentaires. Cependant, elle sont aussi deux à deux orthogonales et a *fortiori* deux à deux distinctes.

Finalement, la matrice A est constituée des colonnes de la matrice I_n dans un certain ordre, il s'agit d'une matrice de permutation [1].

Inversement, une telle matrice est effectivement orthogonale et à coefficients positifs.

Exercice 9 ***

Soit E un espace euclidien de dimension $n \geqslant 1$. On appelle *matrice de Gram* d'une famille $u = (u_1, \dots, u_p)$ de vecteurs de E, la matrice $G_u \in \mathcal{M}_p(\mathbb{R})$ de coefficient général $\langle u_i, u_j \rangle$.

(a) Montrer que la famille $u = (u_1, \dots u_p)$ et la matrice G_u ont le même rang.

(b) Soit $u = (u_1, \dots, u_p)$ et $v = (v_1, \dots, v_p)$ deux familles de vecteurs de E telles que $G_u = G_v$. Montrer qu'il existe une isométrie f de E vérifiant $f(u_i) = f(v_i)$ pour tout indice $i \in [\![1\,;p]\!]$.

Solution

(a) Introduisons l'espace $F = \mathrm{Vect}(u_1, \dots, u_p)$ engendré par la famille u.

méthode

On étudie le rang [2] de l'application $\varphi : F \to \mathbb{R}^p$ déterminée par

$$\varphi(x) = \big(\langle x, u_1\rangle, \dots, \langle x, u_p\rangle\big).$$

L'application φ est linéaire et injective. En effet, un vecteur x de son noyau est nécessairement nul car orthogonal à chaque vecteur u_j, il appartient donc à la fois à F et à $F^\perp$. Par conservation du rang par une application linéaire injective

$$\mathrm{rg}(u_1, \dots, u_p) = \mathrm{rg}\big(\varphi(u_1), \dots, \varphi(u_p)\big).$$

Or les vecteurs $\varphi(u_1), \dots, \varphi(u_p)$ constituent les colonnes de la matrice G_u et l'on peut donc affirmer

$$\mathrm{rg}(u_1, \dots, u_p) = \mathrm{rg}(G_u).$$

1. Voir sujet 22 p. 91.

2. En s'inspirant du sujet 27 p. 322, on peut aussi établir $G_u = {}^tAA$ avec A la matrice des coordonnées des vecteurs u_j dans une base orthonormale de F. On conclut alors grâce à l'égalité $\mathrm{rg}({}^tAA) = \mathrm{rg}(A)$ (voir sujet 21 p. 269).

(b) **méthode**

‖ On commence par résoudre le cas où la famille u est libre.

Cas : La famille u libre. On complète celle-ci en une base $u' = (u_1, \dots, u_n)$ à l'aide de vecteurs unitaires deux à deux orthogonaux choisis dans l'orthogonal de $\mathrm{Vect}(u_1, \dots, u_p)$. La matrice de Gram de la famille u' se déduit de celle de u :

$$G_{u'} = \begin{pmatrix} G_u & 0 \\ 0 & \mathrm{I}_{n-p} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

Par égalité des matrices G_u et G_v la famille v est aussi libre et peut être complétée en une base $v' = (v_1, \dots, v_n)$ comme au-dessus de sorte que $G_{v'} = G_{u'}$.

Considérons alors l'endomorphisme f de E déterminé par $f(u_k) = v_k$ pour tout k de $[\![1\,;n]\!]$. Vérifions que celui-ci conserve la norme.

Soit x un vecteur de E. On peut écrire $x = \lambda_1 u_1 + \cdots + \lambda_n u_n$ car la famille u' est une base de E. Par développement du calcul de la norme

$$\|x\|^2 = \langle x, x\rangle = \Big\langle \sum_{i=1}^{n} \lambda_i u_i, \sum_{j=1}^{n} \lambda_j u_j \Big\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \langle u_i, u_j\rangle.$$

Par définition de l'application linéaire f, on a $f(x) = \lambda_1 v_1 + \cdots + \lambda_n v_n$ et un calcul de norme analogue au précédent donne

$$\big\|f(x)\big\|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \langle v_i, v_j\rangle.$$

Or $G_{u'} = G_{v'}$ et donc $\big\|f(x)\big\| = \|x\|$: l'endomorphisme f est une isométrie qui résout le problème posé.

Cas général : Posons r le rang de la famille $(u_1, \dots, u_p)$. Quitte à permuter les vecteurs, supposons que les r premiers vecteurs de la famille u sont indépendants. On permute de la même façon les vecteurs $(v_1, \dots, v_p)$ afin de conserver l'hypothèse $G_u = G_v$. Par l'étude qui précède, on peut introduire une isométrie f de E envoyant les r premiers vecteurs de la famille u sur les vecteurs correspondants de la famille v. Étudions les images des vecteurs restants $u_{r+1}, \dots, u_p$.

Soit $k \in [\![r+1\,;p]\!]$. Le vecteur u_k est combinaison linéaire des vecteurs $u_1, \dots, u_r$ ce qui permet d'écrire $u_k = \lambda_1 u_1 + \cdots + \lambda_r u_r$. On a alors, pour tout $i \in [\![1\,;n]\!]$,

$$\big\langle \underbrace{u_k - (\lambda_1 u_1 + \cdots + \lambda_r u_r)}_{=0_E}, u_i \big\rangle = 0.$$

Sachant $G_u = G_v$, on a aussi

$$\big\langle v_k - (\lambda_1 v_1 + \cdots + \lambda_r v_r), v_i \big\rangle = 0.$$

Le vecteur $v_k - (\lambda_1 v_1 + \cdots + \lambda_r v_r)$ est alors combinaison linéaire des v_i tout en étant orthogonal à chacun d'entre eux : c'est le vecteur nul. On en déduit

$$v_k = \lambda_1 v_1 + \cdots + \lambda_r v_r = \lambda_1 f(u_1) + \cdots + \lambda_r f(u_r) = f(u_k).$$

Finalement, f est une isométrie solution.

7.4.2 Rotations de l'espace

Exercice 10 *

Soit E un espace euclidien orienté de dimension 3, r une rotation de E et s une symétrie orthogonale de E. Caractériser l'application $s \circ r \circ s$.

Solution

Dans une base orthonormale directe $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ adaptée à son axe, la matrice de la rotation r est de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \in \mathbb{R}.$$

L'endomorphisme r se décrit alors comme la rotation d'axe dirigé et orienté par $\vec{u}$ et d'angle θ.

méthode

Si g est un automorphisme, la matrice d'un endomorphisme f dans une base e est aussi celle de $g \circ f \circ g^{-1}$ dans la base $g(e)$.

Pour tout $\vec{x} \in E$, on remarque $(s \circ r \circ s)\big(s(\vec{x})\big) = s\big(r(\vec{x})\big)$. Dans la base orthonormale $s(\mathcal{B}) = \big(s(\vec{u}), s(\vec{v}), s(\vec{w})\big)$ la matrice de $s \circ r \circ s$ est donc

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

méthode

Avant de conclure, il faut connaître l'orientation de la base $s(\mathcal{B})$.

Distinguons deux cas :

Cas : $\det(s) = 1$. La base $s(\mathcal{B})$ est directe et $s \circ r \circ s$ est la rotation d'axe dirigé et orienté par $s(\vec{u})$ et d'angle θ.

Cas : $\det(s) = -1$. La base $s(\mathcal{B})$ est indirecte. Considérons alors la base orthonormale directe $\big(s(\vec{u}), s(\vec{v}), -s(\vec{w})\big)$. La matrice de $s \circ r \circ s$ dans celle-ci est

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(-\theta) & -\sin(-\theta) \\ 0 & \sin(-\theta) & \cos(-\theta) \end{pmatrix}.$$

L'endomorphisme $s \circ r \circ s$ est alors la rotation [1] d'axe dirigé et orienté par $s(\vec{u})$ et d'angle $-\theta$.

1. C'est aussi la rotation d'axe dirigé et orienté par $-s(\vec{u})$ et d'angle θ comme on peut le voir en figurant l'endomorphisme dans la base orthonormale directe $\big(-s(\vec{u}), s(\vec{v}), s(\vec{w})\big)$.

Exercice 11 **

Soit f une rotation d'un espace euclidien E orienté de dimension 3.

(a) On suppose qu'il existe $\vec{x} \neq \vec{0}$ tel que $f(\vec{x}) = -\vec{x}$. Montrer que f est un retournement, c'est-à-dire une symétrie orthogonale par rapport à une droite.

(b) En déduire que toute rotation f peut s'écrire comme produit de deux retournements.

Solution

(a) **méthode**

‖ On étudie les valeurs propres de la rotation f.

Dans une base orthonormale directe $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ adaptée à son axe, la matrice de la rotation f est de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad \text{avec} \quad \theta \in \mathbb{R}.$$

Son polynôme caractéristique est donc $\chi_f = (X-1)\big(X^2 - 2\cos(\theta)X + 1\big)$.

S'il existe $\vec{x} \neq \vec{0}$ tel que $f(\vec{x}) = -\vec{x}$ alors -1 est valeur propre de f donc racine de χ_f. On en déduit $\cos\theta = -1$ et la matrice de f dans $\mathcal{B}$ est alors

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

C'est la matrice d'une symétrie orthogonale par rapport à la droite $D = \mathrm{Vect}(\vec{u})$.

Notons aussi que le vecteur propre $\vec{x}$ est orthogonal à D car il s'agit d'un vecteur changé en son opposé donc d'un vecteur de $\mathrm{Ker}(f + \mathrm{Id}_E) = D^\perp$.

(b) Soit f une rotation de E d'axe $D = \mathrm{Vect}(\vec{u})$.

méthode

‖ On compose f par un retournement g d'axe Δ orthogonal à D.

Par opération dans le groupe $\mathrm{SO}(E)$ des rotations de E, l'endomorphisme $h = g \circ f$ est une rotation de E. De plus,

$$h(\vec{u}) = (g \circ f)(\vec{u}) = g(\vec{u}) = -\vec{u} \quad \text{car} \quad \vec{u} \in D \subset \Delta^\perp.$$

La résolution de la question précédente assure que la rotation h est un retournement et alors $f = g^{-1} \circ h = g \circ h$ est le produit de deux retournements. Au surplus, on peut souligner que g et h sont d'axes orthogonaux à D.

Exercice 12 **

Soit f et g deux rotations d'un espace euclidien orienté E de dimension 3.

À quelle(s) condition(s) a-t-on $g \circ f = f \circ g$?

Solution

Si l'un des endomorphismes f ou g est égal à l'identité, on a assurément $g \circ f = f \circ g$. Supposons désormais les rotations f et g distinctes de Id_E.

méthode

Si f et g commutent, les sous-espaces propres de l'un sont stables par l'autre.

L'axe $D = \mathrm{Vect}(\vec{u})$ de la rotation f est le sous-espace propre de f associé à la valeur propre 1. Il est stable par g car

$$f\big(g(\vec{u})\big) = g\big(f(\vec{u})\big) = g(\vec{u}).$$

Par conséquent, $\vec{u}$ est vecteur propre[1] de g. Or g est une isométrie et ses seules valeurs propres possibles sont 1 et -1. Distinguons alors deux cas

Cas : $g(\vec{u}) = \vec{u}$. Les rotations f et g ont le même axe. Inversement, on sait que deux rotations de même axe commutent.

Cas : $g(\vec{u}) = -\vec{u}$. La rotation g est alors un retournement[2] d'axe $\Delta = \mathrm{Vect}(\vec{v})$ orthogonal à $\vec{u}$. De plus, un raisonnement symétrique donne que $\vec{v}$ est vecteur propre de f et, comme celui-ci n'appartient pas à l'axe D, la valeur propre associée ne peut être que -1. Les rotations f et g sont alors des retournements d'axes orthogonaux. Inversement, de telles rotations commutent car sont figurées par des matrices diagonales dans une base orthonormale commençant par $\vec{u}$ et $\vec{v}$.

En résumé, deux rotations autres que l'identité commutent si, et seulement si, il s'agit de deux rotations de même axe ou de deux retournements d'axes orthogonaux.

Exercice 13 ***

Soit a, b et c trois réels. On étudie la matrice

$$A = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}.$$

(a) Montrer que A est une matrice orthogonale positive[3] si, et seulement si, a, b, c sont les trois racines d'un polynôme de la forme $X^3 - X^2 + k$ avec $k \in [0\,;4/27]$.

(b) Décrire la transformation géométrique associée aux matrices correspondantes.

Solution

(a) A est une matrice orthogonale positive si, et seulement si, ses colonnes sont unitaires, deux à deux orthogonales et son déterminant vaut 1. Ces conditions sont réunies

1. Rappelons qu'un vecteur est vecteur propre d'un endomorphisme si, et seulement si, il engendre une droite vectorielle stable par cet endomorphisme.
2. Voir le sujet précédent.
3. Autrement dit, $A \in \mathrm{SO}_3(\mathbb{R})$.

dans le système suivant [1] :

$$(\Sigma)\colon \begin{cases} a^2+b^2+c^2=1 \\ ab+bc+ca=0 \\ a^3+b^3+c^3-3abc=1. \end{cases}$$

méthode

Les réels a, b, c sont les trois racines du polynôme

$$(X-a)(X-b)(X-c)=X^3-\sigma_1 X^2+\sigma_2 X-\sigma_3$$

où l'on introduit les expressions symétriques élémentaires

$$\sigma_1=a+b+c, \quad \sigma_2=ab+bc+ca \quad \text{et} \quad \sigma_3=abc.$$

Supposons (a, b, c) solution réelle du système (Σ) et déterminons les valeurs de σ_1 et σ_2. Par la satisfaction de (Σ), on sait déjà $\sigma_2=0$ et l'on a

$$S_2=a^2+b^2+c^2=1 \quad \text{et} \quad S_3=a^3+b^3+c^3=1+3\sigma_3.$$

Par développement

$$\sigma_1^3=(a+b+c)^3=S_3+3t+6\sigma_3$$

avec $t=a^2b+ab^2+b^2c+bc^2+c^2a+ca^2$. Aussi

$$\sigma_1\sigma_2=(a+b+c)(ab+bc+ca)=t+3\sigma_3=0.$$

On a donc $t=-3\sigma_3$ puis $\sigma_1^3=S_3-3\sigma_3=1$. On en déduit $\sigma_1=1$. Ainsi, a, b, c sont les trois racines réelles d'un polynôme de la forme X^3-X^2+k avec $k=-\sigma_3$.

Inversement, le polynôme X^3-X^2+k admet trois racines complexes a, b, c pour lesquelles $\sigma_1=1$ et $\sigma_2=0$. On vérifie alors

$$S_2=\sigma_1^2-2\sigma_2=1$$

puis, en reprenant le calculs précédents en sens inverse, on obtient $S_3-3\sigma_3=1$. Ainsi, (a, b, c) est solution complexe du système (Σ).

Il ne reste plus qu'à étudier à quelle condition sur k les trois racines de P sont réelles. On étudie pour cela les variations du polynôme P. Pour $x\in\mathbb{R}$, on a $P'(x)=3x^2-2x$ et l'on peut dresser le tableau des variations de P :

x	$-\infty$		0		$2/3$		$+\infty$
$P'(x)$		$+$	0	$-$	0	$+$	
$P(x)$	$-\infty$	↗	$P(0)$	↘	$P(2/3)$	↗	$+\infty$

1. Le déterminant est calculé en développant directement selon une rangée.

Une condition nécessaire et suffisante pour que le polynôme P possède trois racines réelles comptées avec multiplicité est [1] que $P(0) \geqslant 0$ et $P(2/3) \leqslant 0$. Ceci conduit à la condition $k \in [0\,;4/27]$.

(b) Soit f l'endomorphisme de $\mathbb{R}^3$ canoniquement associé à A. L'endomorphisme f est une rotation de $\mathbb{R}^3$ muni de sa structure euclidienne orientée canonique. Déterminons son axe. La symétrie dans la répartition des coefficients de A invite [2] à considérer le vecteur $\vec{u} = (1,1,1)$. On constate $f(\vec{u}) = \vec{u}$ car $a+b+c = 1$. L'endomorphisme f est donc une rotation autour de la droite $D = \text{Vect}(\vec{u})$. Orientons celle-ci par le vecteur $\vec{u}$ et déterminons l'angle θ de la rotation f. La trace de A donne $2\cos\theta + 1 = 3a$ et le signe de $\sin\theta$ est celui de

$$\big[\vec{u}, \vec{\imath}, f(\vec{\imath})\big] = \begin{vmatrix} 1 & 1 & a \\ 1 & 0 & b \\ 1 & 0 & c \end{vmatrix} = b - c \quad \text{avec} \quad \vec{\imath} = (1,0,0).$$

On peut conclure que f est la rotation d'axe dirigé et orienté par $\vec{u} = (1,1,1)$ et d'angle

$$\theta = \begin{cases} \arccos\left(\frac{3a-1}{2}\right) & \text{si } b \geqslant c \\ -\arccos\left(\frac{3a-1}{2}\right) & \text{si } b < c. \end{cases}$$

7.4.3 Endomorphismes symétriques

Exercice 14 *

Soit E un espace euclidien de dimension $n \geqslant 2$, a un vecteur unitaire de E et k un réel.

(a) Montrer que l'on définit un endomorphisme symétrique f de E en posant

$$f(x) = x + k\langle a, x\rangle a \quad \text{pour tout } x \in E.$$

(b) Déterminer les éléments propres de f.

Solution

(a) L'application f est bien définie de E vers E. Vérifions qu'il s'agit d'une application linéaire. Pour $\lambda, \mu \in \mathbb{R}$ et $x, y \in E$, on a par linéarité du produit scalaire en sa deuxième variable

$$\begin{aligned} f(\lambda x + \mu y) &= \lambda x + \mu y + k\langle a, \lambda x + \mu y\rangle a \\ &= \lambda x + \lambda k\langle a, x\rangle a + \mu y + \mu k\langle a, y\rangle a = \lambda f(x) + \mu f(y). \end{aligned}$$

1. Si $P(2/3) < 0 < P(0)$, les trois racines sont distinctes et figurent dans chacun des intervalles $]-\infty\,;0[$, $]0\,;2/3[$ et $]2/3\,;+\infty[$. Si $P(0) = 0$, 0 est racine double et il existe une troisième racine réelle dans $]2/3\,;+\infty[$. Si $P(2/3) = 0$, c'est analogue avec $2/3$ racine double.

2. Il s'agit d'une matrice magique dans le sens du sujet 54 p. 190.

L'application f est donc un endomorphisme de E.

méthode

On vérifie que l'endomorphisme f est symétrique[1] en observant

$$\langle f(x), y\rangle = \langle x, f(y)\rangle \quad \text{pour tout } (x,y) \in E^2.$$

Soit x et y deux vecteurs de E. Par linéarité du produit scalaire en la première variable

$$\langle f(x), y\rangle = \big\langle x + \underbrace{k\langle a, x\rangle}_{\in\mathbb{R}}\, a, y\big\rangle = \langle x, y\rangle + k\langle a, x\rangle\langle a, y\rangle.$$

Dans le dernier membre, les vecteurs x et y jouent des rôles symétriques et l'on obtient donc par un calcul analogue $\langle x, f(y)\rangle = \langle f(x), y\rangle$: l'endomorphisme f est symétrique[2].

(b) Pour $x \in \big(\mathrm{Vect}(a)\big)^\perp$, on observe $f(x) = x$. On en déduit que 1 est valeur propre de f et que le sous-espace propre associé contient l'hyperplan de vecteur normal a.

méthode

Un endomorphisme symétrique est diagonalisable et ses sous-espaces propres sont deux à deux orthogonaux (Th. 5 p. 291).

La valeur propre « restante[3] » est donc associée au vecteur a. On a $f(a) = (1+k)a$ et l'on conclut en distinguant deux cas.

Cas : $k = 0$. L'endomorphisme f est l'identité, 1 est sa seule valeur propre, l'espace propre associé est E.

Cas : $k \neq 0$. L'endomorphisme f possède deux valeurs propres distinctes 1 et $1+k$. Les sous-espaces propres associés sont

$$E_1(f) = \big(\mathrm{Vect}(a)\big)^\perp \quad \text{et} \quad E_{1+k}(f) = \mathrm{Vect}(a).$$

En effet, ce qui précède donne les inclusions $\big(\mathrm{Vect}(a)\big)^\perp \subset E_1(f)$ et $\mathrm{Vect}(a) \subset E_{1+k}(f)$ et ces inclusions se transforment en égalité car la somme des dimensions des sous-espaces propres est égale à la dimension n de l'espace E.

1. Cette propriété entraîne que f est linéaire (voir sujet 5 p. 252) ce qui rend caduc le calcul de linéarité qui précède. Notons que lorsque $k = -1$, l'endomorphisme f est une projection orthogonale sur $\{a\}^\perp$ et, lorsque $k = -2$, c'est la symétrie orthogonale par rapport à $\{a\}^\perp$.

2. Avec un peu d'intuition, on peut penser figurer f dans une base orthonormale dont le premier vecteur est a : la représentation est diagonale donc symétrique et l'endomorphisme f est symétrique (Th. 3 p. 290). L'étude des éléments propres de f devient aussi immédiate.

3. L'expression de l'endomorphisme f est de la forme $f(x) = x + \varphi(x)a$ avec φ une forme linéaire : on peut former un polynôme annulateur de f de degré 2 ce qui révèle les valeurs propres possibles de f.

Exercice 15 **

Soit u un endomorphisme symétrique d'un espace euclidien E de dimension $n \geqslant 1$. On pose

$$\lambda_{\min} = \min_{\lambda \in \mathrm{Sp}(u)} \lambda \quad \text{et} \quad \lambda_{\max} = \max_{\lambda \in \mathrm{Sp}(u)} \lambda.$$

Montrer que, pour tout vecteur x de E,

$$\lambda_{\min} \|x\|^2 \leqslant \langle u(x), x \rangle \leqslant \lambda_{\max} \|x\|^2.$$

Solution

méthode

Par le théorème spectral (Th. 5 p. 291), on sait que les endomorphismes symétriques sont diagonalisables en base orthonormale.

Il existe une base orthonormale $e = (e_1, \ldots, e_n)$ dans laquelle la matrice de l'endomorphisme symétrique u est diagonale de la forme

$$\begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \quad \text{avec} \quad \lambda_1, \ldots, \lambda_n \text{ les valeurs propres de } u.$$

Pour $x \in E$, on peut écrire $x = x_1 e_1 + \cdots + x_n e_n$ avec $x_1, \ldots, x_n$ les coordonnées du vecteur x dans e. On a alors

$$\begin{aligned} u(x) &= x_1 u(e_1) + \cdots + x_n u(e_n) \\ &= \lambda_1 x_1 e_1 + \cdots + \lambda_n x_n e_n. \end{aligned}$$

Or la base e est orthonormale et l'on peut donc calculer norme et produit scalaire à l'aide des coordonnées dans e. Ceci donne

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 \quad \text{et} \quad \langle u(x), x \rangle = \sum_{i=1}^{n} \lambda_i x_i^2.$$

Pour tout $i \in [\![1\,;n]\!]$, on a $\lambda_{\min} \leqslant \lambda_i \leqslant \lambda_{\max}$ donc

$$\lambda_{\min} x_i^2 \leqslant \lambda_i x_i^2 \leqslant \lambda_{\max} x_i^2 \quad \text{car} \quad x_i^2 \geqslant 0$$

puis en sommant

$$\lambda_{\min} \underbrace{\sum_{i=1}^{n} x_i^2}_{=\|x\|^2} \leqslant \underbrace{\sum_{i=1}^{n} \lambda_i x_i^2}_{=\langle u(x), x \rangle} \leqslant \lambda_{\max} \underbrace{\sum_{i=1}^{n} x_i^2}_{=\|x\|^2}$$

ce qui donne l'encadrement voulu.

Exercice 16 **

Soit $(e_1, \dots, e_n)$ une base d'un espace euclidien E de dimension $n \geqslant 1$.

(a) Montrer que l'endomorphisme f défini par

$$f(x) = \sum_{i=1}^{n} \langle e_i, x \rangle e_i$$

est symétrique à valeurs propres strictement positives.

(b) Montrer qu'il existe un endomorphisme symétrique g de E tel que $g^2 = f^{-1}$.

(c) Établir que la famille $\big(g(e_1), \dots, g(e_n)\big)$ est une base orthonormale de E.

Solution

(a) Soit x et y dans E. Par linéarité du produit scalaire en la première variable

$$\big\langle f(x), y \big\rangle = \Big\langle \sum_{i=1}^{n} \underbrace{\langle e_i, x \rangle}_{\in \mathbb{R}} e_i, y \Big\rangle = \sum_{i=1}^{n} \langle e_i, x \rangle \langle e_i, y \rangle. \qquad (*)$$

Dans le dernier membre, x et y jouent des rôles identiques et donc $\big\langle f(x), y \big\rangle = \big\langle x, f(y) \big\rangle$: l'endomorphisme f est symétrique. Étudions ses valeurs propres.

Soit $\lambda \in \mathbb{R}$ une valeur propre de f et x un vecteur propre associé. On a $f(x) = \lambda x$ avec x non nul.

méthode

‖ On détermine le signe de λ en calculant $\big\langle f(x), x \big\rangle$.

D'une part, $\big\langle f(x), x \big\rangle = \big\langle \lambda x, x \big\rangle = \lambda \|x\|^2$. D'autre part, $(*)$ donne

$$\big\langle f(x), x \big\rangle = \sum_{i=1}^{n} \langle e_i, x \rangle^2 \geqslant 0.$$

De plus, si $\big\langle f(x), x \big\rangle = 0$, on a $\langle e_i, x \rangle = 0$ pour tout indice $i \in [\![1\,;n]\!]$ et x appartient à $\big(\mathrm{Vect}(e_1, \dots, e_n)\big)^{\perp} = \{0_E\}$ ce qui est exclu. On a donc $\big\langle f(x), x \big\rangle > 0$ puis

$$\lambda = \frac{\big\langle f(x), x \big\rangle}{\|x\|^2} > 0.$$

(b) Notons que l'endomorphisme f est inversible puisque 0 n'en est pas valeur propre.

méthode

‖ On résout matriciellement le problème posé.

Par le théorème spectral (Th. 5 p. 291), il existe une base orthonormale $e' = (e'_1, \dots, e'_n)$ de E dans laquelle la matrice de f est

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}.$$

Les λ_i pour $i \in [\![1\,;n]\!]$ étant les valeurs propres de f, ce sont des réels strictement positifs et on peut introduire les nombres $1/\sqrt{\lambda_i}$. Considérons alors l'endomorphisme g représenté dans la base e' par la matrice ci-dessous

$$\Delta = \begin{pmatrix} 1/\sqrt{\lambda_1} & & (0) \\ & \ddots & \\ (0) & & 1/\sqrt{\lambda_n} \end{pmatrix}.$$

méthode

Un endomorphisme est symétrique si, et seulement si, sa représentation dans une base orthonormale est une matrice symétrique (Th. 3 p. 290).

La matrice Δ est diagonale donc symétrique et figure l'endomorphisme g dans la base orthonormale e', celui-ci est donc symétrique. De plus, $\Delta^2 = D^{-1}$ et donc $g^2 = f^{-1}$.

(c) Pour tous les indices i et j de $[\![1\,;n]\!]$, on peut écrire par symétrie de l'endomorphisme g

$$\langle g(e_i), g(e_j)\rangle = \langle e_i, g^2(e_j)\rangle = \langle e_i, f^{-1}(e_j)\rangle.$$

Or la relation $f\big(f^{-1}(e_j)\big) = e_j$ donne l'identité

$$\sum_{i=1}^{n} \langle e_i, f^{-1}(e_j)\rangle e_i = \sum_{i=1}^{n} \delta_{i,j} e_j \quad \text{avec} \quad \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases}$$

La famille $(e_1, \dots, e_n)$ étant libre, on peut identifier[1] les scalaires des combinaisons linéaires des deux membres et écrire, pour tous $i, j \in [\![1\,;n]\!]$,

$$\langle g(e_i), g(e_j)\rangle = \langle e_i, f^{-1}(e_j)\rangle = \delta_{i,j}.$$

La famille $\big(g(e_1), \dots, g(e_n)\big)$ est donc orthonormale. C'est alors une famille libre et puisqu'elle est de longueur $n = \dim E$, c'est une base orthonormale de E.

1. Par différence de membres, on obtient une combinaison linéaire nulle donc des scalaires tous nuls.

Exercice 17 * (Théorème de Courant-Fischer)**

Soit u un endomorphisme symétrique d'un espace euclidien E de dimension $n \geqslant 1$.

On note $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ les valeurs propres de u comptées avec multiplicité, S la sphère unité de E et, pour tout $p \in [\![1\,;n]\!]$, $\mathcal{F}_p$ l'ensemble de tous les sous-espaces vectoriels de E dimension p.

Soit $p \in [\![1\,;n]\!]$. Établir

$$\lambda_p = \sup_{F\in\mathcal{F}_p}\left(\inf_{x\in F\cap S}\langle u(x),x\rangle\right) = \inf_{F\in\mathcal{F}_{n+1-p}}\left(\sup_{x\in F\cap S}\langle u(x),x\rangle\right).$$

Solution

L'endomorphisme u étant symétrique, l'espace E est la somme directe orthogonale de ses sous-espaces propres. Il existe donc une base orthonormale $e = (e_1,\dots,e_n)$ de E telle que $u(e_i) = \lambda_i e_i$ pour tout $i \in [\![1\,;n]\!]$.

Pour $x \in E$ de coordonnées $x_1,\dots,x_n$ dans la base orthonormale e, on a par des calculs identiques à ceux déjà vus dans le sujet 15 p. 307

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 \quad \text{et} \quad \langle u(x),x\rangle = \sum_{i=1}^{n}\lambda_i x_i^2.$$

Soit F un sous-espace vectoriel de E de dimension p.

méthode

‖ L'intersection de F et $G = \mathrm{Vect}(e_p,\dots,e_n)$ n'est pas réduite au vecteur nul.

L'espace F est de dimension p tandis que G est de dimension $n+1-p$. Par la formule de Grassmann

$$\dim(F\cap G) = \dim F + \dim G - \underbrace{\dim(F+G)}_{\leqslant n} \geqslant 1.$$

Il existe donc un vecteur y appartenant à $F\cap G\cap S$ de coordonnées $y_1,\dots,y_n$ dans la base e. Ce vecteur appartenant à G, les premières coordonnées $y_1,\dots,y_{p-1}$ sont nulles et donc

$$\langle u(y),y\rangle = \sum_{i=1}^{p-1}\underbrace{\lambda_i y_i^2}_{=0} + \sum_{i=p}^{n}\lambda_i y_i^2 = \sum_{i=p}^{n}\lambda_i y_i^2 \leqslant \lambda_p\sum_{i=p}^{n} y_i^2 = \lambda_p.$$

Par conséquent,

$$\inf_{x\in F\cap S}\langle u(x),x\rangle \leqslant \langle u(y),y\rangle \leqslant \lambda_p.$$

Cette comparaison valant pour tout sous-espace vectoriel F de $\mathcal{F}_p$, on a

$$\sup_{F\in\mathcal{F}_p}\left(\inf_{x\in F\cap S}\langle u(x),x\rangle\right) \leqslant \lambda_p. \qquad (*)$$

Considérons ensuite le sous-espace vectoriel $V = \mathrm{Vect}(e_1, \ldots, e_p)$ qui est de dimension p. Pour tout $x \in V \cap S$ de coordonnées $x_1, \ldots, x_n$ dans la base e

$$\langle u(x), x\rangle = \sum_{i=1}^{p} \lambda_i x_i^2 + \sum_{i=p+1}^{n} \underbrace{\lambda_i x_i^2}_{=0} = \sum_{i=1}^{p} \lambda_i x_i^2 \geqslant \lambda_p \sum_{i=1}^{p} x_i^2 = \lambda_p$$

donc

$$\inf_{x \in V \cap S} \langle u(x), x\rangle \geqslant \lambda_p.$$

Le sous-espace vectoriel V figurant parmi les sous-espaces vectoriels de $\mathcal{F}_p$, on obtient

$$\sup_{F \in \mathcal{F}_p} \left(\inf_{x \in F \cap S} \langle u(x), x\rangle \right) \geqslant \inf_{x \in V \cap S} \langle u(x), x\rangle \geqslant \lambda_p. \qquad (**)$$

Les deux inégalités $(*)$ et $(**)$ se complètent pour donner la première égalité [1]

$$\lambda_p = \sup_{F \in \mathcal{F}_p} \left(\inf_{x \in F \cap S} \langle u(x), x\rangle \right).$$

On peut obtenir la deuxième égalité en adaptant le raisonnement précédent ou bien en considérant l'endomorphisme $v = -u$ dont les valeurs propres $\mu_1, \ldots, \mu_n$ classées en ordre décroissant sont données par $\mu_{n+1-k} = -\lambda_k$. On a alors

$$\begin{aligned}
\lambda_p = -\mu_{n+1-p} &= - \sup_{F \in \mathcal{F}_{n+1-p}} \left(\inf_{x \in F \cap S} \langle v(x), x\rangle \right) \\
&= \inf_{F \in \mathcal{F}_{n+1-p}} - \left(\inf_{x \in F \cap S} \langle v(x), x\rangle \right) \\
&= \inf_{F \in \mathcal{F}_{n+1-p}} \left(\sup_{x \in F \cap S} - \langle v(x), x\rangle \right) \\
&= \inf_{F \in \mathcal{F}_{n+1-p}} \left(\sup_{x \in F \cap S} \langle u(x), x\rangle \right).
\end{aligned}$$

Exercice 18 ***

Soit p et q des projecteurs orthogonaux d'un espace euclidien E de dimension $n \geqslant 1$.

(a) Montrer que $p \circ q \circ p$ est diagonalisable.

(b) Déterminer $\big(\mathrm{Im}(p) + \mathrm{Ker}(q)\big)^{\perp}$.

(c) En déduire que $p \circ q$ est diagonalisable.

(d) Établir que les valeurs propres de $p \circ q$ sont comprises entre 0 et 1.

1. La borne supérieure est donc un max atteint en $V = \mathrm{Vect}(e_1, \ldots, e_p)$. Au surplus, la borne inférieure est un min en tant que minimum d'une fonction réelle continue définie sur un compact non vide : ce « sup d'inf » est un « max de min ».

Solution

(a) **méthode**

‖ Les projections orthogonales sont des endomorphismes symétriques.

Montrons que $p \circ q \circ p$ est un endomorphisme symétrique. Soit x et y deux éléments de E. En exploitant successivement les symétries des projections orthogonales p, q et encore p, il vient

$$\langle p \circ q \circ p(x), y\rangle = \langle q \circ p(x), p(y)\rangle = \langle p(x), q \circ p(y)\rangle = \langle x, p \circ q \circ p(y)\rangle.$$

L'endomorphisme $p \circ q \circ p$ est symétrique donc diagonalisable.

(b) **méthode**

‖ Si F et G sont deux sous-espaces vectoriels d'un espace euclidien, on sait [1]

$$(F+G)^\perp = F^\perp \cap G^\perp \quad \text{et} \quad (F \cap G)^\perp = F^\perp + G^\perp.$$

Par passage à l'orthogonal d'une somme

$$\big(\mathrm{Im}(p) + \mathrm{Ker}(q)\big)^\perp = \big(\mathrm{Im}(p)\big)^\perp \cap \big(\mathrm{Ker}(q)\big)^\perp = \mathrm{Ker}(p) \cap \mathrm{Im}(q)$$

car image et noyau d'une projection orthogonale sont l'orthogonal l'un de l'autre.

(c) L'endomorphisme $p \circ q \circ p$ est diagonalisable et l'espace $\mathrm{Im}(p)$ est stable par celui-ci. Il existe donc une base $(e_1, \ldots, e_r)$ de $\mathrm{Im}(p)$ diagonalisant l'endomorphisme induit par $p \circ q \circ p$. On a alors, pour tout $i \in [\![1\,;r]\!]$,

$$(p \circ q \circ p)(e_i) = \lambda_i e_i \quad \text{avec} \quad \lambda_i \in \mathbb{R}.$$

Or e_i est un vecteur de l'image de la projection p, il est donc invariant par p et par conséquent

$$(p \circ q)(e_i) = (p \circ q \circ p)(e_i) = \lambda_i e_i.$$

Ainsi, la famille $(e_1, \ldots, e_r)$ est constituée de vecteurs propres de $p \circ q$. Si besoin, on complète cette famille par des éléments $(e_{r+1}, \ldots, e_s)$ de $\mathrm{Ker}(q)$ afin de former une base de l'espace $F = \mathrm{Im}(p) + \mathrm{Ker}(q)$. Les vecteurs ainsi introduits annulent q et donc aussi $p \circ q$: ce sont des vecteurs propres associés à la valeur propre 0.

Enfin, on complète cette dernière famille par des éléments $(e_{s+1}, \ldots, e_n)$ de $F^\perp$ afin de former une base de E. Puisque $F^\perp = \big(\mathrm{Im}(p) + \mathrm{Ker}(q)\big)^\perp$ est l'intersection de $\mathrm{Ker}(p)$ et $\mathrm{Im}(q)$, les vecteurs $e_{s+1}, \ldots, e_n$ sont invariants par q et annulent p, ils annulent donc $p \circ q$ et sont aussi des vecteurs propres associés à la valeur propre 0.

En résumé, les vecteurs $e_1, \ldots, e_n$ constituent une base de vecteurs propres de $p \circ q$ associés aux valeurs propres $\lambda_1, \ldots, \lambda_r, 0, \ldots, 0$. L'endomorphisme $p \circ q$ est diagonalisable.

1. Voir sujet 6 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

(d) **méthode**

Si p est une projection orthogonale, on sait [1]

$$0 \leqslant \langle p(x), x\rangle \leqslant \|x\|^2 \quad \text{pour tout } x \in E.$$

Soit λ une valeur propre de $p \circ q$ et x un vecteur propre associé : $(p \circ q)(x) = \lambda x$ avec $x \neq 0_E$. Si λ est nul, c'est évidemment un élément de $[0\,;1]$. Sinon, le vecteur x est élément de l'image de p, il est donc invariant par p et l'on obtient par la propriété de symétrie

$$\langle \lambda x, x\rangle = \big\langle (p\circ q)(x), x\big\rangle = \big\langle q(x), \underbrace{p(x)}_{=x}\big\rangle = \big\langle q(x), x\big\rangle.$$

Puisque q est une projection orthogonale

$$0 \leqslant \underbrace{\big\langle q(x), x\big\rangle}_{=\lambda\|x\|^2} \leqslant \|x\|^2$$

et donc $\lambda \in [0\,;1]$ car $\|x\|^2 > 0$.

7.4.4 Matrices symétriques réelles

Exercice 19 *

On considère la matrice

$$A = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 2 & 2 \\ 2 & 2 & -1 \end{pmatrix}.$$

Déterminer $P \in \mathrm{O}_3(\mathbb{R})$ et $D \in \mathcal{M}_3(\mathbb{R})$ diagonale telle que $A = PD\,{}^tP$.

Solution

méthode

Les matrices symétriques réelles sont diagonalisables par l'intermédiaire d'une matrice de passage orthogonale (Th. 6 p. 291).

Afin de diagonaliser A, on détermine ses éléments propres. On commence par le calcul du polynôme caractéristique [2] $\chi_A = (X+3)(X-3)^2$.

La résolution de l'équation matricielle $AX = 3X$ d'inconnue $X \in \mathcal{M}_{3,1}(\mathbb{R})$ détermine le sous-espace propre associé à la valeur propre 3 : on obtient le plan $P\colon x + y - 2z = 0$.

méthode

Les sous-espaces propres d'une matrice symétrique réelle sont deux à deux orthogonaux.

1. Si a désigne le projeté orthogonal $p(x)$ et si $b = x - p(x)$, les vecteurs a et b sont orthogonaux de sorte que $\langle p(x), x\rangle = \|a\|^2 \leqslant \|a\|^2 + \|b\|^2 = \|x\|^2$.

2. On peut amorcer le calcul de $\chi_A(\lambda)$ par l'opération $C_1 \leftarrow C_1 + C_2 + C_3$ afin de faire apparaître un premier facteur $\lambda - 3$.

Sans calculs, on peut affirmer que le sous-espace propre associé à la valeur propre -3 est la droite D normale[1] au plan P : $D = \mathrm{Vect}(1,1,-2)$.

La détermination de bases de ces deux sous-espaces propres suffit à produire une matrice de passage diagonalisant A. Cependant, pour obtenir une matrice de passage orthogonale, il faut choisir des bases orthonormales de ces deux espaces propres.

On obtient une base orthonormale du plan $E_3(A) = P$ en considérant les vecteurs

$$\frac{1}{\sqrt{3}}(1,1,1) \quad \text{et} \quad \frac{1}{\sqrt{2}}(1,-1,0).$$

On obtient une base orthonormale de la droite $E_{-3}(A) = D$ avec le vecteur

$$\frac{1}{\sqrt{6}}(1,1,-2).$$

Ces vecteurs déterminent alors les colonnes d'une matrice de passage orthogonale P convenable

$$A = PD\,{}^tP \quad \text{avec} \quad P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Exercice 20 *

Soit $A = (a_{i,j}) \in \mathcal{S}_n(\mathbb{R})$ de valeurs propres $\lambda_1, \ldots, \lambda_n$. Établir

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}^2 = \sum_{i=1}^{n} \lambda_i^2.$$

Solution

méthode

On introduit le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ défini par

$$\langle A, B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}b_{i,j} = \mathrm{tr}\big({}^tAB\big).$$

On remarque

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}^2 = \langle A, A\rangle = \mathrm{tr}\big({}^tAA\big) = \mathrm{tr}\big(A^2\big).$$

La matrice A est symétrique réelle donc diagonalisable par une matrice de passage orthogonale et l'on peut écrire $A = PD\,{}^tP$ avec $P \in \mathrm{O}_n(\mathbb{R})$ et $D \in \mathcal{M}_n(\mathbb{R})$ diagonale de coefficients diagonaux les valeurs propres $\lambda_1, \ldots, \lambda_n$. On a alors

$$\mathrm{tr}\big(A^2\big) = \mathrm{tr}\big(PD\underbrace{{}^tPP}_{=I_n}D\,{}^tP\big) = \mathrm{tr}\big(PD^2\,{}^tP\big) = \mathrm{tr}\big(PD^2P^{-1}\big).$$

1. Pour le produit scalaire canonique, un vecteur normal à un plan d'équation $ax + by + cz = 0$ (avec a, b, c non tous nuls) est (a, b, c).

Deux matrices semblables ayant la même trace, on conclut

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}^2 = \operatorname{tr}(A^2) = \operatorname{tr}(D^2) = \sum_{i=1}^{n} \lambda_i^2.$$

Exercice 21 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$. On note α et β les plus petite et plus grande valeurs propres de la matrice S déterminée par

$$S = \frac{1}{2}(A + {}^tA).$$

(a) Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. Comparer ${}^tXAX$ et ${}^tXSX$.

(b) Montrer que pour toute colonne $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$\alpha\,{}^tXX \leqslant {}^tXAX \leqslant \beta\,{}^tXX.$$

(c) En déduire que les valeurs propres de A sont comprises entre α et β.

Solution

Notons que la matrice S est symétrique réelle. Elle est donc diagonalisable et par conséquent admet des valeurs propres réelles. Celles-ci sont en nombre fini ce qui assure l'existence de

$$\alpha = \min_{\lambda \in \mathrm{Sp}(S)} \lambda \quad \text{et} \quad \beta = \max_{\lambda \in \mathrm{Sp}(S)} \lambda.$$

(a) **méthode**

${}^tXAX$ désigne un nombre réel.

La matrice X étant une colonne, le produit de la ligne ${}^tX$ par la colonne AX détermine un nombre réel. Celui-ci est inchangé par transposition et donc [1]

$${}^tXAX = {}^t\big({}^tXAX\big) = {}^tX\,{}^tAX.$$

Par conséquent,

$${}^tXSX = \frac{1}{2}\,{}^tX(A + {}^tA)X = \frac{1}{2}\big({}^tXAX + {}^tX\,{}^tAX\big) = {}^tXAX.$$

(b) **méthode**

On encadre [2] ${}^tXSX$ en diagonalisant S par une matrice de passage orthogonale.

La matrice S est symétrique réelle donc orthogonalement diagonalisable : on peut écrire $S = PD\,{}^tP$ avec $P \in \mathrm{O}_n(\mathbb{R})$ et $D \in \mathcal{M}_n(\mathbb{R})$ diagonale de coefficients diagonaux les valeurs propres $\lambda_1, \ldots, \lambda_n$ de S.

1. Plus généralement, en introduisant le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$, on a pour toutes colonnes X et Y la formule $\langle X, AY\rangle = {}^tXAY = \langle {}^tAX, Y\rangle$.

2. On peut rapprocher cette étude de celle menée dans le sujet 15 p. 307.

En introduisant la colonne $Y = {}^tPX$ de coefficients $y_1, \dots, y_n$, on a

$$ {}^tXSX = {}^tX(PD{}^tP)X = ({}^tXP)D({}^tPX) = {}^tYDY = \sum_{i=1}^{n} \lambda_i y_i^2. $$

Les valeurs propres λ_i étant comprises entre α et β, on peut écrire l'encadrement

$$ \alpha \sum_{i=1}^{n} y_i^2 \leqslant {}^tXSX = \sum_{i=1}^{n} \lambda_i y_i^2 \leqslant \beta \sum_{i=1}^{n} y_i^2 \quad \text{avec} \quad \sum_{i=1}^{n} y_i^2 = {}^tYY. $$

Or la matrice P est orthogonale et donc ${}^tYY = {}^tXP{}^tPX = {}^tXX$. Il suffit ensuite de combiner l'ensemble des résultats précédents pour conclure

$$ \alpha {}^tXX \leqslant {}^tXAX \leqslant \beta {}^tXX. $$

(c) Soit λ une valeur propre de A et X un vecteur propre associé. On a $AX = \lambda X$ avec X une colonne non nulle. Par l'encadrement précédent, on obtient

$$ \alpha {}^tXX \leqslant {}^tXAX = \lambda {}^tXX \leqslant \beta {}^tXX. $$

Sachant la colonne X non nulle, le réel ${}^tXX$ est strictement positif car c'est le carré de la norme euclidienne de X. Après simplification, on conclut $\lambda \in [\alpha\,;\beta]$.

Exercice 22 ***

On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{R})$ dont les valeurs propres sont toutes positives. Soit $A \in \mathcal{S}_n^+(\mathbb{R})$. On veut montrer qu'il existe une unique matrice $B \in \mathcal{S}_n^+(\mathbb{R})$ telle que $B^2 = A$.

(a) Montrer l'existence d'une telle matrice.

(b) Soit $B \in \mathcal{S}_n^+(\mathbb{R})$ vérifiant $B^2 = A$. Établir que, pour tout $\lambda \in \mathbb{R}_+$,

$$ \operatorname{Ker}\left(B - \sqrt{\lambda}\mathrm{I}_n\right) = \operatorname{Ker}(A - \lambda \mathrm{I}_n). $$

(c) Justifier l'unicité d'une matrice solution.

Solution

(a) La matrice A est symétrique réelle donc orthogonalement diagonalisable. On peut alors écrire $A = PDP^{-1}$ avec $P \in \mathrm{O}_n(\mathbb{R})$ et

$$ D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} $$

où les réels $\lambda_1, \dots, \lambda_n$ désignent les valeurs propres de la matrice A. Celles-ci étant toutes positives, on peut introduire

$$ \Delta = \begin{pmatrix} \sqrt{\lambda_1} & & (0) \\ & \ddots & \\ (0) & & \sqrt{\lambda_n} \end{pmatrix} $$

et considérer la matrice $B = P\Delta P^{-1}$.

Puisque $\Delta^2 = D$, on a $B^2 = A$. De plus, la matrice B est symétrique car

$$ {}^tB = {}^t(P^{-1})\Delta\,{}^tP = P\Delta P^{-1} = B \quad \text{car} \quad {}^tP = P^{-1}. $$

Enfin, les valeurs propres de B sont les coefficients diagonaux de Δ, elles sont donc toutes positives.

Finalement, B est une matrice de $\mathcal{S}_n^+(\mathbb{R})$ de carré égal à A.

(b) Soit λ un réel positif.

méthode

‖ On exploite le lemme de décomposition des noyaux lorsque $\lambda > 0$.

Cas : $\lambda > 0$. On peut écrire la factorisation,

$$ X^2 - \lambda = \big(X - \sqrt{\lambda}\big)\big(X + \sqrt{\lambda}\big) $$

avec $X - \sqrt{\lambda}$ et $X + \sqrt{\lambda}$ polynômes premiers entre eux car $\sqrt{\lambda} \neq 0$. Le lemme de décomposition des noyaux donne alors

$$ \operatorname{Ker}(A - \lambda \mathrm{I}_n) = \operatorname{Ker}\big(B - \sqrt{\lambda}\mathrm{I}_n\big) \oplus \operatorname{Ker}\big(B + \sqrt{\lambda}\mathrm{I}_n\big). \qquad (*) $$

Or le noyau de $B + \sqrt{\lambda}\mathrm{I}_n$ est réduit à l'élément nul car les valeurs propres de B sont supposées positives. La relation $(*)$ se simplifie donc en

$$ \operatorname{Ker}(A - \lambda \mathrm{I}_n) = \operatorname{Ker}\big(B - \sqrt{\lambda}\mathrm{I}_n\big). $$

Cas : $\lambda = 0$. La matrice B étant diagonalisable, on sait [1] $\operatorname{Ker}(B) = \operatorname{Ker}\big(B^2\big)$ et on a donc $\operatorname{Ker}(B) = \operatorname{Ker}(A)$.

(c) Soit B et C deux matrices de $\mathcal{S}_n^+(\mathbb{R})$ vérifiant $B^2 = A$ et $C^2 = A$.

méthode

‖ On résout [2] la question posée dans le cadre vectoriel.

Soit a, b et c les endomorphismes de $E = \mathbb{R}^n$ canoniquement associés aux matrices respectives A, B et C. La matrice A étant diagonalisable, l'endomorphisme a l'est aussi et l'on a la décomposition en somme directe

$$ E = \bigoplus_{\lambda \in \operatorname{Sp}(A)} \operatorname{Ker}(a - \lambda \mathrm{Id}_E). $$

Pour tout $\lambda \in \operatorname{Sp}(A) \subset \mathbb{R}_+$, la résolution de la question précédente permet d'affirmer

$$ \operatorname{Ker}\big(b - \sqrt{\lambda}\mathrm{Id}_E\big) = \operatorname{Ker}(a - \lambda \mathrm{Id}_E) = \operatorname{Ker}\big(c - \sqrt{\lambda}\mathrm{Id}_E\big). $$

On a donc $b(x) = \sqrt{\lambda}x = c(x)$ pour tout $x \in \operatorname{Ker}(a - \lambda \mathrm{Id}_E)$.

Les applications linéaires b et c sont alors égales sur chaque espace d'une décomposition en somme directe, elles sont donc égales sur E. On peut conclure $B = C$.

1. Voir sujet 7 p. 136.

2. On peut aussi mener une démonstration assez analogue en restant dans le cadre matriciel : on établit $BX = CX$ pour toute colonne X en décomposant celle-ci sur les sous-espaces propres de A.

7.4.5 Matrices antisymétriques réelles

Exercice 23 **

Montrer que toute matrice antisymétrique réelle est de rang pair.

Solution

Soit $A \in \mathcal{M}_n(\mathbb{R})$ antisymétrique. Elle vérifie ${}^tA = -A$.

Cas : A est inversible. Le déterminant de A est non nul. Or

$$\det(A) = \det({}^tA) = \det(-A) = (-1)^n \det(A)$$

et donc $(-1)^n = 1$. On en déduit que $n = \mathrm{rg}(A)$ est un entier pair.

Cas : A n'est pas inversible. On introduit l'endomorphisme a de $\mathbb{R}^n$ canoniquement associé à la matrice A.

méthode

‖ On figure l'endomorphisme a dans une base orthonormale adaptée à son noyau.

Soit $e = (e_1, \ldots, e_n)$ une base orthonormale de $\mathbb{R}^n$ adaptée au noyau de a. Les premières colonnes de la matrice A' figurant l'endomorphisme a dans la base e sont nulles et donc

$$A' = \begin{pmatrix} 0 & B \\ 0 & C \end{pmatrix} \quad \text{avec} \quad C \in \mathcal{M}_r(\mathbb{R}) \text{ et } r = \mathrm{rg}(A') = \mathrm{rg}(A).$$

Vérifions ensuite que cette matrice A' est antisymétrique. Introduisons P la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base e. La matrice P est orthogonale en tant que matrice de passage entre deux bases orthonormales. La formule de changement de base s'écrit alors

$$A' = P^{-1}AP \quad \text{avec} \quad P^{-1} = {}^tP$$

et l'on vérifie :

$${}^tA' = {}^tP\,{}^tA\,{}^tP^{-1} = P^{-1}(-A)P = -P^{-1}AP = -A'.$$

La matrice A' est antisymétrique. Le bloc B est donc nul [1] tandis que le bloc C est antisymétrique. Puisque le bloc B est nul, le rang de A' est égal au rang de C. La matrice antisymétrique C est alors carrée de taille r et de rang r donc inversible. On est ramené au cas précédent et l'on peut conclure que r est un entier pair.

Exercice 24 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice antisymétrique.

(a) Quelles sont les valeurs propres réelles possibles de A ?

(b) En déduire que le déterminant de A est un réel positif.

(c) Montrer que les valeurs propres complexes de A sont imaginaires pures.

1. La nullité de ce bloc pouvait aussi être acquise en observant que $(\mathrm{Ker}(a))^\perp$ est stable par a.

Solution

(a) Soit λ une valeur propre réelle de A et $X \in \mathcal{M}_{n,1}(\mathbb{R})$ un vecteur propre associé. On a $AX = \lambda X$ avec X colonne non nulle.

méthode

‖ On étudie ${}^tXAX$.

D'une part, on a directement ${}^tXAX = \lambda\, {}^tXX$. D'autre part, on peut écrire en exploitant $A = -{}^tA$

$${}^tXAX = -{}^tX\,{}^tAX = -{}^t(AX)X = -\lambda\, {}^tXX.$$

Sachant ${}^tXX = \|X\|^2 \neq 0$ car X non nulle, on obtient $\lambda = -\lambda$ et donc $\lambda = 0$. Ainsi, la seule valeur propre réelle possible d'une matrice antisymétrique est 0.

(b) Le déterminant de A est le produit des valeurs propres complexes de A comptées avec multiplicité. Or ces dernières sont deux à deux conjuguées et l'absence de valeurs propres réelles non nulles entraîne que le déterminant de A est un produit de facteurs $\lambda\overline{\lambda}$ et d'éventuels 0. On en déduit que $\det(A)$ est un réel positif[1].

(c) Soit λ une valeur propre complexe de A et $X \in \mathcal{M}_{n,1}(\mathbb{C})$ un vecteur propre associé. On a $AX = \lambda X$ avec X colonne non nulle.

méthode

‖ On étudie ${}^t\overline{X}AX$.

D'une part, un calcul direct donne ${}^t\overline{X}AX = \lambda\, {}^t\overline{X}X$. D'autre part, il est possible d'écrire $A = -{}^t\overline{A}$ car la matrice A est antisymétrique et réelle. On a alors aussi

$${}^t\overline{X}AX = -{}^t\overline{X}\,{}^t\overline{A}X = -{}^t(\overline{AX})X = -{}^t(\overline{\lambda X})X = -\overline{\lambda}\,{}^t\overline{X}X.$$

Ainsi, on obtient $\lambda\, {}^t\overline{X}X = -\overline{\lambda}\, {}^t\overline{X}X$. Or, en notant $x_1, \ldots, x_n$ les coefficients complexes constituant la colonne non nulle X, on remarque

$${}^t\overline{X}X = \sum_{k=1}^{n} |x_k|^2 > 0.$$

On en déduit $\overline{\lambda} = -\lambda$ donc $\lambda \in i\mathbb{R}$.

Exercice 25 ***

Soit $A \in \mathcal{M}_n(\mathbb{R})$ antisymétrique. Montrer que, par le biais d'une matrice de passage orthogonale, la matrice A est semblable[2] à une matrice diagonale par blocs avec sur la diagonale des zéros et/ou différents blocs M_α avec

$$M_\alpha = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} \quad \text{avec} \quad \alpha \in \mathbb{R}_+^*.$$

1. On trouvera dans le sujet 29 p. 164 une démarche alternative.
2. Ce résultat de réduction des matrices antisymétriques résout aussi les deux sujets précédents.

Solution

On raisonne par récurrence forte sur $n \geqslant 1$.

Pour $n = 1$, une matrice antisymétrique est nulle et la propriété est vérifiée.

Pour $n = 2$, une matrice antisymétrique de taille 2 est nulle ou de la forme

$$\begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} \quad \text{avec} \quad \alpha \neq 0.$$

Si $\alpha > 0$, la propriété est immédiatement vérifiée. Si $\alpha < 0$, on emploie une matrice de permutation :

$$P^{-1}\begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} P = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix} \quad \text{avec} \quad P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{O}_2(\mathbb{R}).$$

Supposons ensuite la propriété établie jusqu'au rang $n - 1 \geqslant 2$. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice antisymétrique. Introduisons l'endomorphisme a de $\mathbb{R}^n$ canoniquement associé à la matrice A et notons $\langle \cdot\,, \cdot \rangle$ le produit scalaire canonique sur $\mathbb{R}^n$. L'hypothèse d'antisymétrie de la matrice A se retraduit en écrivant

$$\langle a(x), y\rangle = -\langle x, a(y)\rangle \quad \text{pour tous } x \text{ et } y \text{ de } \mathbb{R}^n. \qquad (*)$$

En effet, si X et Y désignent les colonnes formées des coefficients constituant x et y,

$$\langle a(x), y\rangle = {}^t(AX)Y = {}^tX\,{}^tAY = -\,{}^tXAY = -\langle x, a(y)\rangle.$$

méthode

Pour obtenir dans la représentation matricielle de a un bloc M_α, on recherche deux vecteurs e_1 et e_2 vérifiant $a(e_1) = \alpha e_2$ et $a(e_2) = -\alpha e_1$. Le vecteur e_1 apparaît alors comme un vecteur propre de a^2.

Étudions l'endomorphisme a^2. Celui-ci est symétrique car figuré par la matrice A^2 dans la base canonique qui est orthonormale avec

$${}^t(A^2) = ({}^tA)^2 = (-A)^2 = A^2$$

Discutons ensuite selon l'éventuelle nullité de a^2.

Cas : $a^2 = 0$. L'endomorphisme a est alors nul car, pour tout $x \in \mathbb{R}^n$,

$$\|a(x)\|^2 = \langle a(x), a(x)\rangle = -\langle x, a(a(x))\rangle = -\langle x, a^2(x)\rangle = 0.$$

La propriété voulue est alors immédiate.

Cas : $a^2 \neq 0$. L'endomorphisme a^2 est symétrique donc diagonalisable. Puisque ce n'est pas l'endomorphisme nul, il possède une valeur propre λ non nulle. Soit e_1 un vecteur propre unitaire associé. Le vecteur $a(e_1)$ ne peut pas être nul et est orthogonal à e_1 car $(*)$ donne

$$\langle a(e_1), e_1\rangle = -\langle e_1, a(e_1)\rangle \quad \text{donc} \quad \langle a(e_1), e_1\rangle = 0.$$

Introduisons alors le vecteur unitaire

$$e_2 = \frac{1}{\|a(e_1)\|} a(e_1).$$

La famille (e_1, e_2) est orthonormale et, par construction, on peut écrire $a(e_1) = \alpha e_2$ avec $\alpha = \|a(e_1)\| > 0$. De plus, $a^2(e_1) = \lambda e_1$ donne

$$a(e_2) = \frac{1}{\alpha} a^2(e_1) = \beta e_1 \quad \text{avec} \quad \beta = \frac{\lambda}{\alpha}.$$

L'égalité $\langle a(e_1), e_2 \rangle = -\langle e_1, a(e_2) \rangle$ montre ensuite $\beta = -\alpha$.

En résumé, on a déterminé un plan $P = \text{Vect}(e_1, e_2)$ stable par a tel que l'endomorphisme induit par a est représenté dans la base orthonormale (e_1, e_2) par la matrice

$$M_\alpha = \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix}.$$

Considérons ensuite une matrice P orthogonale dont les deux premières colonnes sont formées [1] par les vecteurs e_1 et e_2. Puisque $P^{-1} = {}^tP$, on vérifie que la matrice $P^{-1}AP$ est antisymétrique et, par construction, celle-ci est de la forme

$$P^{-1}AP = \begin{pmatrix} M_\alpha & 0 \\ 0 & A' \end{pmatrix}.$$

La matrice A' de $\mathcal{M}_{n-2}(\mathbb{R})$ est antisymétrique et l'on peut employer l'hypothèse de récurrence pour introduire une matrice P' de $\mathrm{O}_{n-2}(\mathbb{R})$ telle que $P'^{-1}A'P'$ est de la forme voulue. Enfin, en considérant la matrice

$$Q = \begin{pmatrix} \mathrm{I}_2 & 0 \\ 0 & P' \end{pmatrix} \in \mathrm{O}_n(\mathbb{R})$$

on obtient $R^{-1}AR$ de la forme souhaitée avec $R = PQ \in \mathrm{O}_n(\mathbb{R})$.

La récurrence est établie.

7.5 Exercices d'approfondissement

Exercice 26 *

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$. On note α et β les plus grandes valeurs propres des matrices ${}^tAA$ et ${}^tBB$. Montrer que les valeurs propres λ de AB vérifient $\lambda^2 \leqslant \alpha\beta$.

1. Il est possible de construire une telle matrice en complétant la famille orthonormale (e_1, e_2) en une base orthonormale de $\mathbb{R}^n$.

Solution

Notons que les matrices ${}^tAA$ et ${}^tBB$ sont symétriques réelles, elles admettent donc des valeurs propres et celles-ci sont en nombre fini ce qui légitime l'introduction de α et β.

Considérons le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ donné par $\langle X, Y\rangle = {}^tXY$.

méthode

‖ On vérifie $\|AX\|^2 \leqslant \alpha \|X\|^2$ pour toute colonne X de $\mathcal{M}_{n,1}(\mathbb{R})$.

Soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$. On a

$$\|AX\|^2 = \langle AX, AX\rangle = {}^t(AX)AX = {}^tX\,{}^tAAX.$$

Or la matrice ${}^tAA$ est symétrique réelle. Par le théorème spectral, on peut donc écrire ${}^tAA = PD\,{}^tP$ avec P matrice orthogonale et D matrice diagonale de coefficients diagonaux les valeurs propres $\lambda_1, \ldots, \lambda_n$ de ${}^tAA$. On poursuit alors le calcul précédent

$$\|AX\|^2 = {}^tXPD\,{}^tPX = {}^tYDY \quad \text{avec} \quad Y = {}^tPX.$$

En notant $y_1, \ldots, y_n$ les coefficients de la colonne Y, on observe

$${}^tYDY = \sum_{i=1}^{n} \lambda_i y_i^2 \leqslant \alpha \sum_{i=1}^{n} y_i^2 = \alpha\, {}^tYY$$

avec

$${}^tYY = {}^tXP\,{}^tPX = {}^tXX.$$

On a donc $\|AX\|^2 \leqslant \alpha \|X\|^2$. On montre de même $\|BX\|^2 \leqslant \beta \|X\|^2$ pour toute colonne X. Considérons alors λ une valeur propre de AB et X un vecteur propre associé. On a $ABX = \lambda X$ avec X colonne non nulle. En appliquant les inégalités qui précèdent successivement avec les matrices A et B, il vient

$$\lambda^2 \|X\|^2 = \|\lambda X\|^2 = \|ABX\|^2 \leqslant \alpha \|BX\|^2 \leqslant \alpha\beta \|X\|^2.$$

On simplifie alors par $\|X\|^2 > 0$ pour conclure $\lambda^2 \leqslant \alpha\beta$.

Exercice 27 **

Soit E un espace vectoriel euclidien de dimension $n \geqslant 1$ et $e = (e_1, \ldots, e_n)$ une base orthonormale de E.

On appelle *matrice de Gram*[1] d'une famille $(x_1, \ldots, x_n)$ de vecteurs de E la matrice $G(x_1, \ldots, x_n) \in \mathcal{M}_n(\mathbb{R})$ de coefficient général $\langle x_i, x_j\rangle$.

(a) On note A la matrice figurant la famille $(x_1, \ldots, x_n)$ dans la base e. Exprimer $G(x_1, \ldots, x_n)$ en fonction des matrices A et ${}^tA$.

(b) Soit $M \in \mathcal{M}_n(\mathbb{R})$. À quelle(s) condition(s) existe-t-il une famille $(x_1, \ldots, x_n)$ de vecteurs de E telle que $M = G(x_1, \ldots, x_n)$?

(c) Application : On suppose $n \geqslant 2$. Pour quelles valeurs de c réelles existe-t-il une famille $(x_1, \ldots, x_n)$ de vecteurs unitaires de E vérifiant $\langle x_i, x_j\rangle = c$ pour tous les indices i et j distincts ?

1. On trouvera une autre application classique des matrices de Gram dans le sujet 33 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

Solution

(a) **méthode**

On sait exprimer le produit scalaire de deux vecteurs à partir de leurs coordonnées dans une base orthonormale.

Les coefficients de la matrice A définissent les coordonnées des vecteurs $x_1, \dots, x_n$ dans la base orthonormale e. Pour $i, j \in [\![1\,;n]\!]$, les coefficients $a_{1,i}, \dots, a_{n,i}$ d'une part, et $a_{1,j}, \dots, a_{n,j}$ d'autre part, donnent les coordonnées des vecteurs x_i et x_j et donc

$$\langle x_i, x_j \rangle = \sum_{k=1}^{n} a_{k,i} a_{k,j} = \sum_{k=1}^{n} [A]_{k,i} [A]_{k,j} = \sum_{k=1}^{n} [{}^tA]_{i,k} [A]_{k,j} = [{}^tAA]_{i,j}.$$

Ainsi[1], $G(x_1, \dots, x_n) = {}^tAA$.

(b) Par représentation d'une famille de vecteurs dans une base, l'existence d'une famille de vecteurs $(x_1, \dots, x_n)$ telle que $M = G(x_1, \dots, x_n)$ équivaut à l'existence d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que $M = {}^tAA$.

méthode

Pour $A \in \mathcal{M}_n(\mathbb{R})$, ${}^tAA$ est une matrice symétrique à valeurs propres positives[2].

Analyse : S'il existe une matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que $M = {}^tAA$, la matrice M est nécessairement symétrique et à valeurs propres positives.

Synthèse : Supposons la matrice M symétrique de valeurs propres $\lambda_1, \dots, \lambda_n$ positives. Par le théorème spectral, on peut écrire $M = PD\,{}^tP$ avec $P \in \mathrm{O}_n(\mathbb{R})$ et

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}.$$

Puisque les valeurs propres $\lambda_1, \dots, \lambda_n$ sont positives, on peut introduire

$$\Delta = \begin{pmatrix} \sqrt{\lambda_1} & & (0) \\ & \ddots & \\ (0) & & \sqrt{\lambda_n} \end{pmatrix}.$$

La matrice[3] $A = \Delta\,{}^tP$ vérifie alors ${}^tAA = M$ et la matrice M est donc la matrice de Gram de la famille de vecteurs figurée par A dans la base e.

1. En conséquence, $\mathrm{rg}\big(G(x_1, \dots, x_n)\big) = \mathrm{rg}({}^tAA) = \mathrm{rg}(A) = \mathrm{rg}(x_1, \dots, x_n)$ (voir sujet 21 p. 269).
2. Voir sujet 6 p. 296.
3. Plus généralement, $A = Q\Delta\,{}^tP$ avec $Q \in \mathrm{O}_n(\mathbb{R})$ convient aussi et l'on peut montrer qu'il n'y en a pas d'autres, par exemple en exploitant le résultat du sujet 9 p. 299.

(c) Il s'agit ici de déterminer pour quelles valeurs de $c \in \mathbb{R}$ la matrice symétrique

$$M_c = \begin{pmatrix} 1 & & (c) \\ & \ddots & \\ (c) & & 1 \end{pmatrix}$$

est à valeurs propres positives.

Les valeurs propres de M_c sont[1] $1-c$ et $1+(n-1)c$. Celles-ci sont positives si, et seulement si,

$$c \in \left[-\frac{1}{n-1}\,;1\right].$$

Exercice 28 * (Pseudo inverse)**

Soit a une application linéaire d'un espace euclidien E vers un espace euclidien E'.

Montrer qu'il existe une unique application linéaire b de E' vers E vérifiant :

(i) $a \circ b$ et $b \circ a$ sont des endomorphismes symétriques ;

(ii) $a \circ b \circ a = a$ et $b \circ a \circ b = b$.

L'application linéaire b est appelée *pseudo-inverse*[2] de a.

Solution

Analyse : Supposons $b \in \mathcal{L}(E', E)$ solution du problème posé.

méthode

Les endomorphismes $a \circ b$ et $b \circ a$ sont des projections orthogonales que l'on peut préciser.

L'endomorphisme $a \circ b$ est symétrique et vérifie $(a \circ b)^2 = a \circ b$, il s'agit d'une projection orthogonale. De plus, $\mathrm{Im}(a \circ b) \subset \mathrm{Im}(a)$ et, inversement, $\mathrm{Im}(a) \subset \mathrm{Im}(a \circ b)$ car on peut écrire $a = (a \circ b) \circ a$. L'endomorphisme $a \circ b$ est donc la projection orthogonale sur $\mathrm{Im}(a)$.

Aussi, $b \circ a$ est symétrique et vérifie $(b \circ a)^2 = b \circ a$, c'est une projection orthogonale. De plus, $\mathrm{Ker}(a) \subset \mathrm{Ker}(b \circ a)$ et aussi $\mathrm{Ker}(b \circ a) \subset \mathrm{Ker}(a)$ car $a = a \circ (b \circ a)$. L'endomorphisme $b \circ a$ est donc la projection orthogonale parallèlement à $\mathrm{Ker}(a)$, c'est-à-dire la projection orthogonale sur $\big(\mathrm{Ker}(a)\big)^\perp$.

Par les études qui précèdent, on peut affirmer que la composée $a \circ b$ est parfaitement déterminée puisqu'il s'agit de la projection orthogonale p sur $\mathrm{Im}(a)$. Pour en déduire b, il suffit de savoir inverser a sur l'espace des valeurs prises par b.

méthode

Une application linéaire définit par restriction un isomorphisme de tout supplémentaire de son noyau vers son image.

L'égalité $b = b \circ a \circ b$ donne par double inclusion $\mathrm{Im}(b) = \mathrm{Im}(b \circ a)$ et on en déduit $\mathrm{Im}(b) = \big(\mathrm{Ker}(a)\big)^\perp$. Or cet espace est un supplémentaire de $\mathrm{Ker}(a)$ et l'on peut donc

1. Voir sujet 25 p. 158.

2. Lorsque l'on veut résoudre l'équation linéaire $a(x) = y$, ni l'existence, ni l'unicité d'une solution ne sont assurées. Introduire le pseudo-inverse b permet de déterminer le vecteur $x_0 = b(y)$ qui est le vecteur de norme minimale tel que $a(x_0)$ soit le plus proche possible de y.

introduire la restriction de $\big(\mathrm{Ker}(a)\big)^\perp$ vers $\mathrm{Im}(a)$ qui est un isomorphisme. Notons a' son isomorphisme réciproque :

$$a' : \mathrm{Im}(a) \to \big(\mathrm{Ker}(a)\big)^\perp.$$

L'égalité $a \circ b = p$ donne alors $a' \circ a \circ b = a' \circ p$ avec $a' \circ a = \mathrm{Id}_{\mathrm{Im}(b)}$ donc $b = a' \circ p$. L'application linéaire b est ainsi déterminée de façon unique.

Synthèse : Soit $b = a' \circ p$ avec a' et p tels que définis au-dessus. L'application linéaire b est bien définie de E' vers E et l'on remarque

$$\mathrm{Im}(b) = \big(\mathrm{Ker}(a)\big)^\perp \quad \text{et} \quad \mathrm{Ker}(b) = \big(\mathrm{Im}(a)\big)^\perp$$

car a' est un isomorphisme. Vérifions que l'application b satisfait aux propriétés requises. On a $a \circ b = a \circ a' \circ p = p$ car $a \circ a'$ est l'identité sur l'image de a. Ainsi, $a \circ b$ est une projection orthogonale et donc un endomorphisme symétrique.

Aussi, $b \circ a = a' \circ p \circ a = a' \circ a$ car p se confond avec l'identité sur l'image de a. Or l'application linaire $a' \circ a$ est nulle sur $\mathrm{Ker}(a)$ et correspond à l'identité sur son orthogonal. L'application $b \circ a$ correspond donc à la projection orthogonale sur l'orthogonal de $\mathrm{Ker}(a)$. Il s'agit encore une fois d'un endomorphisme symétrique.

Enfin, $(a \circ b)^2 = a \circ b$ donne $(a \circ b \circ a - a) \circ b = 0$ donc

$$\big(\mathrm{Ker}(a)\big)^\perp = \mathrm{Im}(b) \subset \mathrm{Ker}(a \circ b \circ a - a).$$

Cependant, on a aussi $\mathrm{Ker}(a) \subset \mathrm{Ker}(a \circ b \circ a - a)$ et donc $a \circ b \circ a - a = 0$.

L'égalité $(a \circ b)^2 = a \circ b$ donne encore $a \circ (b \circ a \circ b - b) = 0$ donc

$$\mathrm{Im}(b \circ a \circ b - b) \subset \mathrm{Ker}(a).$$

Cependant, $\mathrm{Im}(b \circ a \circ b - b) \subset \mathrm{Im}(b) = \big(\mathrm{Ker}(a)\big)^\perp$ et donc $b \circ a \circ b - b = 0$.

Finalement, l'application linéaire b est solution.

Exercice 29 * (Décomposition polaire)**

Soit $S \in \mathcal{M}_n(\mathbb{R})$ telle que $\mathrm{tr}(SU) \leqslant \mathrm{tr}(S)$ pour toute matrice $U \in \mathrm{O}_n(\mathbb{R})$.

(a) Soit $A \in \mathcal{M}_n(\mathbb{R})$ antisymétrique. Montrer $\exp(A) \in \mathrm{O}_n(\mathbb{R})$.

(b) En considérant les matrices orthogonales $\exp(tA)$ pour t réel et A matrice antisymétrique, établir que S est une matrice symétrique.

(c) Montrer que les valeurs propres de S sont positives.

(d) Application : Montrer que pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$, il existe $S \in \mathcal{M}_n(\mathbb{R})$ symétrique à valeurs propres positives et $\Omega \in \mathrm{O}_n(\mathbb{R})$ telles que $M = S\Omega$.

Solution

(a) L'application de transposition est linéaire au départ d'un espace de dimension finie donc continue. On obtient alors par passage à la limite des sommes partielles

$$^t\big(\exp(A)\big) = \exp\big(^tA\big) = \exp(-A).$$

méthode

Si $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifient $AB = BA$, on a $\exp(A)\exp(B) = \exp(A+B)$.

Puisque les matrices A et $-A$ commutent

$$^t(\exp A)\exp(A) = \exp(-A)\exp(A) = \exp(-A+A) = \exp \mathrm{O}_n = \mathrm{I}_n.$$

La matrice $\exp A$ est donc orthogonale.

(b) Soit A une matrice antisymétrique de $\mathcal{M}_n(\mathbb{R})$. La fonction f qui à un réel t associe $\mathrm{tr}\big(S\exp(tA)\big)$ admet par hypothèse un maximum en 0.

méthode

Une fonction réelle dérivable admettant un extremum en un point de part et d'autre duquel elle est définie, est de dérivée nulle en ce point.

Étudions la dérivabilité de la fonction $t \mapsto \exp(tA)$. Celle-ci est la somme de la série de fonctions $\sum u_k$ avec

$$u_k(t) = \frac{1}{k!}t^k.A^k.$$

Chacune de ces fonctions est de classe $\mathcal{C}^1$ de $\mathbb{R}$ vers $\mathcal{M}_n(\mathbb{R})$ avec

$$u_0'(t) = 0 \quad \text{et} \quad u_k'(t) = \frac{1}{(k-1)!}t^{k-1}A^k \quad \text{pour} \quad k \geqslant 1.$$

La série $\sum u_k$ converge simplement sur $\mathbb{R}$. Montrons que la série $\sum u_k'$ converge uniformément sur tout segment $[-r\,;r]$ de $\mathbb{R}$. Pour $k \geqslant 1$ et $t \in [-r\,;r]$ on a

$$\big\|u_k'(t)\big\| \leqslant \frac{1}{(k-1)!}r^{k-1}\big\|A^k\big\|$$

avec $\|\cdot\|$ la norme euclidienne associée au produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$. Or on sait [1] $\|AB\| \leqslant \|A\|\,\|B\|$ pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{R})$. On en déduit

$$\big\|u_k'(t)\big\| \leqslant \frac{1}{(k-1)!}r^{k-1}\,\|A\|^k = \|A\|\,\alpha_k \quad \text{avec} \quad \alpha_k = \frac{1}{(k-1)!}r^{k-1}\,\|A\|^{k-1}.$$

La série $\sum \alpha_k$ est convergente car il s'agit de la série exponentielle en le réel $r\,\|A\|$ et l'on peut donc affirmer que la série $\sum u_k'$ converge normalement, donc uniformément, sur tout segment $[-r\,;r]$ de $\mathbb{R}$. Par théorème, on sait alors que la fonction somme $t \mapsto \exp(tA)$ est de classe $\mathcal{C}^1$ et

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\exp(tA)\big) = \sum_{k=0}^{+\infty} u_k'(t) = \sum_{k=1}^{+\infty} \frac{1}{(k-1)!}t^{k-1}A^k.$$

1. La norme euclidienne sur $\mathcal{M}_n(\mathbb{R})$ est sous-multiplicative, voir sujet 10 p. 257.

Par continuité de l'application linéaire $X \mapsto AX$ sur $\mathcal{M}_n(\mathbb{R})$, on peut factoriser A de la somme infinie et poursuivre

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\exp(tA)\big) = A\sum_{k=1}^{+\infty}\frac{1}{(k-1)!}t^{k-1}A^{k-1} = A\exp(tA).$$

Par opérations sur les fonctions dérivables, la fonction $f\colon t \mapsto \mathrm{tr}\big(S\exp(tA)\big)$ est alors dérivable sur $\mathbb{R}$ avec

$$f'(t) = \mathrm{tr}\big(SA\exp(tA)\big).$$

Or cette fonction admet un maximum en 0 et donc $f'(0) = 0$ ce qui donne $\mathrm{tr}(SA) = 0$.

Introduisons le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ défini par $\langle M, N\rangle = \mathrm{tr}\big({}^tMN\big)$. Le résultat qui précède se relit $\langle {}^tS, A\rangle = 0$ pour toute matrice antisymétrique A. La matrice ${}^tS$ appartient donc à l'orthogonal de l'espace des matrices antisymétriques qui est l'espace des matrices symétriques [1]. Finalement, la matrice S est symétrique.

(c) Par le théorème spectral, on peut écrire $S = PD\,{}^tP$ avec $P \in \mathrm{O}_n(\mathbb{R})$ et D matrice diagonale de coefficients diagonaux $\lambda_1, \ldots, \lambda_n$. Considérons la matrice V diagonale de coefficients diagonaux $\varepsilon_1, \ldots, \varepsilon_n$ avec $\varepsilon_i = \pm 1$ déterminé de sorte que $\varepsilon_i\lambda_i = |\lambda_i|$. La matrice V est orthogonale, donc aussi l'est $U = PV\,{}^tP$. Or

$$\begin{aligned}\mathrm{tr}(SU) &= \mathrm{tr}(SPV\,{}^tP) = \mathrm{tr}\big((SPV)\,{}^tP\big) = \mathrm{tr}\big({}^tP(SPV)\big) = \mathrm{tr}({}^tPSPV)\\ &= \mathrm{tr}(DV) = |\lambda_1| + \cdots + |\lambda_n|\end{aligned}$$

et

$$\mathrm{tr}(S) = \lambda_1 + \cdots + \lambda_n.$$

La propriété $\mathrm{tr}(SU) \leqslant \mathrm{tr}(S)$ entraîne $\lambda_i \geqslant 0$ pour tout $i \in [\![1\,;n]\!]$. La matrice S est donc à valeurs propres positives.

(d) Soit $M \in \mathcal{M}_n(\mathbb{R})$.

méthode

Toute fonction réelle définie et continue sur un compact non vide présente un minimum et un maximum.

La fonction réelle $\varphi\colon U \mapsto \mathrm{tr}(MU)$ est définie et continue sur $\mathcal{M}_n(\mathbb{R})$ car linéaire. Sa restriction au départ du compact non vide $\mathrm{O}_n(\mathbb{R})$ présente un maximum en une certaine matrice U_0 de $\mathrm{O}_n(\mathbb{R})$. On a alors, pour toute matrice U de $\mathrm{O}_n(\mathbb{R})$,

$$\mathrm{tr}(MU) \leqslant \mathrm{tr}(MU_0).$$

Posons ensuite $S = MU_0$. Pour toute matrice V de $\mathrm{O}_n(\mathbb{R})$, on peut écrire

$$\mathrm{tr}(SV) = \mathrm{tr}(MU_0V) \leqslant \mathrm{tr}(MU_0) = \mathrm{tr}(S) \quad \text{car} \quad U_0V \in \mathrm{O}_n(\mathbb{R}).$$

La matrice S est alors symétrique à valeurs propres positives et l'on peut écrire $M = S\,\Omega$ avec $\Omega = U_0^{-1}$ une matrice orthogonale.

1. Voir sujet 18 du chapitre 11 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

Exercice 30 ***

Soit $M \in \mathcal{M}_n(\mathbb{R})$ vérifiant[1] ${}^tMM = M\,{}^tM$. Montrer que M est orthogonalement semblable[2] à une matrice diagonale par blocs, dont les blocs diagonaux sont de taille 1 et/ou de taille 2 de la forme[3]

$$M_{\alpha,\beta} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \quad \text{avec} \quad \alpha, \beta \in \mathbb{R}.$$

Solution

Montrons la propriété en raisonnant par récurrence double sur la taille $n \geqslant 1$ de la matrice M.

Pour $n = 1$, c'est immédiat. Pour $n = 2$, une matrice

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

commute avec sa transposée si, et seulement si,

$$\begin{cases} b^2 = c^2 \\ ac + bd = ab + cd. \end{cases}$$

Distinguons alors deux cas :

Cas : $b = c$. La matrice M est symétrique réelle donc orthogonalement diagonalisable.

Cas : $b \neq c$. On a $b = -c$ (avec $b \neq 0$) et $a = d$. La matrice M est alors égale à $M_{a,c}$.

Supposons la propriété établie aux rangs $n-2$ et $n-1$ avec $n \geqslant 3$. Considérons une matrice M de $\mathcal{M}_n(\mathbb{R})$ commutant avec sa transposée et u l'endomorphisme de $\mathbb{R}^n$ qui lui est canoniquement associé.

méthode

Un endomorphisme d'un espace réel de dimension finie non nulle admet au moins une droite ou un plan stable[4].

Soit F un sous-espace vectoriel de dimension 1 ou 2 stable par u et e une base orthonormale de E adaptée à ce sous-espace vectoriel. La matrice de u dans la base orthonormale e est de la forme

$$N = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \quad \text{avec} \quad A \in \mathcal{M}_{\dim F}(\mathbb{R}).$$

méthode

On montre que le bloc B est nul et que les matrices carrées A et C commutent avec leurs transposées respectives.

1. On dit que M est une *matrice normale*.
2. C'est-à-dire semblable par l'intermédiaire d'une matrice de passage orthogonale.
3. La matrice $M_{\alpha,\beta}$ étant diagonalisable dans $\mathbb{C}$, on peut affirmer que toute matrice commutant avec sa transposée est diagonalisable sur $\mathbb{C}$.
4. Voir sujet 33 p. 232.

Les matrices M et N figurent le même endomorphisme dans des bases orthonormales, elles sont donc orthogonalement semblables ce qui permet d'écrire

$$M = PNP^{-1} = PN\,{}^tP \quad \text{avec} \quad P \in \mathrm{O}_n(\mathbb{R}).$$

L'identité ${}^tMM = M\,{}^tM$ entraîne alors ${}^tNN = N\,{}^tN$. Or

$${}^tNN = \begin{pmatrix} {}^tAA & {}^tAB \\ {}^tBA & {}^tBB + {}^tCC \end{pmatrix} \quad \text{et} \quad N\,{}^tN = \begin{pmatrix} A\,{}^tA + B\,{}^tB & B\,{}^tC \\ C\,{}^tB & C\,{}^tC \end{pmatrix}. \qquad (*)$$

On en déduit en particulier

$${}^tAA = A\,{}^tA + B\,{}^tB.$$

En considérant la trace des deux membres, on obtient

$$\mathrm{tr}({}^tAA) = \mathrm{tr}(A\,{}^tA) + \mathrm{tr}(B\,{}^tB) = \mathrm{tr}({}^tAA) + \mathrm{tr}({}^tBB).$$

Après simplification, il vient[1] $\|B\|^2 = \mathrm{tr}({}^tBB) = 0$ ce qui entraîne que la matrice B est nulle.

Finalement, la matrice N est de la forme

$$N = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}.$$

De plus, $(*)$ donne ${}^tAA = A\,{}^tA$ et ${}^tCC = C\,{}^tC$.

Par l'étude des cas $n = 1$ et $n = 2$, la matrice A est semblable à une matrice de la forme voulue par une matrice de passage orthogonale Q'. Par l'hypothèse de récurrence, la matrice C est aussi semblable à une matrice de la forme souhaitée par une matrice orthogonale Q''. Considérons alors la matrice Q définie par blocs

$$Q = \begin{pmatrix} Q' & 0 \\ 0 & Q'' \end{pmatrix}.$$

Celle-ci est orthogonale et l'on vérifie par produit par blocs que $R^{-1}MR$ est de la forme attendue avec $R = PQ \in \mathrm{O}_n(\mathbb{R})$.

La récurrence est établie.

1. On considère ici la norme euclidienne associée au produit scalaire $\langle A, B\rangle = \mathrm{tr}({}^tAB)$.

CHAPITRE 8

Probabilités

8.1 Ensembles dénombrables

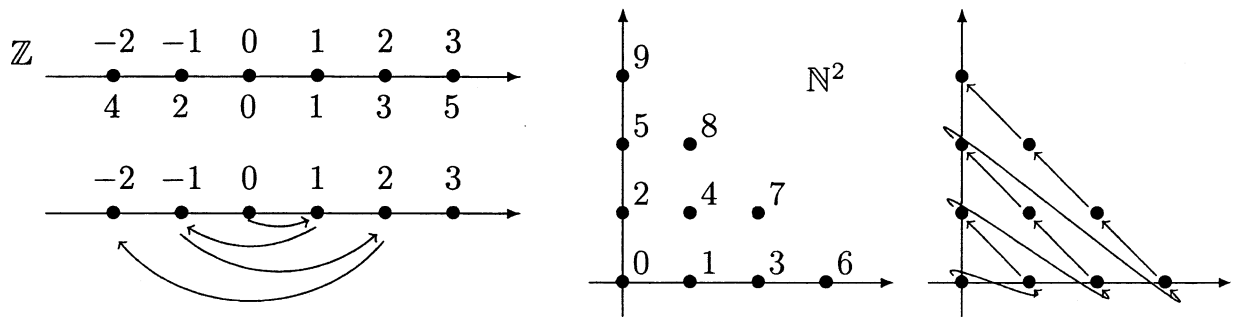
8.1.1 Définition

Définition

Un ensemble est dit *dénombrable* s'il est en bijection avec $\mathbb{N}$.

Lorsqu'un ensemble E est dénombrable, on peut introduire une application $\varphi\colon \mathbb{N} \to E$ bijective. En posant $x_n = \varphi(n)$ pour tout $n \in \mathbb{N}$, la suite (x_n) est constituée de tous les éléments de E sans répétitions. On dit que la suite (x_n) constitue une *énumération de* E.

Les figures ci-dessous illustrent des énumérations des ensembles dénombrables $\mathbb{Z}$ et $\mathbb{N}^2$:



La droite réelle $\mathbb{R}$ n'est pas un ensemble dénombrable : l'infinité des éléments de $\mathbb{R}$ est trop grande pour que l'on puisse énumérer tous les réels. Aussi, ni $\wp(\mathbb{N})$, ni $\{0,1\}^{\mathbb{N}}$ (qui est facilement en bijection avec $\wp(\mathbb{N})$) ne sont des ensembles dénombrables.

8.1.2 Ensembles au plus dénombrables

Définition

Un ensemble est dit *au plus dénombrable* lorsqu'il est fini ou dénombrable.

On peut énumérer les éléments d'un ensemble E au plus dénombrable, soit par une suite finie $(x_k)_{k\in[\![1;n]\!]}$ avec $n = \text{Card}(E)$, soit par une suite infinie $(x_n)_{n\in\mathbb{N}}$.

Un ensemble est au plus dénombrable si, et seulement si, il est en bijection avec une partie de $\mathbb{N}$.

Théorème 1

Toute partie d'un ensemble dénombrable est au plus dénombrable.

8.1.3 Opérations

Théorème 2

Un produit cartésien fini d'ensembles dénombrables est dénombrable.

On retrouve que $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ est dénombrable.

Théorème 3

Une réunion finie ou dénombrable d'ensembles dénombrables est dénombrable.

L'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable car c'est une réunion dénombrable d'ensembles dénombrables :

$$\mathbb{Q} = \bigcup_{q\in\mathbb{N}^*} A_q \quad \text{avec} \quad A_q = \left\{ \frac{p}{q} \,\middle|\, p \in \mathbb{Z} \right\} \text{ dénombrable.}$$

8.2 Espaces probabilisables

8.2.1 Univers

Définition

L'ensemble des résultats possibles d'une expérience aléatoire est appelé *univers*, il est généralement noté Ω.

Les éléments ω de Ω constituent les *issues* (ou *réalisations*) de l'expérience aléatoire. L'univers Ω peut être un ensemble fini, dénombrable ou infini non dénombrable.

8.2.2 Tribu

Définition

On appelle *tribu* sur un ensemble Ω toute partie $\mathcal{T}$ de $\wp(\Omega)$ vérifiant

1) $\Omega \in \mathcal{T}$;

2) $\forall A \in \mathcal{T},\ \overline{A} \in \mathcal{T}$;

3) $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}},\ \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

La dernière propriété s'appelle la *stabilité par réunion dénombrable.*

Une tribu contient alors nécessairement l'ensemble $\emptyset$ et est stable par intersection dénombrable

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{T} \quad \text{pour tout } (A_n)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}}.$$

Une tribu est aussi stable par union et intersection finie.

$\mathcal{T} = \wp(\Omega)$ est un exemple de tribu sur Ω, c'est la *tribu discrète.*

8.2.3 Événements

Définition

On appelle *espace probabilisable* tout couple $(\Omega, \mathcal{T})$ constitué d'un ensemble Ω et d'une tribu $\mathcal{T}$ sur Ω.

$\big(\Omega, \wp(\Omega)\big)$ est un espace probabilisable.

Définition

Si $(\Omega, \mathcal{T})$ est un espace probabilisable, les parties A de Ω éléments de la tribu $\mathcal{T}$ sont appelées *événements* de l'espace $(\Omega, \mathcal{T})$.

L'événement Ω est appelé *événement certain* alors que $\emptyset$ désigne l'*événement impossible.* Lors de l'étude d'une expérience aléatoire, un événement s'exprime en langage naturel et se comprend comme un ensemble élément de la tribu $\mathcal{T}$. Les opérations sur les événements se traduisent par des opérations sur les ensembles. Par exemple, $\overline{A}$ désigne l'*événement contraire* de l'événement A.

Définition

On dit que deux événements A et B sont *incompatibles* lorsque leur intersection est l'événement impossible, c'est-à-dire lorsque $A \cap B = \emptyset$.

8.3 Probabilités

$(\Omega, \mathcal{T})$ désigne un espace probabilisable.

8.3.1 Définition

Définition

On appelle *probabilité* sur l'espace $(\Omega, \mathcal{T})$ toute application $\mathrm{P} : \mathcal{T} \to [0\,;1]$ vérifiant $\mathrm{P}(\Omega) = 1$ et, pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles,

$$\sum \mathrm{P}(A_n) \text{ converge} \quad \text{et} \quad \mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathrm{P}(A_n).$$

Cette dernière propriété est l'*additivité par union dénombrable.*

Cette notion étend aux univers infinis le concept de probabilité défini dans le cours de première année limité aux univers finis. Il n'est cependant plus possible de calculer la probabilité de n'importe quelle partie de Ω : on ne peut calculer la probabilité que d'un événement, c'est-à-dire d'un élément de la tribu $\mathcal{T}$.

Définition

On appelle *espace probabilisé* tout triplet $(\Omega, \mathcal{T}, \mathrm{P})$ formé d'un ensemble Ω, d'une tribu $\mathcal{T}$ sur Ω et d'une probabilité P sur $(\Omega, \mathcal{T})$.

8.3.2 Probabilité sur un univers au plus dénombrable

Soit Ω un univers fini ou dénombrable muni de la tribu discrète $\mathcal{T} = \wp(\Omega)$. Pour une issue $\omega \in \Omega$, on note p_ω la probabilité de l'*événement élémentaire* $\{\omega\}$. La famille $(p_\omega)_{\omega\in\Omega}$ est une famille de réels positifs, sommable et de somme égale à 1. Inversement :

Théorème 4

Si $(p_\omega)_{\omega\in\Omega}$ est une famille de réels positifs, sommable et de somme égale à 1, il existe une unique probabilité P sur $\big(\Omega, \wp(\Omega)\big)$ telle que $p_\omega = \mathrm{P}(\{\omega\})$ pour tout $\omega \in \Omega$.
Celle-ci est déterminée par

$$\mathrm{P}(A) = \sum_{\omega\in A} p_\omega \quad \text{pour tout } A \in \wp(\Omega).$$

Lorsque l'univers Ω n'est plus fini ou dénombrable, il est beaucoup plus difficile de définir une probabilité sur Ω et la tribu des événements est rarement $\wp(\Omega)$.

8.3.3 Propriétés

Soit $(\Omega, \mathcal{T}, \mathrm{P})$ un espace probabilisé.
Les propriétés calculatoires déjà vues en première année restent valides : lorsque A et B sont deux événements

$$A \cap B = \emptyset \implies \mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B).$$

Plus généralement, on a l'égalité

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

On dispose aussi d'une propriété de *croissance*

$$A \subset B \implies \mathrm{P}(A) \leqslant \mathrm{P}(B).$$

Enfin, on a l'identité $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$ qui est utile, notamment pour calculer une probabilité par considération de l'événement contraire.

Ces propriétés se complètent des résultats de *continuité monotone* qui suivent :

Théorème 5 (Continuité croissante)
Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements (c'est-à-dire $A_n \subset A_{n+1}$ pour tout entier $n \in \mathbb{N}$) alors la suite $\big(\mathrm{P}(A_n)\big)$ converge et

$$\mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n).$$

Ce résultat est utile pour calculer la probabilité d'une union dénombrable comme limite des probabilités des unions partielles :

$$\mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathrm{P}\left(\bigcup_{k=0}^{n} A_k\right).$$

En particulier, on peut généraliser l'*inégalité de Boole* :

$$\mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n=0}^{+\infty} \mathrm{P}(A_n)$$

quitte à considérer le second membre égal à $+\infty$ si la série associée diverge.

Théorème 6 (Continuité décroissante)
Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements (c'est-à-dire $A_{n+1} \subset A_n$ pour tout entier $n \in \mathbb{N}$) alors la suite $\big(\mathrm{P}(A_n)\big)$ converge et

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n).$$

Ce résultat permet de calculer la probabilité d'une intersection dénombrable par limite des probabilités des intersections partielles :

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathrm{P}\left(\bigcap_{k=0}^{n} A_k\right).$$

8.3.4 Événements négligeables, événements presque sûrs

Soit $(\Omega, \mathcal{T}, \mathrm{P})$ un espace probabilisé.

Définition

On dit qu'un événement A est *négligeable* si $\mathrm{P}(A) = 0$.

L'événement impossible est négligeable. Un événement négligeable n'est pas pour autant impossible.

Un événement inclus dans un événement négligeable est *a fortiori* négligeable.

Une réunion finie ou dénombrable d'événements négligeables est négligeable.

Définition

On dit qu'un événement A est *presque sûr*[1] si $\mathrm{P}(A) = 1$.

Ceci signifie encore que l'événement contraire $\overline{A}$ est négligeable.

L'événement certain est presque sûr...

Un événement contenant un événement presque sûr est presque sûr. Une intersection finie ou dénombrable d'événements presque sûrs est presque sûre.

Pour montrer qu'un événement A est presque sûr, il est fréquent d'étudier $\overline{A}$ car celui est souvent plus facile à décrire puisque « petit ».

8.4 Probabilités conditionnelles et indépendance

$(\Omega, \mathcal{T}, \mathrm{P})$ désigne un espace probabilisé.

8.4.1 Événements indépendants

Définition

On dit que deux événements A et B de l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ sont *indépendants* si $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$.

Il ne faut pas confondre l'indépendance et l'incompatibilité : deux événements incompatibles sont rarement indépendants ! L'indépendance permet de calculer la probabilité d'une intersection tandis que l'incompatibilité sert au calcul de la probabilité d'une union.

Définition

Soit $(A_i)_{i\in I}$ une famille d'événements de l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ avec I un ensemble d'indexation quelconque. On dit que les événements de la famille $(A_i)_{i\in I}$ sont *mutuellement indépendants* si, pour tout $m \in \mathbb{N}^*$ et tous $i_1, \ldots, i_m \in I$ deux à deux distincts,

$$\mathrm{P}\left(\bigcap_{k=1}^{m} A_{i_k}\right) = \prod_{k=1}^{m} \mathrm{P}(A_{i_k}).$$

L'indépendance mutuelle d'une famille d'événements est souvent une conséquence de la modélisation choisie de l'expérience étudiée. On la rencontre lors de la répétition d'expériences : lancers successifs d'une pièce, tirage avec remise.

1. On parle parfois d'événement *quasi certain*.

8.4.2 Probabilités conditionnelles

Soit A un événement de Ω vérifiant $\mathrm{P}(A) > 0$.

Définition

Pour tout événement B de Ω, on définit la probabilité conditionnelle de B sachant A par [1]

$$\mathrm{P}(B \mid A) \overset{\text{déf}}{=} \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)}.$$

Si A et B sont indépendants, on remarque que $\mathrm{P}(B \mid A) = \mathrm{P}(B)$.

Une probabilité conditionnelle $\mathrm{P}(B \mid A)$ est aussi notée $\mathrm{P}_A(B)$ ce qui introduit l'application

$$\mathrm{P}_A : \begin{cases} \mathcal{T} \to [0\,;1] \\ B \mapsto \mathrm{P}_A(B) = \mathrm{P}(B \mid A). \end{cases}$$

Celle-ci est une probabilité sur $(\Omega, \mathcal{T})$ ce qui autorise pour les probabilités conditionnelles les propriétés calculatoires vues précédemment pour les probabilités.

Lorsque A est un événement tel que $\mathrm{P}(A) = 0$, on ne peut pas définir la probabilité d'un événement B sachant A par le quotient précédent. Il est usuel de poser $\mathrm{P}(B \mid A) = 0$ dans ce cas. L'application P_A ne désigne alors pas une probabilité.

8.4.3 Formule des probabilités composées

Théorème 7 (Formule des probabilités composées)

Si A et B sont deux événements de Ω,

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B \mid A).$$

Lorsque deux événements A et B ne sont pas indépendants mais que l'un peut être mesuré lorsque l'autre est réalisé, la formule des probabilités composées permet le calcul de $\mathrm{P}(A \cap B)$. Ceci est notamment utile pour les expériences présentant une succession d'épreuves non indépendantes.

Plus généralement, si $A_1, \ldots, A_n$ sont des événements de Ω,

$$\mathrm{P}(A_1 \cap \ldots \cap A_n) = \mathrm{P}(A_1)\,\mathrm{P}(A_2 \mid A_1) \ldots \mathrm{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1}).$$

8.4.4 Formule des probabilités totales

Soit $(A_i)_{i \in I}$ une famille d'événements de l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ avec I un ensemble fini ou dénombrable.

Définition

On dit que la famille $(A_i)_{i \in I}$ est un *système complet*[2] *d'événements* si les événements A_i sont deux à deux incompatibles et de réunion Ω.

1. La notation $\mathrm{P}(B|A)$ peut prêter à confusion : $(B|A)$ ne désigne pas un événement dont nous calculons la probabilité par P !

Théorème 8 (Formule des probabilités totales)
Si $(A_i)_{i\in I}$ est un système complet[3] d'événements de l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ alors, pour tout événement B, la famille $\big(\mathrm{P}(B\,|\,A_i)\,\mathrm{P}(A_i)\big)_{i\in I}$ est sommable et

$$\mathrm{P}(B) = \sum_{i\in I} \mathrm{P}(B\,|\,A_i)\,\mathrm{P}(A_i).$$

La formule des probabilités totales permet les études exhaustives : les A_i déterminent tous les cas possibles et les probabilités conditionnelles estiment la réalisation de B dans chaque cas. Cette formule est utile pour évaluer la probabilité d'un événement qui apparaît sur plusieurs feuilles d'une expérience visualisée par un arbre.

8.4.5 Formule de Bayes

Théorème 9 (Formule de Bayes)
Si A et B sont deux événements avec B de probabilité non nulle,

$$\mathrm{P}(A\,|\,B) = \frac{\mathrm{P}(B\,|\,A)\,\mathrm{P}(A)}{\mathrm{P}(B)}.$$

La formule de Bayes est utile pour les raisonnements « rétroactifs » : lorsque l'on sait la réalisation de B, elle permet d'estimer la probabilité que A en soit la cause.

8.5 Exercices d'apprentissage

Exercice 1
Soit Ω un ensemble. On introduit

$$\mathcal{T} = \big\{A \subset \Omega \,\big|\, A \text{ ou } \overline{A} \text{ est au plus dénombrable}\big\}.$$

(a) Vérifier que $\mathcal{T}$ est une tribu sur Ω.

(b) On suppose Ω infini non dénombrable.
Montrer que l'on définit une probabilité P sur $(\Omega, \mathcal{T})$ en posant, pour tout A de $\mathcal{T}$,

$$\mathrm{P}(A) = \begin{cases} 0 & \text{si } A \text{ au plus dénombrable} \\ 1 & \text{si } \overline{A} \text{ au plus dénombrable.} \end{cases}$$

2. On parle aussi de *système quasi complet d'événements* lorsque les A_i sont deux à deux incompatibles et que leur union est un événement presque sûr.
3. Ou plus généralement, un système quasi complet d'événements.

Solution

(a) **méthode**

> On vérifie que $\mathcal{T}$ est une partie[1] de $\wp(\Omega)$ contenant Ω, stable par passage au complémentaire et stable par union dénombrable.

Par définition $\mathcal{T}$ est formée de parties de Ω et donc $\mathcal{T} \subset \wp(\Omega)$. La partie Ω appartient à $\mathcal{T}$ car $\overline{\Omega} = \emptyset$ est une partie finie donc au plus dénombrable.

Si $A \in \mathcal{T}$, on a A ou $\overline{A}$ au plus dénombrable donc $\overline{A}$ ou $\overline{\overline{A}}$ est au plus dénombrable. Ainsi, $\overline{A} \in \mathcal{T}$ et l'ensemble $\mathcal{T}$ est stable par passage au complémentaire. Il reste à vérifier la stabilité par union dénombrable. Étudions

$$A = \bigcup_{n\in\mathbb{N}} A_n \quad \text{avec} \quad (A_n)_{n\in\mathbb{N}} \text{ une suite d'éléments de } \mathcal{T}.$$

Cas : Tous les A_n sont au plus dénombrables. La partie A est une union dénombrable d'ensembles au plus dénombrables, c'est une partie au plus dénombrable (Th. 3 p. 332).

Cas : L'un des A_n est infini non dénombrable. Notons n_0 l'indice correspondant. La partie A_{n_0} étant élément de $\mathcal{T}$, on a nécessairement $\overline{A_{n_0}}$ au plus dénombrable. Or

$$\overline{A} = \overline{\bigcup_{n\in\mathbb{N}} A_n} \subset \bigcap_{n\in\mathbb{N}} \overline{A_n} \subset \overline{A_{n_0}}$$

donc $\overline{A}$ est au plus dénombrable car inclus dans une partie qui l'est (Th. 1 p. 332).

Dans les deux cas, on peut affirmer que l'union des A_n est élément de $\mathcal{T}$.

(b) **méthode**

> On observe que P est une application de $\mathcal{T}$ vers $[0\,;1]$ vérifiant $\mathrm{P}(\Omega) = 1$ et la propriété d'additivité dénombrable.

Commençons par souligner que l'application P est bien définie à valeurs dans $[0\,;1]$ car une partie A de Ω ne peut vérifier A au plus dénombrable et $\overline{A}$ au plus dénombrable puisque $\Omega = A \cup \overline{A}$ est infini non dénombrable. Ainsi, l'application P est définie sans ambiguïté.

On a immédiatement $\mathrm{P}(\Omega) = 1$ car Ω est infini non dénombrable.

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{T}$ deux à deux disjoints. Vérifions

$$\mathrm{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathrm{P}(A_n) \qquad (*)$$

avec convergence de la série de terme général $\mathrm{P}(A_n)$.

Cas : Tous les A_n sont au plus dénombrables.

L'union des A_n est alors au plus dénombrable et l'égalité $(*)$ se relit $0 = 0$.

1. Une tribu est un ensemble d'ensembles.

Cas : L'un des A_n est infini non dénombrable. Notons n_0 l'indice correspondant. On sait alors $\overline{A_{n_0}}$ au plus dénombrable. Puisque les A_n sont deux à deux disjoints, on a $A_n \subset \overline{A_{n_0}}$ pour tout n distinct de n_0 et donc A_n est au plus dénombrable. On en déduit que seul A_{n_0} est infini non dénombrable et l'égalité $(*)$ se relit $1 = 1$.

Finalement, P est une probabilité sur $(\Omega, \mathcal{T})$.

Exercice 2

À quelles conditions sur la suite réelle $(a_n)_{n\in\mathbb{N}}$ existe-t-il une probabilité P sur l'espace probabilisé $\big(\mathbb{N}, \wp(\mathbb{N})\big)$ vérifiant

$$\mathrm{P}\big(\{n, n+1, \ldots\}\big) = a_n \quad \text{pour tout } n \in \mathbb{N}.$$

Solution

Posons $A_n = \{n, n+1, \ldots\}$ pour tout $n \in \mathbb{N}$.

méthode

Une probabilité sur un univers dénombrable Ω muni de la tribu $\wp(\Omega)$ est entièrement déterminée par la connaissance de la famille $\big(\mathrm{P}(\{\omega\})\big)_{\omega\in\Omega}$ qui est une famille de réels positifs sommable et de somme 1 (Th. 4 p. 334).

Analyse : Supposons l'existence de la probabilité P telle que voulue. Puisque $A_0 = \mathbb{N}$, on a $a_0 = \mathrm{P}(\mathbb{N}) = 1$. De plus, on détermine $\mathrm{P}(\{n\})$ pour $n \in \mathbb{N}^*$ en écrivant la réunion disjointe[1] $A_n = \{n\} \cup A_{n+1}$. On a donc par additivité

$$\mathrm{P}(A_n) = \mathrm{P}\big(\{n\}\big) + \mathrm{P}(A_{n+1})$$

ce qui donne $\mathrm{P}(\{n\}) = a_n - a_{n+1}$. Une probabilité étant à valeurs positives, la suite (a_n) est nécessairement décroissante. Enfin, par continuité décroissante (Th. 6 p. 335)

$$\mathrm{P}(\emptyset) = \mathrm{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} a_n \quad \text{car} \quad A_{n+1} \subset A_n \text{ pour tout } n \in \mathbb{N}.$$

On en déduit que la suite (a_n) est de limite nulle. Vérifions que ces conditions sont suffisantes.

Synthèse : Considérons (a_n) une suite de réels décroissante de limite nulle et vérifiant $a_0 = 1$. Posons $p_n = a_n - a_{n+1}$ pour tout $n \in \mathbb{N}$. Les p_n sont des réels positifs et par calcul d'une somme télescopique

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} \big(a_k - a_{k+1}\big) = a_0 - a_{n+1} \xrightarrow[n\to+\infty]{} a_0 = 1.$$

La série $\sum p_n$ est donc convergente de somme égale à 1. La famille $(p_n)_{n\in\mathbb{N}}$ est donc une famille de réels positifs sommable et de somme 1 : il existe une probabilité P sur l'espace probabilisable $\big(\mathbb{N}, \wp(\mathbb{N})\big)$ vérifiant $\mathrm{P}(\{n\}) = p_n$ pour tout $n \in \mathbb{N}$.

1. On dit d'une réunion qu'elle est *disjointe* lorsque les ensembles réunis sont deux à deux disjoints. On utilise quelquefois les symboles $\sqcup$ ou $\uplus$ pour écrire cette opération.

Il reste à vérifier que cette probabilité convient ce qui se résout par un calcul télescopique analogue au précédent. Pour tout $n \in \mathbb{N}$, les $\{k\}$ avec $k \geqslant n$ étant deux à deux incompatibles, on a par additivité dénombrable

$$\mathrm{P}(A_n) = \mathrm{P}\left(\bigcup_{k\geqslant n} \{k\}\right) = \sum_{k=n}^{+\infty} \mathrm{P}(\{k\}) = \sum_{k=n}^{+\infty} p_k = \sum_{k=n}^{+\infty} (a_k - a_{k+1}) = a_n.$$

Exercice 3

On lance indéfiniment un dé équilibré à six faces et l'on admet [1] qu'il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ qui permet d'étudier la succession des valeurs obtenues.

(a) Déterminer la probabilité d'obtenir un 'six' pour la première fois lors du n-ième lancer.

(b) Montrer qu'il est presque sûr d'obtenir un 'six'.

(c) Montrer qu'il est presque sûr d'obtenir une infinité de 'six'.

Solution

(a) Pour tout $n \in \mathbb{N}^*$, on introduit les événements :

A_n = « On obtient pour la première fois la valeur 'six' lors du n-ième lancer »,

D_n = « On obtient la valeur 'six' lors du n-ième lancer ».

Puisque le dé est supposé équilibré, on sait $\mathrm{P}(D_n) = 1/6$.

La question posée consiste à calculer la probabilité de A_n : on exprime [2] A_n en fonction des événements D_n

$$A_n = \overline{D_1} \cap \ldots \cap \overline{D_{n-1}} \cap D_n.$$

Sans que l'hypothèse soit explicite, on suppose les lancers indépendants ce qui permet d'affirmer l'indépendance mutuelle des événements $\overline{D_1}, \ldots, \overline{D_{n-1}}$ et D_n et de terminer le calcul

$$\mathrm{P}(A_n) = \mathrm{P}(\overline{D_1}) \times \cdots \times \mathrm{P}(\overline{D_{n-1}}) \times \mathrm{P}(D_n) = \frac{1}{6}\left(\frac{5}{6}\right)^{n-1}.$$

(b) On étudie :

A = « On obtient au moins un 'six' lors de l'expérience ».

On peut exprimer A qui est de nature existentielle par une réunion dénombrable ce qui assure qu'il s'agit d'un événement :

$$A = \bigcup_{n\in\mathbb{N}^*} D_n.$$

1. L'ensemble Ω des suites d'entiers compris entre 1 et 6 n'est pas dénombrable, il n'est alors pas immédiat de définir une tribu et une probabilité permettant d'étudier l'expérience en cours : le Th. 13 p. 377 du chapitre suivant assure l'existence de l'univers introduit.

2. Sachant que les D_n sont des événements, l'écriture qui suit assure que A_n en est aussi un par opérations dans la tribu $\mathcal{T}$. Lorsque $n = 1$, cette écriture se comprend $A_1 = D_1$.

Les événements D_n n'étant pas incompatibles, on ne peut calculer directement la probabilité de cette union. On peut alors exprimer A comme la réunion disjointe des A_n mais il est plus commode d'étudier l'événement contraire :

$$\overline{A} = \text{« On n'obtient jamais de 'six' lors de l'expérience »}.$$

Cet événement de nature universelle s'exprime par une intersection

$$\overline{A} = \bigcap_{n\in\mathbb{N}^*} \overline{D_n}.$$

méthode

On calcule la probabilité d'une intersection dénombrable par continuité monotone (Th. 6 p. 335).

Par continuité décroissante

$$\mathrm{P}(\overline{A}) = \lim_{n\to+\infty} \mathrm{P}\left(\bigcap_{k=1}^{n} \overline{D_k}\right).$$

Les événements intersectés étant indépendants, la probabilité de l'intersection est le produit de leur probabilité

$$\mathrm{P}(\overline{A}) = \lim_{n\to+\infty} \prod_{k=1}^{n} \mathrm{P}(\overline{D_k}) = \lim_{n\to+\infty} \left(\frac{5}{6}\right)^n = 0.$$

L'événement $\overline{A}$ est négligeable et il est donc presque sûr [1] d'obtenir un 'six' lors de la succession de lancers.

(c) On étudie l'expression contraire :

$$F = \text{« On n'obtient qu'un nombre fini de 'six' lors de l'expérience »}.$$

Ceci signifie encore que les lancers réalisés ne donnent plus de 'six' à partir d'un certain rang,

méthode

On exprime F par opérations ensembliste dans la tribu $\mathcal{T}$.

L'événement « On obtient plus de 'six' à partir du rang N » s'exprime par l'intersection dénombrable

$$\bigcap_{n\geqslant N} \overline{D_n}.$$

Comme au-dessus, on obtient par continuité décroissante que cette intersection est un événement négligeable.

1. Il existe des successions de lancers qui ne produisent jamais de 'six' mais on vient d'établir que l'ensemble de celles-ci est de probabilité nulle.

F traduisant l'existence d'un rang à partir duquel il n'y a plus de 'six' s'exprime par une réunion dénombrable

$$F = \bigcup_{N\in\mathbb{N}^*} \bigcap_{n\geqslant N} \overline{D_n}.$$

Cette écriture justifie que F est bien un événement et qu'il est négligeable en tant que réunion dénombrable d'événements qui le sont.

Exercice 4

Une urne contient une boule blanche. Un joueur lance un dé équilibré à six faces. S'il obtient un 'six', il tire une boule dans l'urne. Sinon, il rajoute une boule rouge dans l'urne et répète la manipulation. On admet l'existence d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ permettant l'étude de cette expérience.

(a) Quelle est la probabilité que le joueur tire la boule blanche ?

On donne

$$\sum_{n=1}^{+\infty} \frac{1}{n}x^n = -\ln(1-x) \quad \text{pour tout } x \in]-1\,;1[.$$

(b) On suppose que le joueur a tiré la boule blanche. Quelle est la probabilité qu'il n'y ait pas d'autres boules dans l'urne ?

Solution

Soit $n \in \mathbb{N}^*$. On introduit les événements suivants :

A_n = « Le joueur obtient pour la première fois un 'six' lors du n-ième lancer »,

A_∞ = « Le joueur n'obtient jamais de 'six' »,

B = « La boule tirée est blanche ».

L'étude menée dans le sujet précédent a donné

$$\mathrm{P}(A_n) = \frac{1}{6}\left(\frac{5}{6}\right)^{n-1} \quad \text{et} \quad \mathrm{P}(A_\infty) = 0.$$

(a) On veut déterminer la probabilité de B ce qui nécessite de connaître la composition de l'urne.

méthode

On utilise la formule des probabilités totales (Th. 8 p. 338) en identifiant un système complet d'événements adapté à l'étude.

Les événements A_n (avec $n \in \mathbb{N}^*$) et A_∞ constituent un système complet d'événements. Par la formule des probabilités totales

$$\mathrm{P}(B) = \sum_{n=1}^{+\infty} \mathrm{P}(B\,|\,A_n)\,\mathrm{P}(A_n) + \mathrm{P}(B\,|\,A_\infty)\,\mathrm{P}(A_\infty).$$

Exercice 6 ** (Inégalités de Fatou)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$.
On introduit

$$A_* = \bigcup_{p\in\mathbb{N}} \bigcap_{n\geqslant p} A_n \text{ et } A^* = \bigcap_{p\in\mathbb{N}} \bigcup_{n\geqslant p} A_n.$$

(a) Vérifier que A_* et A^* sont des événements. Comment interpréter simplement ceux-ci ?

(b) Montrer $A_* \subset A^*$.

(c) Montrer les inégalités

$$\mathrm{P}(A_*) \leqslant \lim_{p\to+\infty} \left(\inf_{n\geqslant p} \mathrm{P}(A_n) \right) \quad \text{et} \quad \lim_{p\to+\infty} \left(\sup_{n\geqslant p} \mathrm{P}(A_n) \right) \leqslant \mathrm{P}(A^*).$$

(d) Déterminer un exemple où ces inégalités sont strictes.

Solution

Pour $p \in \mathbb{N}$. Introduisons

$$B_p = \bigcap_{n\geqslant p} A_n \quad \text{et} \quad C_p = \bigcup_{n\geqslant p} A_n.$$

(a) Pour tout $p \in \mathbb{N}$, on peut affirmer que B_p est un événement car intersection dénombrable d'événements. On en déduit que A_* est un événement par réunion dénombrable d'événements. L'argumentation est analogue pour A^*.

méthode

Une réunion d'événements traduit une quantification existentielle, une intersection traduit une quantification universelle.

L'événement B_p signifie la réalisation [1] de tous les A_n au delà du rang p. La réunion définissant A_* signifie donc l'existence d'un rang au delà duquel les événements A_n sont tous réalisés. En résumé, A_* signifie que les événements A_n sont tous réalisés à partir d'un certain rang. L'événement C_p signifie l'existence d'au moins un A_n réalisé parmi ceux d'indices supérieurs à p. L'intersection définissant A^* signifie que ceci a lieu pour tout p... L'événement contraire est cependant plus simple [2] à comprendre

$$\overline{A^*} = \bigcap_{p\in\mathbb{N}} \bigcup_{n\geqslant p} \overline{A_n}$$

signifie que les A_n ne sont plus réalisés au delà d'un certain rang. L'événement A^* signifie alors qu'il y a une infinité de A_n réalisés.

1. Dire qu'un événement est *réalisé* signifie que l'issue de l'expérience aléatoire en est élément. Étudier la réalisation d'un événement est une façon de décrire celui-ci.

2. Ce qui précède peut néanmoins s'interpréter : il existe des rangs arbitrairement grands pour lesquels A_n est réalisé.

(b) Si une issue ω est élément de A_*, il existe $p \in \mathbb{N}$ tel que ω est élément de tous les A_n d'indices supérieurs à p. Pour tout $q \in \mathbb{N}$, l'issue ω est alors élément de $A_{\max(p,q)}$ et donc élément de la réunion des A_m pour $m \geqslant q$. C'est par conséquent un élément de A^* ce qui justifie l'inclusion [1] $A_* \subset A^*$.

(c) Commençons par justifier l'existence des limites introduites en constatant que les suites associées sont monotones et bornées. Pour tout $p \in \mathbb{N}$, on a

$$\{\mathrm{P}(A_n) \mid n \geqslant p+1\} \subset \{\mathrm{P}(A_n) \mid n \geqslant p\}.$$

Une borne inférieure du second ensemble est donc un minorant du premier ce qui entraîne

$$\inf_{n \geqslant p+1} \mathrm{P}(A_n) \geqslant \inf_{n \geqslant p} \mathrm{P}(A_n).$$

La suite des bornes inférieures est donc croissante et puisqu'elle est majorée par 1, c'est une suite convergente. De même, on établit la convergence de la suite des bornes supérieures qui est décroissante et minorée par 0.

méthode

Par continuité monotone (Th. 5 et Th. 6 p. 335) , la probabilité d'une union croissante dénombrable (ou d'une intersection décroissante dénombrable) est la limite des probabilités des événements considérés.

Soit $p \in \mathbb{N}$. Pour tout $n \geqslant p$,

$$\mathrm{P}(B_p) \leqslant \mathrm{P}(A_n) \quad \text{car} \quad B_p = \bigcap_{m \geqslant p} A_m \subset A_n.$$

Une borne inférieure étant le plus grand des minorants, on obtient

$$\mathrm{P}(B_p) \leqslant \inf_{n \geqslant p} \mathrm{P}(A_n).$$

Enfin, les événements B_p constituent une suite croissante car $B_p \subset B_{p+1}$ et donc, par continuité monotone,

$$\mathrm{P}(A_*) = \lim_{p \to +\infty} \mathrm{P}(B_p) \leqslant \lim_{p \to +\infty} \left(\inf_{n \geqslant p} \mathrm{P}(A_n) \right).$$

La deuxième inégalité se traite par une étude analogue. On commence par observer que C_p contient tous les A_n d'indices supérieurs à p ce qui entraîne

$$\sup_{n \geqslant p} \mathrm{P}(A_n) \leqslant \mathrm{P}(C_p).$$

On conclut à l'inégalité voulue par continuité monotone sachant que la suite (C_p) est décroissante.

1. Si les A_n sont réalisés à partir d'un certain rang, il y a une infinité de A_n réalisés.

(d) On considère l'univers $\Omega = \{0,1\}$ muni de la tribu discrète et de la probabilité uniforme. Pour tout $k \in \mathbb{N}$, on introduit les événements

$$A_{2k} = \{0\} \quad \text{et} \quad A_{2k+1} = \{1\}.$$

On a $\mathrm{P}(A_n) = 1/2$ pour tout $n \in \mathbb{N}$ donc

$$\lim_{p\to+\infty} \left(\inf_{n\geqslant p} \mathrm{P}(A_n) \right) = \lim_{p\to+\infty} \left(\sup_{n\geqslant p} \mathrm{P}(A_n) \right) = \frac{1}{2}$$

tandis que

$$\mathrm{P}(A_*) = \mathrm{P}(\emptyset) = 0 \quad \text{et} \quad \mathrm{P}\big(A^*\big) = \mathrm{P}(\Omega) = 1.$$

Exercice 7 * (Lemme de Borel-Cantelli)**

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$. On considère l'événement

$$A^* = \bigcap_{p\in\mathbb{N}} \bigcup_{n\geqslant p} A_n.$$

(a) On suppose la convergence de la série $\sum \mathrm{P}(A_n)$. Montrer $\mathrm{P}\big(A^*\big) = 0$.

(b) Inversement, on suppose la divergence de la série $\sum \mathrm{P}(A_n)$ ainsi que l'indépendance mutuelle des A_n. Montrer $\mathrm{P}\big(A^*\big) = 1$.

Solution

(a) Pour $p \in \mathbb{N}$, introduisons

$$C_p = \bigcup_{n\geqslant p} A_n.$$

La suite d'événements (C_p) est décroissante car $C_{p+1} \subset C_p$. On a donc par continuité monotone

$$\mathrm{P}\big(A^*\big) = \lim_{p\to+\infty} \mathrm{P}(C_p). \qquad (*)$$

La probabilité de l'union dénombrable définissant C_p est la limite des probabilités des unions partielles

$$\mathrm{P}(C_p) = \lim_{N\to+\infty} \mathrm{P}\left(\bigcup_{n=p}^{N} A_n \right).$$

Or, l'inégalité de Boole donne

$$\mathrm{P}\left(\bigcup_{n=p}^{N} A_n \right) \leqslant \sum_{n=p}^{N} \mathrm{P}(A_n) \quad \text{et par passage à la limite} \quad \mathrm{P}(C_p) \leqslant \sum_{n=p}^{+\infty} \mathrm{P}(A_n).$$

Le majorant est ici le reste d'une série convergente, il est donc de limite nulle lorsque p tend vers l'infini. Par théorème de convergence par encadrement, on conclut $\mathrm{P}\big(A^*\big) = 0$.

(b) **méthode**

‖ On étudie l'événement contraire.

En passant à l'événement contraire l'égalité $(*)$, on obtient

$$\mathrm{P}\left(\overline{A^*}\right) = \lim_{p\to+\infty} \mathrm{P}(\overline{C_p}) \quad \text{avec} \quad \overline{C_p} = \bigcap_{n\geqslant p} \overline{A_n}.$$

On montre que l'événement $\overline{C_p}$ est négligeable en étudiant les probabilités des intersections partielles. Par continuité monotone, on a

$$\mathrm{P}(\overline{C_p}) = \lim_{N\to+\infty} \mathrm{P}\left(\bigcap_{n=p}^{N} \overline{A_n}\right).$$

Soit $N \geqslant p$. On sait par indépendance mutuelle[1]

$$\mathrm{P}\left(\bigcap_{n=p}^{N} \overline{A_n}\right) = \prod_{n=p}^{N} \mathrm{P}(\overline{A_n}) = \prod_{n=p}^{N} (1 - \mathrm{P}(A_n)).$$

méthode

‖ On emploie l'inégalité de convexité $1 - x \leqslant \mathrm{e}^{-x}$ valable pour tout $x \in \mathbb{R}$.

On obtient

$$0 \leqslant \mathrm{P}\left(\bigcap_{n=p}^{N} \overline{A_n}\right) \leqslant \prod_{n=p}^{N} \mathrm{e}^{-\mathrm{P}(A_n)} = \exp\left(-\sum_{n=p}^{N} \mathrm{P}(A_n)\right).$$

Par la divergence de la série à termes positifs $\sum \mathrm{P}(A_n)$, il vient

$$\sum_{n=p}^{N} \mathrm{P}(A_n) \xrightarrow[N\to+\infty]{} +\infty$$

et donc, par théorème de convergence par encadrement, $\mathrm{P}(\overline{C_p}) = 0$. On peut alors conclure $\mathrm{P}\left(\overline{A^*}\right) = 0$ puis $\mathrm{P}\left(A^*\right) = 1$.

8.6.2 Tribus

Exercice 8 **

Soit B un événement d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$. Montrer que l'ensemble des événements indépendants de B forme une tribu sur Ω.

1. L'indépendance mutuelle des événements entraîne celle des événements contraires : voir sujet 13 du chapitre 12 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

Solution

Notons $\mathcal{A}$ l'ensemble des événements de Ω indépendants de B, c'est-à-dire l'ensemble des $A \in \mathcal{T}$ vérifiant $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$.

méthode

On vérifie que $\mathcal{A}$ possède Ω et est stable par passage au complémentaire et par réunion dénombrable.

L'événement certain Ω est indépendant de B car $\mathrm{P}(\Omega \cap B) = \mathrm{P}(B) = \mathrm{P}(\Omega)\,\mathrm{P}(B)$.

Si A est un événement indépendant de B, on peut écrire par additivité puis indépendance

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}\big((A \cap B) \cup (\overline{A} \cap B)\big)\\ &= \mathrm{P}(A \cap B) + \mathrm{P}(\overline{A} \cap B)\\ &= \mathrm{P}(A)\,\mathrm{P}(B) + \mathrm{P}(\overline{A} \cap B)\end{aligned}$$

On en déduit

$$\begin{aligned}\mathrm{P}(\overline{A} \cap B) &= \mathrm{P}(B) - \mathrm{P}(A)\,\mathrm{P}(B)\\ &= \big(1 - \mathrm{P}(A)\big)\,\mathrm{P}(B)\\ &= \mathrm{P}(\overline{A})\,\mathrm{P}(B).\end{aligned}$$

Ainsi, l'événement contraire $\overline{A}$ appartient à $\mathcal{A}$.

Enfin, soit (A_n) une suite d'événements indépendants de B deux à deux incompatibles. Par distributivité de l'intersection sur l'union

$$\mathrm{P}\left(\left(\bigcup_{n\in\mathbb{N}} A_n\right) \cap B\right) = \mathrm{P}\left(\bigcup_{n\in\mathbb{N}} (A_n \cap B)\right).$$

Les événements $A_n \cap B$ sont deux à deux incompatibles et l'on a donc par additivité dénombrable

$$\mathrm{P}\left(\bigcup_{n\in\mathbb{N}} (A_n \cap B)\right) = \sum_{n=0}^{+\infty} \mathrm{P}(A_n \cap B) \quad \text{avec convergence de } \sum \mathrm{P}(A_n \cap B).$$

Or $\mathrm{P}(A_n \cap B) = \mathrm{P}(A_n)\,\mathrm{P}(B)$ et donc

$$\mathrm{P}\left(\left(\bigcup_{n\in\mathbb{N}} A_n\right) \cap B\right) = \left(\sum_{n=0}^{+\infty} \mathrm{P}(A_n)\right) \mathrm{P}(B) = \mathrm{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \mathrm{P}(B).$$

Ainsi, la réunion des A_n est élément de $\mathcal{A}$ et l'on peut conclure que $\mathcal{A}$ est une tribu sur Ω.

Exercice 9 ***

Soit Ω un univers.

(a) Soit $(\mathcal{A}_i)_{i\in I}$ une famille de tribus sur Ω. Montrer que $\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$ détermine une tribu sur Ω.

(b) Soit $\mathcal{S}$ une partie de $\wp(\Omega)$. Montrer qu'il existe une unique tribu contenant $\mathcal{S}$ et incluse dans toutes les tribus contenant $\mathcal{S}$. Celle-ci est notée $\sigma(\mathcal{S})$.

On suppose l'univers Ω muni d'une probabilité P définie sur une tribu $\mathcal{T}$. On dit que deux tribus $\mathcal{A}$ et $\mathcal{B}$ incluses dans $\mathcal{T}$ sont *indépendantes* lorsque

$$\mathrm{P}(A\cap B) = \mathrm{P}(A)\,\mathrm{P}(B) \quad \text{pour tout } (A,B)\in \mathcal{A}\times\mathcal{B}.$$

(c) Soit $\mathcal{S}_1$ et $\mathcal{S}_2$ deux ensembles constitués d'événements de $(\Omega,\mathcal{T})$. On suppose que les événements de $\mathcal{S}_1$ sont indépendants de ceux de $\mathcal{S}_2$. Montrer que les tribus $\sigma(\mathcal{S}_1)$ et $\sigma(\mathcal{S}_2)$ sont indépendantes.

Solution

(a) L'univers Ω appartient [1] à chaque tribu $\mathcal{A}_i$ donc aussi à $\mathcal{A}$.

Si A est un événement élément de $\mathcal{A}$, celui-ci appartient à chaque tribu $\mathcal{A}_i$ et son événement contraire $\overline{A}$ aussi. Ainsi, $\overline{A}$ est élément de $\mathcal{A}$.

Enfin, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, c'est une suite d'éléments de chaque tribu $\mathcal{A}_i$ et la réunion des A_n est élément de chaque $\mathcal{A}_i$ donc aussi de $\mathcal{A}$.

Finalement, $\mathcal{A}$ est une tribu sur Ω.

(b) **méthode**

‖ On considère l'intersection de toutes les tribus contenant $\mathcal{S}$.

Existence : Introduisons l'intersection $\sigma(\mathcal{S})$ de toutes les tribus contenant $\mathcal{S}$. Celle-ci est une tribu en vertu de l'étude ci-dessus, elle contient $\mathcal{S}$ et est incluse dans toutes les tribus contenant $\mathcal{S}$. Ceci établit l'existence.

Unicité : Soit $\mathcal{A}$ une tribu contenant $\mathcal{S}$ et incluse dans toutes les tribus contenant $\mathcal{S}$. La tribu $\mathcal{A}$ figure parmi les tribus intersectées pour définir $\sigma(\mathcal{S})$ et donc $\sigma(\mathcal{S})\subset\mathcal{A}$. Aussi, $\sigma(\mathcal{S})$ est une tribu contenant $\mathcal{S}$ et donc $\mathcal{A}\subset\sigma(\mathcal{S})$ puis $\mathcal{A}=\sigma(\mathcal{S})$.

(c) Commençons par noter que les tribus $\sigma(\mathcal{S}_1)$ et $\sigma(\mathcal{S}_2)$ sont incluses dans $\mathcal{T}$ ce qui autorise à parler de leur indépendance.

Si A est un événement de $(\Omega,\mathcal{T},\mathrm{P})$, on a vu dans le sujet précédent que l'ensemble des événements de $(\Omega,\mathcal{T},\mathrm{P})$ indépendants de A forme une tribu, notons celle-ci $\mathcal{I}_A$. Par intersection d'une famille de tribus

$$\mathcal{I}_{\mathcal{S}_1} \overset{\text{déf}}{=} \bigcap_{A\in\mathcal{S}_1} \mathcal{I}_A \quad \text{est aussi une tribu.}$$

1. Rappelons qu'une tribu est un ensemble d'ensembles : on dit donc que l'ensemble Ω *appartient* à une tribu et non qu'il est inclus dans une tribu.

Cette tribu contient par hypothèse $\mathcal{S}_2$ et donc aussi $\sigma(\mathcal{S}_2)$. Ainsi, les éléments de $\mathcal{S}_1$ sont indépendants de tous les éléments $\sigma(\mathcal{S}_2)$. On peut alors mener un raisonnement symétrique et affirmer

$$\mathcal{I}_{\sigma(\mathcal{S}_2)} \overset{\text{déf}}{=} \bigcap_{B\in\sigma(\mathcal{S}_2)} \mathcal{I}_B$$

est une tribu contenant $\mathcal{S}_1$ donc aussi $\sigma(\mathcal{S}_1)$. Ainsi, les tribus $\sigma(\mathcal{S}_1)$ et $\sigma(\mathcal{S}_2)$ sont indépendantes [1].

8.6.3 Calcul de probabilités

Dans chaque étude qui suit, on admet l'existence d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ permettant de modéliser l'expérience.

Exercice 10 *

Deux archers tirent à tour de rôle sur une cible. Le premier qui touche a gagné. Le joueur qui commence a la probabilité p_1 de toucher à chaque tir et le second la probabilité p_2 (avec $p_1, p_2 \in]0\,;1]$).

(a) Quelle est la probabilité que le premier archer gagne ?

(b) Montrer qu'il est quasi certain que le tournoi se termine.

(c) Pour quelles valeurs de p_1 existe-t-il une valeur de p_2 pour laquelle le tournoi est équitable ?

Solution

(a) Pour $n \in \mathbb{N}^*$, on introduit les événements :

A_n = « La cible est touchée lors de la n-ième tentative »,

V_n = « La cible est touchée pour la première fois lors de la n-ième tentative ».

L'événement J_1 = « Le premier archer gagne » est la réunion des événements deux à deux incompatibles V_{2k+1} pour k parcourant $\mathbb{N}$. Par additivité dénombrable, on a donc

$$\mathrm{P}(J_1) = \sum_{k=0}^{+\infty} \mathrm{P}(V_{2k+1}).$$

Or $V_{2k+1} = \overline{A_1} \cap \ldots \cap \overline{A_{2k}} \cap A_{2k+1}$ et l'expérience laisse supposer que les différentes tentatives des archers sont indépendantes, donc

$$\mathrm{P}(V_{2k+1}) = \mathrm{P}(\overline{A_1}) \times \cdots \times \mathrm{P}(\overline{A_{2k}}) \times \mathrm{P}(A_{2k+1}) = p_1\big((1-p_1)(1-p_2)\big)^k.$$

1. Ce résultat est souvent utilisé sans être explicitement signifié. Si l'on considère une suite d'expériences aléatoires indépendantes, les événements qui s'expriment en fonction des résultats des premières expériences sont indépendants de ceux qui s'expriment en fonction des résultats des suivantes.

Par calcul d'une somme géométrique de raison strictement inférieure à 1, on conclut

$$\mathrm{P}(J_1) = \sum_{k=0}^{+\infty} p_1\big((1-p_1)(1-p_2)\big)^k = \frac{p_1}{p_1+p_2-p_1p_2}.$$

(b) **méthode**

‖ On calcule la somme[1] des probabilités des victoires de chacun des archers.

L'événement J_2 traduisant la victoire du deuxième archer est la réunion des événements V_{2k} pour $k \in \mathbb{N}^*$. Par des calculs analogues aux précédents

$$\mathrm{P}(V_{2k}) = \mathrm{P}\big(\overline{A_1} \cap \ldots \cap \overline{A_{2k-1}} \cap A_{2k}\big) = p_2(1-p_1)\big((1-p_1)(1-p_2)\big)^{k-1}.$$

On obtient alors

$$\mathrm{P}(J_2) = \sum_{k=1}^{+\infty} p_2(1-p_1)\big((1-p_1)(1-p_2)\big)^{k-1} = \frac{p_2-p_1p_2}{p_1+p_2-p_1p_2}.$$

On vérifie alors $\mathrm{P}(J_1)+\mathrm{P}(J_2)=1$ ce qui signifie qu'il est presque sûr que le tournoi se termine.

(c) Le tournoi est équitable lorsque $\mathrm{P}(J_1) = 1/2$, c'est-à-dire $2p_1 = p_1+p_2-p_1p_2$. Cette équation en l'inconnue p_2 possède une solution $p_2 = p_1/(1-p_1)$ lorsque $p_1 < 1$. Cette solution est élément de $]0\,;1]$ si, et seulement si, $p_1 \leqslant 1/2$.

Exercice 11 *

Une urne contient une boule blanche et une boule rouge. On tire dans cette urne une boule, on note sa couleur et on la remet dans l'urne accompagnée de deux autres boules de la même couleur. On répète cette opération indéfiniment.

(a) Quelle est la probabilité que les n premières boules tirées soient rouges ?

(b) Justifier qu'il est presque sûr que la boule blanche initiale sera tirée.

Solution

(a) On pose $A_0 = \Omega$ et, pour $n \in \mathbb{N}^*$, on introduit l'événement :

$$A_n = \text{« Les } n \text{ premières boules tirées sont rouges »}.$$

On a $\mathrm{P}(A_0) = 1$ et, pour tout $n \geqslant 1$,

$$\mathrm{P}(A_n \,|\, A_{n-1}) = \frac{2n-1}{2n}.$$

1. On peut aussi étudier par continuité décroissante la limite de la probabilité de l'événement : « le tournoi dure au moins k parties ».

En effet, si l'événement A_{n-1} est réalisé, l'urne est constituée d'une boule blanche et de $2n-1$ boules rouges lors du n-ième tirage. Sachant $A_n = A_0 \cap A_1 \cap \cdots \cap A_n$, la formule des probabilités composées (Th. 7 p. 337) donne

$$\mathrm{P}(A_n) = \mathrm{P}(A_0) \prod_{k=1}^{n} \mathrm{P}(A_k \,|\, A_1 \cap \cdots \cap A_{k-1}) = \prod_{k=1}^{n} \mathrm{P}(A_k \,|\, A_{k-1}) = \prod_{k=1}^{n} \frac{2k-1}{2k}.$$

Enfin, en exprimant, le produit des entiers pairs et des entiers impairs à l'aide de nombres factoriels [1], on obtient

$$\mathrm{P}(A_n) = \frac{(2n)!}{2^{2n}(n!)^2} = \frac{1}{4^n}\binom{2n}{n}^{-1}.$$

(b) **méthode**

On étudie l'événement contraire en déterminant la limite de $\mathrm{P}(A_n)$ par la formule de Stirling.

Les événements A_n constituent une suite décroissante et l'on a par continuité monotone (Th. 6 p. 335)

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n).$$

À l'aide de la formule de Stirling, on obtient

$$\mathrm{P}(A_n) \underset{n\to+\infty}{\sim} \frac{1}{\sqrt{\pi n}} \xrightarrow[n\to+\infty]{} 0 \quad \text{car} \quad n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n.$$

L'événement « Tirer indéfiniment des boules rouges » est donc négligeable. En considérant l'événement contraire, il est presque sûr que la boule blanche initiale sera tirée.

Exercice 12 **

On effectue une suite de lancers indépendants d'une pièce équilibrée et l'on désigne par a_n la probabilité de ne pas avoir obtenu trois côtés 'piles' consécutifs lors des n premiers lancers.

(a) Calculer a_1, a_2 et a_3.

(b) Pour $n \geqslant 4$, exprimer a_n en fonction de a_{n-1}, a_{n-2} et a_{n-3}.

(c) Déterminer la limite de la suite $(a_n)_{n\geqslant 1}$.

Solution

Pour $n \in \mathbb{N}^*$, on introduit les événements

P_n = « La pièce tombe côté 'pile' lors du n-ième lancer », $F_n = \overline{P_n}$ et

A_n = « On n'a pas obtenu trois 'piles' consécutifs lors des n premiers lancers ».

1. Voir sujet 5 du chapitre 2 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

(a) Les événements A_1 et A_2 sont certains et donc $a_1 = a_2 = 1$. L'événement A_3 est l'événement contraire de [1] $P_1P_2P_3$. Par indépendance,

$$a_3 = \mathrm{P}(A_3) = 1 - \mathrm{P}(P_1P_2P_3) = 1 - \left(\frac{1}{2}\right)^3 = \frac{7}{8}.$$

(b) **méthode**

On forme un système complet d'événements en discutant selon les résultats des premiers lancers.

Soit $n \geqslant 4$. Obtenir un côté 'face' lors des premiers lancers « réinitialise le décompte ». Ceci invite à introduire le système complet constitué des événements suivants

$$F_1,\ P_1F_2,\ P_1P_2F_3 \text{ et } P_1P_2P_3.$$

La formule des probabilités totales donne alors

$$\begin{aligned}\mathrm{P}(A_n) = {} & \mathrm{P}(A_n \,|\, F_1)\,\mathrm{P}(F_1) + \mathrm{P}(A_n \,|\, P_1F_2)\,\mathrm{P}(P_1F_2) \\ & + \mathrm{P}(A_n \,|\, P_1P_2F_3)\,\mathrm{P}(P_1P_2F_3) + \mathrm{P}(A_n \,|\, P_1P_2P_3)\,\mathrm{P}(P_1P_2P_3).\end{aligned} \qquad (*)$$

Si lors du premier lancer la pièce tombe côté 'face', ne pas obtenir trois 'piles' consécutifs lors des n premiers lancers revient à ne pas obtenir trois 'piles' consécutifs lors des $n-1$ lancers qui suivent le premier. On a donc

$$\mathrm{P}(A_n \,|\, F_1) = \mathrm{P}(A_{n-1}).$$

On justifie de même que $\mathrm{P}(A_n \,|\, P_1F_2) = \mathrm{P}(A_{n-2})$ et $\mathrm{P}(A_n \,|\, P_1P_2F_3) = \mathrm{P}(A_{n-3})$ tandis que $\mathrm{P}(A_n \,|\, P_1P_2P_3) = 0$. L'égalité $(*)$ donne alors

$$a_n = \frac{1}{2}a_{n-1} + \frac{1}{4}a_{n-2} + \frac{1}{8}a_{n-3}.$$

(c) **méthode**

On vérifie que la suite (a_n) converge avant de passer la relation de récurrence précédente à la limite.

La suite $(a_n)_{n\geqslant 1}$ est décroissante car l'événement A_{n+1} est inclus dans l'événement A_n. La suite $(a_n)_{n\geqslant 1}$ est de surcroît minorée par 0, elle est donc convergente. En notant ℓ sa limite, la relation de récurrence précédente donne

$$\ell = \frac{1}{2}\ell + \frac{1}{4}\ell + \frac{1}{8}\ell.$$

On en déduit [2] $\ell = 0$.

1. Par soucis de légèreté, on élude le symbole d'intersection en écrivant $P_1P_2P_3$ au lieu de $P_1 \cap P_2 \cap P_3$.
2. L'étude de cette expérience est approfondie dans le sujet 20 p. 366.

Exercice 13 **

Deux joueurs J_1 et J_2 s'affrontent lors d'un tournoi constitué de parties indépendantes. Initialement, les joueurs possèdent à eux deux N jetons. À chaque tour, le joueur J_1 a la probabilité $p \in]0\,;1[$ d'emporter la partie et le joueur J_2 la probabilité complémentaire $q = 1 - p$. Le joueur perdant cède alors un jeton au vainqueur. La succession de parties continue jusqu'à ce que l'un des deux joueurs ne possède plus de jetons, l'autre est alors déclaré vainqueur du tournoi.

Soit $n \in [\![0\,;N]\!]$. On note a_n la probabilité que le joueur J_1 gagne le tournoi lorsqu'il possède initialement n jetons.

(a) Déterminer a_0 et a_N et établir, pour tout $n \in [\![1\,;N-1]\!]$, l'égalité

$$a_n = pa_{n+1} + qa_{n-1}.$$

(b) Calculer a_n en introduisant $u_n = a_n - a_{n-1}$ pour $n \in [\![1\,;N]\!]$.

(c) Montrer que le tournoi s'arrête presque sûrement.

Solution

(a) Les conditions de l'expérience menée entraînent $a_0 = 0$ car le joueur J_1 a immédiatement perdu le tournoi s'il ne possède pas de jetons. On a aussi $a_N = 1$ car cette fois-ci c'est le joueur J_2 qui ne possède pas de jetons.

méthode

On établit l'égalité par la formule des probabilités totales en discutant selon le résultat de la première partie.

Pour $n \in [\![0\,;N]\!]$, on introduit les événements

A_n = « J_1 gagne le tournoi lorsque sa fortune initiale vaut n »,

A'_n = « J_1 gagne le tournoi lorsque sa fortune après la première partie vaut n »,

V = « J_1 gagne la première partie ».

Soit $n \in [\![1\,;N-1]\!]$. Le couple $(V, \overline{V})$ est un système complet d'événements et la formule des probabilités totales (Th. 8 p. 338) donne

$$\mathrm{P}(A_n) = \mathrm{P}(A_n \,|\, V)\,\mathrm{P}(V) + \mathrm{P}(A_n \,|\, \overline{V})\,\mathrm{P}(\overline{V}). \qquad (*)$$

Or si le joueur J_1 a gagné le premier tour de jeu, sa fortune après la première partie vaut $n+1$ et donc $A_n \cap V = A'_{n+1} \cap V$. Cependant, les événements A'_{n+1} et V sont indépendants[1] car il est supposé que les joueurs J_1 et J_2 s'affrontent en des parties indépendantes et donc

$$\mathrm{P}(A_n \,|\, V) = \frac{\mathrm{P}(A_n \cap V)}{\mathrm{P}(V)} = \frac{\mathrm{P}(A'_{n+1} \cap V)}{\mathrm{P}(V)} = \frac{\mathrm{P}(A'_{n+1})\,\mathrm{P}(V)}{\mathrm{P}(V)} = \mathrm{P}(A'_{n+1}).$$

1. L'événement A'_{n+1} s'exprime en fonction des résultats des parties du tournoi sauf la première, sans le dire on emploie ici le résultat du sujet 9 p. 350.

Enfin, $\mathrm{P}(A'_{n+1}) = \mathrm{P}(A_{n+1})$ car la probabilité de victoire du joueur J_1 dépend du nombre de jetons qu'il possède et non du tour de jeu où l'on dénombre ceux-ci. Ainsi [1], on obtient $\mathrm{P}(A_n \mid V) = \mathrm{P}(A_{n+1})$. On étudie de même la probabilité de A_n sachant $\overline{V}$ et l'équation $(*)$ devient $a_n = pa_{n+1} + qa_{n-1}$.

(b) Sachant $p + q = 1$, on peut écrire $a_n = pa_n + qa_n$ et l'égalité précédente donne

$$p(a_{n+1} - a_n) = q(a_n - a_{n-1}).$$

c'est-à-dire $pu_{n+1} = qu_n$ pour tout $n \in [\![1\,;N-1]\!]$. La suite $(u_n)_{1\leqslant n\leqslant N}$ est donc géométrique de raison q/p et l'on peut exprimer son terme général

$$u_n = u_1\left(\frac{q}{p}\right)^{n-1} \quad \text{pour tout } n \in [\![1\,;N]\!].$$

Par calcul de somme télescopique, on en déduit

$$a_n = \underbrace{a_0}_{=0} + \sum_{k=1}^{n}(a_k - a_{k-1}) = \sum_{k=1}^{n} u_1\left(\frac{q}{p}\right)^{k-1} = u_1 \sum_{\ell=0}^{n-1}\left(\frac{q}{p}\right)^{\ell}.$$

Pour poursuivre le calcul, on discute selon la valeur de la raison de la somme géométrique.

Cas : $p = q$. On obtient $a_n = nu_1$ et puisque $a_N = 1$, on conclut

$$a_n = \frac{n}{N} \quad \text{pour tout } n \in [\![0\,;N]\!].$$

Cas : $p \neq q$. On obtient par somme géométrique de raison différente de 1

$$a_n = \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \frac{q}{p}} u_1.$$

Sachant $a_N = 1$, on conclut [2]

$$a_n = \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \left(\frac{q}{p}\right)^N} = \frac{p^n - q^n}{p^N - q^N} p^{N-n} \quad \text{pour tout } n \in [\![0\,;N]\!].$$

(c) Un calcul symétrique détermine la probabilité b_n que le joueur J_2 gagne le tournoi lorsqu'il possède initialement n jetons :

$$b_n = \frac{p^n - q^n}{p^N - q^N} q^{N-n} \quad \text{si } p \neq q \quad \text{et} \quad b_n = \frac{n}{N} \quad \text{sinon.}$$

Dans le cas $p = q$ tout comme dans le cas $p \neq q$, on constate que $a_n + b_{N-n} = 1$ ce qui assure que presque sûrement l'un des deux joueurs gagne le tournoi.

1. L'égalité $\mathrm{P}(A_n \mid V) = \mathrm{P}(A_{n+1})$ est de bon sens mais il peut être intéressant de détailler...

2. La suite (a_n) est une suite récurrente linéaire d'ordre 2 : on pouvait directement calculer son terme général via l'introduction d'une équation caractéristique.

Exercice 14 **

Trois joueurs J_1, J_2 et J_3 participent à un tournoi selon les règles qui suivent. À chaque partie, deux joueurs entrent en concurrence et chacun peut gagner l'affrontement avec la même probabilité. Le gagnant d'une partie affronte à la partie suivante le joueur n'ayant pas participé. Le tournoi s'arrête lorsque l'un des joueurs gagne deux parties consécutives. Celui-ci est alors déclaré vainqueur.

(a) Établir que le tournoi s'arrête presque sûrement.

(b) Les joueurs J_1 et J_2 s'affrontent en premier. Quelles sont les probabilités de gagner le tournoi de chaque joueur ?

Solution

Pour $n \in \mathbb{N}^*$, introduisons les événements

$$A_n = \text{« Le tournoi dure au moins } n \text{ affrontements »},$$
$$B_n = \text{« Le tournoi s'arrête lors du } n\text{-ième affrontement »}.$$

L'événement A_n est la réunion des événements incompatibles A_{n+1} et B_n.

(a) La suite des événements A_n est décroissante et l'événement « Le tournoi ne s'arrête pas » se comprend comme $\bigcap_{n\in\mathbb{N}^*} A_n$. Par continuité décroissante

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \lim_{n\to+\infty} \mathrm{P}(A_n).$$

Or, pour $n \geqslant 2$, le tournoi s'arrête à la n-ième confrontation si A_n est réalisé et que le joueur gagnant à la n-ième confrontation est celui ayant gagné la précédente. Les deux joueurs s'affrontant ayant la même probabilité de gagner la partie, on a

$$\mathrm{P}(B_n) = \frac{1}{2}\mathrm{P}(A_n) \quad \text{et} \quad \mathrm{P}(A_{n+1}) = \frac{1}{2}\mathrm{P}(A_n).$$

Sachant $\mathrm{P}(A_2) = 1$ (il faut au moins deux affrontements pour déclarer un vainqueur), on obtient

$$\mathrm{P}(A_n) = \frac{1}{2^{n-2}} \quad \text{pour tout } n \geqslant 2.$$

On en déduit que le jeu s'arrête presque sûrement :

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \lim_{n\to+\infty} \frac{1}{2^{n-2}} = 0.$$

(b) **méthode**

Le vainqueur du tournoi se déduit du rang auquel celui-ci s'arrête.

Pour $i = 1, 2, 3$, notons V_i l'événement « le joueur J_i gagne le tournoi ». Introduisons aussi P_1 l'événement « J_1 remporte la première confrontation » et $P_2 = \overline{P_1}$ traduisant la victoire de J_2 au premier affrontement.

Lorsque l'on sait que P_1 est réalisé, l'alternance des vainqueurs lors des confrontations jusqu'à l'arrêt du tournoi entraîne que :

— le joueur J_1 gagne le tournoi lorsque la partie s'arrête à un rang $n \equiv 2$ [3] ;
— le joueur J_2 gagne le tournoi lorsque la partie s'arrête à un rang $n \equiv 1$ [3] ;
— le joueur J_3 gagne le tournoi lorsque la partie s'arrête à un rang $n \equiv 0$ [3].

Si P_2 est réalisé, on a un résultat analogue où les rangs associés à J_1 et J_2 sont échangés : on observe que la probabilité de victoire du joueur 3 ne dépend pas du vainqueur de la première partie et

$$\mathrm{P}(V_3) = \sum_{k=1}^{+\infty} \mathrm{P}(B_{3k})$$

avec

$$\mathrm{P}(B_n) = \mathrm{P}(A_{n+1}) = \frac{1}{2^{n-1}} \quad \text{pour tout } n \geqslant 2.$$

On a donc

$$\mathrm{P}(V_3) = \sum_{k=1}^{+\infty} \frac{1}{2^{3k-1}} = \sum_{\ell=0}^{+\infty} \frac{1}{2^{3\ell+2}} = \frac{1}{4} \cdot \frac{1}{1-\frac{1}{8}} = \frac{2}{7}.$$

De plus, la symétrie de l'étude donne $\mathrm{P}(V_1) = \mathrm{P}(V_2)$ et, puisque la somme des probabilités des événements V_i est égale à 1, on conclut

$$\mathrm{P}(V_1) = \mathrm{P}(V_2) = \frac{5}{14}.$$

8.6.4 Ensembles dénombrables

Exercice 15 *

(a) Montrer que l'ensemble des parties finies de $\mathbb{N}$ est dénombrable.

(b) Montrer que l'ensemble des parties de $\mathbb{N}$ n'est pas dénombrable. On pourra raisonner par l'absurde et introduire $\{n \in \mathbb{N} \mid n \notin \varphi(n)\}$ lorsque φ désigne une bijection de $\mathbb{N}$ vers $\wp(\mathbb{N})$.

Solution

(a) **méthode**

On exprime l'ensemble des parties finies de $\mathbb{N}$ comme une réunion dénombrable d'ensembles au plus dénombrables (Th. 3 p. 332).

Toute partie finie de $\mathbb{N}$ est majorée et est donc une partie de $[\![0\,;N]\!]$ pour $N \in \mathbb{N}$ assez grand. L'ensemble E des parties finies de $\mathbb{N}$ peut donc s'écrire comme la réunion suivante

$$E = \bigcup_{N\in\mathbb{N}} E_N \quad \text{avec} \quad E_N = \wp([\![0\,;N]\!]).$$

Les ensembles E_N sont finis et la réunion porte sur une indexation dénombrable, on peut donc affirmer que E est dénombrable [1] car c'est un ensemble infini égal à une réunion dénombrable de parties au plus dénombrables.

(b) Par l'absurde, supposons qu'il existe une bijection φ de $\mathbb{N}$ vers $\wp(\mathbb{N})$. Introduisons l'ensemble $A = \{n \in \mathbb{N} \mid n \notin \varphi(n)\}$. Celui-ci est une partie de $\mathbb{N}$ et il existe donc $n_0 \in \mathbb{N}$ tel que $A = \varphi(n_0)$. On s'interroge alors sur l'appartenance de n_0 à A.

Cas : $n_0 \in A$. L'appartenance de n_0 à A signifie $n_0 \notin \varphi(n_0)$ et donc que n_0 n'appartient pas à A puisque $A = \varphi(n_0)$.

Cas : $n_0 \notin A$. Puisque $A = \varphi(n_0)$, on a donc $n_0 \notin \varphi(n_0)$ ce qui entraîne que n_0 appartient à A.

Dans les deux cas, on conclut à une absurdité [2].

Exercice 16 **

Soit $(u_n)_{n\in\mathbb{N}^*}$ une suite d'éléments de $[0\,;1]$. Pour tout $n \in \mathbb{N}^*$, on pose

$$I_n = \left[u_n - \frac{1}{2^{n+1}}\,;u_n + \frac{1}{2^{n+1}}\right].$$

(a) Montrer que, pour tout entier naturel $n \geqslant 1$, le segment $[0\,;1]$ n'est pas inclus dans la réunion $I_1 \cup \ldots \cup I_n$.

On peut alors construire une suite $(x_n)_{n\in\mathbb{N}^*}$ d'éléments de $[0\,;1]$ choisis tels que x_n n'appartienne pas à $I_1 \cup \ldots \cup I_n$.

(b) Établir que l'intervalle $[0\,;1]$ n'est pas dénombrable [3].

Solution

(a) La somme des longueurs des intervalles réunis vaut

$$L = \sum_{k=1}^{n} \frac{1}{2^k} < \sum_{k=1}^{+\infty} \frac{1}{2^k} = \frac{1}{2} \cdot \frac{1}{1-1/2} = 1.$$

La réunion des intervalles $I_1 \cup \ldots \cup I_n$ ne peut donc recouvrir [4] le segment $[0\,;1]$.

(b) Par l'absurde, supposons l'intervalle $[0\,;1]$ dénombrable et introduisons $(u_n)_{n\in\mathbb{N}^*}$ une énumération de ses éléments. On introduit alors la suite $(x_n)_{n\in\mathbb{N}^*}$ proposée par l'énoncé.

1. On peut aussi proposer un dénombrement explicite : si A désigne une partie finie de $\mathbb{N}$, on pose $n(A)$ égal à la somme des 2^k pour k parcourant A. L'existence et l'unicité de l'écriture d'un entier en somme de puissances de 2 assure la bijectivité de cette association.

2. On retrouvera cette démonstration en situation générale dans le sujet 35 du chapitre 1 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI.*

3. *A fortiori* la droite réelle n'est pas non plus dénombrable. On peut montrer que $\mathbb{R}$ est en bijection avec $\wp(\mathbb{N})$ mais cela n'a rien d'évident.

4. On peut traduire cet argument *géométrique* par un calcul intégral : l'intégrale de la somme S des fonctions indicatrices des intervalles $I_1, \ldots, I_n$ est strictement inférieure à 1 et il existe donc dans $[0\,;1]$ un réel t tel que $S(t) < 1$, c'est-à-dire tel que $S(t) = 0$.

méthode

On étudie une valeur d'adhérence de la suite (x_n).

La suite (x_n) est bornée et possède donc une valeur d'adhérence ℓ qui appartient à $[0\,;1]$. La suite (u_n) énumérant tous les éléments de $[0\,;1]$, il existe $n_0 \geqslant 1$ tel que $\ell = u_{n_0}$. Or par la suite extraite de (x_n) convergeant vers ℓ, on peut affirmer l'existence d'une infinité de termes de (x_n) qui appartiennent à l'intervalle I_{n_0} centré en ℓ. C'est absurde car, à partir du rang n_0, aucun terme de la suite (x_n) n'est élément de cet intervalle !

8.7 Exercices d'approfondissement

Exercice 17 **

Soit $f\colon \mathbb{R} \to \mathbb{R}$ croissante. Montrer que l'ensemble des points de discontinuité de f est au plus dénombrable.

Solution

Puisque la fonction f est croissante, elle admet une limite à droite et une limite à gauche en tout point x de $\mathbb{R}$. De plus, en notant $f(x^+)$ et $f(x^-)$ ces limites, on sait

$$f(x^-) \leqslant f(x) \leqslant f(x^+).$$

Par conséquent, la fonction f est continue en x si, et seulement si, $f(x^-) = f(x^+)$. Les points de discontinuité de f constituent donc l'ensemble

$$D = \left\{x \in \mathbb{R} \mid f(x^-) < f(x^+)\right\}.$$

méthode

On montre que l'ensemble des points de discontinuité est la réunion des x pour lesquels $f(x^+) - f(x^-) \geqslant \frac{1}{n}$ avec $n \in \mathbb{N}^*$.

Soit $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}$. Considérons l'ensemble

$$E_{n,k} = \left\{x \in [k\,;k+1[\;\middle|\; f(x^+) - f(x^-) \geqslant \frac{1}{n}\right\}.$$

Cet ensemble est fini car, en chaque point de celui-ci, la fonction f progresse d'au moins $1/n$ mais ne peut varier en tout que de $f(k^-)$ à $f(k+1)$. Puisque D est la réunion des ensembles $E_{n,k}$ pour tous $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}$, on peut affirmer que D est une réunion dénombrable d'ensembles finis, c'est donc un ensemble au plus dénombrable[1].

1. On peut aussi proposer une alternative « moins géométrique » : pour chaque x appartenant à D, on peut déterminer un nombre rationnel r strictement compris entre $f(x^-)$ et $f(x^+)$. Ceci définit une injection de D dans $\mathbb{Q}$ ce qui assure que l'ensemble D est au plus dénombrable.

Exercice 18 **

Soit s un réel strictement supérieur à 1.

(a) Pour quelle valeur de $\lambda \in \mathbb{R}$, existe-t-il une probabilité P sur l'espace $\big(\mathbb{N}^*, \wp(\mathbb{N}^*)\big)$ telle que

$$\mathrm{P}\big(\{n\}\big) = \frac{\lambda}{n^s} \quad \text{pour tout } n \in \mathbb{N}^* \,?$$

Pour $p \in \mathbb{N}^*$, on introduit l'événement $A_p = \big\{n \in \mathbb{N}^* \,\big|\, p \text{ divise } n\big\}$ et l'on note $\mathcal{P}$ l'ensemble des nombres premiers.

(b) Montrer que la famille des événements A_p pour p parcourant $\mathcal{P}$ est constituée d'événements mutuellement indépendants.

(c) En étudiant $\mathrm{P}\big(\{1\}\big)$, établir

$$\sum_{n=1}^{+\infty} \frac{1}{n^s} = \prod_{p\in\mathcal{P}} \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

Solution

(a) *Analyse :* Si P est une probabilité solution, on sait par additivité dénombrable

$$\mathrm{P}\big(\mathbb{N}^*\big) = \mathrm{P}\Big(\bigcup_{n\in\mathbb{N}^*} \{n\}\Big) = \sum_{n=1}^{+\infty} \mathrm{P}\big(\{n\}\big) = \sum_{n=1}^{+\infty} \frac{\lambda}{n^s} = 1.$$

Ceci suffit à déterminer la valeur de λ :

$$\lambda = \left(\sum_{n=1}^{+\infty} \frac{1}{n^s}\right)^{-1}.$$

Synthèse : Pour la valeur de λ obtenue ci-dessus, on peut affirmer que la famille des p_n, avec $p_n = \lambda/n^s$ pour tout $n \in \mathbb{N}^*$, est une famille de réels positifs, sommable et de somme égale 1. Il existe donc une unique probabilité P sur l'univers dénombrable $\mathbb{N}^*$ vérifiant $\mathrm{P}\big(\{n\}\big) = p_n$ pour tout $n \geqslant 1$. Au surplus, on sait que pour toute partie A incluse dans $\mathbb{N}^*$

$$\mathrm{P}(A) = \sum_{n\in A} p_n$$

(b) Soit $m \in \mathbb{N}^*$ et $p_1, \ldots, p_m$ des nombres premiers deux à deux distincts.

méthode

On vérifie $\mathrm{P}(A_{p_1} \cap \ldots \cap A_{p_m}) = \mathrm{P}(A_{p_1}) \times \cdots \times \mathrm{P}(A_{p_m})$ en commençant par décrire simplement l'événement $A_{p_1} \cap \ldots \cap A_{p_m}$.

Par définition d'une intersection

$$A_{p_1} \cap \ldots \cap A_{p_m} = \big\{n \in \mathbb{N}^* \,\big|\, \forall k \in [\![1\,;m]\!],\ p_k \mid n\big\}.$$

Les p_k étant des nombres premiers deux à deux distincts, on a la propriété arithmétique

$$\left(\forall k \in [\![1\,;m]\!],\ p_k \mid n\right) \iff p_1 \dots p_m \mid n$$

et donc

$$A_{p_1} \cap \dots \cap A_{p_m} = A_{p_1 \dots p_m}.$$

Il reste à calculer les probabilités des événements A_p. Pour tout $p \in \mathbb{N}^*$, les éléments de A_p sont les kp avec k parcourant $\mathbb{N}^*$ et donc

$$\mathrm{P}(A) = \sum_{n \in A_p} p_n = \sum_{k=1}^{+\infty} \frac{\lambda}{(pk)^s} = \frac{\lambda}{p^s} \sum_{k=1}^{+\infty} \frac{1}{k^s} = \frac{1}{p^s}.$$

L'égalité $(p_1 \dots p_m)^s = p_1^s \dots p_m^s$ donne alors immédiatement

$$\mathrm{P}(A_{p_1} \cap \dots \cap A_{p_m}) = \mathrm{P}(A_{p_1 \dots p_m}) = \frac{1}{(p_1 \dots p_m)^s} = \mathrm{P}(A_{p_1}) \times \dots \times \mathrm{P}(A_{p_m}).$$

On peut conclure que les événements A_p pour p parcourant $\mathcal{P}$ sont mutuellement indépendants.

(c) On a

$$\{1\} = \bigcap_{p \in \mathcal{P}} \overline{A_p}$$

car tout entier naturel supérieur à 2 est divisible par un nombre premier. Énumérons les nombres premiers : $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc. On peut écrire par continuité décroissante et indépendance [1]

$$\mathrm{P}(\{1\}) = \lim_{N \to +\infty} \mathrm{P}\left(\bigcap_{k=1}^{N} \overline{A_{p_k}}\right) = \lim_{N \to +\infty} \prod_{k=1}^{N} \mathrm{P}(\overline{A_{p_k}}) = \lim_{N \to +\infty} \prod_{k=1}^{N} \left(1 - \frac{1}{p_k^s}\right).$$

Or $\mathrm{P}(\{1\}) = \lambda$ et donc

$$\lambda = \lim_{N \to +\infty} \prod_{k=1}^{N} \left(1 - \frac{1}{p_k^s}\right).$$

Après passage à l'inverse, ceci fournit la relation demandée sous réserve de comprendre [2] :

$$\prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p^s}\right) \overset{\text{déf}}{=} \lim_{N \to +\infty} \prod_{k=1}^{N} \left(1 - \frac{1}{p_k^s}\right).$$

1. L'indépendance des événements $A_{p_1}, \dots, A_{p_m}$ entraîne celle des événements $\overline{A_{p_1}}, \dots, \overline{A_{p_m}}$ voir sujet 13 du chapitre 12 de l'ouvrage *Exercices d'algèbre et de probabilités MPSI*.

2. Par analogie avec les familles sommables, on peut établir par passage au logarithme que le résultat de ce produit de facteurs strictement positifs ne dépend pas de l'énumération choisie.

Exercice 19 * (Retour à l'origine)**

Un individu se déplace sur l'axe des $\mathbb{Z}$. Il est initialement en 0 et à chaque instant $n \in \mathbb{N}$, il a la probabilité $d \in]0;1[$ de se déplacer sur la droite et $g = 1 - d$ de se déplacer sur la gauche. Pour $n \in \mathbb{N}$, on pose p_n la probabilité que l'individu soit en 0 à l'instant n. Pour $n > 0$, on pose q_n la probabilité qu'il soit une première fois de retour en 0 à l'instant n. Enfin, on introduit les fonctions

$$P : t \mapsto \sum_{n=0}^{+\infty} p_n t^n \quad \text{et} \quad Q : t \mapsto \sum_{n=1}^{+\infty} q_n t^n.$$

(a) Montrer que la fonction P est définie sur $[0;1[$ et calculer $P(t)$ pour $t \in [0;1[$.
On donne :

$$\frac{1}{\sqrt{1-x}} = \sum_{n=0}^{+\infty} \frac{1}{2^{2n}} \binom{2n}{n} x^n \quad \text{pour tout } x \in]-1;1[.$$

(b) Montrer que la fonction Q est définie et continue sur $[0;1]$.

(c) Vérifier que $P(t) = 1 + P(t)Q(t)$ pour tout $t \in [0;1[$.

(d) Calculer la probabilité de l'événement « L'individu repasse par 0 ».

(e) Calculer la probabilité de « L'individu repasse au moins deux fois par 0 ».

(f) On suppose $d = g$. Établir que l'individu repasse presque sûrement une infinité de fois par 0.

Solution

Pour $n \in \mathbb{N}^*$, introduisons les événements :

A_n = « L'individu revient en 0 à l'instant n »,
B_n = « L'individu revient une première fois en 0 à l'instant n ».

Par définition, $p_n = \mathrm{P}(A_n)$, $q_n = \mathrm{P}(B_n)$ et $p_0 = 1$.

(a) Soit $t \in [0;1[$. Étudions la convergence de la série de terme général $p_n t^n$. Pour tout $n \in \mathbb{N}$, on a $p_n \in [0;1]$ puisqu'il s'agit d'une probabilité et donc $0 \leqslant p_n t^n \leqslant t^n$. On sait que la série géométrique $\sum t^n$ converge et l'on en déduit la convergence de la série de terme général $p_n t^n$ par comparaison de séries à termes positifs. Ceci assure la définition de la fonction P sur [1] $[0;1[$.

méthode

On calcule p_n en dénombrant les successions de déplacements qui ramènent en 0.

Soit $n \in \mathbb{N}$. Un retour en 0 à l'instant n correspond à une succession de déplacements comportant autant de déplacements sur la droite que de déplacements sur la gauche.

1. Plus rapidement, on peut aussi dire que la fonction P est la somme d'une série entière et celle-ci est de rayon de convergence au moins égal à 1 car la suite de ses coefficients est bornée.

Cas : n est impair. Il est impossible qu'un retour en 0 ait lieu à l'instant n : $p_{2k+1} = 0$ pour tout $k \in \mathbb{N}$.

Cas : n est pair. On écrit $n = 2k$. Un retour en 0 à l'instant n est constitué de k déplacements sur la gauche et d'autant de déplacements sur la droite. Il y a $\binom{2k}{k}$ combinaisons possibles distinctes de ces déplacements et chacun a la même probabilité $d^k g^k$. On en déduit

$$p_n = p_{2k} = \binom{2k}{k} d^k g^k \quad \text{pour tout } k \in \mathbb{N}.$$

On peut alors calculer $P(t)$ pour tout $t \in [0\,;1[$ par la formule proposée dans l'énoncé après simplification des termes d'indices impairs

$$P(t) = \sum_{n=0}^{+\infty} p_n t^n = \sum_{k=0}^{+\infty} \binom{2k}{k} d^k g^k t^{2k} = \frac{1}{\sqrt{1-4dgt^2}}$$

car $4dgt^2 = 4d(1-d)t^2 \in [0\,;1[$ puisque l'inégalité $d(1-d) \leqslant 1/4$ est bien connue.

(b) Pour tout $n \geqslant 1$, introduisons les fonctions $u_n : x \mapsto q_n x^n$ définies et continues sur le segment $[0\,;1]$. La fonction Q apparaît comme la somme de la série de fonctions $\sum u_n$. On a

$$\sup_{x\in[0;1]} |u_n(x)| = \sup_{x\in[0;1]} q_n x^n = q_n.$$

Or la famille $(B_n)_{n\geqslant 1}$ est une suite d'événements deux à deux incompatibles et la série de terme général $q_n = \mathrm{P}(B_n)$ est donc convergente de somme au plus égale à 1. On en déduit que la série de fonctions $\sum u_n$ converge normalement sur $[0\,;1]$. La fonction Q est donc définie et continue sur $[0\,;1]$.

(c) **méthode**

‖ On exprime A_n en discutant selon le rang du premier retour 0.

Les événements $B_1, \ldots, B_n$ sont deux à deux incompatibles et A_n est inclus dans l'union de ceux-ci. On peut donc décomposer A_n en la réunion qui suit

$$A_n = \bigcup_{k=1}^{n} (A_n \cap B_k) \quad \text{avec} \quad A_n \cap B_k \text{ deux à deux incompatibles.}$$

Par additivité

$$p_n = \mathrm{P}(A_n) = \sum_{k=1}^{n} \mathrm{P}(A_n \cap B_k).$$

Cependant, par la formule des probabilités composées,

$$\mathrm{P}(A_n \cap B_k) = \mathrm{P}(A_n \,|\, B_k)\,\mathrm{P}(B_k).$$

La probabilité d'un retour en 0 à l'instant n sachant que l'on est en 0 à l'instant k s'identifie à la probabilité d'un retour en 0 à l'instant $n-k$ car on peut faire se correspondre

les successions de déplacements associées aux chemins considérés. Ainsi, on obtient [1]

$$p_n = \sum_{k=1}^{n} p_{n-k} q_k$$

Enfin, pour $t \in [0\,;1[$, on obtient par produit de Cauchy de deux séries à termes positifs convergentes

$$P(t)Q(t) = \left(\sum_{j=0}^{+\infty} p_j t^j\right)\left(\sum_{k=1}^{+\infty} q_k t^k\right) = \sum_{n=1}^{+\infty}\left(\sum_{\substack{j+k=n \\ k\geqslant 1}} p_j t^j q_k t^k\right) = \sum_{n=1}^{+\infty}\left(\sum_{k=1}^{n} p_{n-k} q_k\right) t^n$$

$$= \sum_{n=1}^{+\infty} p_n t^n = \sum_{n=0}^{+\infty} p_n t^n - p_0 = P(t) - 1.$$

(d) L'événement R = « L'individu repasse par 0 » est la réunion dénombrable des événements deux à deux incompatibles B_n pour $n \in \mathbb{N}^*$. On a donc

$$\mathrm{P}(R) = \sum_{n=1}^{+\infty} q_n = Q(1) = \lim_{t\to 1^-} Q(t) = \lim_{t\to 1^-} \frac{P(t)-1}{P(t)}.$$

Pour $t \in [0\,;1[$, on sait exprimer $P(t)$ et l'on obtient

$$\frac{P(t)-1}{P(t)} = 1 - \sqrt{1-4dgt^2} \xrightarrow[t\to 1^-]{} 1 - \sqrt{1-4dg} = \mathrm{P}(R).$$

(e) Introduisons l'événement S = « L'individu repasse au moins deux fois par 0 » et, pour tout $n \in \mathbb{N}^*$, C_n = « L'individu repasse une deuxième fois par 0 en l'instant n ». On écrit

$$C_n = \bigcup_{k=1}^{n} (C_n \cap B_k) = \bigcup_{k=1}^{n-1} (C_n \cap B_k) \quad \text{car} \quad C_n \cap B_n = \emptyset.$$

En observant que $\mathrm{P}(C_n \mid B_k) = \mathrm{P}(B_{n-k})$, un étude analogue à celle de la question (c) donne

$$\mathrm{P}(C_n) = \sum_{k=1}^{n-1} \mathrm{P}(C_n \mid B_k)\,\mathrm{P}(B_k) = \sum_{k=1}^{n-1} q_{n-k} q_k.$$

On en déduit

$$\sum_{n=0}^{+\infty} \mathrm{P}(C_n) t^n = Q(t)^2 \quad \text{pour tout } t \in [0\,;1].$$

1. On a ici reproduit la preuve de la formule des probabilités totales : la famille des B_k ne constitue pas un système complet d'événements mais A_n est inclus dans sa réunion.

Enfin, l'événement S est la réunion dénombrable des événements C_n deux à deux incompatibles et donc

$$\mathrm{P}(S) = \sum_{n=0}^{+\infty} \mathrm{P}(C_n) = Q(1)^2 = \left(1 - \sqrt{1-4dg}\right)^2.$$

(f) On peut généraliser l'étude ci-dessus et affirmer que, pour tout $k \in \mathbb{N}^*$, la probabilité de l'événement T_k signifiant que l'individu repasse au moins k fois par 0 vaut

$$\mathrm{P}(T_k) = \left(1 - \sqrt{1-4dg}\right)^k.$$

En particulier, lorsque $d = g$, c'est-à-dire quand $d = 1/2$, cette probabilité est égale à 1. Par continuité décroissante

$$\mathrm{P}\left(\bigcap_{k\in\mathbb{N}^*} T_k\right) = \lim_{k\to+\infty} \mathrm{P}(T_k) = 1.$$

Il est donc presque sûr que l'individu repasse une infinité [1] de fois par 0 lorsque $d = g$.

Exercice 20 * (Succès consécutifs)**

On effectue une succession de lancers indépendants d'une pièce ayant la probabilité p de tomber côté 'pile' et $1-p$ de tomber côté 'face'. Pour $n \in \mathbb{N}^*$, on introduit les événements $P_n =$ « La pièce tombe côté 'pile' au n-ième tirage » et $F_n = \overline{P_n}$.

Soit $r \in \mathbb{N}^*$. On s'intéresse à l'obtention d'une série de r côtés 'piles' consécutifs. Pour $n \in \mathbb{N}^*$, on introduit l'événement $A_n =$ « Au n-ième tirage, on obtient pour la première fois une série de r côtés 'piles' consécutifs » dont on note a_n la probabilité. On convient que a_0 est nul.

(a) Calculer $a_1, \ldots, a_{r-1}$ et a_r.

(b) Soit $n \in \mathbb{N}^*$. Exprimer l'événement A_{n+r} à l'aide des événements $A_1, \ldots, A_{n-1}$ et d'événements F_k et P_k d'indices bien choisis. En déduire

$$a_{n+r} = (1-p)p^r\left(1 - \sum_{k=0}^{n-1} a_k\right).$$

On introduit la série entière $\sum a_n x^n$ dont on note G la somme.

(c) Montrer que la fonction G est bien définie sur $]-1\,;1[$ et vérifier

$$\frac{G(x)}{1-x} = \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n} a_k\right)x^n \quad \text{pour tout } x \in]-1\,;1[.$$

(d) Exprimer $G(x)$.

1. Lorsque $d \neq g$, le même calcul assure qu'il est presque sûr que l'individu ne passe qu'un nombre fini de fois par 0.

Solution

(a) Les premières valeurs $a_1, \ldots, a_{r-1}$ sont nulles car il n'y a pas encore de successions de r lancers. Puisque $A_r = P_1 \cap \ldots \cap P_r$, on a par indépendance $a_r = p^r$.

(b) **méthode**

> On exprime par les opérations ensemblistes usuelles l'événement A_{n+r} comme la réalisation d'une série de r cotés 'piles' consécutifs au rang $n+r$ non précédée de la réalisation d'une série à un rang inférieur.

Pour que l'événement A_{n+r} soit réalisé, il faut que l'on ait des lancers côtés 'piles' aux rangs $n+1, \ldots, n+r$ mais pas au rang n. On a ainsi une première inclusion

$$A_{n+r} \subset F_n \cap P_{n+1} \cap \ldots \cap P_{n+r}.$$

Pour que l'événement A_{n+r} soit réalisé, il faut aussi que les événements $A_1, \ldots, A_{n+r-1}$ ne le soient pas ce qui produit une seconde inclusion

$$A_{n+r} \subset \overline{A_1} \cap \ldots \cap \overline{A_{n+r-1}}.$$

On en déduit

$$A_{n+r} \subset \left(F_n \cap P_{n+1} \cap \ldots \cap P_{n+r}\right) \cap \left(\overline{A_1} \cap \ldots \cap \overline{A_{n+r-1}}\right).$$

L'inclusion réciproque est immédiate et l'on a donc l'égalité.

Or, pour tout $k \in [\![0\,;r-1]\!]$, on remarque $F_n \subset \overline{A_{n+k}}$ car l'événement A_{n+k} ne peut être réalisé lorsque l'on a obtenu un lancer côté 'face' au rang n. On peut donc simplifier le calcul qui précède pour écrire seulement

$$A_{n+r} = \left(F_n \cap P_{n+1} \cap \ldots \cap P_{n+r}\right) \cap \left(\overline{A_1} \cap \ldots \cap \overline{A_{n-1}}\right).$$

Les différents lancers étant supposés indépendants, on a l'indépendance des événements F_n, $P_{n+1}, \ldots, P_{n+r}$ et $B = \overline{A_1} \cap \ldots \cap \overline{A_{n-1}}$ notamment car ce dernier ne se définit qu'à partir des résultats des $n-1$ premiers lancers. On a donc

$$a_{n+r} = \mathrm{P}(F_n)\,\mathrm{P}(P_{n+1}) \ldots \mathrm{P}(P_{n+r})\,\mathrm{P}(B). \qquad (*)$$

Or, par passage à l'événement contraire et parce que les $A_1, \ldots, A_{n-1}$ sont deux à deux incompatibles, on obtient

$$\mathrm{P}(B) = 1 - \mathrm{P}\left(A_1 \cup \ldots \cup A_{n-1}\right) = 1 - \sum_{k=1}^{n-1} a_k.$$

Sachant de plus $a_0 = 0$, la relation $(*)$ devient

$$a_{n+r} = (1-p)p^r \left(1 - \sum_{k=0}^{n-1} a_k\right).$$

(c) Les a_k constituent une suite de réels compris entre 0 et 1. La série entière définissant $G(x)$ est donc de rayon de convergence au moins égale à 1 et converge par conséquent en tout [1] point de $]-1\,;1[$. Par produit de deux séries entières [2] on vérifie l'identité proposée.

(d) Pour tout $n \in \mathbb{N}$, on a obtenu

$$a_{n+r+1} = (1-p)p^r\left(1 - \sum_{k=0}^{n} a_k\right).$$

méthode

‖ On multiple cette identité par x^{n+r+1} et l'on somme pour $n \in \mathbb{N}$.

Soit $x \in]-1\,;1[$. On peut écrire avec convergence des séries introduites

$$\sum_{n=0}^{+\infty} a_{n+r+1}x^{n+r+1} = (1-p)p^r x^{r+1}\sum_{n=0}^{+\infty}\left(1 - \sum_{k=0}^{n} a_k\right)x^n. \qquad (**)$$

D'une part, le premier membre s'écrit

$$\sum_{n=0}^{+\infty} a_{n+r+1}x^{n+r+1} = \sum_{n=0}^{+\infty} a_n x^n - \sum_{n=0}^{r} a_n x^n = G(x) - a_r x^r = G(x) - p^r x^r.$$

D'autre part, le second membre s'écrit

$$(1-p)p^r x^{r+1}\sum_{n=0}^{+\infty}\left(1 - \sum_{k=0}^{n} a_k\right)x^n = (1-p)p^r\frac{1-G(x)}{1-x}x^{r+1}.$$

L'égalité $(**)$ donne alors après résolution [3]

$$G(x) = \frac{p^r x^r(1-px)}{1-x+(1-p)p^r x^{r+1}}.$$

1. Les événements A_n étant deux à deux incompatibles, la série de terme général a_n converge et la série définissant $G(x)$ converge aussi pour $x = 1$ et $x = -1$.

2. Voir sujet 14 du chapitre 9 de l'ouvrage *Exercices d'analyse MP*.

3. L'étude de la limite de $G(x)$ quand x tend vers 1^- permet d'établir qu'il est presque sûr qu'au moins une série de r côtés 'piles' sera réalisée. La fonction G est en fait la fonction génératrice (notion qui est présentée au chapitre suivant) de la variable aléatoire qui détermine le premier rang en lequel apparaît une succession de r côtés 'piles' consécutifs. L'obtention de cette fonction génératrice permet un calcul rapide de l'espérance et de la variance de la variable aléatoire associée.

CHAPITRE 9

Variables aléatoires

$(\Omega, \mathcal{T}, \mathrm{P})$ désigne un espace probabilisé.

9.1 Variables aléatoires discrètes

9.1.1 Définition

Définition

On appelle *variable aléatoire discrète* définie sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ et à valeurs dans un ensemble[1] E toute application $X : \Omega \to E$ vérifiant :

1) l'ensemble $X(\Omega)$ des valeurs prises par X est fini ou dénombrable ;
2) pour tout[2] $x \in E$, l'ensemble

$$(X = x) \stackrel{\text{déf}}{=} X^{-1}(\{x\})$$

des antécédents de la valeur x est élément de la tribu $\mathcal{T}$.

L'ensemble E peut se limiter à l'ensemble $X(\Omega)$ des valeurs prises par X ou seulement le contenir. Sans perte de généralité, on peut supposer si besoin que E est fini ou dénombrable.

Une fonction constante définie sur Ω est une variable aléatoire discrète.
Si $A \in \mathcal{T}$, la fonction indicatrice $\mathbf{1}_A$ est une variable aléatoire discrète.

1. Lorsque l'ensemble d'arrivée E est inclus dans $\mathbb{R}$, on dit que X est une *variable aléatoire réelle*.
2. On peut limiter l'étude aux x éléments de $X(\Omega)$ car pour tout $x \notin X(\Omega)$, l'ensemble $X^{-1}(\{x\})$ est vide donc élément de $\mathcal{T}$.

Théorème 1
Soit $X : \Omega \to E$ une variable aléatoire discrète. Pour toute partie A de E, l'ensemble

$$X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$$

désigne un événement de $(\Omega, \mathcal{T}, \mathrm{P})$.

Définition
Pour toute partie A de E, l'événement $X^{-1}(A)$ est noté $(X \in A)$ ou $\{X \in A\}$. Ainsi,

$$(X \in A) = \{\omega \in \Omega \mid X(\omega) \in A\}.$$

En particulier, si x est une valeur de E et si A désigne $\{x\}$, l'événement $(X \in A)$ correspond [1] à $(X = x)$.

9.1.2 Opérations sur les variables aléatoires

Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ à valeurs dans un ensemble E et f une application définie sur E et à valeurs dans un ensemble F.

Définition
On appelle *variable aléatoire composée* de X par f, l'application $Y = f(X) : \Omega \to F$ déterminée par

$$Y(\omega) = f\big(X(\omega)\big) \quad \text{pour tout } \omega \in \Omega.$$

Théorème 2
L'application $Y = f(X)$ définit une variable aléatoire discrète sur $(\Omega, \mathcal{T}, \mathrm{P})$.

Si la fonction f présente une notation usuelle, celle-ci est reprise pour la description de la variable aléatoire $Y = f(X)$. C'est ainsi que l'on peut écrire $Y = X^2, \sqrt{X}, |X|, \ldots$

Plus généralement, si $X_1, \ldots, X_m$ sont des variables aléatoires discrètes sur $(\Omega, \mathcal{T}, \mathrm{P})$, on peut donner un sens à la variable aléatoire composée $Y = f(X_1, \ldots, X_m)$ sous réserve que f soit définie sur l'ensemble des valeurs prises par $\omega \mapsto \big(X_1(\omega), \ldots, X_m(\omega)\big)$.

En considérant la fonction somme $f : (x, y) \mapsto x + y$ définie sur $\mathbb{R}^2$, on peut parler de la somme $X + Y$ de deux variables alcatoires réelles. On peut aussi parler du produit de deux variables réelles et l'on a le résultat :

Théorème 3
L'ensemble des variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ est une sous-algèbre de l'algèbre $\mathcal{F}(\Omega, \mathbb{R})$.

1. Si X est une variable aléatoire réelle et $x \in \mathbb{R}$, on peut aussi introduire des événements $(X \leqslant x)$, etc.

9.1.3 Loi d'une variable aléatoire discrète

Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ et E un ensemble[1] contenant l'ensemble $X(\Omega)$ des valeurs prises par X.
Pour toute partie A de E, on peut calculer la probabilité $\mathrm{P}(X \in A)$ car $(X \in A)$ est un événement.

Définition

On appelle *loi sur* E de la variable aléatoire X l'application

$$\mathrm{P}_X : \wp(E) \to [0\,;1]$$

définie par $\mathrm{P}_X(A) \overset{\text{déf}}{=} \mathrm{P}(X \in A)$ pour toute partie A de E.

Théorème 4

La loi P_X définit une probabilité sur l'espace $\big(E, \wp(E)\big)$.

Définition

On appelle *probabilités élémentaires* de la variable X les nombres

$$p_x = \mathrm{P}_X(\{x\}) = \mathrm{P}(X = x) \quad \text{avec} \quad x \in E.$$

Théorème 5

Lorsque E est fini ou dénombrable, la loi P_X est entièrement déterminée par les probabilités élémentaires de X :

$$\mathrm{P}_X(A) = \mathrm{P}(X \in A) = \sum_{x \in A} p_x \quad \text{pour toute partie } A \text{ de } E.$$

Ces probabilités élémentaires constituent alors une famille de réels positifs $(p_x)_{x \in E}$ sommable et de somme égale à 1. On peut figurer la loi d'une variable aléatoire à l'aide d'un diagramme en bâton dont les hauteurs correspondent aux probabilités p_x.

La loi de X suffit à déterminer la loi de toute variable composée $Y = f(X)$.

9.1.4 Lois usuelles

Définition

Soit X et Y deux variables aléatoires discrètes sur Ω à valeurs dans un ensemble E. On dit que les variables X et Y *suivent la même loi* lorsque $\mathrm{P}_X = \mathrm{P}_Y$. On note alors $X \sim Y$. S'il est usuel de désigner par $\mathcal{L}$ une loi donnée, on écrit $X \hookrightarrow \mathcal{L}$ pour signifier que la variable X suit la loi $\mathcal{L}$.

1. E peut être exactement l'ensemble $X(\Omega)$ mais ce peut aussi être un ensemble plus grand. Si besoin, on peut supposer E au plus dénombrable.

Définition

Soit $p \in [0\,;1]$. On note $\mathcal{B}(p)$ la *loi de Bernoulli* de paramètre p. Une variable X suit cette loi si $X(\Omega) \subset \{0,1\}$ et

$$\mathrm{P}(X=0) = 1-p \quad \text{et} \quad \mathrm{P}(X=1) = p.$$

Les lois des Bernoulli servent à modéliser les épreuves à deux issues : succès ou échec.

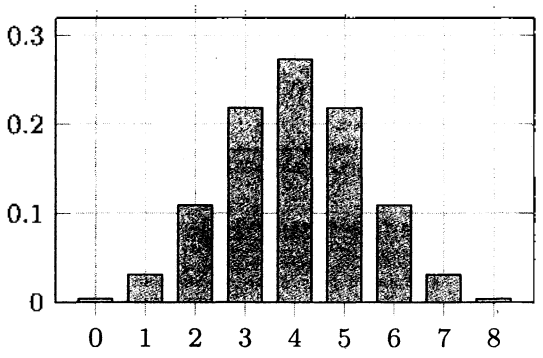
Définition

Soit $n \in \mathbb{N}$ et $p \in [0\,;1]$. On note $\mathcal{B}(n,p)$ la *loi binomiale* de paramètres n et p. Une variable X suit la loi $\mathcal{B}(n,p)$ si $X(\Omega) \subset [\![0\,;n]\!]$ et

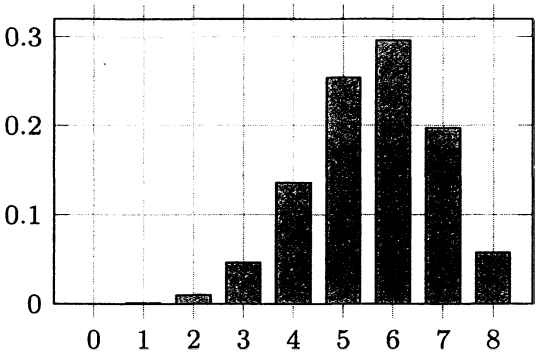
$$\mathrm{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{pour tout } k \in [\![0\,;n]\!].$$

Les lois binomiales servent à modéliser le nombre de succès lors de la répétition d'épreuves de Bernoulli indépendantes. En particulier, il a été vu en premier année que la somme de n variables de Bernoulli indépendantes de même paramètre p suit une loi binomiale de paramètres n et p.

Loi binomiale $n = 8$ et $p = 0{,}5$



Loi binomiale $n = 8$ et $p = 0{,}7$



Définition

Soit λ un réel strictement positif. On note $\mathcal{P}(\lambda)$ la *loi de Poisson* de paramètre λ. Une variable X suit cette loi si $X(\Omega) = \mathbb{N}$ et

$$\mathrm{P}(X=k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} \quad \text{pour tout } k \in \mathbb{N}.$$

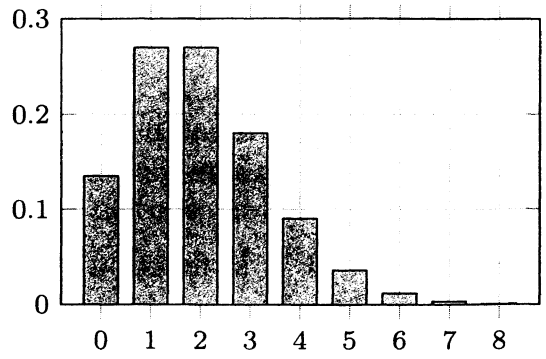
Les lois de Poisson servent à modéliser le nombre de succès lors d'une grande répétition d'épreuves ayant une faible probabilité de réussite, on parle d'*événements rares*. Cette interprétation s'explique par le résultat suivant :

Théorème 6

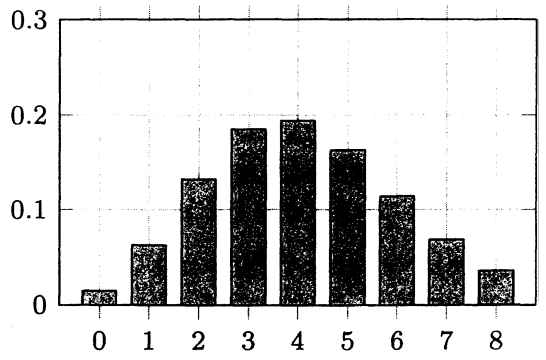
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. Si $X_n \hookrightarrow \mathcal{B}(n,p_n)$ pour tout $n \in \mathbb{N}$ et si [1] $np_n \xrightarrow[n\to+\infty]{} \lambda$ alors, pour tout $k \in \mathbb{N}$,

$$\mathrm{P}(X_n=k) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}.$$

Loi de Poisson $\lambda = 2$



Loi de Poisson $\lambda = 4{,}2$



Définition

Soit $p \in]0\,;1[$. On note $\mathcal{G}(p)$ la *loi géométrique* de paramètre p. Une variable X suit cette loi si[2] $X(\Omega) = \mathbb{N}^*$ et

$$\mathrm{P}(X = k) = (1-p)^{k-1}p \quad \text{pour tout } k \in \mathbb{N}^*.$$

Une loi géométrique permet de modéliser le *temps d'attente* du premier succès dans la répétition d'épreuves de Bernoulli indépendantes et de même paramètre p.

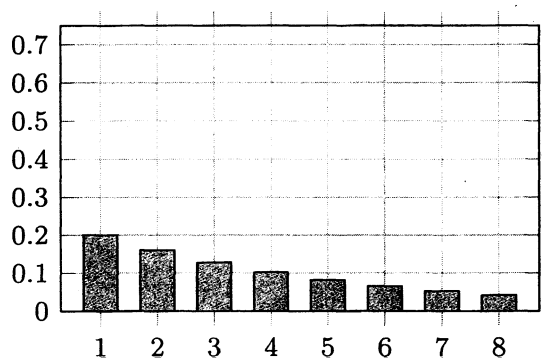
Théorème 7 (Processus sans mémoire)

Si $X \hookrightarrow \mathcal{G}(p)$, la variable X est *sans mémoire* dans le sens où

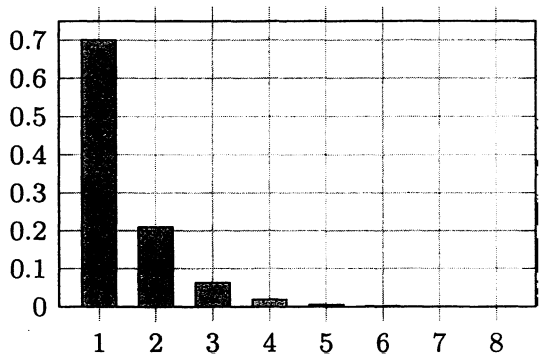
$$\mathrm{P}(X > n+k \,|\, X > n) = \mathrm{P}(X > k) \quad \text{pour tout } (n,k) \in \mathbb{N}^2.$$

Inversement, une variable aléatoire à valeurs dans $\mathbb{N}^*$ vérifiant la propriété précédente suit une loi géométrique.

Loi géométrique $p = 0{,}2$



Loi géométrique $p = 0{,}7$



1. La quantité np_n correspond au nombre moyen de succès pour la variable X_n. L'hypothèse np_n proche de λ pour n grand, signifie que l'expérience est répétée un grand nombre de fois avec une faible probabilité de réussite mais produit en moyenne un total de λ succès.

2. On peut autoriser X à prendre la valeur $+\infty$, l'événement correspondant étant alors nécessairement négligeable.

9.2 Vecteurs aléatoires

9.2.1 Loi conjointe

Soit X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$. On note E et F des ensembles finis ou dénombrables contenant les valeurs prises par X et Y.

Définition

On appelle *couple* défini par les variables aléatoires X et Y la fonction $Z = (X, Y)$ de Ω vers $E \times F$ déterminée par

$$Z(\omega) = \big(X(\omega), Y(\omega)\big) \quad \text{pour tout } \omega \in \Omega.$$

L'application Z est une variable aléatoire discrète sur $(\Omega, \mathcal{T}, \mathrm{P})$.

Définition

La loi du couple $Z = (X, Y)$ est appelée *loi conjointe* des deux variables aléatoires X et Y.

La loi conjointe P_Z est entièrement déterminée par la connaissance des probabilités élémentaires

$$\mathrm{P}\big(Z = (x, y)\big) = \mathrm{P}(X = x, Y = y) \quad \text{pour tout } (x, y) \in E \times F.$$

Lorsque les ensembles E et F sont finis, on peut visualiser une loi conjointe en figurant ses probabilités élémentaires dans un tableau.

9.2.2 Lois marginales

Soit Z une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ à valeurs dans un produit cartésien [1] $E \times F$. Sans perte de généralité, on suppose $E \times F$ fini ou dénombrable.

Pour chaque issue ω de Ω, $Z(\omega)$ désigne un couple élément de $E \times F$. On note $X(\omega) \in E$ et $Y(\omega) \in F$ les deux éléments de ce couple ce qui définit des variables aléatoires X et Y sur $(\Omega, \mathcal{T}, \mathrm{P})$.

Définition

Les lois des variables aléatoires X et Y sont appelées *les lois marginales* du couple de variables aléatoires $Z = (X, Y)$.

Théorème 8

La loi de Z détermine entièrement ses lois marginales puisque :

$$\mathrm{P}(X = x) = \sum_{y \in F} \mathrm{P}\big(Z = (x, y)\big) \quad \text{pour tout } x \in E$$

$$\mathrm{P}(Y = y) = \sum_{x \in E} \mathrm{P}\big(Z = (x, y)\big) \quad \text{pour tout } y \in F.$$

1. On dit parfois que Z est un *vecteur aléatoire*.

Dans un tableau visualisant la loi conjointe, les lois marginales s'obtiennent en sommant sur les rangées.

9.2.3 Lois conditionnelles

Soit X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$. On note E et F des ensembles finis ou dénombrables contenant les valeurs prises par X et Y.

Définition

Soit $x \in E$ tel que $\mathrm{P}(X = x) > 0$. On appelle *loi conditionnelle* de Y sachant $(X = x)$ la loi de la variable aléatoire Y pour la probabilité conditionnelle $\mathrm{P}(.\,|\,X = x)$.

Celle-ci est entièrement déterminée par les probabilités élémentaires

$$\mathrm{P}_{(X=x)}(Y = y) = \mathrm{P}(Y = y\,|\,X = x) \quad \text{pour tout } y \in F.$$

Lorsque $(X = x)$ est un événement négligeable, on pose $\mathrm{P}(Y = y\,|\,X = x) = 0$.

Théorème 9

La loi de X et les lois conditionnelles de Y connaissant les valeurs prises par X déterminent entièrement la loi conjointe des variables X et Y et donc la loi de Y :

$$\mathrm{P}(Y = y) = \sum_{x \in E} \mathrm{P}(Y = y\,|\,X = x)\,\mathrm{P}(X = x) \quad \text{pour tout } y \in F.$$

9.2.4 Vecteurs aléatoires

Soit $X_1, \ldots, X_n$ des variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$.

Définition

On appelle *vecteur aléatoire discret* défini à partir des variables aléatoires $X_1, \ldots, X_n$ la variable aléatoire discrète Z donnée par

$$Z(\omega) = \big(X_1(\omega), \ldots, X_n(\omega)\big) \quad \text{pour tout } \omega \in \Omega.$$

La loi de la variable Z est appelée la *loi conjointe* des variables $X_1, \ldots, X_n$ tandis que les lois de $X_1, \ldots, X_n$ sont les *lois marginales* de Z.

La loi conjointe détermine les lois marginales, mais l'inverse n'est pas vrai.

9.2.5 Couple de variables indépendantes

Soit X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ à valeurs dans des ensembles E et F.

Définition

On dit que les deux variables aléatoires X et Y sont *indépendantes* si, pour toute partie A de E et toute partie B de F, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants :

$$\mathrm{P}\big((X, Y) \in A \times B\big) = \mathrm{P}(X \in A)\,\mathrm{P}(Y \in B).$$

Théorème 10

Les variables aléatoires X et Y sont indépendantes si, et seulement si,

$$\mathrm{P}(X = x, Y = y) = \mathrm{P}(X = x)\,\mathrm{P}(Y = y) \quad \text{pour tout } (x, y) \in X(\Omega) \times Y(\Omega).$$

Deux événements A et B sont indépendants si, et seulement si, les fonctions indicatrices 1_A et 1_B définissent des variables aléatoires indépendantes.

Si X et Y sont deux variables indépendantes alors, pour toutes fonctions f et g définies sur E et F, les variables aléatoires composées $f(X)$ et $g(Y)$ sont aussi indépendantes.

9.2.6 Famille finie de variables mutuellement indépendantes

Soit $(X_i)_{i\in I}$ une famille de variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$. On note E_i l'ensemble d'arrivée de la variable X_i.

Définition

On dit que la famille $(X_i)_{i\in I}$ est constituée de *variables mutuellement indépendantes* si, pour toute famille $(A_i)_{i\in I}$ avec $A_i \subset E_i$, les événements $(X_i \in A_i)$ sont mutuellement indépendants.

Les sous-familles d'une telle famille sont aussi constituées de variables mutuellement indépendantes.

Théorème 11

Les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes si, et seulement si,

$$\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathrm{P}(X_1 = x_1) \times \cdots \times \mathrm{P}(X_n = x_n)$$

pour tout $(x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$.

Théorème 12 (Indépendance par paquets [1])

Si $X_1, \ldots, X_p$ et $X_{p+1}, \ldots, X_n$ sont des variables mutuellement indépendantes alors, pour toutes fonctions f et g définies sur $E_1 \times \cdots \times E_p$ et $E_{p+1} \times \cdots \times E_n$, les variables

$$X = f(X_1, \ldots, X_p) \quad \text{et} \quad Y = g(X_{p+1}, \ldots, X_n)$$

sont indépendantes.

9.2.7 Suites infinies d'épreuves

Afin d'assurer l'existence d'un cadre probabiliste permettant l'étude de la répétition indépendante et infinie d'une même expérience, nous disposons du résultat :

1. Aussi appelé *lemme de regroupement*.

Théorème 13
Si $\mathcal{L}$ désigne une loi d'une variable aléatoire discrète, il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ sur lequel existe une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires mutuellement indépendantes et suivant chacune la loi $\mathcal{L}$.

Définition
On dit alors que $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires *indépendantes et identiquement distribuées*[1] selon la loi $\mathcal{L}$.

En particulier, il existe un cadre probabiliste permettant de modéliser la répétition à l'infini d'épreuves de Bernoulli indépendantes.

9.3 Espérance d'une variable aléatoire réelle

Les variables aléatoires introduites ici sont toutes supposées discrètes et définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$.

9.3.1 Définition

Soit X une variable aléatoire réelle.
On note E un ensemble fini ou dénombrable contenant les valeurs prises par X.

Définition
On dit que la variable X *admet une espérance finie* lorsque $\big(x\,\mathrm{P}(X=x)\big)_{x\in E}$ est une famille sommable[2]. Sa somme définit alors l'*espérance* de la variable X

$$\mathrm{E}(X) \overset{\text{déf}}{=} \sum_{x\in E} x\,\mathrm{P}(X=x).$$

L'espérance de X définit la moyenne probabiliste de la variable X, elle ne dépend que de la loi de la variable X.

Les espérances des lois usuelles sont regroupées dans le tableau p. 383.

9.3.2 Propriétés

Soit X et Y deux variables aléatoires réelles.

Théorème 14
Si les variables X et Y admettent des espérances finies alors, pour tout $\lambda \in \mathbb{R}$, les variables λX et $X+Y$ admettent chacune une espérance finie avec

$$\mathrm{E}(\lambda X) = \lambda\,\mathrm{E}(X) \quad \text{et} \quad \mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y).$$

1. On utilise souvent l'abréviation « i.i.d. ».
2. Si la variable X est à valeurs positives et si la famille $\big(x\,\mathrm{P}(X=x)\big)_{x\in E}$ n'est pas sommable, on pose $\mathrm{E}(X) = +\infty$.

L'ensemble $\mathrm{L}^1(\Omega)$ des variables aléatoires réelles discrètes définies sur $(\Omega, \mathcal{T}, \mathrm{P})$ admettant une espérance finie est un sous-espace vectoriel de l'espace des fonctions de Ω vers $\mathbb{R}$. De plus, l'application espérance y définit une forme linéaire.

Théorème 15
Si la variable X est à valeurs positives alors $\mathrm{E}(X) \geqslant 0$.
Si de plus $\mathrm{E}(X) = 0$ alors $X = 0$ presque sûrement [1].

On en déduit la croissance de l'espérance : si X et Y sont d'espérances finies

$$X \leqslant Y \implies \mathrm{E}(X) \leqslant \mathrm{E}(Y).$$

Théorème 16 (Domination)
Si $|X| \leqslant Y$ et si Y admet une espérance finie alors X aussi.

Si une variable aléatoire X est bornée, elle est d'espérance finie.

9.3.3 Formule de transfert

Soit X une variable aléatoire prenant ses valeurs dans un ensemble [2] E au plus dénombrable.

Théorème 17 (Formule de transfert)
Si f est une fonction à valeurs réelles au moins définie sur E, la variable $f(X)$ admet une espérance finie si, et seulement si, la famille $\big(f(x)\,\mathrm{P}(X = x)\big)_{x \in E}$ est sommable. De plus, on a alors

$$\mathrm{E}\big(f(X)\big) = \sum_{x \in E} f(x)\,\mathrm{P}(X = x).$$

Ainsi, et sous réserve de sommabilité, on peut écrire pour une variable réelle X

$$\mathrm{E}(X^k) = \sum_{x \in E} x^k\,\mathrm{P}(X = x), \quad \mathrm{E}\big(\mathrm{e}^X\big) = \sum_{x \in E} \mathrm{e}^x\,\mathrm{P}(X = x), \ldots$$

9.3.4 Variables indépendantes

Soit X et Y deux variables aléatoires réelles.

Théorème 18
Si les variables X et Y sont indépendantes et admettent chacune une espérance finie alors XY admet une espérance finie et $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

La réciproque est fausse : on peut avoir $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$ sans pour autant avoir l'indépendance des variables X et Y.

1. Autrement dit, l'événement $(X = 0)$ est quasi certain.
2. Ici, E n'est pas nécessairement une partie de $\mathbb{R}$. Par exemple, ce peut être une partie de $\mathbb{R}^2$ ce qui autorise l'usage de la formule de transfert avec un X vecteur aléatoire.

9.4 Variance d'une variable aléatoire réelle

Les variables aléatoires introduites ici sont toutes supposées à valeurs réelles, discrètes et définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$.

9.4.1 Moments

Définition

On dit qu'une variable aléatoire réelle X *admet un moment d'ordre* $k \in \mathbb{N}$ si la variable X^k admet une espérance finie. On peut alors introduire le *moment d'ordre* k de la variable X :

$$m_k = \mathrm{E}(X^k).$$

Si la variable X admet un moment d'ordre n alors X admet [1] un moment d'ordre k pour tout $k \in [\![0\,;n]\!]$. En particulier, une variable admettant un moment d'ordre 2 admet une espérance.

9.4.2 Espace des variables possédant un moment d'ordre 2

L'ensemble $\mathrm{L}^2(\Omega)$ des variables admettant un moment d'ordre 2 est un sous-espace vectoriel de l'espace $\mathrm{L}^1(\Omega)$ des variables admettant une espérance finie.

Théorème 19 (Inégalité de Cauchy-Schwarz)

Si X et Y sont des variables aléatoires réelles admettant chacune un moment d'ordre 2 alors XY est d'espérance finie et

$$\mathrm{E}(XY)^2 \leqslant \mathrm{E}(X^2)\,\mathrm{E}(Y^2).$$

9.4.3 Variance et écart-type

Définition

Si une variable aléatoire réelle X admet un moment d'ordre 2, on appelle *variance* de la variable X le réel

$$\mathrm{V}(X) \overset{\text{déf}}{=} \mathrm{E}\Big(\big(X - \mathrm{E}(X)\big)^2\Big) \in \mathbb{R}_+.$$

On introduit aussi l'*écart-type* [2] de X par $\sigma(X) \overset{\text{déf}}{=} \sqrt{\mathrm{V}(X)}$.

Variance et écart-type mesurent la dispersion de la variable X autour de sa moyenne. Si $\mathrm{V}(X) = 0$, la variable X est presque sûrement constante [3].

Les variances des lois usuelles sont regroupées dans le tableau p. 383.

1. Voir sujet 7 p. 390.
2. L'espérance et l' écart-type s'expriment dans la même unité que les valeurs de la variable X.
3. Autrement dit, il existe une constante C telle que l'événement $(X = C)$ est quasi certain : cette constante C est l'espérance de la variable X.

Théorème 20 (Formule de Huygens)
Si X admet un moment d'ordre 2,

$$\mathrm{V}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2.$$

9.4.4 Variable centrée réduite

Théorème 21
Si X admet un moment d'ordre 2 alors, pour tous a et $b \in \mathbb{R}$,

$$\mathrm{V}(aX + b) = a^2\,\mathrm{V}(X).$$

Définition
Lorsqu'une variable X admet une espérance finie et que celle-ci est nulle, on dit que la variable X est *centrée*. Si de plus celle-ci admet une variance égale à 1, on la qualifie de *réduite*.

Si X est une variable admettant un moment d'ordre 2 alors, en introduisant son espérance μ et son écart-type σ (que l'on suppose non nul), la variable

$$Y = \frac{X - \mu}{\sigma}$$

est centrée réduite.

9.4.5 Covariance

Soit X et Y deux variables aléatoires réelles.

Définition
Si X et Y admettent chacune un moment d'ordre 2, on introduit leur *covariance* :

$$\mathrm{Cov}(X,Y) = \mathrm{E}\Big(\big(X - \mathrm{E}(X)\big)\big(Y - \mathrm{E}(Y)\big)\Big) \in \mathbb{R}.$$

En particulier, $\mathrm{V}(X) = \mathrm{Cov}(X,X)$ et l'inégalité de Cauchy-Schwarz donne [1]

$$\big|\mathrm{Cov}(X,Y)\big| \leqslant \mathrm{V}(X)\,\mathrm{V}(Y).$$

Théorème 22 (Formule de Huygens)
Si les variables X et Y admettent chacune un moment d'ordre 2,

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y).$$

Si les variables X et Y sont indépendantes, on a $\mathrm{Cov}(X,Y) = 0$. La réciproque est fausse.

1. Lorsque les variances des variables X et Y sont non nulles, on introduit leur *coefficient de corrélation* : $\mathrm{cor}(X,Y) \overset{\text{déf}}{=} \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{V}(X)\,\mathrm{V}(Y)}} \in [-1\,;1]$. Si les variables sont indépendantes, celui-ci est nul. Si les variables évoluent « dans le même sens », ce coefficient est proche de 1. Il est proche de -1 lorsqu'elles évoluent « en sens inverse ».

9.4.6 Variance d'une somme

La covariance définit une forme bilinéaire symétrique sur l'espace $L^2(\Omega)$.

Théorème 23
Si X et Y sont des variables aléatoires admettant chacune un moment d'ordre 2,

$$V(X+Y) = V(X) + 2\,\mathrm{Cov}(X,Y) + V(Y).$$

Si les variables sont indépendantes, $V(X+Y) = V(X) + V(Y)$. Plus généralement,

Théorème 24
Si $X_1, \ldots, X_n$ sont des variables aléatoires admettant chacune un moment d'ordre 2,

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

Si les variables X_1,..., X_n sont deux à deux indépendantes

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i).$$

Cette égalité est *a fortiori* vérifiée si les variables sont mutuellement indépendantes.

9.4.7 Inégalités de concentration

Théorème 25 (Inégalité de Markov)
Si X est une variable à valeurs positives admettant une espérance finie,

$$a\,P(X \geqslant a) \leqslant E(X) \quad \text{pour tout } a \in \mathbb{R}_+.$$

Théorème 26 (Inégalité de Bienaymé-Tchebychev)
Si X est une variable aléatoire admettant un moment d'ordre 2,

$$P\big(|X - E(X)| \geqslant a\big) \leqslant \frac{V(X)}{a^2} \quad \text{pour tout } a > 0.$$

Cette inégalité permet de mesurer l'écart entre « l'expérimentation et l'espérance ».
Cette inégalité explique aussi pourquoi la variance mesure la dispersion de la variable aléatoire.

9.4.8 Loi faible des grands nombres

Théorème 27 (Loi faible des grands nombres)
Si $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires deux à deux indépendantes et suivant une même loi admettant une espérance μ et un moment d'ordre 2,

$$\mathrm{P}\left(\left|\frac{1}{n}S_n - \mu\right| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0 \quad \text{avec} \quad S_n = \sum_{k=1}^{n} X_k.$$

Ce théorème [1] montre que la moyenne expérimentale tend à se rapprocher de la moyenne probabiliste (c'est-à-dire de l'espérance).

9.5 Fonctions génératrices

Les variables aléatoires introduites ici sont toutes supposées définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$.

9.5.1 Définition

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

Définition
On appelle *série génératrice* de la variable X la série entière

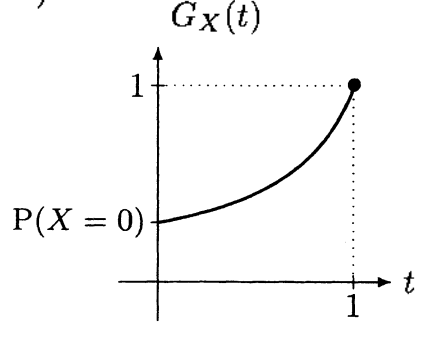
$$\sum \mathrm{P}(X=n)t^n.$$

Cette série entière est de rayon de convergence supérieur ou égal à 1 et converge normalement sur le segment $[-1\,;1]$.

Définition
On appelle *fonction génératrice* de la variable X la somme de sa série génératrice

$$G_X(t) \overset{\text{déf}}{=} \sum_{n=0}^{+\infty} \mathrm{P}(X=n)t^n = \mathrm{E}\left(t^X\right).$$

Celle-ci est définie et continue au moins [2] sur $[-1\,;1]$. C'est une fonction croissante sur $[0\,;1]$ qui prend la valeur 1 en 1 et, plus généralement, c'est une fonction bornée par 1 sur $[-1\,;1]$.
La fonction génératrice de X est entièrement déterminée par sa loi. Inversement, la fonction génératrice caractérise la loi de X puisque



1. Sous les mêmes hypothèses, la loi forte des grands nombres affirme $\mathrm{P}(S_n/n \xrightarrow[n\to+\infty]{} \mu) = 1$. Voir sujet 35 p. 431.

2. Si le rayon de convergence R de la série génératrice est strictement supérieur à 1, la fonction génératrice peut être définie sur $]-R\,;R[$ mais l'étudier sur $[-1\,;1]$, voire sur $[0\,;1]$, est souvent suffisant en pratique.

$$\mathrm{P}(X=n)=\frac{G_X^{(n)}(0)}{n!} \quad \text{pour tout } n\in\mathbb{N}.$$

Les expressions des fonctions génératrices des lois usuelles sont regroupées dans le tableau à la fin de ces rappels de cours.

9.5.2 Calcul d'espérances et de variances

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

Théorème 28
La variable aléatoire X admet une espérance finie si, et seulement si, sa fonction génératrice G_X est dérivable en 1 et alors

$$\mathrm{E}(X)=G_X'(1).$$

La variable aléatoire X admet un moment d'ordre 2 si, et seulement si, sa fonction génératrice G_X est deux fois dérivable en 1 et alors

$$\mathrm{V}(X)=G_X''(1)+G_X'(1)-\big(G_X'(1)\big)^2.$$

Ce résultat peut notamment être employé lorsque la série génératrice de X est de rayon de convergence strictement supérieur à 1.

9.5.3 Fonction génératrice d'une somme

Théorème 29
Si X et Y sont deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$,

$$G_{X+Y}(t)=G_X(t)\times G_Y(t) \quad \text{pour tout } t\in[-1\,;1].$$

Ce résultat se généralise à la somme de plusieurs variables aléatoires sous réserve que celles-ci soient mutuellement indépendantes.

9.6 Lois usuelles

Loi	Espérance	Variance	Fonction génératrice	
$\mathcal{B}(p)$	p	$p(1-p)$	$t\mapsto(1-p)+pt$	$p\in[0\,;1]$
$\mathcal{B}(n,p)$	np	$np(1-p)$	$t\mapsto\big((1-p)+pt\big)^n$	$n\in\mathbb{N}$, $p\in[0\,;1]$
$\mathcal{P}(\lambda)$	λ	λ	$t\mapsto \mathrm{e}^{\lambda(t-1)}$	$\lambda>0$
$\mathcal{G}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$t\mapsto\frac{pt}{1-(1-p)t}$	$p\in]0\,;1[$

9.7 Exercices d'apprentissage

9.7.1 Variables aléatoires

Exercice 1

Pour quels $a \in \mathbb{R}$, existe-t-il une variable aléatoire X à valeurs dans $\mathbb{N}^*$ telle que

$$\mathrm{P}(X = n) = \frac{a}{n(n+1)} \quad \text{pour tout } n \in \mathbb{N}^* \text{ ?}$$

Solution

méthode

Les probabilités élémentaires d'une variable aléatoire forment une famille de réels positifs sommable et de somme égale à 1.

Analyse : Soit X une variable aléatoire sur un univers $(\Omega, \mathcal{T}, \mathrm{P})$ à valeurs dans $\mathbb{N}^*$ dont les probabilités élémentaires sont celles proposées. L'univers Ω est la réunion dénombrable des événements $(X = n)$ pour n parcourant $\mathbb{N}^*$. Par additivité dénombrable

$$\mathrm{P}(\Omega) = \sum_{n=1}^{+\infty} \mathrm{P}(X = n) = \sum_{n=1}^{+\infty} \frac{a}{n(n+1)}.$$

Par décomposition en éléments simples

$$\frac{a}{n(n+1)} = \frac{a}{n} - \frac{a}{n+1}$$

puis par calcul télescopique

$$\mathrm{P}(\Omega) = \lim_{N \to +\infty} \sum_{n=1}^{N} \left(\frac{a}{n} - \frac{a}{n+1} \right) = \lim_{N \to +\infty} \left(a - \frac{a}{N+1} \right) = a.$$

On en déduit $a = 1$.

Synthèse : Supposons $a = 1$. Les calculs qui précèdent assure que la famille de réels positifs $\frac{1}{n(n+1)}$ avec $n \in \mathbb{N}^*$ est sommable et de somme égale à 1. On peut alors affirmer qu'il existe une variable aléatoire X à valeurs dans $\mathbb{N}^*$ pour laquelle

$$\mathrm{P}(X = n) = \frac{1}{n(n+1)} \quad \text{pour tout } n \in \mathbb{N}^*.$$

En effet[1], on peut munir l'espace $\left(\mathbb{N}^*, \wp(\mathbb{N}^*)\right)$ d'une probabilité P déterminée par (Th. 4 p. 334)

$$\mathrm{P}\left(\{n\}\right) = \frac{1}{n(n+1)} \quad \text{pour tout } n \in \mathbb{N}^*.$$

L'application $X = \mathrm{Id}_{\mathbb{N}^*}$ est alors une variable aléatoire telle que voulue.

1. Cette dernière étape est rarement précisée.

Exercice 2

On lance indéfiniment et indépendamment un dé équilibré. Pour $n \in \mathbb{N}^*$, on note X_n la variable aléatoire définie par la valeur du n-ième lancer. On introduit le temps d'attente du premier 'six' :

$$T = \min\Big(\big\{n \in \mathbb{N}^* \,\big|\, X_n = 6\big\} \cup \{+\infty\}\Big).$$

Montrer que T est une variable aléatoire discrète à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$ et déterminer sa loi.

Solution

Notons $(\Omega, \mathcal{T}, \mathrm{P})$ un espace probabilisé [1] modélisant l'expérience.

L'application $T\colon \Omega \to \mathbb{N}^* \cup \{+\infty\}$ est bien définie car le min existe dans $\mathbb{N}^* \cup \{+\infty\}$. Plus précisément, pour une issue ω donnée, s'il existe $n \in \mathbb{N}^*$ tel que $X_n(\omega) = 6$, $T(\omega)$ correspond au plus petit élément d'une partie de $\mathbb{N}^*$. Celui-ci existe et appartient à $\mathbb{N}^*$. En revanche, s'il n'existe pas de $n \in \mathbb{N}^*$ tel que $X_n(\omega) = 6$, le min définissant $T(\omega)$ vaut $+\infty$. Montrons ensuite que l'application T est une *variable aléatoire discrète.*

méthode

Une variable aléatoire discrète sur Ω est une application prenant ses valeurs dans un ensemble E au plus dénombrable et telle que les antécédents de toute valeur de E constituent un événement.

La variable T prend ses valeurs dans l'ensemble $\mathbb{N}^* \cup \{+\infty\}$ qui est dénombrable. Pour toute valeur n de $\mathbb{N}^*$, on a [2]

$$\begin{aligned}(T = n) &= \big\{\omega \in \Omega \,\big|\, X_1(\omega) \neq 6, \ldots, X_{n-1}(\omega) \neq 6, X_n(\omega) = 6\big\}\\ &= \overline{(X_1 = 6)} \cap \ldots \cap \overline{(X_{n-1} = 6)} \cap (X_n = 6).\end{aligned}$$

Puisque les X_k sont des variables aléatoires sur Ω, chaque condition $(X_k = 6)$ détermine un événement. Par opérations dans la tribu $\mathcal{T}$, $(T = n)$ est bien un événement de Ω.

Pour la valeur $+\infty$, on a [3]

$$(T = +\infty) = \big\{\omega \in \Omega \,\big|\, \forall n \in \mathbb{N}^*,\ X_n(\omega) \neq 6\big\} = \bigcap_{n \in \mathbb{N}^*} \overline{(X_n = 6)}.$$

Par intersection dénombrable d'événements, $(T = +\infty)$ est bien un événement de Ω.

Ainsi, T est une variable aléatoire à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$. Il reste à déterminer sa loi. Soit $n \in \mathbb{N}^*$. Par indépendance des variables X_n

$$\begin{aligned}\mathrm{P}(T = n) &= \mathrm{P}\big(\overline{(X_1 = 6)} \cap \ldots \cap \overline{(X_{n-1} = 6)} \cap (X_n = 6)\big)\\ &= \mathrm{P}\big(\overline{(X_1 = 6)}\big) \times \ldots \times \mathrm{P}\big(\overline{(X_{n-1} = 6)}\big) \times \mathrm{P}\big((X_n = 6)\big) = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}.\end{aligned}$$

1. Le Th. 13 p. 377 assure l'existence d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ permettant de modéliser cette expérience.

2. Lorsque $n = 1$, ce qui suit se comprend $(T = 1) = (X_1 = 6)$.

3. On peut aussi dire que $(T = +\infty)$ est l'événement contraire de la réunion dénombrable des événements $(T = n)$.

Aussi, par continuité décroissante[1] (Th. 6 p. 335),

$$\mathrm{P}(T=+\infty)=\mathrm{P}\left(\bigcap_{n\in\mathbb{N}^*}\overline{(X_n=6)}\right)=\lim_{n\to+\infty}\mathrm{P}\left(\bigcap_{k=1}^{n}\overline{(X_k=6)}\right)=\lim_{n\to+\infty}\left(\frac{5}{6}\right)^n=0.$$

Finalement, la variable aléatoire T suit une loi géométrique de paramètre $p=1/6$.

Exercice 3

Soit X et Y deux variables aléatoires discrètes indépendantes. On suppose que X et Y suivent des lois de Poisson de paramètres λ et μ strictement positifs.

Déterminer la loi suivie par $X+Y$.

Solution

méthode

Une variable aléatoire X suit une loi de Poisson de paramètre $\lambda>0$ lorsque celle-ci est à valeurs dans $\mathbb{N}$ et vérifie :

$$\mathrm{P}(X=n)=\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\quad\text{pour tout } n\in\mathbb{N}.$$

En tant que somme de deux variables aléatoires, $Z=X+Y$ est bien une variable aléatoire. De plus, X et Y sont à valeurs dans $\mathbb{N}$ et donc Z est aussi à valeurs dans[2] $\mathbb{N}$. Pour déterminer sa loi, il suffit de calculer $\mathrm{P}(Z=n)$ pour tout $n\in\mathbb{N}$.

Soit $n\in\mathbb{N}$. L'événement $(Z=n)$ est la réunion des événements deux à deux incompatibles $(X=k,Y=\ell)$ pour $(k,\ell)\in\mathbb{N}^2$ vérifiant $k+\ell=n$. On a donc

$$\mathrm{P}(X+Y=n)=\sum_{k+\ell=n}\mathrm{P}(X=k,Y=\ell)=\sum_{k=0}^{n}\mathrm{P}(X=k,Y=n-k).$$

Par indépendance des variables X et Y, on écrit

$$\mathrm{P}(X=k,Y=n-k)=\mathrm{P}(X=k)\,\mathrm{P}(Y=n-k).$$

Les lois des variables X et Y étant connues, on concrétise le calcul

$$\mathrm{P}(X+Y=n)=\sum_{k=0}^{n}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\mathrm{e}^{-\mu}\frac{\mu^{n-k}}{(n-k)!}=\mathrm{e}^{-(\lambda+\mu)}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda^k\mu^{n-k}$$

Par la formule du binôme, on conclut

$$\mathrm{P}(X+Y=n)=\mathrm{e}^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!}.$$

Finalement, la variable $Z=X+Y$ suit une loi de Poisson[3] de paramètre $\lambda+\mu$.

1. Ce dernier calcul n'est pas nécessaire, les valeurs de $\mathrm{P}(T=n)$ obtenues au-dessus déterminent une loi géométrique dont la somme des probabilités élémentaires vaut 1 : ceci ne laisse « plus de place » pour l'événement $(T=+\infty)$.

2. Ceci ne signifie pas pour autant que Z prend assurément toutes les valeurs de $\mathbb{N}$: sans l'indépendance de X et Y ceci pourrait ne pas être vrai, c'est le cas par exemple lorsque $X=Y$.

3. On trouvera une démonstration alternative par les fonctions génératrices dans le sujet 9 p. 392.

Exercice 4

Soit X et Y deux variables aléatoires indépendantes suivant des lois géométriques de paramètres p et q éléments de $]0\,;1[$.

(a) Calculer $\mathrm{P}(X > n)$ pour $n \in \mathbb{N}$.

(b) Déterminer la loi de $Z = \min(X, Y)$.

Solution

(a) **méthode**

Une variable aléatoire X suit une loi géométrique de paramètre $p \in]0\,;1[$ lorsque celle-ci est à valeurs dans $\mathbb{N}^*$ (ou $\mathbb{N}^* \cup \{+\infty\}$) et vérifie :

$$\mathrm{P}(X = n) = (1-p)^{n-1}p \quad \text{pour tout } n \in \mathbb{N}^*.$$

Soit $n \in \mathbb{N}$. L'événement $(X > n)$ est la réunion dénombrable des événements deux à deux incompatibles $(X = k)$ pour $k = n+1$, $n+2$, etc. On a donc par additivité dénombrable

$$\mathrm{P}(X > n) = \sum_{k=n+1}^{+\infty} \mathrm{P}(X = k).$$

On exprime la loi de X afin de concrétiser le calcul

$$\mathrm{P}(X > n) = \sum_{k=n+1}^{+\infty} (1-p)^{k-1}p.$$

Par un glissement d'indice puis une sommation géométrique de raison $1-p \in]0\,;1[$, on obtient

$$\mathrm{P}(X > n) = (1-p)^n p \sum_{\ell=0}^{+\infty} (1-p)^{\ell} = (1-p)^n p \times \frac{1}{1-(1-p)} = (1-p)^n.$$

Ce résultat peut facilement se retenir : la variable X s'interprète comme le temps d'attente du premier succès lors de la répétition d'épreuves de Bernoulli indépendantes et de même paramètre p. L'événement $(X > n)$ signifie que la succession d'épreuves débutent par une série de n échecs.

(b) Puisque X et Y sont des variables aléatoires, Z est aussi une variable aléatoire car se déduit de la composition du vecteur aléatoire (X, Y) avec la fonction min définie sur $\mathbb{R}^2$. De plus, la variable aléatoire Z est à valeurs dans $\mathbb{N}^*$.

méthode

On exprime facilement $(Z > n)$ à l'aide de $(X > n)$ et $(Y > n)$ car la variable Z est définie par un min.

Soit $n \in \mathbb{N}$. L'événement $(Z > n)$ est la conjonction des deux événements $(X > n)$ et $(Y > n)$. Or les variables X et Y sont indépendantes et les événements précédents le

sont donc aussi. Ainsi,

$$\mathrm{P}(Z > n) = \mathrm{P}(X > n, Y > n) = \mathrm{P}(X > n)\,\mathrm{P}(Y > n) = (1-p)^n(1-q)^n.$$

méthode

L'événement $(Z = n)$ se déduit [1] des événements $(Z > n)$ et $(Z > n-1)$.

Soit $n \in \mathbb{N}^*$. L'événement $(Z > n-1)$ est la réunion des événements incompatibles $(Z = n)$ et $(Z > n)$. On a donc par additivité

$$\mathrm{P}(Z > n-1) = \mathrm{P}(Z = n) + \mathrm{P}(Z > n)$$

puis

$$\mathrm{P}(Z = n) = \mathrm{P}(Z > n-1) - \mathrm{P}(Z > n) = (p+q-pq)\big((1-p)(1-q)\big)^{n-1}.$$

En posant $r = p + q - pq$, on observe $\mathrm{P}(Z = n) = r(1-r)^{n-1}$. La variable Z suit donc une loi géométrique [2] de paramètre r.

Exercice 5

Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que la loi conjointe de X et Y vérifie :

$$\mathrm{P}(X = j, Y = k) = \frac{1}{3\mathrm{e}^2} \cdot \frac{j+k+1}{j!k!} \quad \text{pour tout } (j,k) \in \mathbb{N}^2.$$

(a) Déterminer les lois des variables X et Y.

(b) Les variables X et Y sont-elles indépendantes ?

Solution

(a) **méthode**

La loi marginale de X s'obtient en sommant sur les valeurs possibles prises par Y (Th. 8 p. 374).

La variable X est à valeurs dans $\mathbb{N}$ et, pour tout $j \in \mathbb{N}$,

$$\mathrm{P}(X = j) = \sum_{k=0}^{+\infty} \mathrm{P}(X = j, Y = k) = \frac{1}{3\mathrm{e}^2} \sum_{k=0}^{+\infty} \frac{j+k+1}{j!k!}.$$

1. Les événements $(Z > n)$ suffisent donc pour caractériser la loi de Z : cet argument peut être employé pour parvenir sans calculs à l'identification de la loi de Z.

2. On peut comprendre cette propriété en termes de temps d'attente. Si X_n et Y_n suivent des lois de Bernoulli indépendantes de paramètres p et q, la variable $Z_n = \max(X_n, Y_n)$ suit une loi de Bernoulli de paramètre r et Z_n est le temps d'attente associé aux épreuves de Bernoulli correspondantes.

Il s'agit d'une somme convergente[1] de termes positifs, on peut séparer celle-ci en deux par linéarité avec convergence des séries à termes positifs introduites

$$\mathrm{P}(X=j)=\frac{1}{3\mathrm{e}^2}\left(\sum_{k=0}^{+\infty}\frac{j+1}{j!k!}+\sum_{k=0}^{+\infty}\frac{k}{j!k!}\right).$$

Dans la première somme, on peut rapidement faire apparaître une somme exponentielle calculée en 1. Dans la seconde, on peut simplifier le terme d'indice $k=0$ avant de faire apparaître à nouveau une somme exponentielle par glissement d'indice[2]

$$\mathrm{P}(X=j)=\frac{1}{3\mathrm{e}^2}\left(\frac{j+1}{j!}\sum_{k=0}^{+\infty}\frac{1}{k!}+\frac{1}{j!}\sum_{\ell=0}^{+\infty}\frac{1}{\ell!}\right)=\frac{1}{3\mathrm{e}}\cdot\frac{j+2}{j!}.$$

Les variables X et Y jouant des rôles symétriques, on obtient aussi

$$\mathrm{P}(Y=k)=\frac{1}{3\mathrm{e}}\frac{k+2}{k!}\quad\text{pour tout } k\in\mathbb{N}.$$

(b) **méthode**

Par le Th. 10 p. 376 on vérifie, ou on contredit, l'indépendance des variables X et Y en étudiant l'égalité :

$$\mathrm{P}(X=j,Y=k)=\mathrm{P}(X=j)\,\mathrm{P}(Y=k).$$

Pour $(j,k)=(0,0)$, on observe

$$\mathrm{P}(X=0,Y=0)=\frac{1}{3\mathrm{e}^2}\neq\frac{2}{3\mathrm{e}}\cdot\frac{2}{3\mathrm{e}}=\mathrm{P}(X=0)\,\mathrm{P}(Y=0).$$

Les variables X et Y ne sont pas indépendantes.

Exercice 6

Une urne contient un dé truqué donnant systématiquement un 'six'. On lance une pièce équilibrée. Si l'on obtient 'face', on ajoute un dé équilibré dans l'urne et l'on relance la pièce. Si l'obtient 'pile', on tire un dé dans l'urne et on lance celui-ci. Déterminer la loi de variable donnant la valeur du dé lancé.

Solution

méthode

On peut déterminer la loi d'une variable Y à partir de la loi d'une variable X et des lois conditionnelles de Y connaissant la valeur prise par X (Th. 9 p. 375).

1. Par hypothèse, la famille des $\mathrm{P}(X=j,Y=k)$ avec $(j,k)\in\mathbb{N}^2$ est sommable (et même de somme égale à 1) car définit la loi du couple (X,Y).

2. On peut vérifier que la somme des $\mathrm{P}(X=j)$ pour j parcourant $\mathbb{N}$ est bien égale à 1 ce qui légitime le facteur $1/3\mathrm{e}^2$ introduit lors de la définition de $\mathrm{P}(X=j,Y=k)$.

Notons X la variable aléatoire donnant le nombre de dés dans l'urne lorsque l'on y pioche un dé et Y la variable aléatoire donnant la valeur du dé lancé.

Le temps d'attente du premier 'pile' donne exactement le nombre de dés figurant dans l'urne au moment où l'on pioche le dé. La pièce étant supposée équilibrée, la variable X suit une loi géométrique[1] de paramètre $p = 1/2$.

Soit $n \in \mathbb{N}^*$. Lorsque l'urne est constituée d'un dé truqué et de $n-1$ dés équilibrés, une application rapide de la formule des probabilités totales permet de calculer la loi de Y

$$\mathrm{P}(Y = 6 \,|\, X = n) = 1 \cdot \frac{1}{n} + \frac{1}{6} \cdot \frac{n-1}{n} = \frac{n+5}{6n} \quad \text{et}$$
$$\mathrm{P}(Y = k \,|\, X = n) = 0 \cdot \frac{1}{n} + \frac{1}{6} \cdot \frac{n-1}{n} = \frac{n-1}{6n} \quad \text{pour } k \in [\![1\,;5]\!].$$

On en déduit

$$\mathrm{P}(Y = 6) = \sum_{n=1}^{+\infty} \mathrm{P}(Y = 6 \,|\, X = n)\,\mathrm{P}(X = n) = \sum_{n=1}^{+\infty} \frac{n+5}{6n} \cdot \frac{1}{2^n}.$$

En exploitant pour $x = 1/2$ l'identité

$$\sum_{n=1}^{+\infty} \frac{1}{n} x^n = -\ln(1-x) \quad \text{valable pour tout } x \in]-1\,;1[$$

on obtient

$$\mathrm{P}(Y = 6) = \frac{1 + 5\ln 2}{6} \simeq 0{,}744 \text{ à } 10^{-3} \text{ près.}$$

Un calcul analogue donne aussi

$$\mathrm{P}(Y = 1) = \cdots = \mathrm{P}(Y = 5) = \frac{1 - \ln 2}{6} \simeq 0{,}051 \text{ à } 10^{-3} \text{ près.}$$

9.7.2 Moments, espérance et variance

Exercice 7

Soit X une variable aléatoire discrète à valeurs réelles.

On suppose que X admet un moment d'ordre $n \in \mathbb{N}$. Montrer que X admet un moment d'ordre k pour tout $k \in [\![0\,;n]\!]$.

1. Il se peut que ce temps d'attente soit infini ce qui signifierait que l'on ne pioche jamais de dé dans l'urne... L'événement correspondant est cependant *négligeable* et n'influence donc pas les calculs.

Solution

Soit $k \in [\![0\,;n]\!]$.

méthode

Pour tout x réel, l'inégalité $|x^k| \leqslant 1 + |x|^n$ est vraie que $|x| \leqslant 1$ ou non.

Pour toute issue ω, on peut écrire par ce qui précède $|X(\omega)^k| \leqslant 1 + |X(\omega)|^n$. On a donc la comparaison de variables aléatoires $|X^k| \leqslant 1 + |X|^n$. Or, la variable aléatoire constante 1 admet une espérance finie et la variable $|X|^n$ aussi par hypothèse. Par domination (Th. 16 p. 378), on peut affirmer que la variable X^k admet aussi une espérance finie, autrement dit, X admet un moment d'ordre k.

Exercice 8 (Loi binomiale négative [1])

Soit $r \in \mathbb{N}^*$, $p \in]0\,;1[$ et X une variable aléatoire à valeurs dans $\mathbb{N}$ telle qu'il existe un réel a pour lequel

$$\mathrm{P}(X = k) = a\binom{k+r-1}{k}(1-p)^k \quad \text{pour tout } k \in \mathbb{N}.$$

Calculer l'espérance et la variance de X.

On rappelle l'identité binomiale [2] :

$$\sum_{k=0}^{+\infty}\binom{k+n}{k}x^k = \frac{1}{(1-x)^{n+1}} \quad \text{pour tous } n \in \mathbb{N} \text{ et } x \in]-1\,;1[.$$

Solution

méthode

La valeur de a ne peut être arbitraire : on commence par déterminer celle-ci.

On calcule la valeur de a en exploitant que la somme des $\mathrm{P}(X = k)$ pour k parcourant $\mathbb{N}$ est égale à 1. Par l'identité binomiale utilisée pour $x = 1 - p$ et $n = r - 1$

$$\sum_{k=0}^{+\infty}\mathrm{P}(X = k) = a\sum_{k=0}^{+\infty}\binom{k+r-1}{k}(1-p)^k = \frac{a}{p^r}.$$

On a donc $a = p^r$.

méthode

On calcule l'espérance de X par dérivation de l'identité binomiale.

L'identité binomiale correspond à un développement en série entière, il est possible de dériver terme à terme celui-ci sur l'intervalle ouvert de convergence. Ceci donne

$$\sum_{k=1}^{+\infty}k\binom{k+r-1}{k}x^{k-1} = \frac{r}{(1-x)^{r+1}} \quad \text{pour tout } x \in]1\,;1[.$$

1. Cette loi étudie le nombre d'échecs précédant le r-ième succès lors de la répétition d'épreuves de Bernoulli indépendantes de même paramètre p.

2. Voir sujet 5 du chapitre 9 de l'ouvrage *Exercices d'analyse MP*. Cette identité est souvent utilisée en calcul des probabilités.

En évaluant cette identité en $x = 1 - p$ et en ajoutant un terme nul d'indice $k = 0$ à la somme, on obtient que X admet une espérance finie et

$$\mathrm{E}(X) = a\sum_{k=0}^{+\infty} k\binom{k+r-1}{k}(1-p)^k = a\frac{r(1-p)}{p^{r+1}} = \frac{r(1-p)}{p}.$$

méthode

La variance de X se déduit de la formule de Huygens (Th. 20 p. 380) après calcul de $\mathrm{E}\big(X(X-1)\big)$

Une nouvelle dérivation de l'identité binomiale donne

$$\sum_{k=2}^{+\infty} k(k-1)\binom{k+r-1}{k}x^{k-2} = \frac{r(r+1)}{(1-x)^{r+2}} \quad \text{pour tout } x \in]-1\,;1[.$$

On en déduit

$$\mathrm{E}\big(X(X-1)\big) = \frac{r(r+1)(1-p)^2}{p^2}.$$

La variable $X(X-1)$ admettant une espérance finie, il en est de même de X^2 car on peut écrire $X^2 = X(X-1) + X$. La variable X admet donc un moment d'ordre 2 et l'on peut calculer sa variance :

$$\mathrm{V}(X) = \mathrm{E}\big(X(X-1)\big) + \mathrm{E}(X) - \mathrm{E}(X)^2 = \frac{r(1-p)}{p^2}.$$

9.7.3 Fonctions génératrices

Exercice 9

(a) Calculer la fonction génératrice d'une variable aléatoire X suivant une loi de Poisson de paramètre $\lambda > 0$

(b) À l'aide de leurs fonctions génératrices, déterminer la loi suivie par la somme de deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et μ strictement positifs.

Solution

(a) **méthode**

La fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est définie par

$$G_X(t) = \mathrm{E}\big(t^X\big) = \sum_{n=0}^{+\infty} \mathrm{P}(X=n)t^n \quad \text{pour } t \text{ réel convenable}^1.$$

1. La fonction génératrice est au moins définie sur $[-1\,;1]$.

Puisque X suit une loi de Poisson, on obtient en reconnaissant une somme exponentielle

$$G_X(t) = \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} t^n = \mathrm{e}^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} t^n = \mathrm{e}^{-\lambda} \mathrm{e}^{\lambda t} = \mathrm{e}^{\lambda(t-1)} \quad \text{pour tout } t \text{ réel.}$$

(b) **méthode**

La fonction génératrice d'une variable aléatoire caractérise sa loi.

Les variables X et Y étant indépendantes, la fonction génératrice de leur somme est le produit de leurs fonctions génératrices (Th. 29 p. 383)

$$G_{X+Y}(t) = G_X(t) \times G_Y(t) = \mathrm{e}^{(\lambda+\mu)(t-1)}.$$

On reconnaît ici la fonction génératrice d'une loi de Poisson de paramètre $\lambda + \mu$, on peut donc affirmer que $X + Y$ suit cette loi [1].

Exercice 10

On considère de nouveau la variable aléatoire X présentée dans le sujet 8 p. 391
Calculer la fonction génératrice de X et retrouver les valeurs de son espérance et de sa variance.

Solution

Par définition de la fonction génératrice d'une variable aléatoire à valeurs dans $\mathbb{N}$

$$G_X(t) = \mathrm{E}(t^X) = \sum_{k=0}^{+\infty} \mathrm{P}(X = k) t^k = \sum_{k=0}^{+\infty} a \binom{k+r-1}{k} \big((1-p)t\big)^k$$

pour tout t convenable. Grâce à l'identité binomiale

$$G_X(t) = \frac{a}{\big(1-(1-p)t\big)^r} \quad \text{pour tout } t \in]-R\,;R[\text{ avec } R = \frac{1}{1-p}\,.$$

Une fonction génératrice prend la valeur 1 et 1 ce qui détermine la valeur de a et permet de conclure

$$G_X(t) = \left(\frac{p}{1-(1-p)t}\right)^r \quad \text{pour tout } t \in]-R\,;R[.$$

méthode

Espérance et variance se déduisent de la fonction génératrice à partir du calcul de $G'_X(1)$ et $G''_X(1)$ lorsque celui-ci est possible (Th. 28 p. 383).

La fonction G_X est de classe C^∞ sur $]-R\,;R[$ (avec $R > 1$) donc au moins dérivable deux fois en 1 avec

$$G'_X(1) = \frac{r(1-p)}{p} \quad \text{et} \quad G''_X(1) = \frac{r(r+1)(1-p)^2}{p^2}.$$

1. Ce résultat a déjà été acquis de façon plus élémentaire dans le sujet 3 p. 386.

On en déduit les valeurs de l'espérance et de la variance de X

$$\mathrm{E}(X) = G'_X(1) = \frac{r(1-p)}{p}$$

$$\mathrm{V}(X) = G''_X(1) + G'_X(1) - \left(G''_X(1)\right)^2 = \frac{r(1-p)}{p^2}.$$

9.8 Exercices d'entraînement

9.8.1 Variables aléatoires

Exercice 11 *

Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et μ. Pour $n \in \mathbb{N}$, identifier la loi de X sachant $(X+Y=n)$.

Solution

Les variables X et Y étant à valeurs dans $\mathbb{N}$, les seules valeurs possibles pour X lorsque $(X+Y=n)$ sont les $k \in [\![0\,;n]\!]$.

méthode

Il s'agit de calculer $\mathrm{P}(X=k \,|\, X+Y=n)$ pour toute valeur $k \in [\![0\,;n]\!]$.

Les variables X et Y sont indépendantes et suivent des lois de Poisson de paramètres λ et μ. Leurs somme $X+Y$ suit alors[1] une loi de Poisson de paramètre $\lambda+\mu$. On sait donc

$$\mathrm{P}(X+Y=n) = \mathrm{e}^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!} \quad \text{pour tout } n \in \mathbb{N}.$$

En particulier, cette probabilité est non nulle et, pour tout $k \in [\![0\,;n]\!]$, la probabilité conditionnelle $\mathrm{P}(X=k \,|\, X+Y=n)$ est définie par

$$\mathrm{P}(X=k \,|\, X+Y=n) = \frac{\mathrm{P}(X=k, X+Y=n)}{\mathrm{P}(X+Y=n)}.$$

Cependant, on a l'égalité d'événements

$$(X=k, X+Y=n) = (X=k, Y=n-k).$$

Par indépendance des variables X et Y,

$$\mathrm{P}(X=k, Y=n-k) = \mathrm{P}(X=k)\,\mathrm{P}(Y=n-k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\mathrm{e}^{-\mu}\frac{\mu^{n-k}}{(n-k)!}.$$

Après simplification,

$$\mathrm{P}(X=k \,|\, X+Y=n) = \frac{n!}{k!(n-k)!}\cdot\frac{\lambda^k\mu^{n-k}}{(\lambda+\mu)^n} = \binom{n}{k}p^k(1-p)^{n-k} \quad \text{avec} \quad p=\frac{\lambda}{\lambda+\mu}.$$

Finalement, on reconnaît une loi binomiale de paramètres n et p.

1. Voir sujet 3 p. 386 et sujet 9 p. 392.

Exercice 12 *

Soit X et Y deux variables aléatoires indépendantes suivant des lois géométriques de paramètres respectifs p et $q \in]0\,;1[$.

Quelle est la probabilité que la matrice réelle suivante soit diagonalisable ?

$$A = \begin{pmatrix} X & -Y \\ Y & -X \end{pmatrix}.$$

Solution

méthode

On traduit l'étude de la diagonalisabilité de A en un événement s'exprimant en fonction des variables X et Y.

La valeur du polynôme caractéristique de la matrice A en $\lambda \in \mathbb{R}$ est

$$\chi_A(\lambda) = \lambda^2 - (X^2 - Y^2).$$

Cas : $(X < Y)$. Le polynôme χ_A n'est pas scindé sur $\mathbb{R}$ et la matrice A n'est pas diagonalisable dans $\mathcal{M}_2(\mathbb{R})$.

Cas : $(X = Y)$. Le polynôme χ_A possède une seule racine 0. La matrice A est diagonalisable si, et seulement si, elle est semblable à la matrice nulle donc égale à la matrice nulle. Ceci est exclu car les variables X et Y prennent leurs valeurs dans $\mathbb{N}^*$.

Cas : $(X > Y)$. Le polynôme χ_A possède deux racines réelles distinctes et la matrice A qui est de taille 2 est donc diagonalisable.

En résumé, l'événement « La matrice A est diagonalisable » correspond à $(X > Y)$.

méthode

On décompose l'événement $(X > Y)$ en une réunion d'événements dont on sait calculer les probabilités.

L'événement $(X > Y)$ est la réunion des événements deux à deux incompatibles $(X > n, Y = n)$ pour n parcourant $n \in \mathbb{N}^*$. Par additivité dénombrable

$$\mathrm{P}(X > Y) = \sum_{n=1}^{+\infty} \mathrm{P}(X > n, Y = n).$$

Par indépendance $\mathrm{P}(X > n, Y = n) = \mathrm{P}(X > n)\,\mathrm{P}(Y = n)$ avec[1] $\mathrm{P}(X > n) = (1-p)^n$. On a donc

$$\mathrm{P}(X > Y) = \sum_{n=1}^{+\infty} (1-p)^n (1-q)^{n-1} q.$$

Après glissement d'indice et calcul d'une somme géométrique de raison $r = (1-p)(1-q)$ avec $r \in]0\,;1[$, on obtient la probabilité que la matrice A soit diagonalisable

$$\mathrm{P}(X > Y) = \frac{(1-p)q}{1-(1-p)(1-q)} = \frac{q-pq}{p+q-pq}.$$

1. L'événement $(X > n)$ correspond à une succession de n échecs, voir sujet 4 p. 387.

Exercice 13 **

Sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$, on considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires telles que, pour tout $n \in \mathbb{N}$, X_n suit une loi binomiale de paramètres n et $p \in]0\,;1[$. On considère aussi une variable aléatoire N indépendante des variables X_n et telle que $N+1$ suit une loi géométrique de paramètre $q \in]0\,;1[$.

Pour toute issue ω de l'univers Ω, on pose $Y(\omega) = X_{N(\omega)}(\omega)$.

Justifier que Y est une variable aléatoire discrète et déterminer sa loi.

On pourra employer l'identité binomiale déjà présentée dans le sujet 8 p. 391.

Solution

méthode

On montre que Y est à valeurs dans un ensemble dénombrable et que, pour chaque valeur y de cet ensemble, $(Y = y)$ constitue un événement.

Pour $n \in \mathbb{N}$, la variable X_n est à valeurs dans $[\![0\,;n]\!] \subset \mathbb{N}$. On peut donc affirmer que la variable Y est à valeurs dans $\mathbb{N}$ qui est évidemment un ensemble dénombrable. Pour tout $k \in \mathbb{N}$ et tout $\omega \in \Omega$,

$$Y(\omega) = k \iff \exists n \in \mathbb{N},\ N(\omega) = n \text{ et } X_n(\omega) = k.$$

On peut donc écrire

$$(Y = k) = \bigcup_{n\in\mathbb{N}} \big((N = n) \cap (X_n = k)\big). \qquad (*)$$

Les ensembles $(N = n)$ et $(X_n = k)$ définissent des événements et donc, par opérations dans la tribu $\mathcal{T}$, $(Y = k)$ est un événement. Ainsi, Y est une variable aléatoire discrète.

Pour déterminer la loi de Y, nous calculons $\mathrm{P}(Y = k)$ pour tout $k \in \mathbb{N}$. La réunion de l'égalité $(*)$ est constituée d'événements deux à deux incompatibles, on a donc par additivité dénombrable

$$\mathrm{P}(Y = k) = \sum_{n=0}^{+\infty} \mathrm{P}(N = n, X_n = k).$$

Par indépendance des variables N et X_n, on peut écrire

$$\mathrm{P}(N = n, X_n = k) = \mathrm{P}(N = n)\,\mathrm{P}(X_n = k).$$

Si $k > n$, cette quantité est nulle car l'événement $(X_n = k)$ est impossible. En simplifiant les termes correspondants de la somme, il reste

$$\mathrm{P}(Y = k) = \sum_{n=k}^{+\infty} q(1-q)^n \binom{n}{k} p^k (1-p)^{n-k} = qp^k(1-q)^k \sum_{n=k}^{+\infty} \binom{n}{k} \big((1-p)(1-q)\big)^{n-k}.$$

Après glissement d'indice, on emploie l'identité binomiale pour achever le calcul

$$\mathrm{P}(Y = k) = qp^k(1-q)^k \sum_{n=0}^{+\infty} \binom{n+k}{k} \big((1-p)(1-q)\big)^n = \frac{qp^k(1-q)^k}{(p+q-pq)^{k+1}}.$$

La variable $Y+1$ suit une loi géométrique de paramètre $r = \frac{q}{p+q-pq}$.

9.8.2 Espérances et variances

Exercice 14 *

Soit X une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$.
Calculer

$$\mathrm{E}\left(\frac{1}{X+1}\right).$$

Solution

méthode

Par la formule de transfert (Th. 17 p. 378), on peut calculer l'espérance d'une variable composée $Y = f(X)$ à partir de la loi de X.

Puisque la variable X suit une loi de Poisson de paramètre λ, elle prend ses valeurs dans $\mathbb{N}$ et

$$\mathrm{P}(X = n) = \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} \quad \text{pour tout } n \in \mathbb{N}.$$

La variable $Y = 1/(X+1)$ étant bornée, elle admet une espérance finie et la formule de transfert donne

$$\mathrm{E}\left(\frac{1}{X+1}\right) = \sum_{n=0}^{+\infty}\frac{1}{n+1}\mathrm{P}(X=n) = \sum_{n=0}^{+\infty}\frac{1}{n+1}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{(n+1)!}.$$

On transforme l'écriture de la somme en dernier membre afin de faire apparaître une somme exponentielle, notamment à l'aide d'un glissement d'indice

$$\sum_{n=0}^{+\infty}\frac{\lambda^n}{(n+1)!} = \frac{1}{\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^{n+1}}{(n+1)!} = \frac{1}{\lambda}\sum_{n=1}^{+\infty}\frac{\lambda^n}{n!} = \frac{1}{\lambda}\left(\mathrm{e}^{\lambda}-1\right).$$

Finalement,

$$\mathrm{E}\left(\frac{1}{X+1}\right) = \frac{1-\mathrm{e}^{-\lambda}}{\lambda}.$$

Exercice 15 *

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On suppose qu'il existe $k \in]0\,;1[$ vérifiant

$$\mathrm{P}(X = n) = k\,\mathrm{P}(X \geqslant n) \quad \text{pour tout } n \in \mathbb{N}.$$

Déterminer la loi de X puis calculer son espérance et sa variance.

Solution

Soit $n \in \mathbb{N}$.

méthode

On exprime $(X = n)$ à l'aide d'événements $(X \geqslant k)$ pour k bien choisis.

L'événement $(X \geqslant n)$ est la réunion des deux événements incompatibles $(X = n)$ et $(X \geqslant n+1)$. Par additivité

$$\mathrm{P}(X \geqslant n) = \mathrm{P}(X = n) + \mathrm{P}(X \geqslant n+1).$$

En multipliant par k cette relation, on obtient

$$\mathrm{P}(X = n) = k\,\mathrm{P}(X = n) + \mathrm{P}(X = n+1)$$

et donc

$$\mathrm{P}(X = n+1) = (1-k)\,\mathrm{P}(X = n).$$

La suite $\big(\mathrm{P}(X = n)\big)_{n\in\mathbb{N}}$ est alors géométrique de raison $1-k$ ce qui permet d'exprimer son terme général à partir de $a = \mathrm{P}(X = 0)$:

$$\mathrm{P}(X = n) = a(1-k)^n \quad \text{pour tout } n \in \mathbb{N}.$$

La somme de toutes les probabilités élémentaires vaut 1, on peut donc déterminer la valeur de a en calculant une somme géométrique de raison $1-k \in]0\,;1[$

$$\mathrm{P}(\Omega) = \sum_{n=0}^{+\infty} \mathrm{P}(X = n) = a\sum_{n=0}^{+\infty}(1-k)^n = \frac{a}{1-(1-k)} = \frac{a}{k}.$$

Ainsi, $a = k$ et la loi de la variable X est donnée par

$$\mathrm{P}(X = n) = k(1-k)^n \quad \text{pour tout } n \in \mathbb{N}.$$

méthode

L'espérance et la variance de X se déduisent des valeurs connues pour les lois géométriques.

La loi de la variable X ressemble à une loi géométrique. Précisément, $Y = 1 + X$ suit une loi géométrique de paramètre k. On sait alors

$$\mathrm{E}(Y) = \frac{1}{k} \quad \text{et} \quad \mathrm{V}(Y) = \frac{1-k}{k^2}.$$

On en déduit

$$\mathrm{E}(X) = \mathrm{E}(Y) - 1 = \frac{1-k}{k} \quad \text{et} \quad \mathrm{V}(X) = \mathrm{V}(Y) = \frac{1-k}{k^2}.$$

Exercice 16 **

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Montrer que X admet une espérance finie si, et seulement si, la série $\sum \mathrm{P}(X > n)$ converge et qu'alors [1]

$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} \mathrm{P}(X > n).$$

Solution

méthode

On décompose $(X > n)$ en une réunion d'événements $(X = k)$ puis on réorganise le calcul de la somme par paquets [2].

L'événement $(X > n)$ est la réunion des événements deux à deux incompatibles $(X = k)$ pour $k = n+1$, $n+2$, etc. Par additivité dénombrable, on peut écrire avec convergence de la série introduite

$$\mathrm{P}(X > n) = \mathrm{P}\left(\bigcup_{k>n} (X = k)\right) = \sum_{k=n+1}^{+\infty} \mathrm{P}(X = k).$$

Ainsi, et sous réserve d'existence,

$$\sum_{n=0}^{+\infty} \mathrm{P}(X > n) = \sum_{n=0}^{+\infty} \left(\sum_{k=n+1}^{+\infty} \mathrm{P}(X = k) \right).$$

Ceci invite à étudier la sommabilité de la famille

$$\big(\mathrm{P}(X = k)\big)_{(k,n)\in I} \quad \text{avec} \quad I = \big\{(k,n) \in \mathbb{N}^2 \;\big|\; k > n\big\}. \tag{Δ}$$

On organise le calcul de cette somme de termes positifs par paquets.

D'une part, I est la réunion disjointe des ensembles [3] $I_n = \big\{(k,n) \;\big|\; k \in \mathbb{N} \text{ et } k > n\big\}$ pour n parcourant $\mathbb{N}$. Par le calcul au-dessus, on est assuré des sommabilités des sous-familles indexées par I_n. Par conséquent, la famille donnée par (Δ) est sommable si, et seulement si, la série $\sum \mathrm{P}(X > n)$ converge et alors

$$\sum_{(k,n)\in I} \mathrm{P}(X = k) = \sum_{n=0}^{+\infty} \left(\sum_{k=n+1}^{+\infty} \mathrm{P}(X = k) \right) = \sum_{n=0}^{+\infty} \mathrm{P}(X > n). \tag{$*$}$$

D'autre part, I est la réunion disjointe des ensembles $J_k = \big\{(k,n) \;\big|\; n \in \mathbb{N} \text{ et } n < k\big\}$ pour $k \in \mathbb{N}^*$. Les sous-familles indexées par les ensembles J_k sont finies, donc sommables, avec

$$\sum_{n\in J_k} \mathrm{P}(X = k) = \sum_{n=0}^{k-1} \underbrace{\mathrm{P}(X = k)}_{=\text{constante}} = k\,\mathrm{P}(X = k).$$

1. Cette identité sera souvent utilisée par la suite.
2. On emploie le Th 10 du chapitre 1 de l'ouvrage *Exercices d'analyse MP*.
3. Dans cette description des éléments de l'ensemble I_n, l'indice n est fixé et seul k varie.

La famille donnée par (Δ) est alors sommable si, et seulement si, la série $\sum k\,\mathrm{P}(X=k)$ converge et alors

$$\sum_{(k,n)\in I}\mathrm{P}(X=k)=\sum_{k=1}^{+\infty}\left(\sum_{n=0}^{k-1}\mathrm{P}(X=k)\right)=\sum_{k=1}^{+\infty}k\,\mathrm{P}(X=k)=\sum_{k=0}^{+\infty}k\,\mathrm{P}(X=k). \qquad (**)$$

Des sommabilités étudiées pour écrire $(*)$ et $(**)$, on conclut que la variable X admet une espérance finie si, et seulement si, la série $\sum \mathrm{P}(X>n)$ converge et [1]

$$\mathrm{E}(X)=\sum_{k=0}^{+\infty}k\,\mathrm{P}(X=k)=\sum_{n=0}^{+\infty}\mathrm{P}(X>n).$$

Exercice 17 **

Soit X et Y deux variables aléatoires indépendantes géométriques de paramètres p et q éléments de $]0\,;1[$. Calculer l'espérance de $Z=\max(X,Y)$.

Solution

méthode

On détermine $\mathrm{P}(Z>n)$ avant de calculer l'espérance de Z par la formule du sujet précédent.

Soit $n\in\mathbb{N}$. On a

$$(Z>n)=(X>n)\cup(Y>n)$$

donc

$$\mathrm{P}(Z>n)=\mathrm{P}(X>n)+\mathrm{P}(Y>n)-\mathrm{P}(X>n,Y>n).$$

Par indépendance des variables X et Y

$$\mathrm{P}(Z>n)=\mathrm{P}(X>n)+\mathrm{P}(Y>n)-\mathrm{P}(X>n)\,\mathrm{P}(Y>n).$$

Enfin, les lois de X et Y étant géométriques, on sait [2]

$$\mathrm{P}(X>n)=(1-p)^n \quad\text{et}\quad \mathrm{P}(Y>n)=(1-q)^n.$$

Ainsi,

$$\mathrm{P}(Z>n)=(1-p)^n+(1-q)^n-\big((1-p)(1-q)\big)^n.$$

Enfin, par calcul [3] de sommes géométriques [4]

$$\mathrm{E}(Z)=\sum_{n=0}^{+\infty}\mathrm{P}(Z>n)=\frac{1}{p}+\frac{1}{q}-\frac{1}{p+q-pq}.$$

1. La variable X étant à valeurs positives, si elle n'est pas d'espérance finie, cette espérance vaut $+\infty$ ce qui correspond à la limite des sommes partielles de la série exprimant le second membre.

2. L'événement $(X>n)$ correspond à une succession de n échecs, voir sujet 4 p. 387.

3. On peut aussi dire que la somme des $(1-r)^n$ se comprend comme le calcul de l'espérance d'une loi géométrique de paramètre r que l'on sait valoir $1/r$.

4. On a $\max(X,Y)=X+Y-\min(X,Y)$. Si l'on sait que $\min(X,Y)$ suit une loi géométrique de paramètre $p+q-pq$ (voir de nouveau le sujet 4 p. 387) on peut retrouver ce résultat plus rapidement.

Exercice 18 **

Soit $N \in \mathbb{N}^*$ et $(U_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi uniforme sur $[\![1\,;N]\!]$. Pour tout $n \in \mathbb{N}^*$, on pose

$$M_n = \max(U_1, \ldots, U_n).$$

Déterminer les limites de $\mathrm{E}(M_n)$ et de $\mathrm{V}(M_n)$ lorsque n tend vers $+\infty$.

Solution

Pour tout $n \in \mathbb{N}^*$, M_n se déduit des variables $U_1, \ldots, U_n$ par composition avec la fonction max, il s'agit bien d'une variable aléatoire. De plus, les variables U_n sont à valeurs dans $[\![1\,;N]\!]$ donc M_n aussi.

méthode

La variable M_n étant définie par un max, on détermine la loi de M_n en calculant $\mathrm{P}(M_n \leqslant k)$ pour $k \in [\![1\,;N]\!]$.

Soit $k \in [\![1\,;N]\!]$. L'événement $(M_n \leqslant k)$ est la conjonction des événements $(U_i \leqslant k)$ pour i allant de 1 à n. Par indépendance des variables $U_1, \ldots, U_n$, il vient

$$\mathrm{P}(M_n \leqslant k) = \mathrm{P}\left(\bigcap_{i=1}^{n}(U_i \leqslant k)\right) = \prod_{i=1}^{n}\mathrm{P}(U_i \leqslant k) = \left(\frac{k}{N}\right)^n.$$

Notons que cette formule est aussi valable pour $k = 0$.

Soit $k \in [\![1\,;n]\!]$. L'événement $(M_n = k)$ se déduit de $(M_n \leqslant k)$ et $(M_n \leqslant k-1)$ de sorte que

$$\mathrm{P}(M_n = k) = \mathrm{P}(M_n \leqslant k) - \mathrm{P}(M_n \leqslant k-1) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n.$$

On peut alors exprimer l'espérance de M_n

$$\mathrm{E}(M_n) = \sum_{k=1}^{N} k\left(\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n\right).$$

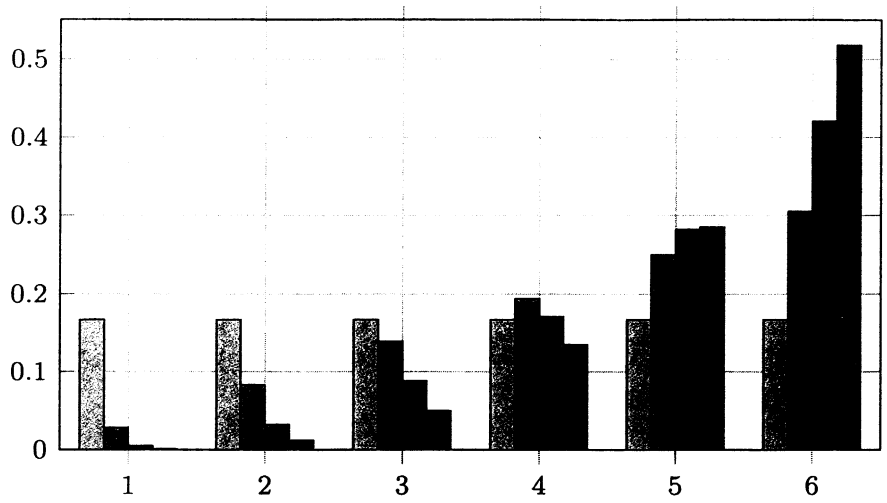
Il n'est pas nécessaire d'exprimer plus simplement [1] cette somme pour en déterminer la limite. En effet, celle-ci comporte un nombre N constant de termes tous de limites nulles sauf le dernier qui tend vers N. On en déduit que $\mathrm{E}(M_n)$ tend vers N quand n tend vers $+\infty$.

À l'aide de la formule de Huygens, la valeur de $\mathrm{V}(M_n)$ se déduit de celle de $\mathrm{E}(M_n^2)$ avec, par des arguments analogues à ceux ci-dessus,

$$\mathrm{E}(M_n^2) = \sum_{k=1}^{N} k^2\left(\left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n\right) \xrightarrow[n\to+\infty]{} N^2.$$

1. En séparant la somme en deux puis en s'aidant d'un glissement d'indice ou, plus rapidement, en employant la formule du sujet 16 p. 399, on obtient l'expression $\mathrm{E}(M_n) = N - \sum_{k=1}^{N-1}(k/N)^n$.

On en déduit que la variance de M_n tend vers 0 quand n tend vers $+\infty$.



Les lois de M_n quand $N = 6$ pour $n = 1, 2, 3, 4$.

Exercice 19 * (Fonction génératrice des moments)**

Soit X une variable aléatoire discrète réelle. On note I_X l'ensemble des $t \in \mathbb{R}$ pour lesquels la variable e^{tX} admet une espérance finie et l'on pose

$$M_X(t) = \mathrm{E}(e^{tX}) \quad \text{pour tout } t \in I_X.$$

(a) Montrer que I_X est un intervalle contenant 0.

(b) On suppose que 0 est intérieur à l'intervalle I_X. Montrer que la variable X admet des moments à tout ordre et que, sur un intervalle centré en 0,

$$M_X(t) = \sum_{n=0}^{+\infty} \frac{\mathrm{E}(X^n)}{n!} t^n.$$

Solution

(a) Il est entendu que 0 est élément de I_X car e^{0X} est la variable constante égale à 1. On peut même écrire $M_X(0) = \mathrm{E}(1) = 1$. Reste à montrer que I_X est un intervalle.

méthode

‖ On vérifie $[a\,;b] \subset I_X$ pour tous a et $b \in I_X$ tels que $a < b$.

Soit $a, b \in I_x$ avec $a < b$ (si de tels éléments existent). Introduisons t élément de $[a\,;b]$.

méthode

‖ On majore e^{tx} selon le signe du réel x.

Si x est positif, $e^{tx} \leqslant e^{bx}$. Si x est négatif, $e^{tx} \leqslant e^{ax}$. Dans les deux cas, on peut écrire $e^{tx} \leqslant e^{ax} + e^{bx}$ et l'on a donc la comparaison de variables aléatoires positives

$$e^{tX} \leqslant e^{aX} + e^{bX}.$$

Or la variable aléatoire en second membre admet une espérance finie et donc, par domination (Th. 16 p. 378), la variable e^{tX} aussi. Ainsi, $t \in I_X$ et l'on peut conclure que I_X est un intervalle.

(b) **méthode**

‖ On réorganise le calcul à l'aide d'une sommation par paquets.

Notons E l'ensemble (au plus dénombrable) des valeurs prises par la variable X. Pour t dans I_X, on peut écrire par la formule de transfert

$$\mathrm{E}(e^{tX}) = \sum_{x \in E} e^{tx} \, \mathrm{P}(X = x)$$

avec sommabilité de la famille des $e^{tx} \, \mathrm{P}(X = x)$ pour x parcourant E. En décomposant le terme exponentiel, on écrit encore

$$\mathrm{E}(e^{tX}) = \sum_{x \in E} \left(\sum_{n=0}^{+\infty} \frac{t^n x^n}{n!} \, \mathrm{P}(X = x) \right)$$

ce qui invite à considérer la famille doublement indexée

$$\left(\frac{t^n x^n}{n!} \, \mathrm{P}(X = x) \right)_{(x,n) \in E \times \mathbb{N}} \qquad (*)$$

Cependant, pour réorganiser le calcul de la somme à l'aide d'une sommation par paquets, il faut savoir si cette famille est sommable ce qui nécessite d'étudier ses termes en valeurs absolues.

Soit α un réel strictement positif tel que α et $-\alpha$ appartiennent à I_X. Les variables $e^{\alpha X}$ et $e^{-\alpha X}$ admettent chacune une espérance finie et donc la variable $e^{|tX|}$ aussi pour tout t de $[-\alpha\,;\alpha]$ car on a la domination

$$e^{|tX|} \leqslant e^{\alpha X} + e^{-\alpha X}.$$

Pour cette valeur de t, on a par des calculs analogues aux précédents la sommabilité de la famille de termes positifs

$$\left(\frac{|t|^n \, |x|^n}{n!} \, \mathrm{P}(X = x) \right)_{(x,n) \in E \times \mathbb{N}}$$

Pour $n \in \mathbb{N}$ fixé, on peut alors affirmer la sommabilité de la sous-famille

$$\left(\frac{|t|^n \, |x|^n}{n!} \, \mathrm{P}(X = x) \right)_{x \in E}$$

En choisissant t non nul dans $[-\alpha\,;\alpha]$, les facteurs $|t|^n$ et $n!$ apparaissent comme des constantes non nulles dans la description de cette famille et l'on peut donc affirmer la

sommabilité de la famille $\left(|x|^n \mathrm{P}(X=x)\right)_{x\in E}$. Ainsi, la variable X admet un moment d'ordre n.

De plus, pour tout $t \in [-\alpha\,;\alpha]$, on peut calculer la somme de la famille décrite dans $(*)$ par deux organisations par paquets qui donnent

$$\sum_{(n,x)\in\mathbb{N}\times E} \frac{t^n x^n}{n!}\mathrm{P}(X=x) = \sum_{n=0}^{+\infty}\left(\sum_{x\in E}\frac{t^n x^n}{n!}\mathrm{P}(X=x)\right) = \sum_{n=0}^{+\infty}\frac{\mathrm{E}(X^n)}{n!}t^n$$

et

$$\sum_{(n,x)\in\mathbb{N}\times E} \frac{t^n x^n}{n!}\mathrm{P}(X=x) = \sum_{x\in E}\left(\sum_{n=0}^{+\infty}\frac{t^n x^n}{n!}\mathrm{P}(X=x)\right) = \sum_{x\in E}\mathrm{e}^{tx}\mathrm{P}(X=x) = \mathrm{E}\left(\mathrm{e}^{tX}\right).$$

On en déduit la relation voulue.

9.8.3 Calcul de loi

Dans chacun des sujets qui suit, on admet l'existence d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ qui permet son étude.

Exercice 20 *

Le nombre quotidien de clients entrant dans une boulangerie suit une loi de Poisson de paramètre $\lambda > 0$. Chaque client a la probabilité $p \in]0\,;1[$ d'acheter des croissants. Sur une journée, on note X_1 le nombre de clients ayant acheté des croissants et X_2 le nombre de ceux qui n'en ont pas achetés.

(a) Déterminer la loi de X_1.

(b) Calculer la covariance de X_1 et X_2.

(c) Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

Solution

(a) Notons X le nombre de clients entrant dans la boulangerie

$$\mathrm{P}(X=n) = \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} \quad \text{pour tout } n \in \mathbb{N}.$$

Lorsque $(X=n)$, la variable aléatoire X_1 suit une loi binomiale de paramètres n et p (le nombre de clients achetant un croissant lorsque n clients entrent dans la boulangerie peut se comprendre comme le nombre de succès dans une série de n épreuves de Bernoulli indépendantes et de même paramètre p).

méthode

La loi de la variable X_1 se déduit de la loi de X et des lois conditionnelles de X_1 connaissant la valeur prise par X (Th. 9 p. 375).

La variable X_1 est à valeurs dans $\mathbb{N}$ et, pour tout $k \in \mathbb{N}$,

$$\mathrm{P}(X_1 = k) = \sum_{n=0}^{+\infty} \mathrm{P}(X_1 = k \,|\, X = n)\,\mathrm{P}(X = n).$$

Lorsque $n < k$, la probabilité $\mathrm{P}(X_1 = k \,|\, X = n)$ est nulle, on peut donc simplifier les premiers termes de la somme et écrire

$$\mathrm{P}(X_1 = k) = \sum_{n=k}^{+\infty} \mathrm{P}(X_1 = k \,|\, X = n)\,\mathrm{P}(X = n) = \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!}.$$

On exprime le coefficient binomial à l'aide de nombres factoriels puis on opère un glissement d'indice

$$\mathrm{P}(X_1 = k) = \mathrm{e}^{-\lambda} \frac{(p\lambda)^k}{k!} \sum_{n=k}^{+\infty} \frac{1}{(n-k)!} (1-p)^{n-k} \lambda^{n-k} = \mathrm{e}^{-\lambda} \frac{(p\lambda)^k}{k!} \sum_{n=0}^{+\infty} \frac{1}{n!} (1-p)^n \lambda^n.$$

Enfin, on reconnaît une somme exponentielle calculée en $(1-p)\lambda$

$$\mathrm{P}(X_1 = k) = \mathrm{e}^{-\lambda} \frac{(p\lambda)^k}{k!} \mathrm{e}^{(1-p)\lambda} = \mathrm{e}^{-p\lambda} \frac{(p\lambda)^k}{k!}.$$

Ainsi, X_1 suit une loi de Poisson de paramètre[1] $p\lambda$.

(b) **méthode**

La covariance de X_1 et X_2 se déduit de la variance de $X_1 + X_2$ (Th. 23 p. 381).

Un calcul analogue au précédent (ou simplement un argument de symétrie) assure que X_2 suit une loi de Poisson de paramètre $(1-p)\lambda$. Les lois X_1 et X_2 admettant chacune un moment d'ordre 2, on a

$$\mathrm{V}(X_1 + X_2) = \mathrm{V}(X_1) + 2\,\mathrm{Cov}(X_1, X_2) + \mathrm{V}(X_2).$$

Puisque la variance d'une loi de Poisson de paramètre λ est égale à λ, on obtient

$$\mathrm{V}(X_1 + X_2) = \mathrm{V}(X) = \lambda \quad \text{et} \quad \mathrm{V}(X_1) + \mathrm{V}(X_2) = p\lambda + (1-p)\lambda = \lambda.$$

On en déduit que la covariance de X_1 et X_2 est nulle.

(c) **méthode**

La nullité de la covariance est une condition nécessaire mais pas suffisante pour affirmer que deux variables sont indépendantes : on étudie plutôt si

$$\mathrm{P}(X_1 = k, X_2 = \ell) = \mathrm{P}(X_1 = k)\,\mathrm{P}(X_2 = \ell).$$

1. Il s'achète en moyenne $p\lambda$ croissants par jour ce qui est conforme à l'intuition.

Soit $(k,\ell) \in \mathbb{N}^2$. Posons $k+\ell = n$. L'événement $(X_1 = k, X_2 = \ell)$ se confond avec $(X_1 = k, X = n)$. Par la formule des probabilités composées

$$\begin{aligned}\mathrm{P}(X_1 = k, X_2 = \ell) &= \mathrm{P}(X_1 = k \,|\, X = n)\,\mathrm{P}(X = n)\\ &= \binom{n}{k} p^k (1-p)^{n-k} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda} \frac{p^k (1-p)^{n-k} \lambda^n}{k!(n-k)!}.\end{aligned}$$

Parallèlement,

$$\mathrm{P}(X_1 = k)\,\mathrm{P}(X_2 = \ell) = \mathrm{e}^{-p\lambda} \frac{(p\lambda)^k}{k!} \mathrm{e}^{-(1-p)\lambda} \frac{\big((1-p)\lambda\big)^\ell}{\ell!} = \mathrm{e}^{-\lambda} \frac{p^k (1-p)^\ell \lambda^{k+\ell}}{k!\ell!}$$

et donc [1]

$$\mathrm{P}(X_1 = k, X_2 = \ell) = \mathrm{P}(X_1 = k)\,\mathrm{P}(X_2 = \ell).$$

Ceci valant pour tout $(k,\ell) \in \mathbb{N}^2$, on peut affirmer que les variables X_1 et X_2 sont indépendantes : savoir combien de clients ont acheté des croissants ne donne pas « d'informations » sur le nombre de clients n'en ayant pas achetés !

Exercice 21 **

Un joueur dispose de N dés équilibrés. Il lance une première fois ceux-ci et met de côté les dés ayant donné un 'six'. S'il en reste, les autres dés sont relancés et l'on répète l'expérience jusqu'à ce que tous les dés aient donné un 'six'. On introduit la variable aléatoire T à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$ donnant le nombre de lancers nécessaires.

(a) Soit $n \in \mathbb{N}^*$. Calculer la probabilité de $(T \leqslant n)$.

(b) Justifier que l'expérience s'arrête presque sûrement.

(c) Vérifier que la variable T admet une espérance finie et donner une formule exprimant celle-ci.

Solution

(a) Distinguons chacun des dés et notons $X_1, \ldots, X_N$ les variables aléatoires donnant le nombre de lancers nécessaires avant que le dé correspondant ne donne un 'six'. Chaque variable X_i suit une loi géométrique de paramètre $p = 1/6$.

méthode

La variable aléatoire T s'exprime comme le maximum des $X_1, \ldots, X_N$.

Soit $n \in \mathbb{N}^*$. L'événement $(T \leqslant n)$ est l'intersection de tous les événements $(X_i \leqslant n)$ pour $i = 1, \ldots, N$. L'énoncé suppose implicitement l'indépendance des différents résultats des dés et donc

$$\mathrm{P}(T \leqslant n) = \mathrm{P}\left(\bigcap_{i=1}^{n} (X_i \leqslant n)\right) = \prod_{i=1}^{n} \mathrm{P}(X_i \leqslant n)$$

1. On a ici mené le calcul inverse de celui du sujet 11 p. 394.

avec $\mathrm{P}(X_i \leqslant n) = 1 - \mathrm{P}(X_i > n) = 1 - (1-p)^n$ car $(X_i > n)$ signifie que les n premiers lancers du dé d'indice i n'ont pas donné de 'six'. On a donc

$$\mathrm{P}(T \leqslant n) = \left(1 - (1-p)^n\right)^N = \left(\frac{6^n - 5^n}{6^n}\right)^N.$$

(b) L'événement $\overline{(T = +\infty)}$ traduit que l'expérience s'arrête. Celui-ci est la réunion pour $n \in \mathbb{N}^*$ des événements $(T \leqslant n)$ qui forment une suite croissante. Par continuité monotone

$$\mathrm{P}\left(\overline{(T = +\infty)}\right) = \lim_{n \to +\infty} \mathrm{P}(T \leqslant n) = \lim_{n \to +\infty} \left(1 - (1-p)^n\right)^N = 1 \quad \text{car} \quad p > 0.$$

(c) **méthode**

On calcule l'espérance de T par la formule[1]

$$\mathrm{E}(T) = \sum_{n=0}^{+\infty} \mathrm{P}(T > n).$$

Soit $n \in \mathbb{N}^*$. Par passage à l'événement contraire

$$\mathrm{P}(T > n) = 1 - \mathrm{P}(T \leqslant n) = 1 - \left(1 - (1-p)^n\right)^N.$$

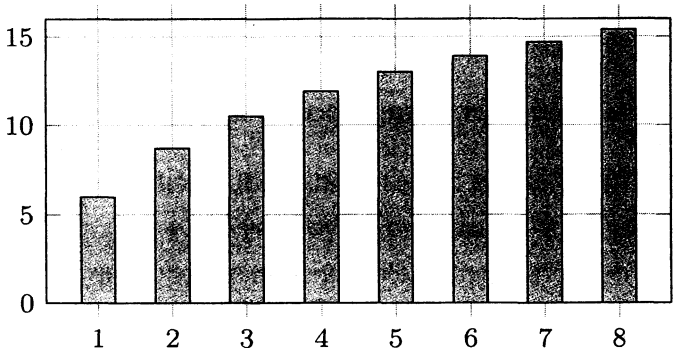
Cette formule est aussi valable lorsque $n = 0$ et l'on a donc à l'aide d'un développement par la formule du binôme

$$\mathrm{E}(T) = \sum_{n=0}^{+\infty} \left(1 - \left(1 - (1-p)^n\right)^N\right) = \sum_{k=1}^{N} \binom{N}{k} (-1)^{k-1} \sum_{n=0}^{+\infty} (1-p)^{nk}.$$

La dernière somme est géométrique de raison $(1-p)^k \in]0\,;1[$ et l'on conclut :

$$\mathrm{E}(T) = \sum_{k=1}^{N} (-1)^{k-1} \binom{N}{k} \frac{1}{1 - (1-p)^k} = \sum_{k=1}^{N} (-1)^{k-1} \binom{N}{k} \frac{6^k}{6^k - 5^k}.$$

Espérance de T en fonction de N



1. Voir sujet 16 p. 399.

Exercice 22 ** (Problème du collectionneur)

Chez un marchand de journaux, on peut acheter des pochettes contenant chacune une image. La collection complète comporte N images distinctes.

(a) Montrer qu'il est presque sûr d'obtenir la collection complète en un nombre fini d'achats.

On limite l'étude à un espace probabilisé dans lequel la collection complète est toujours obtenue en un nombre fini d'achats. Pour $k \in [\![1\,;N]\!]$, on note $X_k \in \mathbb{N}^*$ le nombre d'achats ayant permis d'obtenir k images distinctes. En particulier, $X_1 = 1$ et X_N est le nombre d'achats nécessaires à l'obtention de la collection complète.

(b) Soit $k \in [\![1\,;N-1]\!]$. Par quelle loi peut-on modéliser la variable $X_{k+1} - X_k$?

(c) En déduire une expression de l'espérance de X_N.

Solution

(a) Numérotons les images de 1 à N et considérons, pour $i \in [\![1\,;N]\!]$, l'événement :

$A_i =$ « L'acheteur n'acquiert pas l'image d'indice i en un nombre fini d'achats ».

La probabilité que l'acheteur n'obtienne pas l'image d'indice i en n achats vaut $\left(\frac{N-1}{N}\right)^n$. Par continuité décroissante,

$$\mathrm{P}(A_i) = \lim_{n\to+\infty}\left(\frac{N-1}{N}\right)^n = 0.$$

L'événement A_i est donc négligeable.

L'événement « Ne pas obtenir la collection complète en un nombre fini d'achats » correspond à la réunion de tous les A_i, il est négligeable en tant que réunion finie d'événements négligeables.

(b) **méthode**

> Lorsque k images distinctes sur N ont été acquises, obtenir une nouvelle image se comprend comme le temps d'attente d'un succès dans une suite d'épreuves de Bernoulli de paramètre $(N-k)/N$.

Si cette interprétation de *bon sens* est sans doute suffisante pour affirmer que la variable $X_{k+1} - X_k$ suit une loi géométrique de paramètre $p = (N-k)/N$, nous allons proposer une justification détaillée de celle-ci. Pour $n \in \mathbb{N}^*$, on introduit l'événement :

$$B_n = \text{« On obtient une image nouvelle lors du } n\text{-ième achat »}.$$

La difficulté de l'étude réside dans la non indépendance de ces différents événements.

Soit $k \in [\![1\,;N-1]\!]$ et $m \in \mathbb{N}^*$. Calculons $\mathrm{P}(X_{k+1}-X_k = m)$. La famille $\big((X_k = n)\big)_{n\geqslant k}$ est un système complet d'événements. Par la formule des probabilités totales

$$\mathrm{P}(X_{k+1} - X_k = m) = \sum_{n=k}^{+\infty} \mathrm{P}(X_{k+1} - X_k = m \,|\, X_k = n)\,\mathrm{P}(X_k = n). \qquad (*)$$

Or, pour $n \geqslant k$,

$$\mathrm{P}(X_{k+1} - X_k = m \,|\, X_k = n) = \mathrm{P}(\overline{B_{n+1}} \cap \ldots \cap \overline{B_{n+m-1}} \cap B_{n+m} \,|\, X_k = n).$$

Sachant $(X_k = n)$, la probabilité de $\overline{B_{n+1}}$ vaut k/N car l'acheteur possède k images sur N lors du $(n+1)$-ième achat. Sachant $(X_k = n)$ et $\overline{B_{n+1}}$, la probabilité de $\overline{B_{n+2}}$ est identique car la configuration est inchangée lors du $(n+2)$-ième achat. On poursuit ce raisonnement jusqu'au calcul de $\mathrm{P}(B_{n+m})$ qui vaut $(N-k)/N$ car il s'agit maintenant d'obtenir une image nouvelle. Par la formule des probabilités composées, on obtient

$$\mathrm{P}(X_{k+1} - X_k = m \,|\, X_k = n) = (1-p)^{m-1}p \quad \text{avec} \quad p = \frac{N-k}{N}.$$

L'égalité $(*)$ donne alors

$$\mathrm{P}(X_{k+1} - X_k = m) = (1-p)^{m-1}p \underbrace{\sum_{n=k}^{+\infty} \mathrm{P}(X_k = n)}_{=1} = (1-p)^{m-1}p.$$

Finalement, $X_{k+1} - X_k$ suit bien une loi géométrique de paramètre p.

(c) **méthode**

‖ On calcule $\mathrm{E}(X_N)$ par linéarité de l'espérance et calcul télescopique.

On peut écrire $X_N = X_1 + (X_2 - X_1) + \cdots + (X_N - X_{N-1})$. Par linéarité de l'espérance, on conclut [1]

$$\mathrm{E}(X_N) = \mathrm{E}(X_1) + \sum_{k=1}^{N-1} \mathrm{E}(X_{k+1} - X_k) = 1 + \sum_{k=1}^{N-1} \frac{N}{N-k} = N \sum_{k=1}^{N} \frac{1}{k}.$$

Exercice 23 * (Loi de Pascal)**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi de Bernoulli de paramètre $p \in]0;1[$.

Pour $r \in \mathbb{N}^*$, on définit une variable aléatoire T_r à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$ en posant

$$T_r = \min\Big(\big\{n \in \mathbb{N}^* \;\big|\; X_1 + \cdots + X_n = r\big\} \cup \{+\infty\}\Big).$$

La variable T_r se comprend comme le temps d'attente du r-ième succès [2].

(a) Pour $n \in \mathbb{N}^*$, calculer $\mathrm{P}(T_r = n)$.

(b) Montrer que l'événement $(T_r = +\infty)$ est négligeable.

(c) Soit $Y_1, \ldots, Y_r$ des variables aléatoires indépendantes suivant chacune une loi géométrique de paramètre p. Montrer que $Y_1 + \cdots + Y_r$ et T_r suivent la même loi.

(d) En déduire l'espérance et la variance de T_r

1. Asymptotiquement, celle-ci est voisine de $N \ln N$.
2. Cette variable est directement liée à la variable X du sujet 8 p. 391 : $T_r = X + r$.

Solution

(a) Soit $n \in \mathbb{N}^*$. Les variables X_n prenant les valeurs 0 ou 1, l'événement $(T_r = n)$ est réalisé lorsque $(X_1 + \cdots + X_n = r)$ et $(X_n = 1)$. Or

$$(X_1 + \cdots + X_n = r) \cap (X_n = 1) = (X_1 + \cdots + X_{n-1} = r - 1) \cap (X_n = 1).$$

Par l'indépendance mutuelle des variables de la suite $(X_n)_{n\in\mathbb{N}^*}$, on a donc

$$\mathrm{P}(T_r = n) = \mathrm{P}(X_1 + \cdots + X_{n-1} = r - 1)\,\mathrm{P}(X_n = 1).$$

Enfin, la variable $X_1 + \cdots + X_{n-1}$ suit une loi binomiale de paramètres $n - 1$ et p en tant que somme de variables de Bernoulli mutuellement indépendantes.

Cas : $n < r$. On obtient $\mathrm{P}(T_r = n) = 0$.

Cas : $n \geqslant r$. On obtient [1]

$$\mathrm{P}(T_r = n) = \binom{n-1}{r-1} p^{r-1}(1-p)^{(n-1)-(r-1)} \times p = \binom{n-1}{r-1} p^r (1-p)^{n-r}.$$

(b) Calculons [2] la probabilité de $(T_r = +\infty)$. Par continuité décroissante

$$\mathrm{P}(T_r = +\infty) = \lim_{n\to+\infty} \mathrm{P}(T_r > n).$$

L'événement $(T_r > n)$ signifie que l'on n'a pas obtenu r succès lors des n premières expériences et donc

$$\mathrm{P}(T_r > n) = \mathrm{P}(X_1 + \cdots + X_n < r) = \sum_{j=0}^{r-1} \binom{n}{j} p^j (1-p)^{n-j}$$

car $X_1 + \cdots + X_n$ suit une loi binomiale de paramètres n et p. Or par croissances comparées

$$\binom{n}{j} p^j (1-p)^{n-j} \xrightarrow[n\to+\infty]{} 0.$$

En effet,

$$\binom{n}{j} = \frac{\overbrace{n(n-1) \times \cdots \times (n-j+1)}^{j \text{ facteurs équivalents à } n}}{j!} \underset{n\to+\infty}{\sim} \frac{n^j}{j!}$$

donc

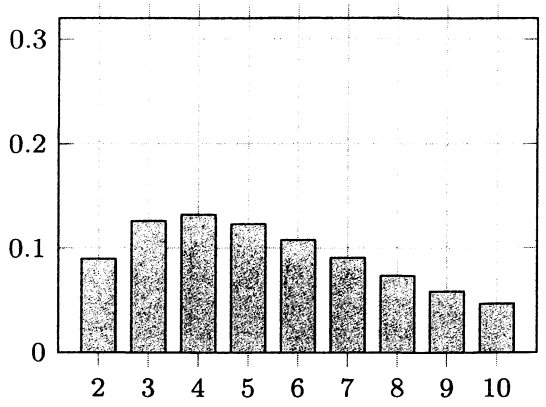
$$\binom{n}{j} p^j (1-p)^{n-j} \underset{n\to+\infty}{\sim} C^{te} n^j (1-p)^n \xrightarrow[n\to+\infty]{} 0.$$

On obtient donc $\mathrm{P}(T_r = +\infty) = 0$: il est presque sûr d'obtenir r succès.

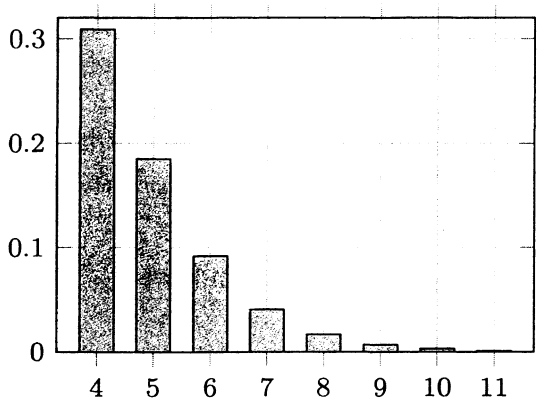
1. Chaque séquence comportant r succès et $n-r$ échecs est de probabilité $p^r(1-p)^{n-r}$ et il y a $\binom{n-1}{r-1}$ séquences qui se terminent par un succès.

2. On peut aussi inclure l'événement $(T_r = +\infty)$ dans celui exprimant que les variables X_n sont toutes nulles au delà d'un certain rang : $(T_r = +\infty)$ se comprend alors comme inclus dans une réunion dénombrable d'événements négligeables.

Loi de T_k pour $p = 0{,}5$ et $k = 2$



Loi de T_k pour $p = 0{,}7$ et $k = 3$



(c) **méthode**

‖ On étudie[1] la fonction génératrice de $Y_1 + \cdots + Y_r$.

Les variables $Y_1, \ldots, Y_r$ étant indépendantes, la fonction génératrice de leur somme est le produit de leurs fonctions génératrices (Th. 29 p. 383) et donc

$$G_{Y_1+\cdots+Y_r}(t) = \left(\frac{pt}{1-(1-p)t}\right)^r \quad \text{pour tout } t \in [-1\,;1].$$

On calcule la loi de $Y_1 + \cdots + Y_r$ en développant sa fonction génératrice en série entière. On part du développement géométrique

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{pour tout } x \in \,]-1\,;1[.$$

On dérive celui-ci à l'ordre $r-1$ en exprimant les termes sommés selon x^{n-r} par anticipation des calculs qui suivent :

$$\frac{(r-1)!}{(1-x)^r} = \sum_{n=r}^{+\infty} \frac{(n-1)!}{(n-r)!} x^{n-r} \quad \text{pour tout } x \in \,]-1\,;1[.$$

En employant cette identité[2] pour $x = (1-p)t$ et multipliant par $p^r t^r$, il vient

$$\left(\frac{pt}{1-(1-p)t}\right)^r = \sum_{n=r}^{+\infty} \binom{n-1}{r-1} p^r (1-p)^{n-r} t^n \quad \text{pour tout } t \in [-1\,;1].$$

Dans cette expression, le coefficient de t^n correspond à la probabilité de $(T_r = n)$ et l'on peut affirmer que les variables $Y_1 + \cdots + Y_r$ et T_r suivent essentiellement[3] la même loi.

1. On peut aussi raisonner par récurrence sur r.

2. On retrouve ici l'identité binomiale (voir sujet 8 p. 391) avec des notations contextualisées. On pouvait aussi employer le développement en série entière de $(1+u)^\alpha$ avec $\alpha = -r$ et $u = (1-p)t$.

3. La variable T_r peut prendre la valeur $+\infty$ ce que ne fait pas a *priori* $Y_1 + \cdots + Y_r$. Cependant, cet événement est négligeable et l'ignorer est donc sans incidence sur la suite des calculs.

(d) **méthode**

Espérance et variance d'une variable aléatoire se calculent à partir de sa loi : les espérances et variances de T_r et $Y_1 + \cdots + Y_r$ sont donc les mêmes.

On sait qu'une variable géométrique de paramètre p est d'espérance $1/p$ et de variance $(1-p)/p^2$. Par linéarité de l'espérance, on a directement

$$\mathrm{E}(T_r) = \mathrm{E}(Y_1 + \cdots + Y_r) = \mathrm{E}(Y_1) + \cdots + \mathrm{E}(Y_r) = \frac{r}{p}.$$

De plus, on peut calculer la variance de la somme par indépendance des variables

$$\mathrm{V}(T_r) = \mathrm{V}(Y_1 + \cdots + Y_r) = \mathrm{V}(Y_1) + \cdots + V(Y_r) = r\frac{1-p}{p^2}.$$

9.8.4 Inégalités de concentration

Exercice 24 *

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que, pour tout $n \in \mathbb{N}$, X_n suit une loi binomiale de paramètres n et $p \in]0\,;1[$. Soit $k \in \mathbb{N}$. Établir

$$\mathrm{P}(X_n \leqslant k) \xrightarrow[n\to+\infty]{} 0.$$

Solution

méthode

On estime $\mathrm{P}(X_n \leqslant k)$ à l'aide de l'inégalité [1] de Bienaymé-Tchebychev (Th. 26 p. 381).

La variable X_n est d'espérance np et de variance $np(1-p)$. Pour tout $\alpha > 0$, l'inégalité de Bienaymé-Tchebychev donne

$$\mathrm{P}\big(|X_n - np| \geqslant \alpha\big) \leqslant \frac{np(1-p)}{\alpha^2}.$$

Choisissons $\alpha > 0$ de sorte que $(X_n \leqslant k) \subset (|X_n - np| \geqslant \alpha)$. Pour n assez grand, la valeur $\alpha = np - k$ convient et l'on a alors

$$\mathrm{P}\big(X_n \leqslant k\big) \leqslant \mathrm{P}\big(|X_n - np| \geqslant \alpha\big) \leqslant \frac{np(1-p)}{(np-k)^2} \underset{n\to+\infty}{\sim} \frac{1-p}{np} \xrightarrow[n\to+\infty]{} 0.$$

1. On peut aussi mener un calcul direct analogue à celui réalisé dans le sujet 23 p. 409.

Exercice 25 ** (Inégalité de Chernoff)

Soit Y une variable aléatoire discrète prenant ses valeurs dans $[\alpha\,;\beta]$ avec $\alpha < \beta$. On suppose que la variable Y est d'espérance nulle.

(a) Soit s un réel. Montrer que

$$(\beta-\alpha)\mathrm{e}^{sy} \leqslant (\beta-y)\mathrm{e}^{s\alpha} + (y-\alpha)\mathrm{e}^{s\beta} \quad \text{pour tout } y \in [\alpha\,;\beta].$$

En déduire $(\beta-\alpha)\,\mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \beta\mathrm{e}^{s\alpha} - \alpha\mathrm{e}^{s\beta}$.

(b) Montrer à l'aide d'une étude de fonction

$$\ln\big(\beta\mathrm{e}^{s\alpha} - \alpha\mathrm{e}^{s\beta}\big) \leqslant \frac{1}{8}(\beta-\alpha)^2 s^2 + \ln(\beta-\alpha) \quad \text{pour tout } s \geqslant 0.$$

En déduire

$$\mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \exp\left(\frac{(\beta-\alpha)^2}{8}s^2\right).$$

Soit $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes prenant leurs valeurs dans $[a\,;b]$ avec $a < b$. On pose S la somme de ces variables et l'on introduit t un réel strictement positif.

(c) Montrer que pour tout $s > 0$

$$\mathrm{P}\big(S - \mathrm{E}(S) \geqslant t\big) \leqslant \mathrm{e}^{-st} \prod_{i=1}^{n} \mathrm{E}\Big(\mathrm{e}^{s\left(X_i - \mathrm{E}(X_i)\right)}\Big) \leqslant \exp\left(-st + n\frac{(b-a)^2}{8}s^2\right).$$

(d) En déduire l'inégalité de Chernoff

$$\mathrm{P}\big(S - \mathrm{E}(S) \geqslant t\big) \leqslant \exp\left(-\frac{2t^2}{n(b-a)^2}\right).$$

Solution

(a) **méthode**

‖ On emploie la convexité de la fonction $x \mapsto \mathrm{e}^{sx}$.

Pour $y \in [\alpha\,;\beta]$, on peut écrire $y = (1-\lambda)\alpha + \lambda\beta$ avec

$$\lambda = \frac{y-\alpha}{\beta-\alpha} \in [0\,;1].$$

L'inégalité de convexité $\mathrm{e}^{s\left((1-\lambda)\alpha+\lambda\beta\right)} \leqslant (1-\lambda)\mathrm{e}^{s\alpha} + \lambda\mathrm{e}^{s\beta}$ donne alors

$$\mathrm{e}^{sy} \leqslant \frac{\beta-y}{\beta-\alpha}\mathrm{e}^{s\alpha} + \frac{y-\alpha}{\beta-\alpha}\mathrm{e}^{s\beta}$$

ce qui conduit à l'inégalité demandée en multipliant par $\beta - \alpha > 0$.

La variable aléatoire Y prenant ses valeurs dans $[\alpha\,;\beta]$, on a la comparaison de variables aléatoires

$$(\beta-\alpha)\mathrm{e}^{sY} \leqslant (\beta-Y)\mathrm{e}^{s\alpha}+(Y-\alpha)\mathrm{e}^{s\beta}.$$

Par croissance et linéarité de l'espérance, on obtient

$$(\beta-\alpha)\,\mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \big(\beta-\mathrm{E}(Y)\big)\mathrm{e}^{s\alpha}+\big(\mathrm{E}(Y)-\alpha\big)\mathrm{e}^{s\beta}.$$

Puisque la variable aléatoire Y est centrée, on conclut

$$(\beta-\alpha)\,\mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \beta\mathrm{e}^{s\alpha}-\alpha\mathrm{e}^{s\beta}.$$

(b) Commençons par observer que le logarithme introduit est bien défini car il porte sur une quantité supérieure à $(\beta-\alpha)\,\mathrm{E}\big(\mathrm{e}^{sY}\big)$ donc strictement positive[1]. Introduisons la fonction différence des deux membres

$$f(s)=\frac{1}{8}(\beta-\alpha)^2s^2+\ln(\beta-\alpha)-\ln\big(\beta\mathrm{e}^{s\alpha}-\alpha\mathrm{e}^{s\beta}\big) \quad \text{avec} \quad s\in\mathbb{R}.$$

La fonction f est indéfiniment dérivable et, pour tout s réel,

$$f'(s)=\frac{1}{4}(\beta-\alpha)^2s-\alpha\beta\frac{\mathrm{e}^{s\alpha}-\mathrm{e}^{s\beta}}{\beta\mathrm{e}^{s\alpha}-\alpha\mathrm{e}^{s\beta}}$$

$$f''(s)=\frac{1}{4}(\beta-\alpha)^2+\alpha\beta\frac{(\beta-\alpha)^2\mathrm{e}^{s(\alpha+\beta)}}{\big(\beta\mathrm{e}^{s\alpha}-\alpha\mathrm{e}^{s\beta}\big)^2}=\left(\frac{\beta-\alpha}{2}\right)^2\left(\frac{\beta\mathrm{e}^{s\alpha}+\alpha\mathrm{e}^{s\beta}}{\beta\mathrm{e}^{s\alpha}-\alpha\mathrm{e}^{s\beta}}\right)^2.$$

Sachant $f(0)=0$ et $f'(0)=0$, on peut dresser le tableau des variations de f et valider l'inégalité proposée.

s	$-\infty$		0		$+\infty$
$f''(s)$		$+$		$+$	
$f'(s)$	$-\infty$	↗	0	↗	$+\infty$
$f'(s)$		$-$	0	$+$	
$f(s)$	$+\infty$	↘	0	↗	$+\infty$

En passant à l'exponentielle, on prolonge l'inégalité obtenue à la question précédente et il vient

$$(\beta-\alpha)\,\mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \mathrm{e}^{(\beta-\alpha)^2s^2/8}(\beta-\alpha).$$

Il suffit alors de simplifier par $\beta-\alpha>0$ pour obtenir la deuxième inégalité.

1. Aussi, la variable Y étant centrée, on a $\alpha\leqslant 0$ et $\beta\geqslant 0$ ce qui donne le signe du contenu du logarithme.

(c) **méthode**

> On applique l'inégalité de Markov (Th. 25 p. 381) à un événement identique à celui étudié.

Soit $s > 0$. Par stricte croissance de la fonction $x \mapsto \mathrm{e}^{sx}$, on a

$$\big(S - \mathrm{E}(S) \geqslant t\big) = \big(\mathrm{e}^{s(S-\mathrm{E}(S))} \geqslant \mathrm{e}^{st}\big).$$

Par l'inégalité de Markov

$$\mathrm{P}\big(S - \mathrm{E}(S) \geqslant t\big) \leqslant \mathrm{e}^{-st}\,\mathrm{E}\big(\mathrm{e}^{s(S-\mathrm{E}(S))}\big). \qquad (*)$$

Or $S - \mathrm{E}(S)$ est la somme des $X_i - \mathrm{E}(X_i)$ donc

$$\mathrm{e}^{s(S-\mathrm{E}(S))} = \prod_{i=1}^{n} \mathrm{e}^{s(X_i-\mathrm{E}(X_i))}.$$

Par l'indépendance mutuelle des variables X_i, on a aussi [1] l'indépendance mutuelle des variables e^{sX_i} et donc

$$\mathrm{E}\Big(\mathrm{e}^{s(S-\mathrm{E}(S))}\Big) = \prod_{i=1}^{n} \mathrm{E}\Big(\mathrm{e}^{s(X_i-\mathrm{E}(X_i))}\Big).$$

La comparaison $(*)$ donne alors la première des deux inégalités demandées. La deuxième se déduit de l'étude initiale en considérant $Y = X_i - \mathrm{E}(X_i)$ qui est une variable centrée à valeurs dans $[\alpha\,;\beta]$ avec $\alpha = a - \mathrm{E}(X_i)$ et $\beta = b - \mathrm{E}(X_i)$. On obtient alors

$$\mathrm{E}\Big(\mathrm{e}^{s(X_i-\mathrm{E}(X_i))}\Big) = \mathrm{E}\big(\mathrm{e}^{sY}\big) \leqslant \exp\left(\frac{(\beta-\alpha)^2}{8}s^2\right) = \exp\left(\frac{(b-a)^2}{8}s^2\right)$$

et donc

$$\mathrm{e}^{-st}\prod_{i=1}^{n}\mathrm{E}\Big(\mathrm{e}^{s(X_i-\mathrm{E}(X_i))}\Big) \leqslant \mathrm{e}^{-st}\prod_{i=1}^{n}\exp\left(\frac{(b-a)^2}{8}s^2\right) = \exp\left(-st + n\frac{(b-a)^2}{8}s^2\right).$$

(d) On optimise l'inégalité précédente en choisissant $s = 4t/\big(n(b-a)^2\big)$ qui minimise [2] le contenu de l'exponentielle et l'on conclut

$$\mathrm{P}\big(S - \mathrm{E}(S) \geqslant t\big) \leqslant \exp\left(-\frac{2t^2}{n(b-a)^2}\right).$$

1. Cette affirmation se justifie aisément par le Th. 11 p. 376.

2. Une fonction $x \mapsto ax^2 + bx + c$ avec $a > 0$ est minimale en $x = -\frac{b}{2a}$ ce qui correspond au sommet de la parabole qui représente cette fonction.

Exercice 26 * (Inégalité de Kolmogorov)**

Sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$, on considère $X_1, \dots, X_n$ des variables aléatoires discrètes réelles indépendantes, d'espérances nulles et admettant chacune un moment d'ordre 2. Pour tout $k \in [\![1\,;n]\!]$, on note S_k la somme des variables $X_1, \dots, X_k$ et l'on introduit t un réel strictement positif.

Pour $k \in [\![1\,;n]\!]$, on introduit l'événement

$$A_k = \bigcap_{j=1}^{k-1} \{|S_j| < t\} \cap \{|S_k| \geqslant t\}.$$

(a) Justifier l'indépendance des variables $\mathbf{1}_{A_k} S_k^2$ et $S_n - S_k$ pour tout $k \in [\![1\,;n]\!]$.

(b) Montrer

$$\mathrm{E}(S_n^2) \geqslant \sum_{k=1}^{n} \mathrm{E}(S_k^2 \mathbf{1}_{A_k}).$$

(c) En déduire

$$\mathrm{P}\left(\max_{1 \leqslant k \leqslant n} |S_k| \geqslant t \right) \leqslant \frac{1}{t^2} \sum_{i=1}^{n} \mathrm{V}(X_i).$$

Solution

(a) **méthode**

> Le théorème d'indépendance par paquets (Th. 12 p. 376) permet de justifier que deux variables aléatoires sont indépendantes lorsque que celles-ci se définissent à partir de deux portions disjointes d'une famille de variables mutuellement indépendantes.

La variable aléatoire $\mathbf{1}_{A_k}$ est définie à partir des variables $S_1, \dots, S_k$ elles-mêmes déterminées à partir de $X_1, \dots, X_k$. La variable aléatoire $\mathbf{1}_{A_k} S_k^2$ est donc uniquement fonction de $X_1, \dots, X_k$. En revanche, $S_n - S_k = X_{k+1} + \cdots + X_n$ est uniquement fonction de $X_{k+1}, \dots, X_n$. Les variables $X_1, \dots, X_n$ étant mutuellement indépendantes, on peut affirmer l'indépendance de $\mathbf{1}_{A_k} S_k^2$ et $S_n - S_k$.

(b) Les événements $A_1, \dots, A_n$ sont deux à deux incompatibles et donc

$$\mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n} \leqslant 1.$$

En multipliant par S_n^2, on obtient par croissance et linéarité de l'espérance

$$\mathrm{E}(S_n^2) \geqslant \sum_{k=1}^{n} \mathrm{E}(\mathbf{1}_{A_k} S_n^2).$$

Or, pour tout $k \in [\![1\,;n]\!]$, on peut écrire

$$\begin{aligned}\mathrm{E}(\mathbf{1}_{A_k} S_n^2) &= \mathrm{E}(\mathbf{1}_{A_k}(S_k + S_n - S_k)^2)\\ &= \mathrm{E}(\mathbf{1}_{A_k} S_k^2) + 2\,\mathrm{E}(\mathbf{1}_{A_k} S_k(S_n - S_k)) + \underbrace{\mathrm{E}(\mathbf{1}_{A_k}(S_n - S_k)^2)}_{\geqslant 0}.\end{aligned}$$

Par l'indépendance acquise à la première question, on a

$$\mathrm{E}(\mathbf{1}_{A_k} S_k(S_n - S_k)) = \mathrm{E}(\mathbf{1}_{A_k} S_k)\,\mathrm{E}(S_n - S_k).$$

Or l'espérance de $S_n - S_k$ est nulle car les variables $X_{k+1}, \ldots, X_n$ sont centrées. On en déduit

$$\mathrm{E}(\mathbf{1}_{A_k} S_n^2) \geqslant \mathrm{E}(\mathbf{1}_{A_k} S_k^2)$$

ce qui permet de conclure à l'inégalité voulue.

(c) Soit $k \in [\![1\,;n]\!]$. Par définition de l'événement A_k, on a $S_k^2 \mathbf{1}_{A_k} \geqslant t^2 \mathbf{1}_{A_k}$ et par croissance de l'espérance

$$\mathrm{E}(S_n^2) \geqslant \sum_{k=1}^{n} \mathrm{E}(S_k^2 \mathbf{1}_{A_k}) \geqslant \sum_{k=1}^{n} t^2\,\mathrm{E}(\mathbf{1}_{A_k}) = \sum_{k=1}^{n} t^2\,\mathrm{P}(A_k).$$

Or

$$\left(\max_{1\leqslant k\leqslant n} |S_k| \geqslant t\right) = \bigcup_{k=1}^{n} A_k \quad \text{avec} \quad A_1, \ldots, A_n \text{ deux à deux incompatibles}$$

donc

$$\mathrm{P}\left(\max_{1\leqslant k\leqslant n} |S_k| \geqslant t\right) = \sum_{k=1}^{n} \mathrm{P}(A_k) \leqslant \frac{1}{t^2}\,\mathrm{E}(S_n^2).$$

Enfin, $\mathrm{E}(S_n^2)$ se confond avec la variance de S_n car cette variable est centrée et celle-ci est la somme des variances des $X_1, \ldots, X_n$ car ces variables sont supposées indépendantes.

9.8.5 Fonctions génératrices

Exercice 27 *

Deux joueurs lancent chacun deux dés équilibrés et l'on veut calculer la probabilité que les sommes des deux jets soient égales. On note X_1 et X_2 les variables aléatoires déterminant les valeurs des dés lancés par le premier joueur et Y_1 et Y_2 celles associées au deuxième joueur.

(a) Montrer que $\mathrm{P}(X_1 + X_2 = Y_1 + Y_2) = \mathrm{P}(14 + X_1 + X_2 - Y_1 - Y_2 = 14)$.

(b) Déterminer la fonction génératrice de la variable $Z = 14 + X_1 + X_2 - Y_1 - Y_2$.

(c) En déduire la valeur de $\mathrm{P}(X_1 + X_2 = Y_1 + Y_2)$.

Solution

(a) Les événements $(X_1 + X_2 = Y_1 + Y_2)$ et $(14 + X_1 + X_2 - Y_1 - Y_2 = 14)$ sont identiques.

(b) **méthode**

La fonction génératrice d'une somme de variables indépendantes est le produit des fonctions génératrices de chacune (Th. 29 p. 383).

On ne peut parler de fonction génératrice que d'une variable aléatoire à valeurs dans $\mathbb{N}$. Le terme 14 permet de comprendre Z comme la somme de quatre variables aléatoires à valeurs naturelles :

$$Z = X_1 + X_2 + (7 - Y_1) + (7 - Y_2).$$

Lorsque X désigne une variable aléatoire suivant une loi uniforme sur $[\![1\,;6]\!]$, sa fonction génératrice est

$$G_X(t) = \frac{1}{6}\left(t + t^2 + \cdots + t^6\right) = \frac{1}{6} \cdot \frac{t - t^7}{1 - t} \quad \text{pour tout } t \neq 1.$$

Les variables aléatoires X_1 et X_2 donnant la valeur d'un dé équilibré, elles ont cette fonction génératrice. Il en est de même des variables $7 - Y_1$ et $7 - Y_2$ qui suivent elles aussi une loi uniforme sur $[\![1\,;6]\!]$. Par indépendance mutuelle de ces différentes variables, on obtient

$$G_Z(t) = \frac{1}{6^4}\left(\frac{t - t^7}{1 - t}\right)^4 \quad \text{pour tout } t \neq 1.$$

(c) On veut calculer la probabilité de l'événement $(Z = 14)$.

méthode

On détermine le coefficient de t^{14} dans le développement en série entière de $G_Z(t)$.

D'une part, on développe le numérateur par la formule du binôme de Newton

$$\left(t - t^7\right)^4 = t^4 - 4t^{10} + 6t^{16} - 4t^{22} + t^{28} \quad \text{pour tout } t \in \mathbb{R}.$$

D'autre part, le développement du dénominateur se déduit de celui de $(1 + x)^\alpha$ avec α égal à -4

$$\frac{1}{(1 - t)^4} = 1 + 4t + \frac{4 \times 5}{2}t^2 + \cdots = \sum_{n=0}^{+\infty} \binom{n+3}{3} t^n \quad \text{pour tout } t \in]-1\,;1[.$$

Le coefficient de t^{14} dans le produit des deux développements en série entière est donc

$$\mathrm{P}(X_1 + X_2 = Y_1 + Y_2) = \frac{1}{6^4}\left(\binom{13}{3} - 4\binom{7}{3}\right) = \frac{146}{6^4} = \frac{73}{648} \simeq 0{,}113 \text{ à } 10^{-3} \text{ près.}$$

Exercice 28 **

Soit A et B deux événements indépendants d'un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ et la variable aléatoire $Z = \mathbf{1}_A + \mathbf{1}_B$.

Montrer que parmi les événements $(Z = 0)$, $(Z = 1)$ et $(Z = 2)$, il y en a au moins un de probabilité supérieure à $4/9$.

Solution

Posons $a = \mathrm{P}(Z = 2)$, $b = \mathrm{P}(Z = 1)$ et $c = \mathrm{P}(Z = 0)$. On a

$$a + b + c = 1. \qquad (*)$$

méthode

On introduit la fonction génératrice de Z.

La fonction génératrice de Z est définie sur $\mathbb{R}$ par

$$G_Z(t) = \mathrm{E}\big(t^Z\big) = at^2 + bt + c.$$

Par indépendance des événements A et B, les variables aléatoires $\mathbf{1}_A$ et $\mathbf{1}_B$ sont elles aussi indépendantes [1] et donc, pour tout $t \in \mathbb{R}$,

$$G_Z(t) = G_{\mathbf{1}_A}(t) G_{\mathbf{1}_B}(t).$$

Sauf si $\mathrm{P}(A) = \mathrm{P}(B) = 0$, auquel cas la propriété voulue est immédiate, au moins l'une des fonctions $G_{\mathbf{1}_A}$ et $G_{\mathbf{1}_B}$ est affine non constante et admet donc une racine réelle. La fonction G_Z possède donc un moins une racine réelle et par conséquent [2]

$$b^2 \geqslant 4ac. \qquad (**)$$

Si $b < 4/9$, on obtient $a + c > 5/9$ par $(*)$ et $ac < 4/81$ par $(**)$. On en déduit

$$\left(a - \frac{4}{9}\right)\left(c - \frac{4}{9}\right) = ac - \frac{4}{9}(a + c) + \frac{16}{81} < 0.$$

Dans ce cas, l'une des deux quantités $a - 4/9$ ou $c - 4/9$ doit être positive (et l'autre est négative). Ainsi, parmi a, b, c, au moins l'une des trois probabilités est supérieure à $4/9$.

Exercice 29 (Processus de Galton-Watson [3])

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon la loi d'une variable X à valeurs dans $\mathbb{N}$. Soit aussi N une variable aléatoire à valeurs dans $\mathbb{N}$ indépendantes des précédentes. On étudie $S = X_1 + \cdots + X_N$.

(a) Justifier que S est une variable aléatoire à valeurs naturelles.

(b) Établir $G_S(t) = G_N\big(G_X(t)\big)$ pour tout t de $[-1\,;1]$.

(c) On suppose que les variables N et X admettent chacune une espérance finie. Établir l'identité de Wald : $\mathrm{E}(S) = \mathrm{E}(N)\,\mathrm{E}(X)$.

1. Ces deux variables suivent des lois de Bernoulli, si $\mathrm{P}(A) = \mathrm{P}(B)$, Z suit une loi binomiale.

2. La propriété $b^2 \geqslant 4ac$ correspond à $\Delta \geqslant 0$ si $a \neq 0$ et est évidente si $a = 0$.

3. Si la variable N détermine le nombre d'individus d'une population et X le nombre de descendants que chaque individu peut engendrer, la variable S correspond à la population à la génération suivante.

Solution

(a) Notons $(\Omega, \mathcal{T}, \mathrm{P})$ l'espace probabilisé sous-jacent à cette étude. La fonction S est définie par [1]

$$S(\omega) = X_1(\omega) + \cdots + X_{N(\omega)}(\omega) \quad \text{pour tout } \omega \in \Omega.$$

Cette fonction est à valeurs dans $\mathbb{N}$ et, pour tout $k \in \mathbb{N}$ et tout $\omega \in \Omega$,

$$S(\omega) = k \iff \exists n \in \mathbb{N},\ N(\omega) = n \text{ et } (X_1 + \cdots + X_n)(\omega) = k.$$

On peut donc écrire

$$(S = k) = \bigcup_{n \in \mathbb{N}} \big((N = n) \cap (X_1 + \cdots + X_n = k)\big). \qquad (*)$$

Pour chaque $n \in \mathbb{N}$, $X_1 + \cdots + X_n$ est une variable aléatoire par somme de variables aléatoires et donc $(X_1 + \cdots + X_n = k)$ est un événement. Il en est évidemment de même de $(N = n)$ et donc, par opérations dans la tribu $\mathcal{T}$, on peut assurer que $(S = k)$ est un événement. Ainsi, S est une variable aléatoire à valeurs dans $\mathbb{N}$.

(b) Soit $k \in \mathbb{N}$. Les événements réunis dans la formule $(*)$ sont deux à deux incompatibles et donc, par additivité dénombrable,

$$\mathrm{P}(S = k) = \sum_{n=0}^{+\infty} \mathrm{P}(N = n, X_1 + \cdots + X_n = k).$$

Par indépendance de la variable N avec les variables de la suite $(X_n)_{n \geqslant 1}$, on peut affirmer pour tout $n \in \mathbb{N}$ l'indépendance de N avec $X_1 + \cdots + X_n$ et poursuivre le calcul

$$\mathrm{P}(S = k) = \sum_{n=0}^{+\infty} \mathrm{P}(N = n)\,\mathrm{P}(X_1 + \cdots + X_n = k).$$

Soit $t \in [-1\,;1]$. La fonction génératrice de S est assurément définie en t et

$$G_S(t) = \sum_{k=0}^{+\infty} \mathrm{P}(S = k)t^k = \sum_{k=0}^{+\infty} \left(\sum_{n=0}^{+\infty} \mathrm{P}(N = n)\,\mathrm{P}(X_1 + \cdots + X_n = k)t^k \right). \qquad (\triangle)$$

méthode

On réorganise le calcul à l'aide d'une sommation par paquets.

Considérons la famille doublement indexée

$$\big(\mathrm{P}(N = n)\,\mathrm{P}(X_1 + \cdots + X_n = k)t^k\big)_{(k,n) \in \mathbb{N}^2}.$$

1. Si $N(\omega) = 0$, il s'agit d'une somme vide dont la valeur est nulle.

Par le calcul de la relation $(\triangle)$ utilisé avec $|t|$ au lieu de t, on peut affirmer que la famille précédente est sommable [1] et réorganiser le calcul de sa somme comme ci-dessous

$$G_S(t) = \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{+\infty} \mathrm{P}(N=n)\,\mathrm{P}(X_1+\cdots+X_n=k)t^k\right)$$
$$= \sum_{n=0}^{+\infty} \mathrm{P}(N=n)\left(\sum_{k=0}^{+\infty} \mathrm{P}(X_1+\cdots+X_n=k)t^k\right).$$

Dans la somme contenue, on reconnaît $G_{X_1+\cdots+X_n}(t)$. Or les variables $X_1,\ldots,X_n$ sont indépendantes et la fonction génératrice d'une somme est alors le produit des fonctions génératrices

$$G_{X_1+\cdots+X_n}(t) = G_{X_1}(t)\times\cdots\times G_{X_n}(t) = \big(G_X(t)\big)^n.$$

Ainsi,

$$G_S(t) = \sum_{n=0}^{+\infty} \mathrm{P}(N=n)\big(G_X(t)\big)^n = \sum_{n=0}^{+\infty} \mathrm{P}(N=n)T^n \quad \text{avec} \quad T = G_X(t) \in [-1\,;1].$$

On conclut $G_S(t) = G_N\big(G_X(t)\big)$.

(c) Si les variables N et X possèdent une espérance finie, les fonctions G_N et G_X sont dérivables en 1 (Th. 28 p. 383). Par composition, G_S est aussi dérivable en 1 avec

$$G'_S(1) = G'_X(1)G'_N\big(G_X(1)\big) = G'_X(1)G'_N(1) \quad \text{car} \quad G_X(1)=1.$$

Ainsi, S admet une espérance finie et $\mathrm{E}(S) = \mathrm{E}(N)\,\mathrm{E}(X)$.

Exercice 30 ***

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi de Bernoulli de paramètre $p \in\,]0\,;1[$. Soit aussi N une variable aléatoire à valeurs dans $\mathbb{N}$ indépendante des précédentes. On considère les variables aléatoires

$$X = \sum_{k=1}^{N} X_k \quad \text{et} \quad Y = \sum_{k=1}^{N}(1-X_k).$$

(a) Pour t et u dans $[-1\,;1]$, exprimer à l'aide de la fonction génératrice de N l'expression

$$G(t,u) = \mathrm{E}\big(t^X u^Y\big).$$

(b) On suppose que N suit une loi de Poisson. Montrer que les variables X et Y sont indépendantes.

(c) Inversement, on suppose que les variables X et Y sont indépendantes. Montrer que N suit une loi de Poisson.

1. La définition de la fonction G_S en t assure la sommabilité de la famille $\big(\mathrm{P}(S=k)t^k\big)_{k\in\mathbb{N}}$ ce qui n'est pas exactement celle considérée ici : ceci explique la nécessité de l'argument.

Solution

(a) Soit $(t,u) \in [-1\,;1]^2$. Les variables X et Y sont à valeurs dans $\mathbb{N}$ et la variable $t^X u^Y$ est alors bornée ce qui assure que son espérance est finie. Par la formule de transfert appliquée à partir de la variable conjointe $Z = (X,Y)$

$$\mathrm{E}\big(t^X u^Y\big) = \sum_{(k,\ell)\in\mathbb{N}^2} t^k u^\ell \,\mathrm{P}(X=k, Y=\ell)$$

avec la famille doublement indexée $\big(t^k u^\ell \,\mathrm{P}(X=k,Y=\ell)\big)_{(k,\ell)\in\mathbb{N}^2}$ sommable.

méthode

On réorganise le calcul de la somme en fonction de la valeur de $X+Y$.

On peut décomposer $\mathbb{N}^2$ en la réunion des ensembles deux à deux disjoints

$$I_n = \big\{(k,n-k) \,\big|\, k \in [\![0\,;n]\!]\big\} \quad \text{avec} \quad n \in \mathbb{N}.$$

Par le théorème de sommation par paquets

$$\mathrm{E}\big(t^X u^Y\big) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^{n} t^k u^{n-k} \,\mathrm{P}(X=k, Y=n-k) \right).$$

Or

$$(X=k, Y=n-k) = (X_1 + \cdots + X_n = k) \cap (N=n)$$

et l'hypothèse d'indépendance donne

$$\mathrm{P}(X_1 + \cdots + X_n = k, N=n) = \mathrm{P}(X_1 + \cdots + X_n = k)\,\mathrm{P}(N=n).$$

Au surplus, la variable $X_1 + \cdots + X_n$ suit une loi binomiale de paramètres n et p car les variables $X_1, \ldots, X_n$ sont indépendantes et suivent une même loi de Bernoulli de paramètre p. Tous ces résultats permettent d'écrire

$$\mathrm{E}\big(t^X u^Y\big) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^{n} t^k u^{n-k} \binom{n}{k} p^k q^{n-k} \,\mathrm{P}(N=n) \right) \quad \text{avec} \quad q = 1-p.$$

Enfin, en appliquant la formule du binôme [1]

$$\mathrm{E}\big(t^X u^Y\big) = \sum_{n=0}^{+\infty} \mathrm{P}(N=n)(pt+qu)^n = G_N(pt+qu).$$

(b) **méthode**

Les fonctions génératrices de X et Y se déduisent de valeurs bien choisies de $G(t,u)$ ce qui permet d'identifier leurs lois.

1. Notons que $pt + qu = \lambda t + (1-\lambda)u$ est élément de $[-1\,;1]$ car t et u le sont et $\lambda = p \in [0\,;1]$.

Si N suit une loi de Poisson de paramètre $\lambda > 0$, on sait $G_N(t) = \mathrm{e}^{\lambda(t-1)}$ et donc

$$G(t,u) = \mathrm{e}^{\lambda p(t-1)} \times \mathrm{e}^{\lambda q(u-1)}.$$

En particulier,

$$G_X(t) = \mathrm{E}\big(t^X\big) = G(t,1) = \mathrm{e}^{\lambda p(t-1)} \quad \text{et} \quad G_Y(t) = \mathrm{E}\big(u^Y\big) = G(1,u) = \mathrm{e}^{\lambda q(u-1)}.$$

La variable X suit une loi de Poisson de paramètre λp tandis que Y suit une loi de Poisson de paramètre λq.

On peut alors vérifier l'indépendance des variables X et Y. Soit $(k,\ell) \in \mathbb{N}^2$ et $n = k+\ell$. En reprenant une partie des calculs qui précèdent

$$\mathrm{P}(X = k, Y = \ell) = \mathrm{P}(X = k, N = n) = \mathrm{P}(X_1 + \cdots + X_n = k)\,\mathrm{P}(N = n)$$

donc

$$\mathrm{P}(X = k, Y = \ell) = \binom{n}{k} p^k q^{n-k} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!}.$$

Parallèlement,

$$\mathrm{P}(X = k)\,\mathrm{P}(Y = \ell) = \mathrm{e}^{-p\lambda} \frac{(p\lambda)^k}{k!} \mathrm{e}^{-q\lambda} \frac{(q\lambda)^\ell}{\ell!}.$$

Ces deux expressions sont identiques car $p + q = 1$ et $k + \ell = n$: on peut conclure que les variables X et Y sont indépendantes[1].

(c) Si les variables X et Y sont indépendantes, les variables composées t^X et u^Y le sont aussi et donc

$$G(t,u) = \mathrm{E}\big(t^X u^Y\big) = \mathrm{E}\big(t^X\big)\,\mathrm{E}\big(u^Y\big) = G(t,1)G(1,u).$$

On en déduit l'égalité

$$G_N(pt + qu) = G_N(pt + q)G_N(p + qu) \quad \text{pour tout } (t,u) \in [-1\,;1]^2.$$

On dérive cette identité en la variable t pour écrire

$$pG'_N(pt + qu) = pG'_N(pt + q)G_N(p + qu).$$

On simplifie par p qui est non nul et l'on évalue en $t = 1$

$$G'_N(s) = G'_N(1)G_N(s) \quad \text{avec} \quad s = p + qu \text{ et } u \in [-1\,;1].$$

Lorsque la variable u parcourt $[-1\,;1]$, la variable $s = p + qu$ décrit $[p - q\,;1]$.

Cas : $p \leqslant 1/2$. L'identité qui précède vaut au moins pour $s \in [0\,;1]$. La résolution de l'équation différentielle associée donne pour tout $s \in [0\,;1]$

$$G_N(s) = C\mathrm{e}^{\lambda s} \quad \text{avec} \quad \lambda = G'_N(1) \text{ et } C \text{ une constante réelle.}$$

1. On retrouve ici un résultat déjà vu dans le sujet 20 p. 404. Nous allons maintenant en étudier la réciproque.

Sachant $G_N(1) = 1$, on détermine C et l'on conclut $G_N(s) = \mathrm{e}^{\lambda(s-1)}$. Cette relation, qui vaut pour tout $s \in [0\,;1]$, suffit pour calculer les coefficients de la série entière définissant G_N, elle caractérise donc la loi de N : la variable N suit une loi de Poisson de paramètre λ.

Cas : $p \geqslant 1/2$. On reprend le calcul initial en échangeant les rôles de t et u.

9.9 Exercices d'approfondissement

Exercice 31 *

Soit $X_1, \ldots, X_n$ des variables aléatoires discrètes réelles admettant chacune un moment d'ordre 2. On appelle *matrice de covariance* de la famille $(X_1, \ldots, X_n)$ la matrice

$$M = \big(\mathrm{Cov}(X_i, X_j)\big)_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R}).$$

Montrer que cette matrice est diagonalisable à valeurs propres positives.

Solution

La matrice M est symétrique réelle donc diagonalisable (Th. 6 p. 291).

Soit λ une valeur propre de M et $X = {}^t\begin{pmatrix} a_1 & \ldots & a_n \end{pmatrix}$ une colonne propre associée. On sait $MX = \lambda X$ avec $X \neq 0$.

méthode

On rapproche le calcul de ${}^tXMX$ de celui de la variance d'une variable aléatoire.

D'une part, ${}^tXMX = \lambda\, {}^tXX = \lambda \|X\|^2$ en introduisant la norme euclidienne canonique sur l'espace des colonnes réelles de hauteur n.

D'autre part,

$${}^tXMX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \,\mathrm{Cov}(X_i, X_j).$$

Par bilinéarité de la covariance, on a encore

$${}^tXMX = \mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right) = \mathrm{V}(Y)$$

avec Y la variable aléatoire[1] $a_1X_1 + \cdots + a_nX_n$.

On en déduit

$$\lambda = \frac{\mathrm{V}(Y)}{\|X\|^2} \in \mathbb{R}_+.$$

1. Rappelons que $L^2(\Omega)$ est un espace vectoriel donc stable par combinaison linéaire.

Exercice 32 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice dont les coefficients sont des variables aléatoires indépendantes et identiquement distribuées selon la loi uniforme sur $\{-1, 1\}$.
Calculer $\mathrm{E}\big(\det(A)\big)$ et $\mathrm{V}\big(\det(A)\big)$.

Solution

Pour $(i,j) \in [\![1\,;n]\!]^2$, notons $A_{i,j}$ la variable aléatoire associée au coefficient d'indice (i,j) de la matrice A. Par hypothèse $A_{i,j}$ est d'espérance nulle et $A_{i,j}^2$ est la variable constante égale à 1.

Par la formule[1] définissant le déterminant et par la linéarité de l'espérance

$$\mathrm{E}\big(\det(A)\big) = \sum_{\sigma\in\mathcal{S}_n}\left(\varepsilon(\sigma)\,\mathrm{E}\left(\prod_{i=1}^{n} A_{\sigma(i),i}\right)\right).$$

Les variables $A_{i,j}$ étant mutuellement indépendantes, on termine le calcul

$$\mathrm{E}\big(\det(A)\big) = \sum_{\sigma\in\mathcal{S}_n}\left(\varepsilon(\sigma)\prod_{i=1}^{n}\underbrace{\mathrm{E}\big(A_{\sigma(i),i}\big)}_{=0}\right) = 0.$$

Par la formule de Huygens

$$\mathrm{V}\big(\det(A)\big) = \mathrm{E}\Big(\big(\det(A)\big)^2\Big) - \Big(\mathrm{E}\big(\det(A)\big)\Big)^2 = \mathrm{E}\Big(\big(\det(A)\big)^2\Big).$$

méthode

‖ On exprime $\big(\det(A)\big)^2$ comme le produit des deux sommes définissant $\det(A)$.

En développant

$$\begin{aligned}\big(\det(A)\big)^2 &= \left(\sum_{\sigma\in\mathcal{S}_n}\left(\varepsilon(\sigma)\prod_{i=1}^{n} A_{\sigma(i),i}\right)\right)\left(\sum_{\sigma'\in\mathcal{S}_n}\left(\varepsilon(\sigma')\prod_{i=1}^{n} A_{\sigma'(i),i}\right)\right)\\ &= \sum_{\sigma\in\mathcal{S}_n}\sum_{\sigma'\in\mathcal{S}_n}\left(\varepsilon(\sigma)\varepsilon(\sigma')\left(\prod_{i=1}^{n} A_{\sigma(i),i}\right)\left(\prod_{i=1}^{n} A_{\sigma'(i),i}\right)\right).\end{aligned}$$

Par linéarité de l'espérance, il s'agit, pour $(\sigma,\sigma') \in \mathcal{S}_n^2$ donné, de calculer l'espérance de

$$P_{\sigma,\sigma'} = \left(\prod_{i=1}^{n} A_{\sigma(i),i}\right)\left(\prod_{i=1}^{n} A_{\sigma'(i),i}\right) = \prod_{i=1}^{n} B_i \quad \text{avec} \quad B_i = A_{\sigma(i),i}A_{\sigma'(i),i}.$$

Par le théorème d'indépendance par paquets (Th. 12 p. 376), la variable B_1, qui est fonction des coefficients de la première colonne de A, est indépendante du produit des

1. Il est aussi possible de raisonner par récurrence en développant selon une rangée.

variables $B_2, \dots, B_n$, qui est fonction des coefficients des autres colonnes. On a donc

$$\mathrm{E}(P_{\sigma,\sigma'}) = \mathrm{E}(B_1)\,\mathrm{E}\left(\prod_{i=2}^{n} B_i\right).$$

Une utilisation répétée de cet argument donne $\mathrm{E}(P_{\sigma,\sigma'}) = \mathrm{E}(B_1) \times \cdots \times \mathrm{E}(B_n)$. On poursuit en distinguant deux cas :

Cas : $\sigma \neq \sigma'$. Il existe un indice $i \in [\![1\,;n]\!]$ tel que $\sigma(i) \neq \sigma'(i)$. Par indépendance des variables $A_{\sigma(i),i}$ et $A_{\sigma'(i),i}$, on a $\mathrm{E}(B_i) = \mathrm{E}\big(A_{\sigma(i),i}\big)\,\mathrm{E}\big(A_{\sigma'(i),i}\big) = 0$ donc $\mathrm{E}(P_{\sigma,\sigma'}) = 0$.

Cas : $\sigma = \sigma'$. Pour tout $i \in [\![1\,;n]\!]$, $B_i = A_{\sigma(i),i}^2 = 1$ donc $\mathrm{E}(B_i) = 1$ puis $\mathrm{E}(P_{\sigma,\sigma'}) = 1$.

Ainsi,

$$\mathrm{E}\Big(\big(\det(A)\big)^2\Big) = \sum_{\sigma\in\mathcal{S}_n}\left(\sum_{\sigma'\in\mathcal{S}_n} \varepsilon(\sigma)\varepsilon(\sigma')\,\mathrm{E}(P_{\sigma,\sigma'})\right) \quad \text{avec} \quad \mathrm{E}(P_{\sigma,\sigma'}) = \begin{cases} 1 & \text{si } \sigma = \sigma' \\ 0 & \text{sinon} \end{cases}$$

et, finalement,

$$\mathrm{V}\big(\det(A)\big) = \sum_{\sigma\in\mathcal{S}_n} 1 = \mathrm{Card}\big(\mathcal{S}_n\big) = n!$$

Exercice 33 ***

Cinq amis prennent place autour d'une table ronde et possèdent chacun deux jetons. On suppose que les jetons sont tous blancs sauf deux bleus qui sont possédés par deux voisins. À chaque tour, chacun distribue arbitrairement ses deux jetons, l'un à son camarade de droite, l'autre à son camarade de gauche. Le jeu s'arrête lorsque l'un des amis prend possession des deux jetons bleus.

Calculer le nombre moyen de tours nécessaires pour que le jeu s'arrête.

Solution

méthode

‖ On étudie si les jetons bleus sont possédés par deux amis voisins ou non.

Pour $n \in \mathbb{N}$, on introduit les événements

A_n = « Au n-ième tour, les jetons bleus sont possédés par deux amis voisins »,

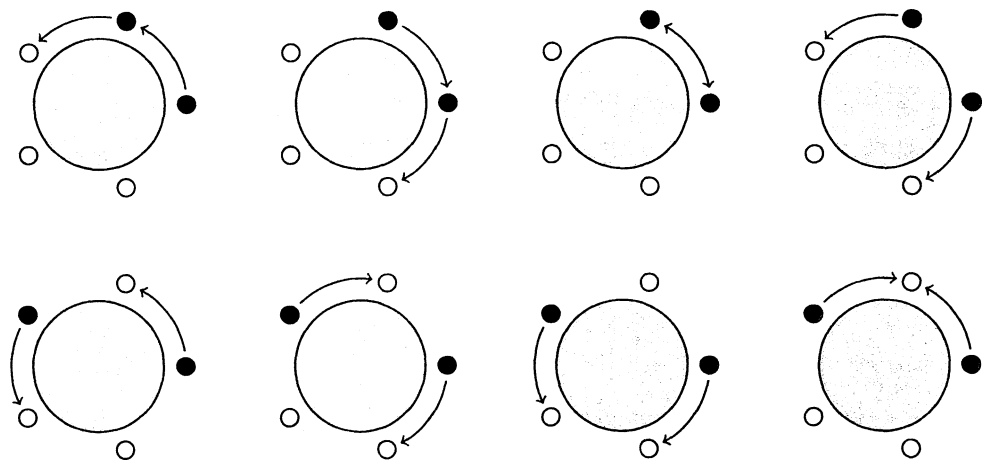
B_n = « Au n-ième tour, les jetons bleus sont possédés par deux amis non voisins ».

On pose $a_n = \mathrm{P}(A_n)$ et $b_n = \mathrm{P}(B_n)$. La configuration initiale donne $a_0 = 1$ et $b_0 = 0$. Enfin, on note X la variable aléatoire déterminant le tour auquel le jeu s'arrête. Il s'agit de calculer $\mathrm{E}(X)$.

Lors d'un tour, les deux jetons bleus peuvent de façon équiprobable évoluer dans le même sens ou en sens contraire autour de la table. Lorsque les jetons évoluent dans le même sens, la configuration globale ne change pas : si les jetons sont voisins, ils restent voisins, s'il ne sont pas voisins, ils restent non voisins. Lorsque les jetons évoluent en sens contraire à partir d'une position où ils sont voisins, ils peuvent se croiser et rester voisins ou, de façon équiprobable, mener à la situation où ils ne sont plus voisins. Aussi,

lorsque les jetons évoluent en sens contraire à partir d'une position où ils ne sont pas voisins, ils peuvent devenir voisins ou entrer en possession de l'ami intercalé entre eux ce qui signe l'arrêt du jeu. Par la formule des probabilités totales, on résume la dynamique qui précède par les équations du système suivant :

$$\begin{cases} a_{n+1} = \frac{3}{4}a_n + \frac{1}{4}b_n & (1) \\ b_{n+1} = \frac{1}{4}a_n + \frac{1}{2}b_n & (2) \end{cases} \quad \text{pour tout } n \in \mathbb{N}. \qquad (*)$$



Nous allons calculer l'espérance de X par la formule du sujet 16 p. 399 :

$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} \mathrm{P}(X > n) \quad \text{avec} \quad \mathrm{P}(X > n) = a_n + b_n. \qquad (**)$$

Pour $n \in \mathbb{N}$, le système $(*)$ s'exprime matriciellement

$$U_{n+1} = AU_n \quad \text{avec} \quad A = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/2 \end{pmatrix}, \; U_{n+1} = \begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} \text{ et } U_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

On a donc $U_n = A^n U_0$ avec $U_0 = {}^t\begin{pmatrix} 1 & 0 \end{pmatrix}$. On peut alors écrire $a_n + b_n = LA^nU_0$ en introduisant la ligne $L = \begin{pmatrix} 1 & 1 \end{pmatrix}$. Par continuité de l'application linéaire $X \mapsto LXU_0$ sur l'espace de dimension finie $\mathcal{M}_2(\mathbb{R})$, on exprime sous réserve de convergence de la série des matrices

$$\sum_{n=0}^{+\infty}(a_n + b_n) = \sum_{n=0}^{+\infty} LA^nU_0 = L\left(\sum_{n=0}^{+\infty} A^n\right)U_0.$$

Dans l'espace $\mathcal{M}_2(\mathbb{R})$, on sait la norme euclidienne $\|\cdot\|$ sous-multiplicative[1]. On vérifie $\|A\| = \sqrt{15/16} < 1$ et l'on peut affirmer que la série $\sum A^n$ converge absolument avec[2]

$$\sum_{n=0}^{+\infty} A^n = (\mathrm{I}_2 - A)^{-1} \quad \text{et après calculs} \quad (\mathrm{I}_2 - A)^{-1} = \begin{pmatrix} 8 & 4 \\ 4 & 4 \end{pmatrix}.$$

1. Voir sujet 10 p. 257.
2. Voir sujet 3 du chapitre 3 de l'ouvrage *Exercices d'analyse MP*.

On en déduit que (A^n) tend vers O_2 ce qui entraîne que $P(X > n) = a_n + b_n$ est de limite nulle. Par continuité monotone, on peut affirmer qu'il est presque sûr que le jeu s'arrête. Ceci légitime l'usage de la formule $(**)$ pour évaluer le nombre moyen de tours nécessaires à l'arrêt du jeu et on obtient

$$E(X) = \sum_{n=0}^{+\infty}(a_n + b_n) = L(I_2 - A)^{-1}U_0 = 12$$

Exercice 34 * (Convergences probabilistes)**

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur ce même espace.

(a) Montrer que l'ensemble des $\omega \in \Omega$ tels que $\big(X_n(\omega)\big)$ tend vers $X(\omega)$ constitue un événement.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ *converge presque sûrement* vers la variable X lorsque

$$P\Big(X_n \xrightarrow[n\to+\infty]{} X\Big) = 1.$$

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité*[1] vers[2] la variable X si, pour tout $\varepsilon > 0$,

$$P\Big(|X_n - X| > \varepsilon\Big) \xrightarrow[n\to+\infty]{} 0.$$

(b) Montrer que la convergence presque sûre entraîne la convergence en probabilité.

(c) On suppose que, pour tout $\varepsilon > 0$, il y a convergence de la série de terme général $P\big(|X_n - X| > \varepsilon\big)$. Montrer à l'aide du résultat du sujet 7 p. 347 que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers la variable X.

(d) Montrer que la convergence en probabilité de $(X_n)_{n\in\mathbb{N}}$ vers X entraîne la convergence presque sûre d'une suite extraite $\big(X_{\varphi(n)}\big)_{n\in\mathbb{N}}$ vers X.

Solution

(a) Les issues ω pour lesquelles $\big(X_n(\omega)\big)$ tend vers $X(\omega)$ sont celles vérifiant

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\quad n \geqslant N \implies \big|X_n(\omega) - X(\omega)\big| \leqslant \varepsilon. \qquad (*)$$

méthode

On traduit cette phrase quantifiée par opérations sur les événements sachant qu'une quantification existentielle s'exprime par une union alors qu'une quantification universelle s'exprime par une intersection.

1. On définit encore d'autres types de convergence. Par exemple, dans l'espace $L^1(\Omega)$, on introduit la *convergence* L^1 par $E(|X_n - X|) \to 0$. Il s'agit presque d'une convergence dans un espace normé, l'application $X \mapsto \|X\|_1 = E(|X_1|)$ étant une semi-norme sur $L^1(\Omega)$: elle vérifie les axiomes définissant une norme sauf celui de séparation. Par l'inégalité de Markov, la convergence L^1 entraîne la convergence en probabilité.

2. Il n'y a pas exactement unicité de la variable X vers laquelle la convergence a lieu. Plus précisément, s'il y a convergence en probabilité vers deux variables X et X' celles-ci ne sont que presque sûrement égales.

Pour $\varepsilon > 0$ et $n \in \mathbb{N}$, l'événement $\big(|X_n - X| \leqslant \varepsilon\big)$ regroupe les issues ω vérifiant $\big|X_n(\omega) - X(\omega)\big| \leqslant \varepsilon$. Pour $N \in \mathbb{N}$, l'intersection

$$\bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)$$

réunit les issues ω vérifiant $\big|X_n(\omega) - X(\omega)\big| \leqslant \varepsilon$ pour tout $n \geqslant N$. Cet ensemble est un événement car c'est une intersection dénombrable d'événements. Aussi, l'union

$$A_\varepsilon = \bigcup_{N \in \mathbb{N}} \bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)$$

regroupe les issues ω vérifiant $\big|X_n(\omega) - X(\omega)\big| \leqslant \varepsilon$ à partir d'un certain rang. Il s'agit encore d'un événement puisque c'est une union dénombrable d'événements. Enfin, l'intersection

$$A = \bigcap_{\varepsilon > 0} \bigcup_{N \in \mathbb{N}} \bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big) = \bigcap_{\varepsilon > 0} A_\varepsilon$$

regroupe tous les ω de Ω vérifiant $(*)$. Il ne s'agit cependant pas d'une intersection dénombrable et c'est la raison pour laquelle nous allons exprimer autrement celle-ci. Les ensembles A_ε constituent une famille croissante : pour tous ε et $\varepsilon' > 0$

$$\varepsilon \leqslant \varepsilon' \implies A_\varepsilon \subset A_{\varepsilon'}.$$

Par conséquent, on peut aussi écrire

$$A = \bigcap_{\varepsilon > 0} A_\varepsilon = \bigcap_{n \in \mathbb{N}^*} A_{1/n}. \qquad (**)$$

Finalement, A est bien un événement car intersection dénombrable d'événements.

(b) Soit $\varepsilon > 0$. L'hypothèse de convergence presque sûre signifie que l'ensemble A introduit ci-dessus est quasi certain. Puisque celui-ci est inclus dans A_ε, ce dernier est aussi quasi certain. Or A_ε est une union d'événements constituant une suite croissante. En effet, pour tous N et $N' \in \mathbb{N}$

$$N \leqslant N' \implies \bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big) \subset \bigcap_{n \geqslant N'} \big(|X_n - X| \leqslant \varepsilon\big).$$

Par continuité monotone, on a donc

$$\lim_{N \to +\infty} \mathrm{P}\left(\bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)\right) = \mathrm{P}(A_\varepsilon) = 1.$$

Or

$$\left(\bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)\right) \subset \big(|X_N - X| \leqslant \varepsilon\big)$$

et donc

$$\mathrm{P}\left(\bigcap_{n \geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)\right) \leqslant \mathrm{P}\big(|X_N - X| \leqslant \varepsilon\big) \leqslant 1$$

Par le théorème de convergence par encadrement, on obtient

$$\lim_{N\to+\infty} \mathrm{P}\big(|X_N - X| \leqslant \varepsilon\big) = 1.$$

Enfin, par passage à l'événement contraire, on conclut

$$\mathrm{P}\big(|X_n - X| > \varepsilon\big) \xrightarrow[n\to+\infty]{} 0.$$

La suite (X_n) converge en probabilité vers X.

(c) Par le lemme de Borel-Cantelli, la convergence de la série $\sum \mathrm{P}\big(|X_n - X| > \varepsilon\big)$ entraîne

$$\mathrm{P}\left(\bigcap_{N\in\mathbb{N}} \bigcup_{n\geqslant N} \big(|X_n - X| > \varepsilon\big)\right) = 0.$$

Par passage à l'événement contraire

$$\mathrm{P}\left(\bigcup_{N\in\mathbb{N}} \bigcap_{n\geqslant N} \big(|X_n - X| \leqslant \varepsilon\big)\right) = 1.$$

En reprenant les notations de la première question, on peut affirmer que l'événement A_ε est presque sûr. À l'aide de $(**)$ et par continuité décroissante, on conclut

$$\mathrm{P}\Big(X_n \xrightarrow[n\to+\infty]{} X\Big) = \lim_{n\to+\infty} \mathrm{P}(A_{1/n}) = 1.$$

(d) On choisit $\varphi(0)$ arbitrairement puis, pour tout $n \in \mathbb{N}$, une fois $\varphi(n)$ déterminé, on choisit $\varphi(n+1) > \varphi(n)$ vérifiant

$$\mathrm{P}\left(|X_{\varphi(n+1)} - X| > \frac{1}{n+1}\right) \leqslant \frac{1}{(n+1)^2}.$$

Ceci est possible car, pour $\varepsilon = 1/(n+1) > 0$, le terme $\mathrm{P}\big(|X_p - X| > \varepsilon\big)$ est de limite nulle quand p tend vers l'infini et il existe des indices p arbitrairement grands pour lesquels cette valeur est inférieure à $1/(n+1)^2$: $\varphi(n+1)$ désigne l'un d'entre eux choisi strictement supérieur à $\varphi(n)$.

Soit $\varepsilon > 0$. On peut déterminer $n_0 \in \mathbb{N}^*$ tel que $1/n_0 \leqslant \varepsilon$ et alors, pour tout $n \geqslant n_0$,

$$\mathrm{P}\big(|X_{\varphi(n)} - X| > \varepsilon\big) \leqslant \mathrm{P}\left(|X_{\varphi(n)} - X| > \frac{1}{n}\right) \leqslant \frac{1}{n^2}.$$

La série de terme général $\mathrm{P}\big(|X_{\varphi(n)} - X| > \varepsilon\big)$ est donc convergente et l'on peut conclure à l'aide de la question précédente.

Exercice 35 * (Loi forte des grands nombres)**

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires deux à deux indépendantes et identiquement distribuées selon une loi admettant une espérance μ et une variance σ^2. On pose

$$S_n = \frac{1}{n}(X_1 + \cdots + X_n)$$

et l'on souhaite établir

$$\mathrm{P}\Big(S_n \xrightarrow[n\to+\infty]{} \mu\Big) = 1.$$

(a) Montrer que l'on peut supposer $\mu = 0$.

On conservera cette hypothèse dans la suite de l'étude.

(b) Soit $\varepsilon > 0$. Montrer

$$\mathrm{P}\big(|S_n| > \varepsilon\big) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

(c) À l'aide du sujet précédent, établir que la suite extraite $(S_{m^2})_{m\in\mathbb{N}^*}$ converge presque sûrement vers 0.

Pour $n \in \mathbb{N}^*$, on pose m égal à la partie entière de la racine carrée de n et l'on introduit

$$T_n = \frac{1}{n}(X_{m^2+1} + \cdots + X_n).$$

(d) Soit $\varepsilon > 0$. Montrer

$$\mathrm{P}\big(|T_n| > \varepsilon\big) \leqslant \frac{2\sigma^2}{n^{3/2}\varepsilon^2}$$

(e) Conclure que la suite $(S_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers 0.

Solution

(a) Pour tout $n \in \mathbb{N}^*$, introduisons $X'_n = X_n - \mu$ et $S'_n = S_n - \mu$ de sorte que

$$S'_n = \frac{1}{n}(X'_1 + \cdots + X'_n) \quad \text{et} \quad \Big(S_n \xrightarrow[n\to+\infty]{} \mu\Big) = \Big(S'_n \xrightarrow[n\to+\infty]{} 0\Big).$$

Les variables X'_n satisfont les mêmes hypothèses que les variables X_n hormis qu'elles sont d'espérances nulles. Si l'on établit le résultat voulu dans le cas des variables d'espérances nulles, on peut affirmer que l'événement $\big(S'_n \xrightarrow[n\to+\infty]{} 0\big)$ est presque sûr et donc $\big(S_n \xrightarrow[n\to+\infty]{} \mu\big)$ l'est aussi.

(b) La variable S_n est d'espérance nulle et admet un moment d'ordre 2. L'inégalité de Bienaymé-Tchebychev donne

$$\mathrm{P}\big(|S_n| > \varepsilon\big) = \mathrm{P}\big(|S_n - \mathrm{E}(S_n)| > \varepsilon\big) \leqslant \frac{\mathrm{V}(S_n)}{\varepsilon^2}.$$

Puisque les variables $X_1, \ldots, X_n$ sont deux à deux indépendantes

$$\mathrm{V}(S_n) = \frac{1}{n^2}\,\mathrm{V}(X_1 + \cdots + X_n) = \frac{1}{n^2}\big(\mathrm{V}(X_1) + \cdots + \mathrm{V}(X_n)\big) = \frac{\sigma^2}{n}.$$

On en déduit la comparaison demandée.

(c) Pour tout $\varepsilon > 0$, la série de terme général $\mathrm{P}\big(|S_{m^2}| > \varepsilon\big)$ est convergente et la dernière question du sujet 34 p. 428 assure que la suite (S_{m^2}) converge presque sûrement vers la variable constante égale à 0.

(d) La variable T_n est d'espérance nulle et admet un moment d'ordre 2. L'inégalité de Bienaymé-Tchebychev donne

$$\mathrm{P}\big(|T_n| > \varepsilon\big) \leqslant \frac{\mathrm{V}(T_n)}{\varepsilon^2} \quad \text{avec} \quad \mathrm{V}(T_n) = \frac{1}{n^2} \sum_{k=m^2+1}^{n} \mathrm{V}(X_k) = \frac{n-m^2}{n^2}\sigma^2.$$

Puisque m désigne la partie entière de $\sqrt{n}$, on a $n < (m+1)^2$ et donc

$$n - m^2 < 2m + 1 \quad \text{puis} \quad n - m^2 \leqslant 2m \leqslant 2\sqrt{n}.$$

On obtient ainsi l'inégalité voulue.

(e) Grâce à l'exposant $3/2 > 1$, on peut affirmer la convergence de la série de terme général $\mathrm{P}\big(|T_n| > \varepsilon\big)$ pour tout $\varepsilon > 0$. La suite (T_n) converge donc presque sûrement vers 0.

Les deux événements

$$A = \Big(S_{m^2} \xrightarrow[m\to+\infty]{} 0\Big) \quad \text{et} \quad B = \Big(T_n \xrightarrow[n\to+\infty]{} 0\Big)$$

sont alors quasi certains et leur intersection l'est aussi.

méthode

‖ On montre que $\big(S_n(\omega)\big)$ tend vers 0 pour toute issue ω de $A \cap B$.

Soit ω élément de $A \cap B$. On a à la fois

$$S_{m^2}(\omega) \xrightarrow[m\to+\infty]{} 0 \quad \text{et} \quad T_n(\omega) \xrightarrow[n\to+\infty]{} 0.$$

Or

$$S_n(\omega) = \frac{m^2}{n} S_{m^2}(\omega) + T_n(\omega) \quad \text{avec} \quad m = \lfloor \sqrt{n} \rfloor .$$

Le terme m^2/n est borné par 1 et les deux autres termes sont de limites nulles quand n tend vers l'infini. On peut donc affirmer que la suite $\big(S_n(\omega)\big)$ tend vers 0. Ceci étant vrai pour toute issue ω de l'événement quasi certain $A \cap B$, on peut conclure

$$\mathrm{P}\Big(S_n \xrightarrow[n\to+\infty]{} 0\Big) = 1.$$

Table des matières