École polytechnique de l'Université de Nantes

***É**lectronique et **T**echnologies **N**umériques*
Troisième année
2009-2010

Accueil scientifique

Algèbre

Laurent Guillopé

Laboratoire de mathématiques Jean Leray
Département de mathématiques, UFR Sciences et techniques
Université de Nantes

TABLE DES MATIÈRES

PROLOGUE

Les étudiants abordant la formation d'ingénieurs à Polytech'Nantes sont d'origines diverses et ont des bases hétérogènes en mathématiques. Ce cours a pour but d'exposer les bases minimales d'algèbre linéaire nécessaires pour suivre les cours liés notamment à l'*analyse fonctionnelle*, aux *transformations de Fourier et de Laplace*, au *traitement du signal et de l'image*, aux *méthodes numériques*, à la *recherche opérationnelle* et aux *techniques de modélisation* : le point de vue linéaire est présent à profusion dans bien d'autres domaines comme les *statistiques*, la *physique*, etc.

Un chapitre préliminaire expose, avec force exemples, les rudiments de théorie des ensembles, complétés par quelques structures algébriques ou combinatoires. Le cœur du cours est consacré aux matrices, à leurs liens avec les structures géométriques linéaires et leur diagonalisation, avant de conclure par un court chapitre sur les propriétés euclidiennes (cadre généralisé plus tard dans les espaces hilbertiens et les approximations quadratiques des fonctions).

À l'issue de ce cours, l'étudiant saura
– manipuler les matrices ;
– faire le lien entre applications linéaires et matrices ;
– résoudre des systèmes linéaires ;
– diagonaliser une matrice.

Ces notions sont utilisées dans de multiples applications, dont deux sont traitées dans ce cours : l'étude de suites vérifiant des relations de récurrences linéaires, les équations différentielles à coefficients constants, deux applications illustrées dans ce prologue par la matrice de Fibonacci diagonalisée

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{5}-1}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \\ \frac{\sqrt{5}+1}{2\sqrt{5}} & -\frac{1}{2\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

permet d'établir simplement la forme générale de suites $\mathbf{u}$ vérifiant une récurrence linéaire ou de fonctions y satisfaisant une équation différentielle linéaire

$\mathbf{u} = (u_n)_{n\in\mathbb{N}}$	$u_{n+2} = u_{n+1} + u_n$	$u_n = U\left(\frac{1+\sqrt{5}}{2}\right)^n + V\left(\frac{1-\sqrt{5}}{2}\right)^n, \quad n \geq 0$
$y(t), t \in \mathbb{R}$	$y''(t) = y'(t) + y(t)$	$y(t) = A\exp\left(\frac{1+\sqrt{5}}{2}t\right) + B\exp\left(\frac{1-\sqrt{5}}{2}t\right), \quad t \in \mathbb{R}$

Nantes, le 30 août 2009
Laurent Guillopé

Certains passages de ce cours sont dans une police réduite ou en note ; ils peuvent être négligés sans nuisance lors de la première lecture. Dans la version en ligne, des liens hypertextes prolongent l'index en renvoyant aux notices biographiques du *MacTutor history of mathematics archive* (Université de St Andrews, Écosse) pour les mathématiciens cités dans le texte.

www.math.sciences.univ-nantes.fr/˜guillope/etn3-alg/

CHAPITRE 1

ENSEMBLES ET STRUCTURES ALGÉBRIQUES

Au début du XXe siècle, les mathématiciens ont dégagé les notions d'ensemble et d'application, puis d'opération et de structure (suivant différents modes : algébrique, topologique, différentiable,...). Puis, dans la dernière moitié du XXe siècle, les informaticiens ont développé la programmation structurée, basée sur des structures de données, plus ou moins complexes par leur taille, leur type et leurs attributs, des procédures (ou fonctions, méthodes) et des algorithmes : s'est dégagé l'intérêt de classer en grandes structures les données manipulées par les ordinateurs, malgré leur grande diversité (tableaux, graphes, arbres, piles, files, fonctions, objets,...). Ces structures sont utilisées bien au-delà des seules mathématique et informatique

Au delà de quelques rappels de théorie des ensembles, ce chapitre a pour objectif de balayer quelques grandes structures algébriques (groupe, corps, espace vectoriel), en insistant tout particulièrement sur celles qui ont des applications en analyse, avec quelques indications sur certaines intervenant aussi en informatique.

1.1. Ensembles et parties

Un *ensemble* E (on parle aussi de *famille*, *univers*, *espace*, *système*, ...) est une collection d'objets, les *éléments* de l'ensemble E : cette collection est définie par une énumération ou une propriété qui caractérise ses éléments.

Voilà quelques exemples d'ensembles remarquables, d'autres seront vus dans la suite de ce cours :

▷ ***Exemples 1.1.***

1. l'ensemble (fini) $\{2, 3, 5, 7, 11, 13, 17, 23, 29, 31\}$ des dix premiers entiers premiers [1] ;
2. l'ensemble $I_n = \{1, 2, \dots, n\}$ des n premiers nombres naturels (non nuls), l'ensemble $\mathbb{N}$ [2] $= \{0, 1, 2, \dots\}$ des entiers naturels, l'ensemble $\mathbb{N}^* = \{1, 2, \dots\}$ des nombres entiers naturels non nuls [3] ;
3. l'ensemble $\overline{\mathbb{N}} = \{0, 1, 2, \dots, \infty\}$ des nombres entiers naturels, complété par un élément ∞ dit *infini* ;

1. Un entier non nul est *premier* s'il a exactement deux diviseurs distincts, 1 et lui-même
2. $\mathbb{N}$ pour *nombres*.
3. Dans les ouvrages anglais, $\mathbb{N}$ désigne l'ensemble des entiers naturels non nuls, alors que l'ensemble des entiers naturels positifs ou nuls est noté $\mathbb{N}_0$. Ce n'est qu'une convention, qu'il est important parfois de garder à l'esprit.

4. l'ensemble $\mathbb{Z}$ [4] $= \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ de tous les nombres entiers, positifs et négatifs ;
5. l'ensemble $\mathbb{D}$ des chiffres décimaux :

$$\mathbb{D} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0.1, 3.1456, 2.71828, \pm n/10^p, \ldots\}$$

   où n et p sont des entiers naturels ;
6. l'ensemble $\mathbb{Q}$ [5] des nombres rationnels, *i. e.* tout nombre que l'on peut mettre sous forme de fraction de nombres entiers. Ainsi $\mathbb{Q} = \{\frac{x}{y}, x \in \mathbb{Z}$ et $y \in \mathbb{N}^*\}$;
7. l'ensemble $\mathbb{R}$ de tous les nombres réels ;
8. l'ensemble $\mathbb{C}$ de tous les nombres complexes ;
9. l'ensemble $C = \{a, b, \ldots, A, B, \ldots, +, ",", <, \ldots\}$ des caractères manipulables par un ordinateur, défini par une propriété commune, celle d'apparaître dans l'alphabet des signes gérés par une famille d'ordinateurs [6] ;
10. L'ensemble des polynômes Poly_∞ à coefficients réels et de degré quelconque

$$\mathrm{Poly}_\infty = \{a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0 ; a_n, a_{n-1}, \ldots, a_1, a_0 \text{ réels}, \ n \text{ entier}\}$$

   On peut aussi considérer l'ensemble $\mathrm{Poly}_\infty^{\mathbb{C}}$ des polynômes à coefficients complexes ;
11. L'ensemble, noté diversement par $\mathbb{F}_2, \mathbb{Z}_2, \mathbb{L}, \mathbb{V}, \ldots$ à deux éléments

$$\mathbb{F}_2 = \mathbb{Z}_2 = \{0, 1\} \text{ ou } \mathbb{L} = \{\texttt{oui}, \texttt{non}\} \text{ ou } \mathbb{V} = \{\texttt{vrai}, \texttt{faux}\}. \qquad \triangleleft$$

Une *partie*, ou *sous-ensemble*, A de E est un ensemble dont les éléments sont dans E.

L'*ensemble vide* $\emptyset$ ne comporte aucun élément et est partie de tout ensemble E. Il n'y a qu'un *ensemble vide*, même s'il a une infinité de présentations : par exemple,

$$\begin{aligned}\emptyset &= \{x | x \text{ est consonne du mot } oui\}\\ &= \{n \text{ entier naturel et } n < 0\}\\ &= \{x \text{ réel et } x^2 = -1\}\end{aligned}$$

Le symbole $\in$ définit l'*appartenance* d'un élément à un ensemble, $\notin$ signifie la non-appartenance.

$\triangleright$ ***Exemples 1.2.***

1. Si $a = 0.3333$ et $\mathbb{D}$ est l'ensemble des décimaux défini ci-dessus, alors $a \in \mathbb{D}$ vu que $a = 3333/10^4$. Mais $\widetilde{a} = 1/3 \notin \mathbb{D}$ [7].
2. Le cercle $\{(x, y) | x^2 + y^2 = 1\}$, partie du plan $P = \{(x, y), x, y \in \mathbb{R}\}$, peut être présenté comme la partie $\{|z| = 1\}$ de l'ensemble des nombres complexes $\mathbb{C}$. $\triangleleft$

Étant donné l'ensemble A, une *partie* ou *sous-ensemble* de A est un ensemble constitué d'éléments de A : on note par $\mathcal{P}(A)$ l'ensemble des parties de l'ensemble A, y compris l'ensemble vide et lui-même.

---

4. $\mathbb{Z}$ pour *Zahlen*, *i. e.* nombres en allemand.
5. $\mathbb{Q}$ pour *quotients*.
6. Cet ensemble varie : de l'ensemble $\mathcal{C}_7$ des caractères ASCII 7 bits des années 70 à l'ensemble $\mathcal{C}_{16}$ des caractères UNICODE codés sur 16 bits d'aujourd'hui, en passant par $\mathcal{C}_8$ des caractères ISOLATIN-1 (ASCII 8 bits) des années 90.
7. Sinon, on aurait $1/3 = n/10^p$, soit l'égalité entre entiers $10^p = 3n$. Mais l'entier 3 ne divise aucune puissance de 10 !

▷ ***Exemple 1.3.*** Si $A = \{1, 2, 3\}$, l'ensemble $\mathcal{P}(A)$ de ses parties compte 8 éléments :
$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}. \quad \triangleleft$$

1.2. Opérations sur les ensembles et leurs parties

Considérons deux ensembles A et B : on peut définir les relations

- $\boxed{\textit{Égalité}}$ deux ensembles sont égaux, si tout élément de l'un est élément de l'autre ;
- $\boxed{\textit{Inclusion}}$ A est inclus dans B, et on écrit $A \subset B$, si $x \in A$ implique $x \in B$; dans ce cas, A est sous-ensemble ou une partie de B. La relation d'inclusion est une relation d'ordre (*cf.* 1.3.1), en particulier elle est transitive : si $A \subset B$ et si $B \subset C$, alors $A \subset C$ et antisymétrique : si $A \subset B$ et $B \subset A$, alors $A = B$.

▷ ***Exemples 1.4.***

1. On a $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ et $\mathbb{N} \subset \overline{\mathbb{N}}$.
2. Notons par Poly_k l'ensemble des polynômes à coefficients réels de degré au plus k
$$\mathrm{Poly}_k = \{a_k X^k + a_{k-1} X^{k-1} + \ldots + a_1 X + a_0 ; a_k, a_{k-1}, \ldots, a_1, a_0 \text{ réels}\}$$
C'est une partie de l'ensemble Poly_∞.
3. Reprenant l'exemple des caractères manipulables par un ordinateur, on a
$$\mathcal{C}_7 \subset \mathcal{C}_8 \subset \mathcal{C}_{16}. \quad \triangleleft$$

On peut définir les opérations :

- $\boxed{\textit{Intersection}}$ l'ensemble C est l'intersection de A et de B, si tout élément de C appartient à A et à B : $C = A \cap B = \{x | x \in A \text{ et } x \in B\}$;
- $\boxed{\textit{Union}}$ l'ensemble C est l'union de A et de B, si tout élément de C appartient à A ou à B : $C = A \cup B = \{x | x \in A \text{ ou } x \in B\}$;
- $\boxed{\textit{Différence}}$ l'ensemble C est la différence de A et de B si tout élément de C appartient à A sans appartenir à B : $C = A \setminus B = \{x | x \in A \text{ et } x \notin B\}$;
- $\boxed{\textit{Différence symétrique}}$ l'ensemble C est la différence symétrique de A et de B si tout élément de C appartient à A ou à B, sans appartenir à l'autre :
$$C = A \Delta B = \{x | (x \in A \text{ et } x \notin B) \text{ ou bien } (x \notin A \text{ et } x \in B)\};$$
- $\boxed{\textit{Complément}}$ Le complément de la partie A de U, noté $\overline{A}$, est la partie [8] des éléments de U n'appartenant pas à A : $\overline{A} = \{x \in U | x \notin A\}$.

On note que ces quatre opérations sont liées aux opérations logiques `ET`, `OU`, `Négation`, `OU exclusif` bien connues : *Union* et `OU` ; *Intersection* et `ET` ; *Différence symétrique* et `OU exclusif` ; *Complément* et `Négation`.

Une *partition* $(A_{i \in I})$ d'un ensemble A est une subdivision de cet ensemble en parties [9] non vides, 2 à 2 disjointes :

$$A = \cup_{i \in I} A_i, \quad A_i \text{ et } A_j \text{ disjoints si } i \neq j. \tag{1}$$

▷ ***Exemples 1.5.***

8. La notation $\overline{A}$ n'indique pas l'ensemble U, dont la nature est en général contextuellement sans ambiguïté : il est parfois utile de bien préciser U : le complémentaire de $\mathbb{N}$ dans $\mathbb{Z}$ est distinct du complémentaire du même $\mathbb{N}$ considéré comme partie de $\mathbb{R}$.

9. Dans (1), l'ensemble d'indice I peut être infini,, ainsi de la partition par des singletons d'une ensemble infini E : $(\{e\}, e \in E)$; : s'il est fini, il peut-être pris égal $I_n = \{1, \ldots, n\}$ et $A = A^n_{i=A}$.

1. $\{\{1\},\{2,3\}\}$ et $\{\{1\},\{2\},\{3\}\}$ sont deux partitions de l'ensemble $A=\{1,2,3\}$.
2. Soit $D \in \mathbb{N}$ non nul. Les parties $D_k=\{k+Dz, z \in \mathbb{Z}\}$ pour $k=0,1,\ldots,D-1$ constituent une partition de $\mathbb{Z}$ en D parties. $\triangleleft$

Le *produit cartésien* $A \times B$ des ensembles A et B est constitué de toutes les paires ordonnées de leurs éléments respectifs :

$$A \times B=\{(x,y); x \in A, y \in B\}.$$

Si $A_1,\ldots,A_n$ sont n ensembles, le produit cartésien $A_1 \times A_2 \times \ldots \times A_n$ est défini comme l'ensemble de n-uplets

$$A_1 \times A_2 \times \ldots \times A_n=\{(x_1,x_2,\ldots,x_n); x_i \in A_i, i=1,\ldots,n\}.$$

On écrit A^2 pour $A \times A$ et A^n pour $A \times \ldots \times A$ (n facteurs).

$\triangleright$ ***Exemples 1.6.***

1. Le produit cartésien de $A=\{a,b\}$ et de $B=\{1,2,3\}$ est l'ensemble à 6 éléments :
$$A \times B=\{(a,1),(a,2),(a,3),(b,1),(b,2),(b,3)\};$$
2. L'espace $\mathbb{R}^n$ est l'ensemble des n-uplets ou n-suites $(x_1,x_2,\ldots,x_n)$ de réels. $\triangleleft$

On peut démontrer simplement les résultats suivants où A et B sont deux parties d'un ensemble U :

– ***Commutativité de l'union et de l'intersection***
$$A \cap B=B \cap A, \quad A \cup B=B \cup A$$
– ***Associativité de l'union et de l'intersection***
$$A \cap (B \cap C)=(A \cap B) \cap C, \quad A \cup (B \cup C)=(A \cup B) \cup C$$
– ***Double distributivité de l'union et de l'intersection***
$$A \cap (B \cup C)=(A \cap B) \cup (A \cap C, \quad A \cup (B \cap C)=(A \cup B) \cap (A \cup C)$$
– ***Idempotence*** $A \cap A=A, \quad A \cup A=A$
– ***Loi de Morgan*** : $\overline{A \cap B}=\overline{A} \cup \overline{B}, \quad \overline{A \cup B}=\overline{A} \cap \overline{B}$
– ***Neutralité de l'intersection*** $A \cup \emptyset=A, \quad A \cap U=A$
– ***Absorption de l'union*** $A \cup U=U, \quad A \cap \emptyset=\emptyset$
– ***Complémentarités*** $A \cup \overline{A}=U, \quad A \cap \overline{A}=\emptyset$; $\overline{\overline{A}}=A$

1.3. Relations, fonctions et opérations

1.3.1. Relations binaires. — Soient deux ensembles A et B, une *relation*, dite *binaire* et notée $\mathcal{R}$, entre A et B (ou plus précisément de A vers B) associe à tout couple (a,b), élément du produit cartésien $A \times B$, la propriété `vrai`, auquel cas l'élément a est en relation avec l'élément b, ou `faux`, et dans ce cas a n'est pas en relation avec b : la définition d'une relation $\mathcal{R}$ est ainsi équivalente à la donnée de la partie, notée encore R, du produit cartésien $A \times B$ constituée des couples (a,b) en relation. Lorsque les deux ensembles A et B sont identiques, on parle de *relation interne* à l'ensemble A.

▷ ***Exemples 1.7.***

1. Soit $\mathcal{F}$ une famille, à plusieurs générations éventuellement. Au sein de l'ensemble F des membres de la famille $\mathcal{F}$, on a les relations *être frère/soeur*, *être enfant de* ou encore *être parent de*.
2. Soient les deux ensembles de mots
$$A = \{\text{ici}, \text{là}, \text{loin}, \text{près}\}, \quad B = \{\text{tôt}, \text{tard}, \text{demain}, \text{hier}\}.$$
La relation $\mathcal{R}$ *a même nombre de lettres* a comme couple d'éléments en relation
$$(\text{ici,tôt}), (\text{loin,tard}), (\text{loin,hier}), (\text{près,tard}), (\text{près,hier})$$
tous les autres couples de mots
$$(\text{ici,tard}), (\text{ici, demain}), (\text{ici,hier}), (\text{là,tôt}), \ldots$$
n'étant pas en relation.
3. Dans $\mathbb{N}$, la propriété *est diviseur de* définit une relation : $2\mathcal{R}3$ est fausse, mais $2\mathcal{R}4$ est vraie. ◁

Une représentation d'une relation sur des ensembles finis peut être fournie par un tableau T prenant les valeurs 0 (pour `faux`) et 1 (pour `vrai`). Conventionnellement, les indices des éléments du premier ensemble sont les indices de lignes et ceux du deuxième les indices de colonnes. Le tableau d'une relation sur le produit cartésien $A \times B$ se définit donc pour $a \in A$ et $b \in B$: $T_{ab} = 1$ si $a\mathcal{R}b$, $T_{ab} = 0$ sinon. Une représentation avec un graphe (*cf.* la section 1.5) est parfois plus visuelle ; une relation sur des ensembles infinis peut aussi être codée par un tableau matriciel (avec des ensembles d'indice infini).

▷ ***Exemples 1.8.***

1. La relation *est diviseur de* de l'ensemble $\{1, 2, 3, 4, 5, 6\}$ vers l'ensemble $\{0, 1, \ldots, 9, 10\}$ a pour tableau représentatif

	0	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1	1
2	1	0	1	0	1	0	1	0	1	0	1
3	1	0	0	1	0	0	1	0	0	1	0
4	1	0	0	0	1	0	0	0	1	0	0
5	1	0	0	0	0	1	0	0	0	0	1
6	1	0	0	0	0	0	1	0	0	0	0

2. La même relation interne sur $\mathbb{N}$ est représenté par un tableau infini qui recouvre le *quadrant* SE. ◁

Une relation $\mathcal{R}$ interne à un ensemble E peut avoir diverses propriétés :
- ***Réflexivité*** tout élément e est en relation avec lui-même ($e\mathcal{R}e$) ;
- ***Antiréflexivité*** aucun élément e de E n'est en relation avec lui-même ;
- ***Symétrie*** si e est en relation avec f ($e\mathcal{R}f$), alors f l'est avec e ($f\mathcal{R}e$) ;
- ***Antisymétrie*** si e est en relation avec f ($e\mathcal{R}f$) et f en relation avec e ($f\mathcal{R}e$), alors $e = f$;
- ***Transitivité*** si e, f, g sont tels que $e\mathcal{R}f$ et $f\mathcal{R}g$, alors $e\mathcal{R}g$.

▷ ***Exemples 1.9.***

1. La relation *être diviseur de* sur les entiers naturels est réflexive, antisymétrique (si a divise b et si b divise a alors $a = b$), non symétrique et transitive.
2. La relation *avoir le même âge que* au sein d'un groupe d'individus est réflexive, symétrique et transitive.
3. Soit $D \in \mathbb{N}^*$. La division euclidienne par D écrit tout entier n sous la forme $n = qD + r$ de manière unique, avec $q \in \mathbb{Z}$ et $r \in \{0, 1, \dots, D-1\}$: on dit qe r est le reste de la division de n par D. Dans l'ensemble $\mathbb{Z}$, la relation *a même reste dans* $\{0, 1, \dots, D-1\}$ *par la division par* D est une relation réflexive, symétrique et transitive. ◁

On a deux grandes classes de relations

– **Relations d'équivalence** Une relation $\mathcal{R}$ sur un ensemble E est une relation d'équivalence si elle possède les trois propriétés de réflexivité, symétrie et transitivité.

– **Relation d'ordre** Une relation $\mathcal{O}$ sur un ensemble E est une relation d'ordre si elle possède les trois propriétés de réflexivité, antisymétrie et transitivité.

▷ ***Exemples 1.10.***

1. La relation d'égalité $(=)$ sur un ensemble quelconque est une relation d'équivalence. Les relations de comparaison sur l'ensemble des nombres réels (ou une de ses parties) $\geq$ ou $\leq$ sont des relations d'ordre.
2. La relation *avoir le même âge que* est une relation d'équivalence.
3. La relation *est diviseur de* sur les entiers naturels $\mathbb{N}$ est une relation d'ordre, mais pas sur l'ensemble des entiers naturels $\mathbb{Z}$ (car 1 et -1 sont diviseurs l'un de l'autre, sans être égaux).
4. Dans l'ensemble $\mathbb{Z}$, la relation *a même reste dans* $\{0, 1, \dots, D-1\}$ *par la division par* D est une relation d'équivalence.
5. Dans l'ensemble des parties $\mathcal{P}(A)$ de A, la relation d'inclusion $\subset$ est une relation d'ordre. ◁

Une relation d'ordre strict est une relation qui est transitive et antiréflexive : la relation de majoration stricte $<$ définit une relation d'ordre strict.

1.3.2. Fonctions. — Une *fonction* [10] $f : A \to B$ associe à tout élément a de A un et un seul élément b de B, élément qui est noté $f(a)$: l'ensemble A est appelé *domaine* (de définition), ou *source*, de f, l'ensemble B en est le *but*, la partie de $A \times B$ constituée des éléments $(a, f(a))$ mis en relation par la fonction f est appelée le *graphe* de la fonction f. Une fonction apparaît ainsi comme un type de relation particulier [11].

▷ ***Exemples 1.11.***

1. Au polynôme X^2 est associé naturellement la fonction $f : x \in \mathbb{R} \to x^2 \in \mathbb{R}$.
2. La fonction $g : x \in \mathbb{R} \to x^2 \in \mathbb{R}^+$ n'a pas même but que la fonction f définie précédemment : les fonctions f et g sont différentes.

10. On parle aussi de *transformation*, de *correspondance*, d'*application*, ...
11. Ou bien : la notion de relation apparaît comme une généralisation de celle de fonction.

3. Quand on parle de fonction *racine carrée* $\sqrt{\ }$, implicitement sa source est $\mathbb{R}^+$ et son but le même $\mathbb{R}^+$: la racine carrée d'un réel négatif n'est pas définie, celle d'un réel positif x est l'unique réel positif r tel que $r^2 = x$.
4. Si S et B sont deux ensembles, notons par $\mathcal{F}(S, B)$ l'ensemble des fonctions de l'ensemble (source) S à valeurs dans l'ensemble (but) B. On peut considérer les ensembles $\mathcal{F}(\mathcal{F}(S, B), B), \mathcal{F}(S, \mathcal{F}(S, B)), \ldots$
5. Soit E un ensemble et $\mathcal{F}(E, \mathbb{R})$ l'ensemble des fonctions de E dans $\mathbb{R}$. Pour A partie de E, notons χ_A la fonction caractéristique de A,*i. e.* la fonction de E dans $\mathbb{R}$ telle que $\chi_A(e) = 1$ si $e \in A$, 0 sinon. On vient donc de définir une fonction de l'ensemble $\mathcal{P}(E)$ des parties de E dans l'ensemble $\mathcal{F}(E, \mathbb{R})$: $A \in \mathcal{P}(E) \to \chi_A \in \mathcal{F}(E)$.
6. L'ensemble $\mathbb{R}^n$ s'identifie à l'ensemble des applications $\mathcal{F}(I_n, \mathbb{R})$ de $I_n = \{1, 2, \ldots, n\}$ dans $\mathbb{R}$: à $(a_1, \ldots, a_n) \in \mathbb{R}^n$ est associée la fonction $i \in \{1, 2, \ldots, n\} \to x_i \in \mathbb{R}$.
7. À un triangle (et de manière générale un polygone) est associé son centre de gravité : cela définit une application de source un ensemble de figures géométriques du plan et de but le plan. ◁

Étant donnée une fonction $f : S \to B$, la principale opération est l'*évaluation* de cette fonction en un point $\sigma \in S$: c'est considérer la valeur $f(\sigma) \in B$.L'*image* de A par f est le sous-ensemble de B, noté $f(A)$, constitué des éléments $f(a)$ où a parcourt le domaine de f

$$f(A) = \{b : b \in B \text{ tel qu'il existe } a \in A\}.$$

Si A' est une partie de A, la *restriction* à A' de la fonction $f : A \to B$ est la fonction $f_{|A'} : A' \to B$ telle que

$$f_{|A'}(a) = f(a), \quad a \in A'.$$

Les fonctions f et $f_{|A'}$ sont différentes, vu qu'elles n'ont pas même ensemble source (on a supposé implicitement $A' \neq A$!).

Dans beaucoup de cas, l'ensemble but B et l'ensemble source A coïncident ; on parlera d'*application* d'un ensemble A dans lui-même.

Une fonction $f : A \to B$ a parfois des propriétés particulières, dont les principales sont

- $\boxed{\textbf{\textit{Injectivité}}}$ pour $a, b \in A$, si $a \neq b$, alors $f(a) \neq f(b)$: les images de deux éléments distincts sont distinctes ; ;
- $\boxed{\textbf{\textit{Surjectivité}}}$ tout élément de B est l'image d'au moins un élément de A ;
- $\boxed{\textbf{\textit{Bijectivité}}}$ la fonction est injective et surjective.

On dit resp. que f est une *injection*, *surjection* ou *bijection*.

▷ ***Exemples 1.12.***

1. L'application $f : \mathbb{Z} \to \mathbb{N}$ qui à n associe $f(n) = 2n$ si $n \geq 0$ et $f(n) = -2n + 1$ si $n < 0$ est injective, mais non surjective : l'entier $k = 1$ n'est l'image d'aucun entier par f. L'application $f : \mathbb{Z} \to \mathbb{N}$ qui à n associe $f(n) = |n|$ est surjective, mais non injective : $k = 1$ et $k = -1$ ont même image.
2. L'application d du produit cartésien $\mathbb{Z} \times \mathbb{N}^*$ sur $\mathbb{Q}$ définie par $d(p, q) = p/q$ pour $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ est surjective, mais non injective : $(2, 1)$ et $(10, 5)$ ont même image (2).
3. Écrivons tout rationnel r de $\mathbb{Q}^*$ sous la forme du quotient p/q avec p et q sans facteur premier commun. L'application $f : p/q \in \mathbb{Q} \cap (0, 1) \to p + q \in \mathbb{N}$ [12] est

12. On note dans ce cours $(0, 1)$ la partie de $\mathbb{R}$ qui est souvent notée suivant $]0, 1[$.

surjective mais non injective : les rationnels $3/1$ et $1/3$ s'appliquent tous deux sur l'entier 4. L'application $f : p/q \in \mathbb{Q} \cap (0,1) \to p + q(q+1) \in \mathbb{N}$ est injective mais non surjective[13].

4. L'application $\varphi : n \in \mathbb{N} \to \{0, 1, \ldots, n\} \in \mathcal{P}(\mathbb{N})$ est injective. On remarque que $m \leq n$ si et seulement si $\varphi(m) \subset \varphi(n)$. $\triangleleft$

Pour une fonction bijective $f : A \to B$, on peut définir de manière unique la *fonction réciproque* $f^{-1} : B \to A$, telle que $f(f^{-1}(b)) = b$ pour $b \in B$ et $f^{-1}(f(a)) = a$ si $a \in A$.

$\triangleright$ ***Exemple 1.13***. La fonction $f : \mathbb{Z} \to \mathbb{N}$ qui à n associe $f(n) = 2n$ si $n \geq 0$ et $f(n) = -2n - 1$ si $n < 0$ est bijective. Son application réciproque g est donnée par $g(2p) = p$ et $g(2p+1) = -p - 1$ où $p \geq 0$. $\triangleleft$

On peut composer une fonction $f : A \to B$ suivie de la fonction $g = B \to C$ pour obtenir une fonction $h : A \to C$ définie par

$$h(a) = g(f(a)), \quad a \in A$$

et notée $g \circ f$ (voire gf en omettant le symbole de composition $> \circ$). Cette opération de composition n'est possible que si le but de f est égal à la source de g. Ainsi, on peut composer deux applications définie sur l'ensemble A

Un ensemble E est *fini* s'il existe un entier n tel que E soit en bijection avec la partie $\{1, 2, \ldots, n\}$ de l'ensemble des entiers naturels $\mathbb{N}$: on dira que E est de *cardinal* n et on écrira $\#E = n$. Un ensemble non fini est dit infini. Un ensemble E infini est dit *dénombrable* s'il est en bijection avec $\mathbb{N}$: les ensembles $\mathbb{N}, \mathbb{D}, \mathbb{Q}$ sont dénombrables, alors que $\mathbb{R}, \mathbb{C}, \mathrm{Poly}_\infty$ ne le sont pas.

1.3.3. Permutations. —

Définition 1.1. On appelle *permutation* de E toute application bijective de E. L'ensemble des permutations sur les n premiers entiers I_n est noté $\mathfrak{S}_n$.

Une permutation change l'ordre d'énumération des nombres de I_n. Ainsi pour $n = 8$, le tableau

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 1 & 3 & 8 & 7 & 5 & 6 \end{pmatrix}$$

représente la permutation $(1 \to 2, 2 \to 4, 3 \to 1, \ldots)$ ou encore $\pi_1(1) = 2, \pi_1(2) = 4, \pi_1(3) = 1, \ldots$. La composition de deux permutations peut se faire suivant la manipulation suivante. On écrit la permutation

$$\pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 3 & 4 & 2 & 6 & 7 & 5 & 8 \end{pmatrix}$$

sous la forme équivalente

$$\pi_2 = \begin{pmatrix} 2 & 4 & 1 & 3 & 8 & 7 & 5 & 6 \\ 3 & 2 & 1 & 4 & 8 & 5 & 6 & 7 \end{pmatrix},$$

13. Étant donné q, un rationnel $r = p/q \in \mathbb{Q} \cap (0,1)$ où p sans facteur premier commun avec q, est envoyé dans l'intervalle $I_q = (q(q+1), q(q+1) + q)$ dont l'extrémité droite $q(q+2)$ est strictement inférieure à l'extrémité gauche $(q+1)(q+2)$ de l'intervalle I_{q+1} : l'injectivité en résulte, ainsi que la non surjectivité puisque l'entier $15 = (q(q+2))_{|q=3}$ n'est pas atteint.

ce qui permet d'obtenir la composée $\pi_2 \circ \pi_1$ en combinant les tableaux décrivant π_1 et π_2

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 1 & 3 & 8 & 7 & 5 & 6 \\ 3 & 2 & 1 & 4 & 8 & 5 & 6 & 7 \end{matrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 2 & 1 & 4 & 8 & 5 & 6 & 7 \end{pmatrix} = \pi_2 \circ \pi_1$$

La permutation identité Id laisse l'ordre des éléments inchangé $=Id : k \to k$. Chaque permutation a une permutation réciproque, qui remet les éléments dans l'ordre initial ; ainsi de

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 1 & 3 & 8 & 7 & 5 & 6 \end{pmatrix} \text{ d'application réciproque } \quad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 1 & 4 & 2 & 7 & 8 & 6 & 5 \end{pmatrix}.$$

Définition 1.2. Une *transposition* de $\mathfrak{S}_n$ est une permutation τ telle que $\tau(k) = k$ sauf pour exactement deux entiers i et j pour lesquels $\tau(i) = j$ et $\tau(j) = i$.

L'ensemble des permutations $\mathfrak{S}_n$ compte $n!$ éléments : pour définir une permutation π de $\mathfrak{S}_n$, on a n choix pour $\pi(1)$, puis $n-1$ choix pour $\pi(2)$ (tout élément de I_n différent de $\pi(1)$), puis $n-2$ choix pour $\pi(3)$; par suite, le nombre de choix pour $\pi \in \mathfrak{S}_n$ est $n(n-1)(n-2)\dots 1 = n!$.

Définition 1.3. Une permutation π est dite *paire* si elle il y a un nombre pair de couple d'entiers j, k de I_n tels que $j < k$ et $\pi(j) > \pi(k)$, *impaire* sinon.

Toute transposition est impaire. On montre qu'une permutation est paire si et seulement si elle est la composée d'un nombre pair de transpositions. La *signature* (parfois nommée simplement *signe*) d'une permutation est défini comme l'entier 1 ou -1 suivant que la permutation est paire ou impaire.

1.3.4. Opérations. —

Définition 1.4. Une *opération* de l'ensemble B sur l'ensemble A est une fonction définie sur le produit $B \times A$ et à valeurs dans A : l'image de $(b, a) \in B \times A$ est notée $b * a$. Si $B = A$ on parle d'opération interne à A.

Le signe d'opération $*$ peut prendre diverses formes : $+, \cdot, \wedge, \vee, \dots$ Cette notion est importante, puisqu'elle inclue les opérations telle qu'addition et multiplication déterminant les grandes structures algébriques que sont les groupes, corps, espaces vectoriels,...

▷ ***Exemples 1.14***.

1. L'addition $+$ et la multiplication $\cdot$ sont des opérations internes dans les ensembles de nombres $\mathbb{N}, \mathbb{Z}, \mathbb{D}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \dots$

$$(a, b) \longrightarrow a + b, \qquad (a, b) \longrightarrow a \cdot b = ab.$$

2. En revanche, dans $\mathbb{Z}^*$, la division n'est pas une opération interne : pour le couple $(1, 4) \in \mathbb{Z}^* \times \mathbb{Z}^*$, le quotient $a : b = 1/4$ n'est pas entier.
3. Sur $\mathbb{F}_2 = \{0, 1\}$, on définit les addition et multiplication suivant les tables

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

Ces opérations vérifient la propriété [14]

$$a(b+c) = ab + ac, \quad a, b, c \in \mathbb{F}_2. \tag{2}$$

4. Si $P = \mathbb{R}^2$ est l'ensemble des vecteurs du plan, le produit du vecteur $v \in P$ par le scalaire réel $\alpha \in \mathbb{R}$ est un vecteur noté αv : on a défini ainsi l'opération externe de multiplication d'un vecteur du plan P par un scalaire $\alpha \in \mathbb{R}$.
5. Si P est un polynôme de Poly_∞ et α un réel, le polynôme αP est obtenu à partir de P en multipliant chaque coefficient de P par α. Cette opération s'étend à l'ensemble des polynômes à coefficients dans $\mathbb{K}$, avec $\mathbb{K} = \mathbb{C}$ ou $\mathbb{F}_2$, α étant pris dans $\mathbb{K}$. ◁

1.4. Monoïdes, groupes, corps

En munissant un ensemble E d'opérations de complexité croissante, on peut définir successivement les structures de :

– **Monoïde** Un monoïde est un ensemble A, doté d'une opération interne notée $*$, qui soit associative au sens où

$$(a * b) * c = a * (b * c), \quad a, b, c \in A$$

et qui admette un *élément neutre*, noté conventionnellement e et dont on montre l'unicité, qui vérifie

$$e * a = a * e = a \text{ pour tout } a \in A.$$

On notera le monoïde $(A, *)$, le signe $*$ de l'opération étant noté additivement $(a+b)$, multiplicativement $(a \times b)$, avec un signe de composition $(a \circ b)$, voire omis (ab).

▷ ***Exemples 1.15.***

1. L'ensemble des entiers naturels $(\mathbb{N}, +)$, forme un monoïde, avec 0 comme élément neutre ; ce même ensemble $\mathbb{N}$, muni de l'opération multiplication $\times$, forme un monoïde avec 1 comme élément neutre. L'opération max qui au couple (a, b) associe $\max(a, b)$ muni $\mathbb{N}$ d'une structure de monoïde avec neutre 0.
2. Soit $\mathcal{A}$ un ensemble de lettres, par exemple les 26 lettres

$$\{\mathrm{a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s,t,u,v,w,x,y,z}\}$$

de l'alphabet romain ou les 24 lettres

$$\{\alpha, \beta, \gamma, \delta, \varepsilon, \eta, \varphi, \iota, \kappa, \lambda, \mu, \nu, o, \pi, \psi, \rho, \sigma, \tau, \upsilon, \theta, \chi, \xi, \zeta, \omega\}$$

de l'alphabet grec. L'ensemble $\mathcal{M}_\mathcal{A}$ des mots construits à partir de l'alphabet $\mathcal{A}$, y compris le mot vide $\emptyset$ de zéro lettre, muni de la concaténation des mots

$$(\texttt{premiermot}, \texttt{secondmot}) \longrightarrow \texttt{premiermotsecondmot}$$

est un monoïde : son neutre est le mot vide $\emptyset$. Ce monoïde est très utilisé en théorie des langages de l'informatique.
3. L'ensemble $\mathcal{F}(A, A)$ des applications de A dans A, muni de la composition, est un monoïde. L'application identique $Id : a \to a$ en est le neutre. ◁

14. Un ensemble A muni d'une loi interne notée additivement $(+)$ et d'une multiplication notée $\cdot$ telles que $(A, +)$ soit un groupe et que $(A, .)$ un monoïde avec élément unité est appelé un *anneau* si la relation (2) est valide.

– **Groupe** Un *groupe* est un monoïde $(A, *)$, dont chaque élément $a \in A$ possède un *inverse* (parfois appelé *opposé*), noté a^{-1}, élément de A tel que $a * a^{-1} = a^{-1} * a = e$ où e est le neutre du monoïde $(A, *)$.

▷ ***Exemples 1.16.***

1. L'ensemble des entiers relatifs $\mathbb{Z}$ avec l'addition $+$ constitue un groupe.
2. L'ensemble des rationnels non nuls $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, muni de la multiplication forme le groupe multiplicatif des rationnels $(\mathbb{Q}^*, \times)$ avec neutre $e = 1$.
3. L'ensemble des applications bijectives d'un ensemble E est un groupe. Considérons le cas d'un ensemble fini, plus particulièrement l'ensemble fini ordonné $I_n = \{1, 2, \ldots, n\}$ des n premiers entiers naturels.
   est évidemment neutre pour la composition : l'ensemble des permutations S_N de l'ensemble I_N forme ainsi un monoïde, c'est aussi un groupe car
4. Les groupes de transformations géométriques sont utilisés en imagerie, où les opérations de rotation, symétrie, projection,... interviennent. Par exemple, l'ensemble $\mathcal{R}$ des rotations $R(\theta)$ d'angle θ du plan P : l'élément neutre est la rotation $R(0)$ d'angle nul; le symétrique de la rotation $R(\theta)$ d'angle θ est la rotation d'angle $-\theta$; de plus ce groupe est commutatif, au sens où $R(\theta) \circ R(\varphi) = R(\varphi) \circ R(\theta)$, vu que ces composées sont égales à $R(\theta + \varphi)$ (*cf.* 2.5.2). ◁

– **Corps** Un *corps* est tout ensemble $\mathbb{K}$ muni de deux opérations internes, nommées conventionnellement *addition* $(+)$et *multiplication* $(*)$ telle que monoïde $(\mathbb{K}, +)$ soit un groupe dont le neutre est noté 0, ainsi que $(\mathbb{K}, \setminus\{0\}, *)$ avec les relations de compatibilité

$$a * (b + c) = (a * b) + (a * c), \quad a, b, c \in \mathbb{K}$$

▷ ***Exemples 1.17.***

1. Les ensembles $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont des corps.
2. L'ensemble $\mathbb{Z}$, muni de ses opérations habituelles, n'est pas un corps : 2 n'a pas d'inverse multiplicatif.
3. L'espace Poly_∞ n'est pas un corps : considérant les quotients de deux polynômes, à la manière de $\mathbb{Q}$ qui apparaît comme ensemble des quotients d'éléments de $\mathbb{Z}$, on introduit l'ensemble des fractions rationnelles $R = \{\ P/Q, P, Q \in \mathrm{Poly}_\infty, Q \neq 0\}$ qui est un corps.
4. L'espace $\mathcal{H}(\Omega)$ des fonctions holomorphes sur un ouvert Ω du plan complexe n'est pas un corps : l'espace des $\mathcal{M}(\Omega)$ constitué des quotients f/g avec $f, g \in \mathcal{H}(\Omega)$ est un corps : ses éléments sont appelés *fonctions méromorphes* (elles sont holomorphes sauf sur un ensemble au plus dénombrable de pôles).
5. L'ensemble $\mathbb{F}_2$, avec les lois introduites dans l'exemple 1.14.3 est un corps, le corps à deux éléments. On vérifie que ces opérations satisfont aux mêmes propriétés que les addition et multiplication dans un des ensembles $\mathbb{K}$ de nombres classiques $(\mathbb{R}, \mathbb{C})$, faisant de $(\mathbb{F}_2, +)$ et de $(\mathbb{F}_2 \setminus \{0\}, \cdot)$ des groupes [15] (où les inverses de a sont notés $-a$ et a^{-1} resp..) avec les relations de compatibilité habituelles valides dans les ensembles $\mathbb{R}$ et $\mathbb{C}$ appelés appelés *corps* des nombres réels et complexes resp.

15. *cf.* la section 1.4 pour la définition d'un groupe.

Cet ensemble $\mathbb{F}_2$, appelé corps à deux éléments, est très présent en informatique. $\triangleleft$

– $\boxed{\textit{Espace vectoriel}}$ Un espace vectoriel E est un ensemble dont les éléments sont appelés *vecteurs*, muni d'une opération interne de type addition

$$(v, w) \in E^2 \longrightarrow v + w \in E$$

qui en fait un groupe avec neutre 0_E (noté 0 s'il n'y a pas d'ambiguïté) appelé *vecteur nul* et d'une opération externe de multiplication par un nombre réel [(16)] :

$$(\alpha, v) \in \mathbb{R} \times E \longrightarrow \alpha v \in E.$$

Les propriétés *habituelles*, *i. e.* comme dans le plan $\mathbb{R}^2$, de ces opérations sont satisfaites (*cf.* Section 3.1.1) :

$$\alpha(v + w) = \alpha v + \alpha w, \quad \alpha(\beta v) = (\alpha\beta)v \quad v, w \in E, \alpha, \beta \in \mathbb{R}.$$

$\triangleright$ ***Exemples 1.18.***

1. L'espace des polynômes Poly_∞, l'espace Poly_k des polynômes de degré au plus k, sont des espaces vectoriels.
2. L'espace E des fonctions de A dans $\mathbb{C}$ est un espace vectoriel : si f, g sont dans E et $\alpha \in R$, on définit $f + g$ et αf suivant

$$(f + g)(a) = f(a) + g(a), \quad (\alpha f)(a) = \alpha f(a), \qquad a \in A. \qquad \triangleleft$$

1.5. Graphes et arbres, piles et files

Diverses structures, interprétables de manière algébrique, géométrique ou algorithmique, sont définies sur les ensembles.

– $\boxed{\textit{Graphe}}$ Un *graphe* G est la donnée d'un ensemble de sommets S, avec certaines paires de sommets $\{s, t\}$ joints par une arête : l'ensemble des arêtes A du graphe G s'identifie à la partie du produit S^2 constituée des couples (s, t) et (t, s) correspondant à l'arête $\{s, t\}$. On peut aussi considérer des graphes orientés où une arête a a une extrémité initiale $\alpha(a)$ et une extrémité finale $\omega(a)$ et sont décrites par un couple $(\alpha(a), \omega(a)) \in S^2$.

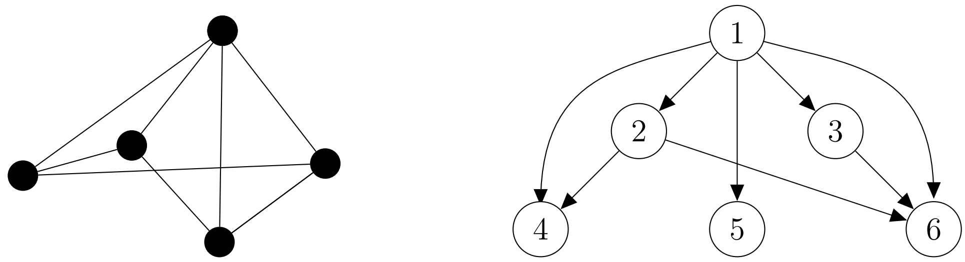
Un chemin c issu du sommet s_1 sur le graphe G est une suite d'arêtes $a_1 a_2 \ldots a_n$ avec $a_i = \{s_i, t_i\}$ telle que $t_i = s_{i+1}, i = 1, \ldots, n-1$, un circuit de G est un chemin $a_1 a_2 \ldots a_n$ tel que $s_1 = t_n$. On peut aussi envisager des chemins orientés, des graphes avec plusieurs arêtes entre deux même sommets et des boucles.

Les graphes et leurs multiples variantes, constituent des ensembles munis d'une relation de précédence (qui vérifie éventuellement les propriétés d'une relation d'ordre), cette relation étant définie sur l'ensemble S des sommets de G par $s\mathcal{R}t$ s'il existe un chemin du sommet s au sommet t.

$\triangleright$ ***Exemples 1.19.***

1. Le réseau des bus et tramways de la métropole nantaise détermine un graphe, avec comme sommets les stations et arêtes les portions de ligne entre deux stations.

16. On parle d'espace vectoriel réel. Si l'opération externe de multiplication est définie sur le corps des complexes $\mathbb{C}$ ou le corps $\mathbb{F}_2$, on parlera de $\mathbb{C}$ ou $\mathbb{F}_2$ espace vectoriel.

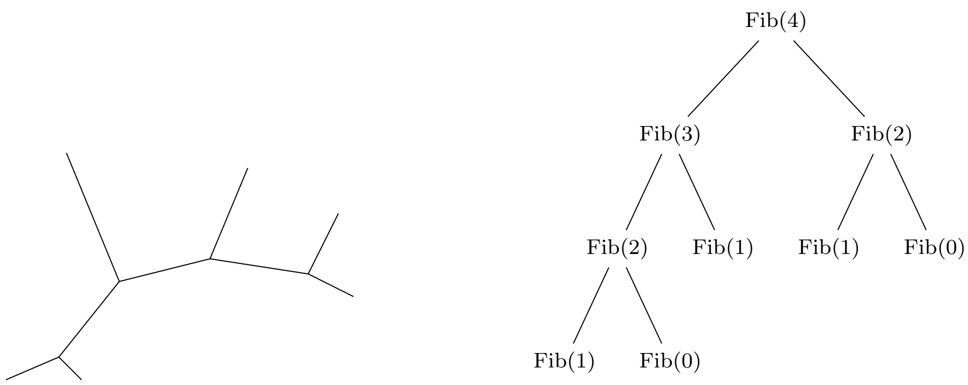


(a) (b)

FIGURE 1 . (a) un graphe (non orienté), (b) le graphe (orienté) de la relation de divisibilité sur $\{1,2,3,4,5,6\}$: il y a une arête du sommet s vers le sommet t si s divise t.

2. Le réseau Internet est un graphe, où les sommets sont les ordinateurs (postes de travail, serveurs, imprimantes, ipod,...) et les arêtes correspondent aux connexions entre ces différents dispositifs.
3. Le réseau d'un opérateur téléphonique est aussi un graphe, avec comme sommets les portables, téléphones fixes ; commutateurs, relais et les arrêtes décrivent les connexions. Ce graphe n'est pas constant au cours du temps, puisque les arrêtes issues d'un mobile varient. ◁

– $\boxed{\textbf{\textit{Arbre}}}$ Un *arbre* est un graphe T qui ne contient pas de circuit. Un arbre enraciné est un arbre, avec un sommet particulier, noté r et appelé *racine* ; il existe alors un chemin unique reliant r à tout autre sommet du graphe G.



(a) (b)

FIGURE 2 . (a) un arbre, (b) l'arbre des appels récursifs de la suite de Fibonacci.

▷ ***Exemples 1.20.***

1. L'arbre généalogique partant d'un ancêtre, muni de la relation "enfant de " : il existe un seul chemin conduisant de l'ancêtre à chacun de ses descendants.
2. Le système de fichiers d'un système LINUX est arborescent, ainsi que celui des processus
3. Une suite de Fibonacci est toute suite qui vérifie

$$u_n = u_{n-1} + u_{n-2}, \quad n \geq 2,$$

déterminée par ses valeurs u_0 et u_1. Le nombre u_n se calcule aisément par le programme récursif

```
def Fib(n)
  if n< 2 :
     return Fib(n)
  return (Fib(n-1)+Fib(n-2))
```

Le calcul récursif de `Fib(4)` est décrit par l'arbre des appels de cette fonction de la Figure 2. ◁

– $\boxed{\textbf{\textit{File et pile}}}$ Les structures de *pile* et *file* sont des exemples de *sacs*, ensembles munis d'opérations d'ajout et d'enlèvement. Ils se distinguent par la relation entre éléments ajoutés et retirés : pour une *pile*, c'est le dernier élément ajouté qui est retiré en premier, alors que pour une *file* c'est le premier élément ajouté qui est retiré.

1.6. Exercices

E1.1. Soit E un ensemble à n éléments. Montrer que l'ensemble des parties $\mathcal{P}(E)$ de E à 2^n éléments.

E1.2. Soit $U = \{1,2,3,4,5,6,7\}$ et ses parties $A = \{1,2,3,4\}, B = \{4,5,6,7\}, C = \{1,3,5,7\}, D = \{2,3,4,6\}$. Calculer $(A \cap B) \cup (C \cap D)$, $(A \cup C) \cap (B \cup D)$ et $\overline{(\overline{A} \cap D)} \cap \overline{B \cup C)}$.

E1.3. Soient $A = \{1,2,3,4,5\}$, $B = \{1,3,5,7,9\}$ et $C = \{2,4,6,8,10\}$ les parties de l'ensemble $\mathbb{N}$. Calculer $A \cup B$, $A \cup C$, $A \setminus C$ et la différence symétrique $A \Delta C$.

E1.4. Soit les fonctions $f : x \in \mathbb{R} \to x^3 - x + 2 \in \mathbb{R}$ et $g : x \in \mathbb{R} \to x^2 - 4 \in \mathbb{R}$. Calculer les fonctions $f \circ g$ et $g \circ f$. ainsi que la fonction $h : x \to f(x)g(x)$.

E1.5. Dessiner les parties suivantes de $\mathbb{R}^2$.

$$\overline{([-1,1] \times [-1,1])}, \quad \overline{[-1,1]} \times \overline{[-1,1]}, \quad \{(a,b) \in \mathbb{Z}^2, a^2 + b^2 \leq 4\}.$$

E1.6. Soit sur $A = [0,1]$ la loi définie par

$$a * b = a + b - ab$$

Vérifier que la loi est bien interne à A, que 0 est élément neutre, que la loi est associative et que tout élément de A a un inverse sauf $a = 1$.

E1.7. Soit $I_5 = \{1,2,3,4,5\}$ et pour i,j deux entiers distincts de I_5 la permutation τ_{ij} de I_5 telle que $\tau_{ij}(i) = j, \tau_{ij}(j) = i$ et $\tau_{ij}(k) = k$ pour k distinct de i et j. Calculer la permutation $\tau_{12}\tau_{23}\tau_{34}\tau_{45}\tau_{51}$.

E1.8. Soit A l'ensemble $A = \{0,1,2,3,4,5\}$. Démontrer que
- la relation a *même parité* est une relation d'équivalence sur A ;
- la relation a *divise* b est une relation d'ordre sur A.

E**1**.9. Soit E l'ensemble des fonctions de A dans $\mathbb{R}$. On définit sur E la relation *est inférieure ou égale* suivant

$$f\mathcal{R}g \text{ si et seulement si } f(a) \leq g(a), \quad a \in A.$$

Montrer que cette relation est une relation d'ordre.
On dit que g *est supérieur ou égal à* f si f est inférieur ou égal à g. Montrer que la fonction $h = \max(f, g)$ définie par

$$h(a) = \max(f(a), g(a)), \quad a \in A,$$

est l'unique élément de E qui soit supérieur à f et à g tout en étant inférieur à tout élément supérieur à f et g. Montrer que $\max(k + f, k + g) = k + \max(f, g)$ et que $\max(f, g) = (f + g + |f - g|)/2$.

E**1**.10. (1) Montrer que l'application

$$(m, n) \in \mathbb{N}^2 \to 2^m 3^n \in \mathbb{N}$$

est injective.
(2) Démontrer que les applications

$$(a, b) \in \mathbb{N}^2 \to |a - b| \in \mathbb{N}$$
$$(p, q) \in \mathbb{Z} \times \mathbb{Z}^* \to p/q \in \mathbb{Q}$$

sont surjectives. Sont-elles injectives ?

E**1**.11. Soit $f : S \to B$ surjective. Montrer que l'ensemble des parties $\{f^{-1}(\{b\}), b \in B\}$ est une partition de S.

CHAPITRE 2

CALCUL MATRICIEL

Il y a deux approches, quasi-orthogonales, pour aborder les propriétés des matrices et leurs applications :

– L'approche numérique ou *calcul matriciel*, où l'on étudie les propriétés des tableaux de nombres que sont les matrices : c'est une approche pragmatique, suffisante pour aborder de nombreux problèmes comme la recherche opérationnelle, la théorie des graphes… Le calcul matriciel est au cœur de ce chapitre.

– L'approche algébrique imprégnée par l'*algèbre linéaire*, où les matrices sont des représentations d'applications linéaires : cette approche permet de mieux comprendre l'origine et la naturalité de propriétés empiriquement *observées* en calcul matriciel. Les bases de l'algèbre linéaire sont abordées dans le chapitre suivant.

2.1. Calcul indiciel

Le calcul matriciel (en algèbre linéaire ou pour la manipulation des tableaux de données en algorithmique) nécessite la manipulation de grandeurs indicées, comme

– les vecteurs, grandeurs [(1)] à un seul indice : $V = (V_i) = (V_1 V_2 \dots V_n)$

– les tableaux bi-dimensionnels ou matrices : $A = (A_{ij})$, où $i \in \{1, \dots, m\}$ est l'indice de ligne et $j \in \{1, \dots, n\}$ celui de colonne ; le tenseur électromagnétisme $(F_{ab} = \partial_a A_b - \partial_b A_a)_{1 \leq a,b \leq 4}$ défini à partir du potentiel électromagnétique $(A_a)_{a=1,\dots,4}$ en relativité restreinte,

– d'autres objets, notamment les tenseurs, qui peuvent avoir plus de 2 indices : le tenseur d'élasticité $(C_{ij}^{vw})_{1 \leq i,j,v,w \leq 3}$ à 81 coefficients, le tenseur de Riemann $(R_{ijk\ell})_{1 \leq i,j,k,ell \leq 4}$ essentiel à la relativité générale,…

2.1.1. Types d'indices. — On distingue deux grandes catégories d'indices :

1. Les indices fixes ont une valeur définie (bien que parfois non connue) ; on ne peut pas changer la valeur d'un indice fixe. Par exemple, dans le raisonnement suivant : *Soit V_k, la k-ième composante d'un vecteur V…*, l'indice k est fixe.

2. Les indices libres ou muets paramètrent des sommations ou des produits ; on peut changer un indice muet sans changer le résultat, à condition de ne pas lui donner le nom d'un autre indice existant.

1. Un informaticien dit *données*.

▷ ***Exemple 2.1.*** Dans l'expression :

$$V_k = \sum_{i=1}^{n} A_{ki} B_i$$

k et n sont des indices fixes ; i est un indice muet ou libre. L'expression précédente ne change pas si l'on fait la transformation i en j :

$$V_k = \sum_{j=1}^{n} A_{kj} B_j.$$

Bien entendu les transformations : $i \to k$ et $i \to n$ sont interdites. ◁

2.1.2. Grandeurs indicées particulières. — Deux symboles particuliers sont très utilisés en calcul indiciel :

- Le symbole de Kronecker δ_{ik} :

$$\delta_{ik} = \begin{cases} 1 & \text{si } i = k, \\ 0 & \text{sinon.} \end{cases}$$

- Le symbole d'antisymétrie $\varepsilon_{ijk\dots}$ qui peut avoir 2, 3, 4,…indices. Il a la valeur nulle dès que 2 indices ont des valeurs égales ; il vaut 1 pour $ijk = 123$ et plus généralement toute permutation paire de ces valeurs ; pour les permutations impaires, il prend la valeur -1. Par exemple, considérons le cas à 3 indices ijk, pour lequel les seules valeurs non nulles sont

$$\varepsilon_{123} = 1, \varepsilon_{312} = 1, \varepsilon_{231} = 1, \varepsilon_{132} = -1, \varepsilon_{213} = -1, \varepsilon_{321} = -1.$$

2.1.3. Opérations sur les grandeurs indicées. — Il y a quatre opérations de base sur les grandeurs indicées :

1. l'affectation et l'enumération donnent les valeurs de chaque composante, comme dans l'expression suivante : $U_i = 2^i, i \in \{0, 1, 2, 3\}$.
2. la sommation est l'opération définie suivant : $\sum_{i=1}^{n} V_i = V_1 + V_2 + V_3 + \ldots + V_n$. La sommation peut porter sur plusieurs indices comme dans $\sum_{i,j=1}^{n} A_{ij} B_{ji}$, somme égale à
$$\sum_{i=1}^{n} \left[\sum_{j=1}^{n} A_{ij} B_{ji} \right] = \sum_{j=1}^{n} \left[\sum_{i=1}^{n} A_{ij} B_{ji} \right]$$
vu les propriétés de commutativité et d'associativité de l'addition.
3. le produit est défini de manière semblable : $\prod_{i=1}^{n} V_i = V_1 V_2 V_3 \cdots V_n$.
4. la contraction supprime, par sommation ou produit, un ou plusieurs indices. Ainsi de l'opération suivante sur la grandeur à trois indices A_{ij}^k : $a_j = \sum_{i=1}^{n} A_{ij}^i$ est une contraction, le résultat ne dépend plus que d'un seul indice.

2.2. Matrices et vecteurs

Définition 2.1. Une *matrice* A est un tableau rectangulaire à m lignes et n colonnes de *coefficients* ou *éléments matriciels* (2) : le coefficient à l'intersection de la i-ème ligne

2. On a mentionné que ces coefficients sont de type *nombre* $(\mathbb{N}, \overline{\mathbb{N}}, \mathbb{Z}, \mathbb{R}, \mathbb{Q}, \ldots)$ le plus souvent dans ce cours, même s'ils peuvent être de types *booléen* (`vrai`/`faux` ou 0/1 comme élément de l'ensemble $\mathbb{V}$ de l'exemple 1.1.11), *chaîne de caractères*, *fonction*, *liste* et même *matrice*. Considérer des coefficients de

$(i = 1, \ldots, m)$ et de la j-ème colonne $(j = 1, \ldots, n)$ sera noté A_{ij}, voire parfois $A(ij)$. La dimension de la matrice A est le couple (m, n) ; si $m = n$, on dira que la matrice est carrée d'ordre n.

$$\begin{pmatrix} A_{11} & \ldots & & \ldots & A_{1n} \\ \vdots & & & & \vdots \\ & & A_{ij} & & \\ \vdots & & & & \vdots \\ A_{m1} & \ldots & & \ldots & A_{mn} \end{pmatrix}$$

Un vecteur ligne L s'identifie à une matrice à 1 ligne et n colonnes, alors que le vecteur colonne C s'identifie à une matrice à m lignes et 1 colonne :

$$L = (\ell_1, \ldots, \ell_n), \qquad C = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}.$$

▷ ***Exemples 2.2.***

1. La matrice nulle, notée toujours par 0 quelle que soit sa dimension est celle dont tous les coefficients A_{ij} sont nuls.
2. Une matrice carrée d'ordre $n = 1$ est identifiée au scalaire qu'est son unique coefficient.
3. Une matrice est diagonale si elle est carrée et a tous ses coefficients non diagonaux $A_{ij}, i \neq j$ nuls.
4. La matrice identité d'ordre n est la matrice diagonale, dont tous les coefficients diagonaux valent 1. Elle est notée I_n.
5. Une matrice carrée U (resp. L) est dite triangulaire supérieure (resp. inférieure) si pour $i > j$ on a $U_{ij} = 0$ (resp. $L_{ji} = 0$). ◁

△ ***Remarque 2.1.*** Une matrice est implicitement un tableau fini. On peut considérer des matrices $A = (A_{ij})_{(i,j) \in I \times J}$ avec des ensembles d'indice I et J quelconques, finis ou infinis : $\mathcal{P}(\{1, \ldots, n\}), \mathbb{N}, \mathbb{R}, \ldots$ Ainsi, une suite $\mathbf{u} = (u_k)_{k \in \mathbb{N}}$ peut être considéré comme une matrice à une ligne et une infinité de colonne, avec $I \times J = \{1\} \times \mathbb{N}$. Ce point de vue ne sera guère utilisé dans ce cours. ▽

L'ensemble $(x_1, \ldots, x_n)$ des composantes d'un vecteur V relativement à une *base* $\mathbf{b} = (b_1, \ldots, b_n)$ (*cf.* Section 3.1.3 ci-après) sera considéré comme une matrice C à une colonne et n lignes.

$$V = x_1 b_1 + \ldots + x_n b_n, \qquad C = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Inversement, la matrice C à une colonne et n lignes

$$C = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

type *tableau* permet de voir une matrice à m lignes et n colonnes comme une liste de m listes (les *lignes*) de n éléments. On peut ainsi facilement introduire des tableaux à longueur de ligne variables, mais nous n'en aurons pas besoin dans ce cours.

s'identifie au vecteur $(x_1, \ldots, x_n)$ de $\mathbb{R}^n$ dont le vecteur colonne des composantes relativement à la base canonique (*cf.* 3.4.1) de $\mathbb{R}^n$ est exactement C : une matrice colonne sera donc parfois considérée soit comme un objet du calcul matriciel, soit comme un objet de l'algèbre linéaire, la cohérence de cette identification étant parfaitement justifiée.

On pourra donc appliquer aux vecteurs colonne les opérations sur les matrices rectangulaires définies par la suite.

2.3. Opérations matricielles

Les opérations de base sur les matrices sont :

- $\boxed{\textit{\textbf{Addition}}}$ Si A, B sont deux matrices de dimension (m, n), leur somme $A + B$ est définie comme la matrice de dimension (m, n) telle que
$$(A + B)_{ij} = A_{ij} + B_{ij}, \quad i \in \{1, \ldots, m\}; j \in \{1, \ldots, n\}.$$
- $\boxed{\textit{\textbf{Multiplication par un scalaire}}}$ Si A est une matrice de dimension (m, n), son produit par le réel α est la matrice αA de dimension (m, n) définie par
$$(\alpha A)_{ij} = \alpha A_{ij} \quad i \in \{1, \ldots, m\}; j \in \{1, \ldots, n\}.$$
- $\boxed{\textit{\textbf{Multiplication matricielle}}}$ Si A est une matrice de dimension (m, n) et B une matrice de dimension (n, p), leur produit AB est la matrice de dimension (m, p) telle que
$$(AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk} \quad i \in \{1, \quad , m\}; k \in \{1, \quad , p\}.$$
- $\boxed{\textit{\textbf{Inversion}}}$ Si elle existe, la matrice inverse, notée A^{-1} de la matrice carrée A d'ordre n est la matrice carrée d'ordre n, qu'on démontre unique, satisfaisant la relation :
$$A^{-1}A = AA^{-1} = I_n.$$
Si l'inverse A^{-1} existe, la matrice A est dite *régulière* ou *inversible*. Une matrice qui n'est pas régulière est dite *singulière*.
- $\boxed{\textit{\textbf{Transposition}}}$ À la matrice A de dimension (m, n) est associée la matrice de dimension (n, m), notée A^t et appelée *transposée*, définie suivant
$$(A^t)_{ij} = A_{ji}, \quad i \in \{1, \ldots, n\}; j \in \{1, \ldots, m\}.$$
La transposée d'un vecteur colonne est un vecteur ligne et réciproquement.
- $\boxed{\textit{\textbf{Produit scalaire}}}$ Le $Y^t X$ pour X, Y deux vecteurs colonnes est un scalaire, dit *produit scalaire* des vecteurs (colonne) X et Y. On dira que les deux vecteurs colonnes X et Y sont orthogonaux si leur produit scalaire est nul. Un vecteur colonne X est dit normé si $X^t X = 1$.

Pour des matrices à coefficients complexes, on a de plus deux opérations :

- $\boxed{\textit{\textbf{Conjugaison}}}$ À la matrice A de dimension (m, n) est associée la matrice notée $\overline{A}$, de même dimension, telle que
$$(\overline{A})_{ij} = \overline{A_{ij}}, \quad i \in \{1, \ldots, m\}; j \in \{1, \ldots, n\}.$$
- $\boxed{\textit{\textbf{Adjonction}}}$. À la matrice A de dimension (m, n) et à coefficients complexes, est associée la matrice, notée A^* et appelée *adjointe*, de dimension (n, m) définie suivant
$$(A^*)_{ij} = \overline{A_{ji}}, \quad i \in \{1, \ldots, n\}; j \in \{1, \ldots, m\}.$$

L'adjointe d'une matrice à coefficients réels n'est rien d'autre que sa transposée : si la matrice A a un coefficient complexe non réel, les matrices transposée A^t et adjointe A^* diffèrent.

△ ***Remarque 2.2.*** On prendra garde à la non commutativité du produit matriciel : si A et B ne sont pas carrées, pour que AB et BA soient définies, il faut que B soit de dimension (n,m) si A est de dimension (m,n) : AB est carrée d'ordre n, alors que BA est carrée d'ordre m. Ainsi, si X est un vecteur colonne de m lignes et $\widetilde{Y}$ un vecteur ligne de m colonnes : le produit $\widetilde{Y}X$, bien défini, est un scalaire (on a identifié une matrice carrée d'ordre 1 à son unique coefficient), alors que $X\widetilde{Y}$ est une matrice carrée d'ordre n. Même si A et B sont carrées de même ordre au moins égal à 2, on n'a pas $AB = BA$ en général, ainsi par exemple

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & 3 \\ 0 & 2 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}. \qquad \triangledown$$

Les deux premières opérations (addition et multiplication par un scalaire) font de l'ensemble des matrices de dimension (m,n) un espace vectoriel (*cf.* définition 1.4). Muni des opérations d'addition et de multiplication des matrices, l'espace des matrices carrées d'ordre n vérifie la propriété (2) : c'est donc un anneau (*cf.* note 14).

Le type matriciel, à coefficients réels ou complexes, avec les opérations précédentes (et bien d'autres...), sont implémentés dans les logiciels de calcul courants : `Maple`, `Matlab`, `Scilab`,....

2.4. Caractéristiques des matrices

Aux opérations précédentes sur les matrices, il faut ajouter les *caractéristiques* d'une matrice carrée A d'ordre n : ce sont des nombres, les principaux introduits dans ce cours étant :

- **_Trace_** c'est la somme des éléments de la diagonale principale : $\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}$
- **_Déterminant_** c'est la valeur en A, considérée comme une collection de n vecteurs lignes ou n vecteurs colonnes, d'une forme multilinéaire sur l'espace des n-uplets de vecteurs, qui se ramène pour les matrices diagonales au produit de ses éléments diagonaux : si $A_{ij} = 0$ lorsque $i \neq j$, alors $\text{dét}\, A = \prod_{i=1}^{n} A_{ii}$.
- **_Valeurs et vecteurs propres_** λ est une *valeur propre*, avec *vecteur propre* associé V (matrice de dimension $(1,n)$) si $AV = \lambda V$.

2.5. Familles de matrices particulières

2.5.1. Matrices en relation avec son adjointe. — On peut classer les matrices selon leurs propriétés par rapport aux diverses opérations définies plus haut :

- **_Matrice symétrique_** la matrice S est symétrique si $S = S^t$;
- **_Matrice antisymétrique_** la matrice A est antisymétrique si $A = -A^t$;
- **_Matrice hermitienne_** la matrice H est hermitienne, si $H = H^*$;
- **_Matrice unitaire et orthogonale_** la matrice carrée Q est unitaire, si son adjointe est égale à son inverse : $Q^* = Q^{-1}$. Pour les matrices réelles, une matrice unitaire, qualifiée d'orthogonale, est une matrice régulière vérifiant $Q^t = Q^{-1}$;
- **_Matrice idempotente_** une matrice A est idempotente si $A^2 = A$.

2.5.2. Matrices de transformations géométriques du plan. — Les opérations basiques de manipulation d'images (objets représentés par des parties du plan) correspondent à des types particuliers de matrices, dont la forme sera justifiée dans le chapitre sur les applications linéaires (*cf.* chapitre 3).

– $\boxed{\textit{Rotation}}$ La matrice de rotation d'angle θ d'un vecteur dans le plan s'écrit

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Les matrices de rotation sont orthogonales, on vérifie :

$$R^{-1}(\theta) = R(\theta)^t = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Le produit de deux matrices de rotation est commutatif

$$\begin{aligned} R(\theta)R(\varphi) &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) \end{pmatrix} \\ &= R(\theta+\varphi) = R(\varphi)R(\theta). \end{aligned}$$

– $\boxed{\textit{Projection}}$ Pour la droite D_θ passant par l'origine et faisant un angle θ avec l'axe Ox, la projection sur la droite D_θ a pour matrice

$$P(\theta) = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

Elle vérifie $P(\theta)^2 = P(\theta)$ et est donc idempotente. Cette matrice n'est pas inversible.

2.5.3. Matrices par blocs. — Terminons ce catalogue par le cas où le type des éléments matriciels est une matrice (disons à coefficients scalaires pour simplifier), qui apparaissent dans les décompositions par blocs. Une matrice $B = (B_{ij})$ par blocs est une matrice dont le coefficient B_{ij} est une matrice de dimension (m_i, n_j) : développant ces blocs matriciels, on obtient une matrice à coefficients scalaires de dimension $(M, N) = (\sum_{i=1}^m m_i, \sum_{j=1}^n n_j)$.

▷ ***Exemple 2.3.*** La matrice B carrée d'ordre 2 à coefficients blocs carrés d'ordre 2 est une matrice carrée d'ordre 4 et inversement toute matrice carrée d'ordre 4 peut être considérée comme matrice carrée de blocs carrés.

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} b_{1111} & b_{1112} & b_{1211} & b_{1212} \\ b_{1121} & b_{1122} & b_{1221} & b_{1222} \\ b_{2111} & b_{2112} & b_{2211} & b_{2212} \\ b_{2111} & b_{2112} & b_{2211} & b_{2212} \end{pmatrix}$$

où

$$B_{ij} = \begin{pmatrix} b_{ij11} & b_{ij12} \\ b_{ij21} & b_{ij22} \end{pmatrix}, \quad i, j = 1, 2.$$ ◁

Une matrice carrée d'ordre 2^n peut être considérée comme une matrice carrée d'ordre 2 à coefficients matrices carrées d'ordre 2 à coefficients matrices carrées d'ordre 2… : cette forme est utilisée dans les calculs de transformation de Fourier rapide.

On peut parler de matrice diagonale par blocs : c'est ainsi qu'est formulée la décomposition de Jordan (*cf.* Section 5.2), forme générale de la diagonalisation des matrices. On peut aussi multiplier des matrices par blocs entre-elles, à supposer que les dimensions respectives soient compatibles : avec les notations précédentes, le produit BC des matrices par bloc $B = (B_{ij})$ de dimension (m, n), $B_{i,j}$ étant de dimension (m_i^B, n_j^B), et $C = (C_{k\ell})$ de dimension (p, q), $C_{k\ell}$ étant de dimension (p_k^C, q_ℓ^C), est effectuable si $n = p$ et $n_j^B = p_j^C$.

2.5.4. Matrices booléennes associées aux graphes. — Une matrice est booléenne si elle n'a pour éléments que des 0 et des 1 : le 0 est associé à la valeur `faux`, le 1 est associé à la valeur `vrai`. Dans la définition des opérations matricielles, on substitue :
- à l'addition des réels $(+)$ l'addition logique `ou_exclusif`, notée $\vee$,
- à la multiplication des réels $\times$ la multiplication logique `et`, notée $\wedge$.

Ces lois, $+/$`ou_exclusif` et $\times/$`et`, sont exactement celles introduites dans l'exemple 1.14.3 de lois sur l'ensemble $\mathbb{F}_2$, incarnation particulière de l'ensemble à deux éléments de l'exemple 1.1.11.

Les matrices booléennes sont importantes dans la théorie des graphes ; soient 4 villes $V_1, \ldots, V_4$, par exemple Paris, Nantes, Cholet, Saint-Nazaire, dont on étudie les liaisons par chemin de fer ; on décrit ces liaisons par une matrice booléenne G, avec $G_{ij} = 1$ s'il y a une liaison directe de V_i vers V_j, $G_{ij} = 0$ sinon. Ainsi, avec notre exemple

$$G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Le produit de G avec elle-même s'écrit : $G^2 = G \wedge G$. Le coefficient $(G^2)_{ij}$ du carré G^2 est nul s'il existe un nombre pair de liaisons ferroviaires avec un changement entre V_i et V_j et vaut 1 si ce nombre est impair.

On peut définir la matrice, à coefficients dans $\overline{\mathbb{N}}$ (introduit dans l'exemple 1.1.3), des distances D qui donne la distance de la liaison ferroviaire directe entre chaque ville, soit dans l'exemple ci-dessus (et approximativement) :

$$D = \begin{pmatrix} 0 & 350 & \infty & 430 \\ 350 & 0 & 90 & 100 \\ \infty & 90 & 0 & \infty \\ 439 & 100 & \infty & 0 \end{pmatrix}$$

La distance d'une ville à elle-même est nulle, ainsi $D_{ii} = 0$. Par ailleurs, pour $i \neq j$, le coefficient D_{ij} est un entier fini non nul s'il existe une liaison entre les deux villes ; $D_{ij} = \infty$ s'il n'existe pas de liaison entre les deux villes. Les opérations associées à ce type de matrice sont redéfinies : on substitue à l'addition des réels $(+)$ la recherche du minimum ; on substitue à la multiplication des réels $(\times)$ l'addition des distances. D'où le produit d'une matrice "distance" par elle-même : $(D^2)_{ij} = \min_{k=1,\ldots,n}(D_{ik} + D_{kj})$, ce coefficient donnant la distance minimum pour aller de V_i à V_j en passant par une troisième ville.

On peut aussi considérer la matrice des débits $\mathcal{D}$ d'une ligne de transmission entre les villes : $D_{ii} = \infty$ vu que le débit de transmission d'une ville avec elle-même est maximal, $D_{ij} > 0$ s'il existe une liaison de débit D_{ij} `k-octets/sec` entre les villes V_i et V_j ;. $D_{ij} = 0$, s'il n'existe pas de liaison entre les villes V_i et V_j.

2.6. Exercices

E**2**.1. Montrer que, si $a \neq 1$,

$$1 + a + \ldots + a^{n-1} = \begin{cases} \frac{a^n - 1}{a-1} & \text{si } a \neq 1, \\ n & \text{sinon.} \end{cases}$$

et en déduire si $k \in \{0, 1, 2, \ldots, n\}$

$$\sum_{\ell=1}^{n} \mathrm{e}^{2\mathrm{i}\ell k/n} = \begin{cases} 0 & \text{si } k \text{ n'est pas divisible par } n, \\ n & \text{sinon.} \end{cases}$$

Soit n entier au moins égal à 2 et $\omega_n = \mathrm{e}^{2\mathrm{i}\pi/n}$ Au n-uplet $x = (x_1, \ldots, x_n\}$, on associe le n-uplet $X = (X_1, \ldots, X_n\}$ défini par

$$X_\ell = \sum_{k=1}^{n} \omega_n^{k\ell} x_k, \quad \ell = 1, \ldots, n.$$

Montrer que

$$x_k = \frac{1}{n} \sum_{\ell=1}^{n} \omega_n^{-k\ell} X_\ell, \quad k = 1, \ldots, n.$$

E**2**.2. Évaluer les sommes

$$S_{kn} = \sum_{i=1}^{n} \delta_{ik}, \quad k \in \mathbb{N}$$

$$S_{jn} = \sum_{i=1}^{n} \delta_{ij} V_i, \quad j \in \mathbb{N}$$

$$\widetilde{S} = \sum_{i,j,k=1}^{n} \varepsilon_{ijk} (V_i + V_j) V_k$$

où $V = (V_1, V_2, \ldots, V_n)$ est un vecteur de $\mathbb{R}^n$.

E**2**.3. Soit la suite $u_0 = 1, u_1 = -\frac{1}{2}\mathrm{i}, \ldots, u_n = (2\mathrm{i})^{-n}, \ldots$ où i est la racine carrée de -1. Séparer la somme de la série $S = \sum_{n=0}^{\infty} u_n$ en partie réelle et imaginaire $S = \Re e\, S + \mathrm{i} \Im m\, S$ en explicitant $\Re e\, S$ et $\Im m\, S$.

E**2**.4. Soit I_n la matrice identité d'ordre n. Démontrer que $AI_n = I_n A = A$ pour toute matrice A carrée d'ordre n.

E**2**.5. Soient A matrice de dimensions (m, n) et B matrice de dimensions (n, m). Montrer l'égalité des traces $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

E**2**.6. La transposée A^t d'une matrice A est définie par : $(A^t)_{ij} = A_{ji}$. Démontrer que, pour α réel, A et B matrices de dimensions telles que les expressions sont bien définies

$$(A^t)^t = A, (A+B)^t = A^t + B^t, (\alpha A)^t = \alpha A^t, (AB)^t = B^t A^t, (A^{-1})^t = (A^t)^{-1}.$$

E**2**.7. Soit A la matrice carrée d'ordre n dont les coefficients sont $A_{ij} = (-1)^{i+j}$. Calculer A^2.

E**2**.8. Soit A la matrice carrée carrée d'ordre 4

$$A = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & -1 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 1 & 1 & 0 & 4 \end{pmatrix}$$

Calculer $S = \sum_{i=1}^{4} \sum_{j=1}^{4} A_{ij}\delta_{i,j+1}$.

E**2**.9. Soit A la matrice

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Calculer A^2 et A^3. Montrer que A et A^3 sont proportionnelles ; en déduire la forme général de A^n.

E**2**.10. A et B étant des matrices carrées, développer et simplifier

$$(2A)(3B) - (A + 2B)^2 + (A - B)(A + B).$$

E**2**.11. La matrice

$$A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix}$$

est une matrice rectangulaire. Calculer sa transposée A^t, puis les produits AA^t et A^tA.

E**2**.12. Soient les trois matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Montrer que $\sigma_1\sigma_2 = -\sigma_2\sigma_1 = \mathrm{i}\sigma_3$, puis une égalité similaire pour $\sigma_3\sigma_1$ et $\sigma_2\sigma_3$. Soient a, b, c des nombres complexes. Montrer l'égalité

$$(a\sigma_1 + b\sigma_2 + c\sigma_3)^2 = (a^2 + b^2 + c^2)I_2.$$

E**2**.13. Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Déterminer toutes les matrices carrées B d'ordre 2 telles que $AB = BA$.

E**2**.14. On considère la matrice

$$J = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Calculer J^2, J^3 et J^4 en fonction de J. Montrer par récurrence que $J^n = 4^{n-1}J$ pour $n \geq 1$.

E**2**.15. On considère la matrice

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

1. Calculer A^2. Expliciter α et β tel que $A^2 = \alpha A + \beta Id$.
2. Montrer par récurrence qu'il existe a_n et b_n tels $A^n = a_n A + b_n Id$
3. Expliciter a_{n+1} et b_{n+1} en fonction de a_n et b_n.
4. Expliciter a_n en fonction de n.
5. Calculer de deux façons distinctes $\sum_{k=0}^{n-1}(b_{k+1} - b_k)$. En déduire l'expression de b_n.
6. Déterminer tous les coefficients de la matrice A^n.

CHAPITRE 3

APPLICATIONS LINÉAIRES ET MATRICES

Les propriétés multiples des tableaux de nombres, que sont les matrices du chapitre précédent, deviennent moins mystérieuses si elles sont interprétées comme représentations particulières d'applications linéaires opérant dans des espaces vectoriels : c'est l'objet de ce chapitre d'introduire à l'algèbre linéaire et d'expliquer le dictionnaire avec les notions du chapitre précédent.

3.1. Espaces vectoriels

On connaît ce qu'est un vecteur dans le plan ou dans l'espace à 3 dimensions, qui permet de visualiser des directions, des forces, des champs,..., ainsi que les opérations possibles sur ces vecteurs : addition, rotation, translation, projection... Le cadre général de l'algèbre linéaire est une généralisation de ces notions intuitives.

3.1.1. Algèbre vectorielle. — Un ensemble E dont les éléments sont nommés *vecteurs* et notés $V, W, \dots$ est un espace vectoriel si :
- deux opérations y sont définies : une loi interne, noté additivement par $+$, qui permet de calculer $V+W$ et qui fait de $(E,+)$ un groupe dont le neutre noté 0_E [(1)] est appelé *vecteur nul*, et une loi externe, noté multiplicativement, qui permet de calculer αW avec α réel,
- ces opérations sont compatibles au sens où les relations suivantes sont vérifiées

$$\alpha(V+W) = \alpha V + \alpha W, \quad \alpha(\beta V) = (\alpha\beta)V, \quad V, W \in E, \alpha, \beta \in \mathbb{R}.$$

▷ ***Exemples 3.1.***

1. L'ensemble $\mathbb{R}^2 = \{(x,y), x, y \in \mathbb{R}\}$ des vecteurs du plan est un espace vectoriel, où

$$(x,y) + (x',y') = (x+x', y+y'), \quad \alpha(x,y) = (\alpha x, \alpha y), \quad 0_{\mathbb{R}^2} = (0,0).$$

2. L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel où les addition et multiplication proviennent des opérations algébriques sur les nombres complexes : le vecteur nul est le complexe nul.
3. L'ensemble des polynômes de degré au plus 3, de la forme : $P = a_0 + a_1X + a_2X^2 + a_3X^3$ avec a_0, a_1, a_2, a_3 réels est un espace vectoriel. Le vecteur nul est le polynôme

1. Dans la pratique, passé le temps des définitions, on omettra l'indice E quasi-systématiquement.

nul $0+0X+0X^2+0X^3$ et les opérations sont

$$\begin{aligned}(a_0+a_1X+a_2X^2+a_3X^3)+(b_0+b_1X+b_2X^2+b_3X^3)\\ = a_0+b_0+(a_1+b_1)X+(a_2+b_2)X^2+(a_3+b_3)X^3,\\ \alpha(a_0+a_1X+a_2X^2+a_3X^3) = (\alpha a_0)+(\alpha a_1)X+(\alpha a_2)X^2+(\alpha a_3)X^3.\end{aligned}$$

4. L'ensemble E_3 des polynômes de degré égal à 3, de la forme $P=a_0+a_1X+a_2X^2+a_3X^3$ avec $a_3\neq 0$ n'est pas un espace vectoriel, car il n'a pas de vecteur nul : $P+Q\neq P$ pour tous polynômes P et Q de E_3.
5. L'ensemble des fonctions de la forme $\alpha\cos+\beta\sin$ avec α,β réels est un espace vectoriel, qui peut être défini comme l'espace des fonctions définies sur $\mathbb{R}$ solutions de l'équation différentielle (linéaire) $y''+y=0$. ◁

Sous-espace vectoriel Un sous-ensemble S d'un espace vectoriel E est lui-même un espace vectoriel si :
- le vecteur nul de E appartient à S,
- la partie S est stable pour les 2 opérations de P : si α est un réel et V,W deux vecteurs de S, alors $V+\alpha W$ est aussi un vecteur de S.

On dit alors que S est un *sous-espace vectoriel* de l'espace vectoriel E.

▷ ***Exemples 3.2.***

1. L'ensemble des polynômes de degré au plus 2 est un sous-espace de l'espace vectoriel Poly_3.
2. L'ensemble F des monômes de la forme a_2x^2 avec $a_2\neq 0$ n'est pas un sous-espace de l'espace vectoriel des polynômes de degré au plus 3, car il ne contient pas le vecteur nul. ◁

3.1.2. Systèmes de vecteurs.

— Les n vecteurs $V_1,\ldots,V_n$ de l'espace vectoriel E sont dits *linéairement indépendants*, si la seule combinaison linéaire nulle :

$$\alpha_1V_1+\alpha_2V_2+\ldots+\alpha_nV_n=0$$

est obtenue avec : $\alpha_1=\alpha_2=\ldots=\alpha_n=0$; sinon, ils sont *linéairement dépendants.*

▷ ***Exemples 3.3.***

1. Dans Poly_3, les polynômes $P=2X$ et $Q=X^2$ sont linéairement indépendants.
2. Dans Poly_3, les polynômes $P=2x, Q=X^2$ et $R(x)=X-X^2$ ne sont pas linéairement indépendants. On a en effet $-P+2Q+2R=0$.
3. Dans le plan $\mathbb{R}^2$, trois vecteurs sont toujours linéairement dépendants [2]. ◁

2. Sinon, on aurait, par ex., $V_1=(x_1,y_1)$ et $V_2=(x_2,y_2)$ indépendants : on ne pourrait écrire $V_2=kV_1$ et donc $x_1/x_2\ /y_1/y_2$ (le cas y_1 ou y_2 nul est laissé au lecteur), *i. e.* $x_1y_1-x_2y_2\neq 0$. L'équation d'inconnues $\alpha\beta$ donnée par $V_3=\alpha V_1+\beta V_2$ s'écrit comme le système $\alpha x_1+\beta x_2=x_3$ et $\alpha y_1+\beta y_2=y_3$, qui a une solution unique grâce à la condition $x_1y_1-x_2y_2\neq 0$. Ainsi V_3 est combinaison linéaire de V_1 et V_2, ce qui est bien ce qu'il fallait prouver.

3.1.3. Base d'un espace vectoriel. —

Définition 3.1. Un ensemble de vecteurs $(V_1, V_2, \ldots, V_n)$ d'un espace vectoriel E est une *base* de E si
- les n vecteurs sont linéairement indépendants,
- les n vecteurs engendrent l'espace vectoriel E, *i. e.* tout vecteur V de E est combinaison linéaire des vecteurs $V_1, \ldots, V_n$: $V = \alpha_1 V_1 + \ldots \alpha_n V_n$.

L'entier n est la *dimension* de l'espace vectoriel E. On démontre que toute base de E a exactement n éléments si E admet une famille particulière de n vecteurs qui en est une base.

$\triangleright$ ***Exemples 3.4.***

1. Dans $\mathbb{R}^n$, soit la famille $\mathbf{c} = (c_1, \ldots, c_n)$ définie par
$$c_1 = (1, 0, \ldots, 0), c_2 = (0, 1, 0, \ldots, 0), \ldots, c_n = (0, \ldots, 0, 1)$$
où le i-ème vecteur c_i a toutes ses coordonnées nulles, sauf la i-ème valant 1. La famille $\mathbf{c}$ est une base, dite *base canonique* de $\mathbb{R}^n$. L'espace $\mathbb{R}^n$ est de dimension n.
2. Dans l'espace Poly_3, les vecteurs $p_0 = 1, p_1 = X, p_2 = X^2, p_3 = X^3$ forment une base de l'espace, qui est de dimension 4.
3. L'espace Poly_∞ des polynômes à coefficients réels n'a pas de base (finie) : s'il en avait une, en désignant par N le plus grand des degrés de ses vecteurs, on pourrait écrire le polynôme $M_{N+1} = x^{N+1}$ comme combinaison linéaire de polynômes de degré strictement plus petits que celui de M_N, ce qui n'est pas possible ! $\triangleleft$

3.1.4. Composantes d'un vecteur. — Lorsque l'on a défini une base d'un espace de dimension n, conventionnellement notée $\mathbf{b} = (b_1, \ldots, b_n)$, tout vecteur V peut s'exprimer de manière unique comme une combinaison linéaire des vecteurs de base :
$$V = \sum_{i=1}^{n} \alpha_i b_i$$
Les coefficients $\alpha_1, \ldots, \alpha_n$ sont les *composantes* du vecteur V dans la base $\mathbf{b}$, qu'on rassemble dans le vecteur colonne
$$X = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$
des *coordonnées* de V relativement à la base $\mathbf{b}$: le vecteur ligne $X^t = (\alpha_1 \alpha_2 \ldots \alpha_n)$ est aussi utilisé.

$\triangleright$ ***Exemple 3.5.*** Dans l'espace Poly_3, le polynôme $Q = 2X^3 - X^2 + 1$ a pour vecteur de coordonnées $X_Q = (1, 0, -1, 2)$ relativement à la base $\mathbf{b} = ((b_i = x^{i-1})_{i=1}^4)$. On prendra garde au décalage des indices ! $\triangleleft$

3.2. Applications linéaires

Rappelons qu'une application $f : A \to B$ fait correspondre à tout élément a de A un seul et unique élément de B, élément noté $f(a)$. L'ensemble des éléments $f(a)$ de B correspondant aux éléments de A constitue l'image de A dans B ; on la note $f(A)$.

3.2.1. Définition. — Soient deux espaces vectoriels E et F, une application[3] $f : E \to F$ est linéaire, si elle est compatible avec la structure d'espace vectoriel au sens où

$$f(\alpha V + \beta W) = \alpha f(V) + \beta f(W), \quad V, W \in E, \alpha, \beta \in \mathbb{R}.$$

▷ ***Exemples 3.6.***

1. Dans l'espace des vecteurs de la géométrie ordinaire, les opérations de rotation, d'homothétie et de projection sont des applications linéaires : par exemple, la projection de la somme de 2 vecteurs sur un axe est égale à la somme des projections de chacun des vecteurs.
2. Dans l'espace Poly_n des polynômes de degré au plus n, la dérivation est une application linéaire, ni injective ($D(0) = D(1) = 0$), ni surjective $D(\mathrm{Poly}_n) = \mathrm{Poly}_{n-1}$. L'intégration $I : P \to \int_0^X p(t)dt$ n'est pas une application linéaire de l'espace Poly_n, puisque l'intégrale d'un polynôme de degré n est un polynôme de degré $n+1$: l'application d'intégration est bien définie comme application de Poly_∞ dans lui même et c'est une application linéaire. ◁

Une application linéaire bijective $f : E \to F$ est appelée un *isomorphisme* de E sur F, un *automorphisme* de E si $E = F$.

3.2.2. Matrices et applications linéaires. — Le postulat de base est qu'une matrice est la représentation d'une application linéaire entre deux espaces vectoriels, chacun pourvu d'une base.

Soient des espaces vectoriels E, F, de dimensions respectives p, n, munis de bases respectives $\mathbf{b} = (b_1, b_2, \ldots, b_p), \mathbf{c} = (c_1, c_2, \ldots, c_n)$. Soit une application linéaire f de E dans F. Il y a deux méthodes pour définir la matrice A de dimension (n, p) associée à cette application :

– Dans la formulation covariante, on décompose le transformé $f(b_j)$ de chaque vecteur de base b_j sur la base $\mathbf{c}$

$$f(b_j) = \sum_{i=1}^{n} A_{ij} c_i, \quad j = 1, \ldots, p..$$

Les composantes du transformé $f(b_j)$ du vecteur b_j par f constituent la j-ème colonne de la matrice A : on prendra garde aux indices de A_{ij}, b_j et c_i et à leur ordre !

– La formulation contravariante considère les composantes des vecteurs source et but dans les bases. Soit un vecteur V, décomposé sur la base $\mathbf{b}$ suivant $V = \sum_{j=1}^{p} \alpha_j b_j$. Son transformé $W = f(V)$ par l'application f s'écrit : $W = \sum_{i=1}^{n} \beta_i c_i$. Les coordonnées du transformé donne, ligne par ligne, la matrice de l'application :

$$\beta_i = \sum_{j=1}^{p} A_{ij} \alpha_j.$$

Ainsi si X (resp. Y) est le vecteur colonne des coordonnées de V relativement à la base $\mathbf{b}$ (resp. $W = f(V)$ dans la base $\mathbf{c}$), soit

$$X = (\alpha_1, \alpha_2, \ldots, \alpha_p)^t, \quad Y = (\beta_1, \beta_2, \ldots, \beta_n)^t,$$

alors

$$Y = AX.$$

3. On parle parfois d'*opérateur (linéaire)*, de transformées, d'*endomorphisme* si $E = F$.

Si f est une application linéaire interne à l'espace vectoriel E, on prend en général implicitement une même base pour les bases $\mathbf{b}$ et $\mathbf{c}$ des espaces source et but.

Cette définition d'une matrice associée à une application linéaire est cohérente avec la définition du produit matriciel et de la composition des applications : si A, B, C représentent les applications linéaires $f, g, f \circ g$, alors $C = AB$.

▷ ***Exemple 3.7.*** Dans l'espace des polynômes de degré au plus 3, calculons la matrice de l'application linaire $t : P(X) \to P(X+1)$ relativement à la base $(1, X, X^2, X^3)$. Suivant la voie contravariante, on a

$$(X+1)^j = \sum_{i=0}^{j} C_j^i X^i,$$

ce qui donne comme matrice de l'application linéaire la matrice $T = (T_{ij})_{1 \leq i,j \leq 4}$ avec $T_{ij} = C_{j-1}^{i-1}$ si $1 \leq i \leq j \leq 4$ et 0 sinon, soit

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$ ◁

Inversement, on interprète une matrice A de dimension (n, p) comme la matrice représentant une application linéaire, notée (encore !) A, de $\mathbb{R}^p$ dans $\mathbb{R}^n$, ces deux espaces vectoriels étant munis de leurs bases canoniques.

La matrice associée à une application linéaire donnée dépend des bases dans lesquelles on la calcule ; cependant toutes les matrices associées à une application auront ses comportements identiques ou similaires, du point de certaines propriétés, comme les caractéristiques (Sect. 2.4) ou de symétrie et antisymétrie (Sect.2.5.1).

On appelle $\boxed{\textit{\textbf{noyau d'une application linéaire}}}$ $f : E \to F$, et on note $\ker f$, l'ensemble des vecteurs de l'espace source de f dont l'image est le vecteur nul

$$\ker f = \{V \in E, f(V) = 0\}.$$

Le noyau d'une application est un sous-espace vectoriel de l'espace source.

L'image $\operatorname{Im} f = f(E)$ est un sous-espace vectoriel de F, de dimension finie si E ou F l'est. Le $\boxed{\textit{\textbf{rang d'une application}}}$ f de E dans F est la dimension de l'espace vectoriel $\operatorname{Im} f = f(E)$. On admet le théorème suivant.

Théorème 3.1. *Soient E et F des espaces vectoriels de dimension finie, f une application linéaire de E dans F. Alors, les dimensions des espaces E, $\ker f$ et $\operatorname{Im} f$ sont liées par la relation :*

$$\dim E = \dim \operatorname{Im} f + \dim \ker f.$$

Le rang *de l'application linéaire f de E dans E vérifie* $\operatorname{rang} f = \dim E - \dim \ker f$.

Si f est un isomorphisme, son application réciproque f^{-1} est une application linéaire.

*La dimension de $\ker f$ est nulle (*i. e. *le noyau est réduit à l'espace vectoriel nul $\{0\}$) si et seulement si et seulement si f est injective.*

Si $E = F$, les propriétés d'injectivité, de surjectivité et de bijectivité sont équivalentes.

3.2.3. Matrices et changement de bases. — Soit un espace vectoriel E de dimension n, pourvu de deux bases, respectivement une base $\mathbf{b} = (b_1, \ldots, b_n)$ dite ancienne et $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_n)$ dite nouvelle. L'application linéaire P qui fait passer de l'ancienne base $\mathbf{b}$ à la nouvelle base $\widetilde{\mathbf{b}}$ telle que $Pb_j = \widetilde{b}_j$ est un isomorphisme. Pour l'exprimer matriciellement, on écrit le vecteur $\widetilde{b}_j$ comme combinaison linéaire des vecteurs b_i

$$\widetilde{b}_j = \sum_{i=1}^{n} P_{ij} b_i. \tag{3}$$

Si le vecteur V a comme composantes (α_i) dans l'ancienne base $\mathbf{b}$ et $(\widetilde{\alpha}_i)$ dans la nouvelle $\widetilde{b}$, alors

$$\alpha_i = \sum_{j=1}^{n} P_{ij} \widetilde{\alpha}_j,$$

soit pour $X = (\alpha_1, \alpha_2, \ldots, \alpha_n)^t$ et $\widetilde{X} = (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots, \widetilde{\alpha}_n)^t$,

$$X = P\widetilde{X}.$$

Démonstration. — On a $V = \sum_{i=1}^{n} \alpha_i b_i = \sum_{j=1}^{n} \widetilde{\alpha}_j \widetilde{b}_j$, d'où en insérant (3) dans le dernier membre

$$\sum_{i=1}^{n} \alpha_i b_i = \sum_{j=1}^{n} \widetilde{\alpha}_j \sum_{i=1}^{n} P_{ij} b_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij} \widetilde{\alpha}_j \right) b_i. \qquad \square$$

Soit P la matrice de changement de base (dite de *matrice de passage*) de à$\tilde{}$, soit A la matrice d'un endomorphisme de E relativement à l'ancienne base $\mathbf{b}$ et $\widetilde{A}$ son expression dans la nouvelle $\widetilde{\mathbf{b}}$, alors

$$\widetilde{A} = P^{-1} A P. \tag{4}$$

Démonstration. — La combinaison des égalités $Y = AX, \widetilde{Y} = \widetilde{A}\widetilde{X}, X = P\widetilde{X}$ et $Y = P\widetilde{Y}$ donne

$$AP\widetilde{X} = AX = Y = P\widetilde{Y} = P\widetilde{A}\widetilde{X}, \quad \widetilde{X} \in \mathbb{R}^n,$$

et par suite $AP = P\widetilde{A}$, d'où l'identité (4) en multipliant à gauche par la matrice P^{-1}. $\square$

On dit que la matrice A est semblable à la matrice $\widetilde{A}$: les matrices A et $\widetilde{A}$ ont mêmes caractéristiques (déterminant, trace, valeurs propres, ...).

3.3. Exercices

E**3**.1. Dans chacun des cas suivants, dire, en justifiant votre réponse, si F_i est ou n'est pas un sous-espace vectoriel de E_i, les divers espaces vectoriels E_i étant munis des opérations usuelles.
Si votre réponse est *non*, précisez cependant s'il y a stabilité pour l'une des deux opérations.

1. $E_1 = \mathbb{R}^3$; $F_1 = \{(x, y, z) \in \mathbb{R}^3 \ ; \ 2x + 3y - 4z = 0\}$;
2. $E_2 = \mathrm{Poly}_\infty$; $F_2 = \{P \in \mathrm{Poly}_\infty \ ; \ \deg P \leq 5\}$;
3. $E_3 = F(\mathbb{R}, \mathbb{C})$; $F_3 = \{f \in F(\mathbb{R}, \mathbb{C}) \ ; \ f(3) = 0\}$;
4. $E_4 = \mathbb{R}^3$; $F_4 = \{(x, y, z) \in \mathbb{R}^3 \ ; \ x^2 + y^2 + z^2 \leq 1\}$;
5. $E_5 = \mathrm{Poly}_8$; $F_5 = \{P \in \mathrm{Poly}_8 \ ; \ \deg P = 4\}$.

E**3**.2. Prouver que la famille $\mathbf{b}$ définie par

$$\mathbf{b} = ((0,1,1),(1,0,1),(1,1,0))$$

est une base de l'espace vectoriel $\mathbb{R}^3$ et calculer les coordonnées du vecteur $V = (a,b,c)$ relativement à la base $\mathbf{b}$.

E**3**.3. Montrer dans chaque cas, si l'application f de l'espace vectoriel E dans l'espace vectoriel F est linéaire ou non.

1. $E = \mathbb{R}$; $F = \mathbb{R}$; $f(x) = x^2$

2. $E = \mathbb{R}$; $F = \mathbb{R}$; $f(x) = (x+1)^2 - (x-1)^2$

3. $E = \mathbb{R}^2$; $F = \mathbb{R}$; $f(x,y) = ax + by + c$; a, b et c réels.

E**3**.4. Soit M_2 l'ensemble des matrices carrées d'ordre 2

$$\mathcal{M}_2 = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a,b,c,d \text{ réels} \right\}$$

des matrices carrées d'ordre 2. Démontrer que $\mathcal{M}_2$ est un espace vectoriel. Donner une base de cet espace. Les quatre matrices

$$A_= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

suivantes forment-elles une base de l'espace $\mathcal{M}_2$?

E**3**.5. Soient les espaces $\mathbb{R}^2$ et $\mathbb{R}^3$ munis de leurs bases canoniques. Soient les applications linéaires f de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par

$$f((x,y,z)) = (x + 2y + 3z, y + 2z), \quad (x,y,z) \in \mathbb{R}^3$$

et g de $\mathbb{R}^2$ dans $\mathbb{R}^3$ définie par :

$$g((x,y)) = (x - y, x - 2y, x - 3y), \quad (x,y) \in \mathbb{R}^2.$$

Donner les matrice M de f et N de g, puis calculer les matrices A de $g \circ f$ et B de $f \circ g$. En déduire les expressions de $(g \circ f)((x,y,z))$ et de $(f \circ g)((x,y))$.
Montrer que l'application $f \circ g$ est bijective. En déduire que la matrice B est inversible et calculer B^{-1}.

E**3**.6. Dans l'espace Poly_n des polynômes de degré au plus n, est-ce que l'application $P(X) \to P(X+a)$ est linéaire ?

E**3**.7. Dans l'espace Poly_4 des polynômes de degré au plus 4, calculer la matrice représentative de l'application de dérivation $P \to P'$ relativement à la base $(X^4, X^3, X^2, X, 1)$. Quel est le noyau de l'application ?

E**3**.8. Dans le plan, calculer la matrice $P(\varphi)$ de la projection sur la droite D passant par l'origine et faisant un angle φ avec l'axe Ox. Calculer $P(\varphi)^2$.

E**3**.9. Dans $\mathbb{R}^3$, on considère un plan P passant par l'origine, d'équation $ux + vy + wz = 0$, avec $u^2 + v^2 + w^2 = 1$. Calculer la matrice P de l'application linéaire qui projette tout vecteur $M = (x,y,z)$ de $\mathbb{R}^3$ sur le plan P. Calculer P^2.

E**3**.10. Soit l'application linéaire f de $\mathbb{R}^2$, représentée dans la base canonique du plan par la matrice

$$A = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{3}} & \frac{5}{2} \end{pmatrix}.$$

Exprimer la matrice $\widetilde{A}$ de cette application relativement à la base obtenue par la rotation d'angle $\pi/6$ de la base canonique.

E**3**.11. Soit L l'ensemble des suites $\mathbf{u} = (u_n)_{n\geq 0}$ de nombres réels, satisfaisant la relation de récurrence

$$u_{n+2} = u_{n+1} + 2u_n, \quad n \geq 0.$$

Montrer que L est un espace vectoriel. Montrer que les suites particulières $\mathbf{v}$ et $\mathbf{w}$ définies par

$$v_n = (-1)^n, \quad w_n = 2^n, \quad n \geq 0$$

appartiennent à L et qu'elles forment une famille libre. Tenant compte qu'une suite $\mathbf{u}$ de L est uniquement déterminée par la connaissance de ses premiers termes u_0 et u_1, montrer que $\mathbf{v}$ et $\mathbf{w}$ forment une base de L. Déterminer la suite $\mathbf{u}$ telle que $u_0 = 1$ et $u_1 = -2$.

E**3**.12. Dans l'espace $\mathbb{R}^3$, on considère l'application f qui fait correspondre au vecteur V de coordonnées (x, y, z) le vecteur $f(V)$ de coordonnées $(5x+y-z, 2x+4y-2z, x-y+3z)$. Écrire la matrice A de f relativement à la base canonique $\mathbf{c} = (c_1 = (1,0,0), c_2 = (0,1,0), c_3 = (0,0,1))$.
Soit la seconde base $\mathbf{b} = (b_1 = (0,1,1), b_2 = (1,0,1), b_3 = (1,1,0))$, Calculer la matrice P de changement de base de $\mathbf{b}$ à $\mathbf{c}$. Calculer la matrice $\widetilde{A}$ représentant f dans la base $\mathbf{b}$. Déduire A^n

E**3**.13. Soit ℓ l'application de $\mathbb{R}^3$ dans $\mathbb{R}^2$ (espaces munis de leurs bases canoniques), qui fait correspondre au vecteur V de coordonnées (x, y, z) le vecteur de $\mathbb{R}^2$ des coordonnées $(2x+y-z, 3x-2y+4z)$. Écrire la matrice A représentant l'application ℓ dans les bases canoniques. Calculer la matrice $\widetilde{A}$ représentant ℓ dans les bases respectives $\mathbf{f} = (f_1 = (1,1,1), f_2 = (1,1,0), f_3 = (1,0,0))$ et $\mathbf{g} = (g_1 = (1,3), g_2 = (1,4))$.

CHAPITRE 4

ÉQUATIONS LINÉAIRES ET DÉTERMINANTS

4.1. Équations linéaires

Étant donnée l'application linéaire A de E dans F et le vecteur B de F, l'équation linéaire

$$AX = B \tag{5}$$

comporte l'inconnue X à déterminer comme vecteur de E. Le vecteur $B \in F$, deuxième membre de l'équation, est parfois nommé terme constant. L'équation $AX = 0$ est appelée *équation homogène* associée.

Avec $A = (a_{ij}), X = (x_1 \dots x_n)^t$ et $B = (b_1 \dots b_m)^t$, écrire l'égalité des coefficients des deux membres de (5) présente cette équation de manière équivalente sous la forme d'un *système linéaire* de $m = \dim F$ équations à $n = \dim E$ inconnues (scalaires) $x_1, \dots, x_n$:

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array}\right. \tag{6}$$

On appelle *rang du système* (6) l'ordre maximum des matrices carrées inversibles que l'on peut extraire de A : on montre que le rang du système linéaire coïncide avec le rang de la matrice A tel que défini juste avant le théorème 3.1.

Les résultats généraux de résolution des équations linéaires prennent diverses formes, suivant qu'on insiste plus sur la formulation géométrique donnée par l'équation linéaire (5) :

Proposition 4.1. *Si l'équation $AX = B$ admet une solution particulière X_0, alors toute solution de l'équation est la somme $X = X_0 + X_1$, où X_1 est une solution de l'équation homogène $AX = 0$.*

Si B n'appartient pas à l'image $\operatorname{Im} A$ de A, l'équation $AX = B$ n'a pas de solution.

Si B appartient à l'image $\operatorname{Im} A$,

- *l'équation $AX = B$ admet une unique solution si et seulement si $\ker A = 0$.*
- *si $\ker A \neq 0$, le système a une infinité de solutions, somme de la solution particulière et d'un vecteur de $\ker A$.*

ou sa version en système linéaire (6) :

Proposition 4.2. *Soit r le rang du système linéaire (6) à m équations et n inconnues.*

Si $r = m = n$*, auquel cas on appelle un tel système un* système de Cramer*, la solution du système* (6) *est unique. Vu que dans ce cas la matrice* A *est inversible, cette unique solution est donnée en terme matriciel sous la forme* $X = A^{-1}B$*.*

Si $r \leq m < n$*, il y a plus d'inconnues que de données (coordonnées de* B*) et d'équations : si l'ensemble solution n'est pas vide, les solutions dépendent de* $n - r$ *paramètres, dits* variables libres.

Si $r \leq n \leq m$ *avec* $m - r > 0$*, le système est surdéterminé (il y a plus d'équations que d'inconnues) et en général n'a pas de solution*[1]*. S'il a une solution, l'ensemble de ses solutions dépend de* $n - r$ *variables.*

Pratiquement, la solution s'obtient par élimination de variables ou par la méthode des pivots de Gauss pour se ramener à un système associé à une matrice triangulaire, système dit *échelonné*, qui est résoluble aisément. La méthode reposant sur les déterminants sera présentée ci-dessous, méthode qui est rarement appliquée dans les systèmes à nombre n d'inconnues important (*i. e.* d'au moins 4) vu la complexité (au sens algorithmique même des formules).

▷ ***Exemples 4.1.***

1. Soit le système :

$$\left\{\begin{aligned} x + y + z &= 1 \\ x - y + z &= 2 \\ 2x - 2y + 3z &= 3 \end{aligned}\right. \tag{7}$$

On écrit les systèmes équivalents successifs

$$\left\{\begin{aligned} x + y + z &= 1 \\ x - y + z &= 2 \\ 2x - 2y + 3z &= 3 \end{aligned}\right. \qquad \left\{\begin{aligned} x + y + z &= 1 \\ 2y &= -1 \\ 2x - 2y + 3z &= 3 \end{aligned}\right.$$

$$\left\{\begin{aligned} y &= -1/2 \\ x + z &= 3/2 \\ 2x + 3z &= 2 \end{aligned}\right. \qquad \left\{\begin{aligned} y &= -1/2 \\ x + z &= 3/2 \\ z &= -1 \end{aligned}\right. \qquad \left\{\begin{aligned} y &= -1/2 \\ x &= 5/2 \\ z &= -1 \end{aligned}\right.$$

On a retranché la première ligne à la seconde, puis substitué la valeur $y = -1/2$ dans les deux équations restantes, puis éliminé x en retranchant deux fois la seconde équation à la troisième,...

Anticipant la section suivante, on peut dire que le déterminant de A vaut -2 : on a un système de Cramer, qui a une solution unique.

2. Soit le système :

$$\left\{\begin{aligned} x + y + z - w &= 1 \\ x - y + z + 2w &= 2 \\ 2x - 2y + 3z - w &= 3 \end{aligned}\right. \tag{8}$$

En mettant les termes en w dans le membre de droite, on retrouve les membres de gauche de l'exemple précédent (7) : les manipulations similaires ramènent le système

1. On obtient des solutions, considérées comme approchées, par des méthodes d'optimisation : on remplace l'équation $AX = 0$ par la recherche du minimum de la fonction positive $X \to \|AX - B\|$ où $\| \quad \|$ est une norme comme définie dans le dernier chapitre de ce cours.

(8) au système

$$\begin{cases} x = 5/2 - 11/2w \\ y = -1/2 + 3/2w \\ z = -1 + 5w \end{cases}$$

Les solutions du système (8) dépendent du paramètre w. Cette solution générale s'exprime sous forme d'une somme : $X = X_0 + wX_1$, avec $X_0 = (5/2, -1/2, -1, 0)$ une solution particulière du système complet et $X_1 = (-11/2, 3/2, 5, 1)$ une solution de l'équation homogène $AX_1 = 0$ et base de l'espace $\ker A$ qui est de dimension 1. C'est le paramètre w qui est libre.

3. Soit le système, dépendant du paramètre a

$$\begin{cases} x + y = 1 \\ x - y = a \\ 2x + y = -1 \end{cases}$$

L'image $\operatorname{Im} A$ est engendrée par les deux vecteurs $V_1 = (1, 1, 2)^t$ et $V_2 = (1, -1, 1)^t$ et contient le second membre $B_a = (1, a, -1)^t$ si et seulement si $a = -5$ et on a $B_{-5} = -2V_1 + 3V_2$. Ainsi, le système n'a de solution que si $a = -5$ et dans ce cas la solution est unique, valant $(x = -2, y = 3)$. $\triangleleft$

4.2. Déterminants

Les déterminants sont des objets mathématiques associés à des systèmes de n vecteur dans un espace de dimension n et donc aussi par suite à des matrices carrées d'ordre n considérées comme une liste de n vecteurs colonne à n lignes, soient des vecteurs de $\mathbb{R}^n$. Malgré leurs propriétés intéressantes et profondes, leur intérêt pratique pour la résolution numérique de problèmes linéaires est modéré, car leur calcul explicite est complexe ; en particulier, l'utilisation des déterminants pour la solution des systèmes d'équations linéaires, dite méthode de Cramer, est en général à proscrire, sauf pour les systèmes d'ordre peu élevé et des visées théoriques.

4.2.1. Définition et propriétés. — Soit E un espace vectoriel de dimension n. On appelle *forme n-linéaire alternée* toute application f de E dans $\mathbf{R}$, qui possède les propriétés suivantes :

– $\boxed{\textit{linéarités partielles}}$

$$f(V_1, \ldots, \alpha V_i + \beta W_i, \ldots, V_n) = \alpha f(V_1, \ldots, V_i, \ldots, V_n) + \beta f(V_1, \ldots, W_i, \ldots, V_n)$$

– $\boxed{\textit{changement de signe par transposition de deux vecteurs}}$

$$f(V_1, \ldots, V_i, \ldots, V_j, \ldots, V_n) = -f(V_1, \ldots, V_j, \ldots, V_i, \ldots, V_n)$$

En particulier, si $V_i = V_j$ avec $i \neq j$, le scalaire $f(V_1, \ldots, V_i, \ldots, V_j, \ldots, V_n)$, égal à son opposé, est nul.

De ces propriétés, on déduit la conséquence importante : si le vecteur V_k est une combinaison linéaire des autres $V_k = \sum_{i=1, i\neq k}^{n} \alpha_i V_i$, le scalaire $f(V_1, \ldots, V_n)$ est nul.

On peut déduire la forme générale des formes n-linéaires alternées, en commençant par les dimensions 2 et 3.

– Pour l'espace $\mathbb{R}^2$, le vecteur V_i s'écrit relativement à la base $\mathbf{b} = (b_1, b_2)$ sous la forme $V_i = \alpha_{i1}b_1 + \alpha_{i2}b_2$, d'où appliquant les remarques précédentes, on obtient :

$$\begin{aligned} f(V_1, V_2) &= f\left(\sum_{j=1}^{2} \alpha_{1j}b_j, \sum_{j=1}^{2} \alpha_{2j}b_j\right) \\ &= \alpha_{11}\alpha_{21}f(b_1, b_1) + \alpha_{11}\alpha_{22}f(b_1, b_2) + \alpha_{12}\alpha_{21}f(b_2, b_1) + \alpha_{12}\alpha_{22}f(b_2, b_2) \\ &= (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})f(b_1, b_2). \end{aligned}$$

– Après développement et disparition des termes nuls, on obtient pour une forme déterminant sur l'espace $\mathbb{R}^3$:

$$\begin{aligned} f(V_1V_2, V_3) &= f\left(\sum_{j=1}^{3} \alpha_1 j b_j, \sum_{j=1}^{3} \alpha_2 j b_j, \sum_{j=1}^{3} \alpha_{3j}b_j\right) \\ &= (\alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{13}\alpha_{21}\alpha_{32} + \alpha_{12}\alpha_{23}\alpha_{31} \\ &\qquad - \alpha_{12}\alpha_{21}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31})f(b_1, b_2, b_3). \end{aligned}$$

Pour mémoriser les signes $\pm$ de cette formule, on a la « règle de Sarrus », propre à la dimension 3. Pour calculer le déterminant

$$D = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}$$

on complète le tableau en lui ajoutant ses deux premières lignes sous la troisième, obtenant ainsi un tableau à 5 lignes et 3 colonnes, puis on prend comme termes positifs les produits provenant de la lecture dans ce tableau des trois parallèles à la diagonale principale et comme termes négatifs ceux provenant des parallèles à la diagonale secondaire (celle portant α_{13}, α_{22} et α_{31})

$$\begin{array}{ccccccc} + & \alpha_{11} & & \alpha_{12} & & \alpha_{13} & - \\ + & \alpha_{21} & \ddots & \alpha_{22} & & \alpha_{23} & - \\ + & \alpha_{31} & \ddots & \alpha_{32} & \ddots & \alpha_{33} & - \\ & \alpha_{11} & \ddots & \alpha_{12} & \ddots & \alpha_{13} & \\ & \alpha_{21} & & \alpha_{22} & \ddots & \alpha_{23} & \end{array}$$

– Pour un espace de dimension n, on obtient de manière analogue pour

$$V_i = \sum_{j=1}^{n} \alpha_{ij}b_j, \quad i = 1, \dots, n,$$

$$(9) \quad f(V_1, V_2, \dots, V_n) = \sum_{\pi \text{ permutation sur } (1,2,3,\dots,n)} \varepsilon_\pi \alpha_{1\pi(1)}\alpha_{2\pi(2)} \dots \alpha_{n\pi(n)} \quad f(b_1, b_2, \dots, b_n)$$

où ε_π est le signe de la permutation π. On retrouve bien entendu les cas particuliers précédents.

On montre que la formule (9) définit bien une n-forme multilinéaire alternée, qui est uniquement déterminée par la valeur $f(b_1, b_2, \dots, b_n)$. On appelle *déterminant* associée à la base $\mathbf{b}$ (qui sera dans le cas de $\mathbb{R}^n$ la base canonique) l'unique forme multilinéaire alternée f telle que $f(b_1, b_2, \dots, b_n) = 1$.

Le déterminant $\det A$ d'une matrice $A = (A_{ij})$ carrée d'ordre n est le déterminant de la famille constituée de ses n vecteurs colonne dans $\mathbb{R}^n$ muni de sa base canonique. On note

$$\det A = \begin{vmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nn} \end{vmatrix}.$$

De ces propriétés fondamentales, on déduit les résultats suivants :

– Soit n vecteurs $(V_i)_{i=1,\dots,n}$, qui se décomposent sur la base $(b_1, \dots, b_n)$ suivant $V_i = \sum_{j=1}^n \alpha_{ij} b_j$. Les vecteurs $V_1, \dots, V_n$ sont linéairement indépendants si et seulement si le déterminant $\det(\alpha_{ij})$ est non nul.
– Si on échange 2 lignes (ou 2 colonnes) d'une matrice, le déterminant change de signe. On peut ajouter à une ligne (une colonne resp.) n'importe quelle combinaison des autres lignes (autres colonnes resp.) sans changer la valeur du déterminant.

4.2.2. Méthodes de calcul. — La définition (9) est intéressante formellement mais, vu sa taille (somme de $n!$ termes) inapplicable pratiquement dès que la matrice est d'ordre au moins 4.

La méthode classique d'évaluation d'un déterminant consiste à développer selon une ligne ou une colonne. Par exemple, en développant selon la j-ème colonne on a

$$\det A = \sum_{i=1}^n (-1)^{i+j} A_{ij} \mathrm{Mineur}_{ij}(A)$$

où $\mathrm{Mineur}_{ij}(A)$ est le déterminant de la matrice obtenue à partir de A en supprimant la i-ème ligne et la j-ème colonne (on l'appelle aussi cofacteur).

▷ ***Exemple 4.2.*** Calculons le déterminant de la matrice $A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & -2 & -2 \end{pmatrix}$.

$$\det A = \begin{vmatrix} 2 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & -2 & -2 \end{vmatrix} = 2 \begin{vmatrix} 0 & 1 \\ -2 & -2 \end{vmatrix} - 2 \begin{vmatrix} -1 & 1 \\ 0 & -2 \end{vmatrix} = 2 \cdot 2 - 2 \cdot 2 = 0. \qquad \triangleleft$$

Le calcul d'un déterminant d'ordre n, se ramène au calcul de n déterminants d'ordre $n-1$, qui pour chacun d'entre eux nécessitent $n-1$ calculs de déterminants d'ordre $n-2$, etc. Le calcul d'un déterminant par la méthode classique est de complexité $n!$. Il est à proscrire dès que $n > 3$ sauf configuration très particulière (beaucoup d'éléments nuls).

La méthode du pivot de Gauss consiste, par ajout de combinaison linéaire de lignes et de colonnes et/ou permutation de lignes et colonnes, à transformer la matrice en une matrice triangulaire supérieure ou inférieure : c'est le pendant pour le calcul des déterminants de la méthode de résolution des systèmes par combinaison des équations scalaires. Pour une matrice triangulaire T, en appliquant la méthode de décomposition sur les colonnes, on on obtient aisément son déterminant comme produit des coefficients diagonaux $\det T = \prod_{i=1}^n T_{ii}$.

4.2.3. Propriétés et applications. —

– ***Produit de matrices*** Soit deux matrices carrées A et B de mêmes dimensions. Alors $\det AB = \det BA = \det A \det B$
– ***Inverse d'une matrice*** A est une matrice inversible si et seulement si $\det A \neq 0$, dans ce cas $\det(A^{-1}) = (\det A)^{-1}$

– $\boxed{\textit{Matrices semblables}}$ Deux matrices semblables A et B ont même déterminant ; en effet, il existe une matrice inversible P, telle que $B = P^{-1}AP$ et par suite

$$\text{dét}\, B = \text{dét}(P^{-1}AP) = (\text{dét}\, P)^{-1}\, \text{dét}\, A\, \text{dét}\, P = \text{dét}\, A.$$

4.3. Résolution des systèmes de Cramer

Soit un système de n équations linéaires à n inconnues mis sous la forme matricielle $AX = B$. On considère A comme une ligne de n vecteurs colonnes : $A = (C_1C_2 \dots C_n)$. Le théorème suivant est attribué à Cramer

Théorème 4.1. *Soit $AX = B$ un système de n équations linéaires à n inconnues avec une matrice carrée A régulière. Alors l'unique solution $X = (x_1x_2 \dots x_n)^t$ est donnée par*

$$x_i = \frac{\text{dét}[C_1 \dots C_{i-1}BC_{i+1} \dots C_n)}{\text{dét}\, A}, \quad i = 1, \dots, n.$$

L'application de cette règle requiert le calcul de n déterminants de matrices carrée d'ordre n : les méthodes de pivot sont beaucoup plus efficaces pour le calcul effectué par un humain !

4.4. Exercices

E4.1. Calculer le déterminant des matrices suivantes, indiquer lesquelles sont inversibles et calculer leur inverse.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 0 & -5 \end{pmatrix}, \begin{pmatrix} 12 & 0 \\ 7 & 3 \end{pmatrix}, \begin{pmatrix} -4 & 3 \\ 8 & -6 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 5 & 3 \\ 1 & 5 & 4 \\ -3 & 7 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 2 \\ 1 & -1 & 2 \\ -3 & 2 & -1 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 2 \\ 0 & -3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

E4.2. Soit $a \in \mathbb{R}^*$. On considère le système

$$(S) : \begin{cases} ay + a^2 z &= a^2 \\ \dfrac{1}{a}x + az &= a \\ \dfrac{1}{a^2}x + \dfrac{1}{a}y &= 1 \end{cases}$$

1. Déterminer trois matrices A, X, B telles que le système (S) est équivalent à l'équation matricielle $AX = B$.
2. Montrer que $A^2 - A - 2I = 0$.
3. En déduire que la matrice A est inversible et calculer A^{-1}.
4. Résoudre le système (S)

E4.3. On considère la matrice

$$J = \begin{pmatrix} 0 & 0 & 1 & -1 \\ -2 & 1 & 2 & -1 \\ 2 & -1 & -1 & 0 \\ 3 & -2 & -2 & 0 \end{pmatrix}$$

1. Calculer J^2, J^3 et J^4. Que peut-on en déduire de J^k pour $k \geq 4$?
2. Développer algébriquement l'expression $(I+J)(I-J+J^2-J^3)$.
3. La matrice $I+J$ est-elle inversible ? Si oui, expliciter cet inverse.
4. Déterminer les solutions du système

$$\left\{ \begin{array}{ccc} a+c-d & = & 2 \\ -2a+2b+2c-d & = & 0 \\ 2a-b & = & 0 \\ 3a-2b-2c+d & = & 3 \end{array} \right.$$

E4.4. Résoudre les systèmes linéaires suivants en utilisant les déterminants :

$$\left\{ \begin{array}{ccccccc} 2x & + & 3y & - & z & = & 1 \\ x & + & 2y & + & 3z & = & 2 \\ 3x & + & 4y & - & 5z & = & -4 \end{array} \right. , \qquad \left\{ \begin{array}{ccccccc} x & - & y & + & 3z & = & -2 \\ x & & & + & z & = & 0 \\ x & + & 4y & - & z & = & 2 \end{array} \right. .$$

E4.5. Soit la matrice A d'ordre 3 définie par :

$$\begin{pmatrix} 1 & 1 & 1 \\ x & 1 & x^2 \\ 1 & x^2 & x^4 \end{pmatrix}$$

Calculer $\det A$. Pour quelles valeurs de x ce déterminant s'annule-t-il ?

E4.6. Soit la matrice $A(x)$ définie par

$$A(x) = \begin{pmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{pmatrix}$$

Calculer le déterminant de $A(x)$. Pour les valeurs de x qui annulent le déterminant $\det A(x)$, calculer l'ordre maximum du mineur non nul.

E4.7. Soit D le déterminant d'une matrice bi-bande : $A_{ij} = 0$, sauf si $i = j$ et $i = j+1$. Montrer que D ne dépend que de sa diagonale principale.

E4.8. Montrer que

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & c & b \\ 1 & c & 0 & a \\ 1 & b & a & 0 \end{vmatrix} = a^2 + b^2 + c^2 - 2ab - 2ac - 2ca.$$

CHAPITRE 5

VALEURS ET VECTEURS PROPRES

Définition 5.1. Étant donné une application linéaire f de l'espace vectoriel E, le vecteur V non nul est dit *vecteur propre* de f, associé à la *valeur propre* (réelle [1]) $\lambda \in \mathbb{R}$ de l'application f si $f(V) = \lambda V$.

Étant donné une matrice A, le vecteur colonne X non nul est dit *vecteur propre* de A avec *valeur propre* λ si $AX = \lambda X$.

Ces deux définitions sont compatibles : si V est un vecteur propre de f, alors, le vecteur X_V des coordonnées de V relativement à une base b est vecteur propre de la matrice A_f représentant f dans cette base b. Par ailleurs, si X est vecteur propre de la matrice A, la matrice colonne X vu comme vecteur de $\mathbb{R}^n$ est vecteur propre de l'application linéaire [2] A de $\mathbb{R}^n$ représentée par la matrice A relativement à la base canonique de $\mathbb{R}^n$.

△ ***Remarque 5.1.***

1. Un vecteur propre est en fait une direction propre, car tout multiple d'un vecteur propre est aussi vecteur propre.
2. Dans la version matricielle, il est aisé d'étendre la définition des *éléments propres* (vecteur propre et valeur propre associée) à des éléments à valeurs complexes.
3. Les logiciels de calcul courants (`Maple`, `Matlab`, `Scilab`,...) ont des fonctions donnant les valeurs propres, avec des vecteurs propres associés. ▽

▷ ***Exemples 5.1.***

1. Si
$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
la matrice A a $\lambda = 2 \pm \sqrt{5}$ comme valeur propre associée au vecteur propre $V = (2, 1 \pm \sqrt{5})^t$, alors que B a comme valeur propre $\pm \mathrm{i}$ avec comme vecteur propres respectifs $(1 \pm \mathrm{i})^t$: la matrice B est réelle, mais a des valeurs propres complexes !

---

1. Même si on ne considère ici principalement que les espaces vectoriels sur les réels, la considération de valeurs propres complexes est quasiment obligée pour la théorie générale : on le comprendra aisément dès que l'on constate que les valeurs propres sont racines du polynôme caractéristique P_A, qui peut avoir des racines complexes !

2. L'usage de la même lettre pour l'application linéaire et la matrice qui la représente, ou à l'inverse une matrice et une application linéaire représentée par la matrice dans des bases convenables, est faite désormais.

2. Si $AX = \lambda X$, alors $A^n X = \lambda^n X$: si X est vecteur propre de A avec valeur propre λ, il est aussi vecteur propre de A^n avec valeur propre λ^n. On montrera ci-après que la connaissance des éléments propres d'une matrice permet de calculer efficacement les puissances d'une matrice : ce calcul (pour des matrices de très grande taille) est utilisé dans les algorithmes d'indexation préférentielle des moteurs de recherche tels `google`.
3. Si D est l'opérateur linéaire de dérivation sur l'espace des fonctions à valeurs complexes indéfiniment dérivables sur $\mathbb{R}$, alors la fonction exponentielle $e_\lambda : t \in \mathbb{R} \to \mathrm{e}^{\lambda t} \in \mathbb{C}$ est vecteur propre de D avec valeur propre λ : $De_\lambda = \lambda e_\lambda$.
4. La projection P_Δ sur la droite Δ passant par l'origine a deux directions propres : d'une part celle des vecteurs V portés par la droite elle-même, correspondant à la valeur propre $\lambda = 1$ d'autre part les vecteurs V perpendiculaires à la droit D, correspondant à la valeur propre $\lambda = 0$.
5. La symétrie σ_Δ par rapport à la droite Δ passant par l'origine : cette application a deux directions propres : celle des vecteurs portés par la droite elle-même pour la valeur propre $\lambda = 1$ et la direction orthogonale avec valeur propre -1 dans ce cas.
6. La rotation plane $\mathcal{R}_\theta$ d'angle θ non nul et non plat n'a pas de vecteurs (réels !) non nuls correspondant à des valeurs propres réelles. Néanmoins, les complexes $\mathrm{e}^{\pm \mathrm{i}\theta}$ sont valeurs propres, avec vecteurs propres respectifs $(1, \mp \mathrm{i})$ de la matrice $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ représentant la rotation $\mathcal{R}_\theta$ dans la base canonique. $\triangleleft$

La question essentielle est de savoir si les vecteurs propres de A sont linéairement indépendants et peuvent former une base de E : l'existence de telles bases permet de simplifier les calculs dans divers problèmes linéaires, comme les applications de la section 5.3 en témoignent.

5.1. Éléments propres

Le vecteur V est vecteur propre avec valeur propre λ si et seulement si V appartient au noyau de l'application $A - \lambda I_n$ puisque l'égalité $(A - \lambda I_n)V = 0$ signifie $V \in \ker(A - \lambda I_n)$.

D'après le Théorème 3.1, pour que le noyau $\ker(A - \lambda I_n)$ soit non réduit au sous-espace nul $\{0\}$, il est nécessaire et suffisant que $A - \lambda I_n$ soit non injectif ou que $A - \lambda I_n$ ne soit pas un isomorphisme, soit

$$\text{dét}(A - \lambda I_n) = 0.$$

Définition 5.2. Le *polynôme caractéristique* P_A d'une matrice carrée A d'ordre n est le polynôme de degré n donné par le déterminant de $A - \lambda I_n$. Le *spectre* de la matrice A est défini comme l'ensemble des valeurs propres (réelles ou complexes) de A : c'est l'ensemble des racines du polynôme caractéristique.

Si le polynôme P_A a n racines réelles

$$\text{dét}(A - \lambda I_n) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$$

ces racines donnent les valeurs propres de A.

Remarquons que les polynômes caractéristiques (et donc leurs spectres) de deux matrices semblables coïncident :

$$\begin{aligned} P_{P^{-1}AP}(\lambda) &= \text{dét}(P^{-1}AP - \lambda I_n) = \text{dét}(P^{-1}(A - \lambda I_n)P) \\ &= \text{dét}(P^{-1})\,\text{dét}(A - \lambda I_n)\,\text{dét}\, P = \text{dét}(A - \lambda I_n) = P_A(\lambda). \end{aligned}$$

Ainsi, deux matrices semblables ont mêmes spectres.

Le calcul des éléments propres débute par le calcul du polynôme caractéristique et la recherche de ses racines. La détermination des vecteurs propres pour la valeur propre λ_i repose sur l'étude du noyau $\ker(A - \lambda_i I_n)$, soit la résolution de l'équation linéaire $(A - \lambda_i)V = 0$ d'inconnue V.

▷ ***Exemple 5.2***. Soit la matrice $A = \begin{pmatrix} 3 & -4 \\ 1 & -2 \end{pmatrix}$. Son polynôme caractéristique est

$$\text{dét}(A - \lambda I_2) = \begin{vmatrix} 3-\lambda & -4 \\ 1 & -2-\lambda \end{vmatrix} = \lambda^2 - \lambda - 2$$

et a pour racines $\lambda_1 = -1, \lambda_2 = 2$. On vérifie au passage les identités (en fait générales) $\operatorname{tr} A = \lambda_1 + \lambda_2 = 1$ et $\text{dét}\, A = \lambda_1\lambda_2 = -2$.

Équivalent à l'équation linéaire $AV = \lambda V$, le système linéaire donnant les vecteurs propres $V = (x, y)$ associés à la valeur propre λ prend la forme

$$\begin{cases} (3-\lambda)x - 4y = 0 \\ x + (-2-\lambda)y = 0 \end{cases}$$

Pour $\lambda_1 = -1$, le système $4x - 4y = 0, x - y = 0$, bien entendu dégénéré, est équivalent à l'unique équation $x = y$, équation de la droite dont on peut prendre comme vecteur directeur $V_1 = (1,1)^t$. Pour $\lambda_2 = 2$, le système est équivalent à l'unique équation $x - 4y = 0$, d'où $V_2 = (4,1)^t$ comme vecteur propre.

Les deux vecteurs propres V_1 et V_2 sont linéairement indépendants et constituent une base du plan, avec matrice de l'opérateur A relativement à cette base qui est diagonale $\widetilde{A} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$. Le changement de la base canonique à la base (V_1, V_2) a pour matrice

$$P = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$$

d'inverse

$$P^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{4}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

Ainsi, la formule $\widetilde{A} = P^{-1}AP$ redonne la forme réduite de l'application de A dans la nouvelle base

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}.$$ ◁

5.2. Diagonalisation

On admettra le théorème suivant :

Théorème 5.1. *Des vecteurs propres correspondant à des valeurs propres distinctes sont linéairement indépendants.*

Pour qu'une matrice A d'ordre n soit semblable à une matrice diagonale, il faut et il suffit que la matrice A ait n vecteurs propres linéairement indépendants.

Corollaire 5.1. *Toute matrice carrée ayant n valeurs propres distinctes est diagonalisable.*

Soit la matrice A carrée d'ordre r avec λ_0 comme (unique) valeur propre dégénérée au sens où λ_0 est racine d'ordre $r > 1$ du polynôme caractéristique de A : alors, si A a r vecteurs propres linéairement indépendants correspondant à la valeur propre λ_0, A est diagonalisable : A est en fait la matrice $\lambda_0 I_r$. Il se peut que A n'ait pas de base de vecteurs propres (et ne soit pas égale à $\lambda_0 I_r$) : une telle matrice non diagonalisable (dans l'espace des matrices complexes) est néanmoins toujours semblable à une forme réduite de Jordan

$$J(\lambda_0) = \begin{pmatrix} J_{r_1}(\lambda_0) & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & J_{r_\ell}(\lambda_0) & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & J_{r_k}(\lambda_0) \end{pmatrix} \tag{10}$$

avec $r = r_1 + \ldots + r_k$ et les matrices $J_{r_j}(\lambda_0)$ du type

$$J_r(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \lambda_0 & 1 \\ 0 & \dots & \dots & 0 & \lambda_0 \end{pmatrix}.$$

On vérifie que le polynôme caractéristique $P_{J_r(\lambda_0)}$ de la matrice $J_r(\lambda_0)$ est $P_{J_r(\lambda_0)}(\lambda) = (\lambda - \lambda_0)^r$. On montre que une matrice A est semblable à une matrice diagonales par blocs, *i. e.* de la forme (10) avec des blocs de Jordan $J_r(\lambda)$ $(r \geq 1)$ [3]

***Théorème 5.2* (Cayley-Hamilton).** *Une matrice A annule son polynôme caractéristique $P_A(\lambda) = \text{dét}(A - \lambda I_n)$, soit $P_A(A) = 0$.*

Démonstration. — Si $B = P^{-1}AP$, alors

$$P_A(A) = \sum_{i=1}^n a_i A^i = \sum_{i=1}^n a_i \left(PBP^{-1}\right)^i = \sum_{i=1}^n a_i PB^iP^{-1} = P\left(\sum_{i=1}^n a_i B^i\right)P^{-1} = PP_A(B)P^{-1} = PP_B(B)P^{-1}$$

vu que les polynômes caractéristiques de A et B coïncident : en admettant que A est semblable à une matrice diagonale de blocs de Jordan, on obtient que $P_A = \prod_{r,\lambda} P_{J_r(\lambda)}$. Le théorème de Cayley-Hamilton résulte alors de la nullité (aisément vérifiée) $P_{J_r(\lambda)}(J_r(\lambda)) = 0$. □

Pour une matrice inversible d'ordre petit, le théorème de Cayley-Hamilton peut fournir un moyen alternatif de calcul son inverse.

▷ ***Exemple 5.3.*** Si le polynôme caractéristique P_A d'une matrice A est $P_A(\lambda) = \lambda^2 - 4\lambda + 1$, alors $P_A(A) = A^2 - 4A + I_2 = 0$, ce qui est équivalent à $I_2 = -A^2 + 4A = (-A + 4I_2)A$ et par suite $A^{-1} = -A + 4I_2$. ◁

Si A est une matrice inversible semblable à une matrice diagonale $D = (D_{ij}\delta_{ij})$, alors $A - \lambda I_n$ est semblable à $D - \lambda I_n = ((D_{ij} - 1)\delta_{ij})$ et si A est inversible, son inverse A^{-1} est semblable à la matrice diagonale D^{-1} de coefficients diagonaux D_{ii}^{-1}.

3. Ce résultat sera admis !

Deux matrice semblables ayant même trace et même déterminant, on déduit que

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i, \quad \operatorname{dét} A = \prod_{i=1}^{n} \lambda_i$$

si A est diagonalisable de spectre $\lambda_1, \dots, \lambda_n$. En fait, ces égalités sont aussi valables lorsque A n'est pas diagonalisable : il suffit de considérer les λ_i comme les n racines (complexes) du polynôme caractéristique P_A de degré n.

Proposition 5.1*. Une matrice A symétrique est diagonalisable, avec valeurs propres réelles. Il existe une matrice de changement de base P orthogonale rendant A diagonale.*

Vu que P est orthogonale, on a $P^t = P^{-1}$: l'orthogonalité de P peut s'interpréter en terme d'orthogonalité [4] des vecteurs propres : deux de vecteurs propres correspondant à des valeurs propres distinctes sont orthogonaux [5] : il existe une base orthonormée de vecteurs propres et, par suite, une matrice P unitaire (ou orthogonale réelle) telle que $P^{-1}AP$ soit diagonale [6].

5.3. Applications

5.3.1. Puissance de matrices. — Si A est semblable à une matrice diagonale D, il existe une matrice inversible P telle que $A = P^{-1}DP$, et aussi $A^p = P^{-1}D^pP$. Comme D^p est aisément calculable, on pourra calculer la puissance A^p facilement :

$$A^p = P^{-1}D^pP,$$

ce qui sera utile dans les études de suites définies par récurrence de la sous-section 5.3.2

Si on définit pour une matrice A carrée d'ordre n son exponentielle suivant la série (convergente)

$$\mathrm{e}^A = I_n + A + \frac{A^2}{2} + \ldots + \frac{A^p}{p!} + \ldots$$

alors pareillement

$$\mathrm{e}^A = P^{-1}\mathrm{e}^D P = P^{-1} \begin{pmatrix} \mathrm{e}^{D_{11}} & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & \mathrm{e}^{D_{ii}} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \mathrm{e}^{D_{nn}} \end{pmatrix} P.$$

---

4. Cette notion est introduite dans le chapitre suivant

5. Soient (λ_1, X_1) et (λ_2, X_2) deux couples de valeurs et vecteurs propres de la matrice symétrique A. Alors $\lambda_1 X_1^t X_2 = (AX_1)^t X_2 = X_1^t A^t X_2 = X_1^t A X_2 = \lambda_2 X_1^t X_2$ et la nullité de $X_1^t X_2$ si $\lambda_1 \neq \lambda_2$: c'est l'orthogonalité de X_1 et X_2. L'existence pour une matrice symétrique d'une base de vecteurs colonnes propres, orthogonaux deux à deux, est admise ici sans démonstration : c'est le caractère diagonalisable de toute matrice réelle symétrique.

6. Une matrice P est orthogonale si et seulement si $P^tP = I_n$: cette égalité est équivalente aux n^2 relations $C_i^t C_j = \delta_{ij}$ sur les vecteurs colonnes C_i de P, *i. e.* l'orthogonalité des vecteurs C_i et C_j pour $i \neq j$ et l'unitarité des vecteurs C_i : c'est exactement la définition d'une base orthonormée (*cf.* Section 6.2).

On montre que l'application $E_A : t \in \mathbb{R} \to \mathrm{e}^{tA} \in M_n(\mathbb{R})$ est dérivable, sa dérivée vérifiant $E_A'(t) = AE_A(t), t \in \mathbb{R}$: il en découle que pour tout vecteur X la fonction $\in \mathbb{R} \to E_A(t)X$ vérifie l'équation différentielle

$$x'(t) = Ax(t)$$

étudiée dans la sous-section 5.3.3.

5.3.2. Suites définies par une relation de récurrence linéaire. — La diagonalisation de matrices permet de décrire les suites récurrentes $\mathbf{u} = (u_n)_{n\geq 0}$ satisfaisant des relations de récurrence linéaire du type

$$u_{n+p} + a_1 u_{n+p-1} + a_2 u_{n+p-2} + \ldots + a_p u_n = 0.$$

Voyons-le sur deux exemples, dont celui de la suite de Fibonacci mentionnée dans le prologue.

▷ ***Exemples 5.4.***

1. Soit le système

(11) $$\begin{cases} u_{n+1} = u_n - v_n \\ v_{n+1} = 2u_n + 4v_n \end{cases}$$

   avec $u_0 = 2$ et $v_0 = 1$. Si on pose $X_n = (u_n, v_n)^t$, le système (6) est équivalent) $X_{n+1} = AX_n$ avec $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$, soit $X_n = A^n X_0$. La matrice A se diagonalise suivant

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & -1 \end{pmatrix}.$$

   Alors, la puissance de A se calcule aisément

$$A^n = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2^{1+n} - 3^n & 2^n - 3^n \\ -2^{1+n} + 2\cdot 3^n & -2^n + 2\cdot 3^n \end{pmatrix}$$

   et finalement

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = X_n = A^n X_0 = \begin{pmatrix} 5\cdot 2^n - 2\cdot 3^n \\ -5\cdot 2^n + 6\cdot 3^n \end{pmatrix}.$$

2. La suite de Fibonacci du prologue et de l'exemple 1.20.3 vérifie

$$u_n = u_{n-1} + u_{n-2}, \quad n \geq 2,$$

   relation de récurrence qu'on peut réécrire, en introduisant $X_n = (u_n, u_{n-1})^t$ et la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, suivant

$$X_n = AX_{n-1} = A^{n-1} X_1, \quad n \geq 1.$$

   Les valeurs propres de A sont $\lambda_\pm = (1 \pm \sqrt{5})/2$, racines du polynôme caractéristique $\text{dét}(A - \lambda I_2) = \lambda^2 - \lambda - 1$. Ainsi, et comme dans l'exemple précédent, on obtient

$$u_n = a_+ \lambda_+^n + a_- \lambda_-^n, n \geq 0$$

   les constantes $a_\pm$ étant déterminées par les conditions initiales u_0 et u_1. ◁

5.3.3. Systèmes différentiels linéaires à coefficients constants. — Le système différentiel linéaire à coefficients constants

$$\begin{cases} x_1'(t) & = a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) + b_1(t) \\ x_2'(t) & = a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) + b_2(t) \\ \vdots & \\ x_i'(t) & = a_{i1}x_1(t) + a_{i2}x_2(t) + \ldots + a_{in}x_n(t) + b_i(t) \\ \vdots & \\ x_n'(t) & = a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) + b_n(t) \end{cases} \tag{12}$$

s'écrit sous la forme concise

$$X'(t) = AX(t) + B(t)$$

avec $X(t) = (x_1(t), \ldots, x_n(t))^t$ le vecteur des n inconnues, $B(t) = (b_1(t), \ldots, b_n(t))^t$ celui de n fonctions de t connues et $A = (a_{ij})$ la matrice carrée d'ordre n des coefficients constants de (12).

Théorème 5.3. *Les solutions du système homogène $Y'(t) = AY(t)$ forment un espace vectoriel de dimension n.*

La solution générale du système (12) est la somme d'une solution particulière $X_0(t)$ de l'équation et de la solution générale de l'équation homogène $Y'(t) = AY(t)$.

Théorème 5.4. *Si la matrice A est diagonalisable, semblable à la matrice diagonale $D = (D_{ii}\delta_{ij})$ par la matrice régulière P,* i. e. $A = PDP^{-1}$, *la solution du système homogène $Y'(t) = AY(t)$ est de la forme*

$$Y(t) = P \begin{pmatrix} C_1 e^{D_{11}t} \\ C_2 e^{D_{22}t} \\ \vdots \\ C_n e^{D_{nn}t} \end{pmatrix}.$$

Démonstration. — Posons $Z = P^{-1}Y$. Alors $Z' = P^{-1}Y' = P^{-1}AY = P^{-1}APZ = DZ$: l'équation différentielle sur le vecteur Z se découple en n équation différentielles scalaires $z_i' = D_{ii}z_i$, dont la solution est $z_i(t) = C_i e^{D_{ii}t}$. On obtient alors Y via la relation $Y = PZ$. □

▷ ***Exemples 5.5.***

1. On considère le système d'équations différentielles homogène :

$$\begin{cases} x_1'(t) - x_1(t) + x_2(t) - 2x_3(t) = 0 \\ x_2'(t) + x_1(t) - x_3(t) = 0 \\ x_3'(t) - 2x_1(t) - x_2(t) - x_3(t) = 0 \end{cases}$$

La matrice A du problème est :

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

Elle est diagonalisable sous la forme : $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = P^{-1}AP$ par la matrice P et son inverse

$$P = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} \frac{1}{6} & -\frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Le système réduit $\widetilde{X}'(t) = D\widetilde{X}(t)$ a pour solution :

$$\widetilde{X}(t) = \begin{pmatrix} C_1\mathrm{e}^t \\ C_2\mathrm{e}^{-2t} \\ C_3\mathrm{e}^{3t} \end{pmatrix}.$$

Le système a donc pour solution générale :

$$X(t) = P\widetilde{X}(t)) = \begin{pmatrix} C_1\mathrm{e}^t + C_2\mathrm{e}^{-2t} + C_3\mathrm{e}^{3t} \\ -2C_1\mathrm{e}^t + C_2\mathrm{e}^{-2t} \\ -C_1\mathrm{e}^t - C_2\mathrm{e}^{-2t} + C_3\mathrm{e}^t \end{pmatrix}.$$

Les conditions initiales constituant le vecteur $X(0)$ déterminent la valeur des constantes $C_i, i = 1, 2, 3$.

2. L'équation différentielle de Fibonnaci $y'' = y' + y$ du prologue s'exprime matriciellement suivant $X'(t) = AX(t)$ avec

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad X = (x_1, x_0)$$

soit

$$\begin{cases} x_1' &= x_1 + x_0 \\ x_0' &= x_1 \end{cases}$$

Le système $Z' = \begin{pmatrix} (1+\sqrt{5})/2 & 0 \\ 0 & (1-\sqrt{5})/2 \end{pmatrix} Z$ se ramène aux deux équations différentielles $z_\pm' = (1 \pm \sqrt{5})/2 z_\pm$ si $Z = (z_+, z_-)$. On en déduit que la solution y est de la forme

$$y = a_+\mathrm{e}^{(1+\sqrt{5})/2t} + a_-\mathrm{e}^{(1-\sqrt{5})/2t}. \qquad \triangleleft$$

Si la matrice A est diagonalisable à valeurs propres complexes, celles-ci sont conjuguées deux à deux ; en posant les constantes correspondantes complexes conjuguées, on obtient des solutions trigonométriques.

$\triangleright$ ***Exemple 5.6***. On considère le système homogène :

$$\begin{cases} x_1'(t) = x_1(t) - x_2(t) \\ x_2'(t) = 5x_1(t) - 3x_2(t) + 4x_3(t) \\ x_3'(t) = 4x_1(t) - 2x_2(t) + 4x_3(t) \end{cases}$$

La forme réduite de la matrice $A = \begin{pmatrix} 1 & -1 & 0 \\ 5 & -3 & 4 \\ 4 & -2 & 4 \end{pmatrix}$ est complexe :

$$\widetilde{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1+\mathrm{i} & 0 \\ 0 & 0 & 1-\mathrm{i} \end{pmatrix}.$$

La solution générale est donc :

$$\begin{cases} x_1(t) = -2C_1 + 2C_2\mathrm{e}^t\cos(t) - 2C_3\mathrm{e}^t\sin(t) \\ x_2(t) = -2C_1 + 2C_3\mathrm{e}^t\cos(t) + 2C_2\mathrm{e}^t\sin(t) \\ x_3(t) = C_1 - 2C_2\mathrm{e}^t\cos(t) + 2C_3\mathrm{e}^t\sin(t) \end{cases}$$ ◁

Si la matrice A n'est pas diagonalisable, on utilise la réduction de Jordan. On montre que les solutions sont des exponentielles polynômes $t^k\mathrm{e}^{\lambda t}$, l'ordre k étant majoré par les ordres des blocs de Jordan élémentaires. Dans beaucoup de cas particuliers, on peut se passer de la théorie générale (qui n'est pas traitée ici).

▷ ***Exemple 5.7.***

$$\begin{cases} x_1'(t) = x_1(t) + x_2(t) \\ x_2'(t) = x_2(t) \end{cases}$$

Il admet la solution immédiate $x_2(t) = C_2\mathrm{e}^t$ qui, portée dans la première équation, donne : $x_1(t) = C_1\mathrm{e}^t + tx_2(t)$. ◁

5.4. Exercices

E5.1. Calculer les valeurs propres, puis les vecteurs propres des matrices suivantes.

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 2 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix}, B = \begin{pmatrix} 5 & -1 & 2 \\ -1 & -3 & -2 \\ 2 & 2 & 2 \end{pmatrix}, C = \begin{pmatrix} -7 & 1 & -1 \\ -7 & 1 & -1 \\ -6 & 2 & -2 \end{pmatrix}, D = \begin{pmatrix} 4 & 1 & -1 \\ 1 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix}.$$

E5.2. Sans faire aucun calcul, dire si les matrices suivantes sont diagonalisables

$$A = \begin{pmatrix} \cos(\pi/7) & -\sin(\pi/7) \\ \sin(\pi/7) & \cos(\pi/7) \end{pmatrix} \quad B = \begin{pmatrix} 2 & -5 & -3 \\ 0 & 2 & 10 \\ 0 & 0 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 2 & -5 & 3 \\ -5 & 3 & 9 \\ 3 & 9 & 5 \end{pmatrix}.$$

E5.3. À quelles conditions la matrice

$$A = \begin{pmatrix} 2 & 2 & 0 \\ a & 0 & 1 \\ 0 & 2a & -2 \end{pmatrix}$$

est-elle diagonalisable ?

E5.4. Soient les suites $\mathbf{u} = (u_n)_{n\geq 0}$ et $\mathbf{v} = (v_n)_{n\geq 0}$ telles que

$$\begin{cases} u_{n+1} = 6u_n - v_n \\ v_{n+1} = u_n + 4v_n \end{cases} \quad n \geq 0.$$

Déterminer une matrice A telle que

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = A \begin{pmatrix} u_n \\ v_n \end{pmatrix}$$

Soit J la matrice telle que $A = 5I+J$. Calculer J^n pour $n \in \mathbb{N}$; en déduire une expression pour A^n, puis les expressions des termes u_n et v_n.

E**5**.5. Soit f l'application linéaire de $\mathbb{R}^2$ définie par

$$f(x,y) = (5x-6y, 4x-5y), \quad (x,y) \in \mathbb{R}^2$$

1. Montrer que 1 et -1 sont valeurs propres de f.
2. Déterminer une base $\mathbf{b}$ de $\mathbb{R}^2$ constituée de vecteurs propres de f et donner la matrice N de f relativement à cette base.
3. Déterminer une relation matricielle entre la matrice N et la matrice M représentant f dans la base canonique de $\mathbb{R}^2$.

E**5**.6. Soit A une matrice carrée d'ordre n avec n valeurs propres réelles distinctes ordonnées de manière décroissante : $\lambda_1 > \lambda_2 > \ldots > \lambda_{n-1} > \lambda_n$, avec $V_1, \ldots, V_n$ des vecteurs propres correspondants. Montrer que, étant donné un vecteur X, on a

$$\lim_{k\to\infty} \frac{A^k X}{\lambda_1^k} = X_\infty$$

où X_∞ est soit nul, soit un vecteur propre de A de valeur propre λ_1.
Indication On écrira $X = X_1 + \ldots + X_n$ avec X_i vecteur nul ou vecteur propre de A pour la valeur propre λ_i.

E**5**.7. Trouver la solution du système d'équations :

$$\begin{aligned} x_1'(t) &= x_1(t) - x_2(t) + 2x_3(t) \\ x_2'(t) &= -x_1(t) + x_3(t) \\ x_3'(t) &= 2x_1(t) + x_2(t) + x_3(t) \end{aligned}$$

avec les conditions initiales : $x_1(0) = 0, x_2(0) = 3, x_3(0) = 0$.

CHAPITRE 6

ESPACES EUCLIDIENS

La plupart des espaces vectoriels utiles tant à l'ingénieur qu'au mathématicien sont munis de normes : on peut y mesurer la longueur d'un vecteur, ce qui permet de quantifier certaines de ses caractéristiques, estimer des erreurs dans des algorithmes numériques. Certaines de ces normes sont construites en généralisant la notion du produit scalaire bien connu dans $\mathbb{R}^3$: la géométrie de ces espaces vectoriels ainsi normés est similaire à celle de la géométrie élémentaire (angles compris).

6.1. Formes quadratiques

Une forme bilinéaire a sur un espace vectoriel E est une application $a : E^2 \to \mathbb{R}$ telle que les applications

$$a_V : W \in E \to a(V,W) \in \mathbb{R}, \quad a_W : V \in E \to a(V,W) \in \mathbb{R}$$

sont linéaires. La forme bilinéaire a est dite symétrique si

$$a(V,W) = a(W,V), \quad V, W \in E.$$

Une forme quadratique q sur l'espace vectoriel E est une application sur E à valeurs réelles telle qu'il existe une forme bilinéaire symétrique a avec

$$q(V) = a(V,V), \quad V \in E.$$

▷ ***Exemples 6.1***.

1. La forme bilinéaire symétrique sur $\mathbb{R}^2$ définie par
$$a_h((x,y),(x',y')) = xy' + yx', \quad (x,y),(x',y') \in \mathbb{R}^2$$
induit la forme quadratique q sur $\mathbb{R}^2$ définie par
$$q_h(V) = 2xy, \quad V = (x,y) \in \mathbb{R}^2.$$
2. La forme bilinéaire symétrique sur $\mathbb{R}^n$ définie par
$$a_n((x_1,\dots,x_n),(y_1,\dots,y_n)) = x_1y_1 + \dots + x_ny_n, \quad (x_1,\dots,x_n),(y_1,\dots,y_n) \in \mathbb{R}^n$$
induit la forme quadratique q_n sur $\mathbb{R}^n$ définie par
$$q_n((x_1,\dots,x_n)) = x_1^2 + \dots + x_n^2, \quad (x_1,\dots,x_n) \in \mathbb{R}^n.$$
3. L'application
$$(P,Q) \in \mathrm{Poly}_\infty \to \int_0^1 P(t)Q(t)dt \in \mathbb{R}, \quad P,Q \in \mathrm{Poly}_\infty$$
est une forme bilinéaire symétrique sur l'espace des polynômes réels Poly_∞.

4. L'application
$$a_h : (A, B) \in M_n(\mathbb{R}) \to \mathrm{tr}(AB) \in \mathbb{R}$$
est une forme bilinéaire symétrique sur l'espace $M_n(\mathbb{R})$ des matrices carrées d'ordre n. Ici, $\mathrm{tr}\, M$ désigne la trace matricielle définie précédemment définie comme la somme des coefficients diagonaux de M. ◁

Par rapport à une base $\mathbf{b} = (b_1, \ldots, b_n)$, la forme quadratique q déterminée par la forme bilinéaire symétrique a possède une expression simple en termes de calcul matriciel : si $V = \sum_{i=1}^n x_i b_i$ et $X = (x_i)^t$ le vecteur colonne des coordonnées de V relativement à la base , alors
$$q(V) = a(V, V) = a\left(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n x_j b_j\right) = \sum_{i,j=1}^n x_i x_j a(b_i, b_j) = X^t A_a X$$
où $A_a = (a(b_i, b_j))$ est une matrice symétrique réelle vu la symétrie de la forme a.

Si la matrice A_a est diagonale (cette propriété dépendant du choix de la *base* $\mathbf{b}$), la forme quadratique q apparaît comme une somme pondérée de carrés
$$q(V) = q\left(\sum_{i=1}^n x_i b_i\right) = \sum_{i=1}^n a_{ii} x_i^2.$$
Réduire une forme quadratique consiste à la transformer en une telle somme de k carrés de coordonnées linéaires. On distingue deux méthodes :

– **Méthode aux valeurs propres** La matrice associée A étant symétrique réelle, la prop. 5.1 énonce l'existence d'une matrice orthogonale P telle que $\widetilde{A} = PAP^t$ soit diagonale : vu que P est orthogonale ($P^t = P^{-1}$), les matrices A et $\widetilde{A}$ sont semblables. Ainsi
$$q(X) = X^t AX = X^t(P^t P)A(P^t P)X = (PX)^t \widetilde{A}(PX) = Y^t \widetilde{A} Y$$
où $Y = PX$ est la colonne des composantes du vecteur X dans la *base* de vecteurs propres de A donnée par les colonnes de la matrice de passage P^t.

▷ ***Exemple 6.2***. Considérons la forme quadratique sur $\mathbb{R}^3$
$$q(x, y, z) = 5x^2 + 4zx - 2yx - 3y^2 + 4zy + 2z^2.$$
Sa matrice associée A est
$$A = \begin{pmatrix} 5 & -1 & 2 \\ -1 & -3 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$
Via la matrice orthogonale
$$P = \begin{pmatrix} -\frac{1}{6\sqrt{6}} & \frac{2}{5\sqrt{5}} & -\frac{1}{30\sqrt{30}} \\ \frac{1}{6\sqrt{6}} & 0 & -\frac{1}{6\sqrt{30}} \\ \frac{1}{2\sqrt{6}} & \frac{1}{5\sqrt{5}} & -\frac{1}{15\sqrt{30}} \end{pmatrix},$$
la matrice A est semblable à la matrice diagonale $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -4 \end{pmatrix}$, ce qui donne pour q la forme réduite, avec $(\widetilde{x}, \widetilde{y}, \widetilde{z}) = \widetilde{X} = PX$,
$$\widetilde{q}(\widetilde{x}, \widetilde{y}, \widetilde{z}) = 2\widetilde{x}^2 + 6\widetilde{y}^2 - 4\widetilde{z}^2.$$ ◁

Cette méthode nécessite le calcul des vecteurs propres, calcul parfois délicat. La méthode suivante est souvent plus rapide.

– $\boxed{\textbf{\textit{Méthode de Gauss}}}$ Elle consiste à construire les carrés de manière progressive. Dans le cas précédent :

$$\begin{aligned}
q(x,y,z) &= 5x^2 + 4zx - 2yx - 3y^2 + 4zy + 2z^2 \\
&= 5\left(x^2 + \frac{4}{5}zx - \frac{2}{5}yx\right) - 3y^2 + 4zy + 2z^2 \\
&= 5\left(x + \frac{2}{5}z - \frac{1}{5}y\right)^2 - 5\left(\frac{2}{5}z - \frac{1}{5}y\right)^2 - 3y^2 + 4zy + 2z^2 \\
&= 5\left(x + \frac{2}{5}z - \frac{1}{5}y\right)^2 - \frac{16}{5}y^2 + \frac{24}{5}zy + \frac{6}{5}z^2 \\
&= 5\left(x + \frac{2}{5}z - \frac{1}{5}y\right)^2 - \frac{16}{5}\left(y^2 - \frac{3}{2}zy\right) + \frac{6}{5}z^2 \\
&= 5\left(x + \frac{2}{5}z - \frac{1}{5}y\right)^2 - \frac{16}{5}\left(y - \frac{3}{4}zy\right)^2 + \frac{42}{5}z^2 \\
&= (QX)^t D(QX),
\end{aligned}$$

où on a commencé par éliminer les termes croisés xy et xz en écrivant les premiers carrés $(x+\ldots)^2$, puis le terme croisé en yz avec un carré $(y+\ldots)^2$ et la forme finale où on a introduit, outre le vecteur coordonnées $X = (x,y,z)^t$, les matrices Q (non orthogonale !) et diagonale D

$$Q = \begin{pmatrix} 1 & \frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & \frac{6}{5} \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -\frac{16}{5} & 0 \\ 0 & 0 & \frac{42}{5} \end{pmatrix}$$

qui vérifient [1]

$$D = Q^t A Q.$$

Cette méthode est rapide, mais conduit à des multiples formes réduites possibles (commencer par chasser les termes rectangles yx et yz par exemple) : en incorporant les coefficients diagonaux de D dans Q, on voit aisément qu'on peut toujours se ramener à une matrice diagonale D avec des 1 ou -1 sur la diagonale.

6.2. Espaces normés

La donnée d'un produit scalaire sur un espace E consiste en la donnée d'une forme bilinéaire symétrique dont la valeur sur deux vecteurs V et W sera notée $\langle V, W\rangle$, ayant les propriétés supplémentaires

– $\boxed{\textbf{\textit{positivité}}}$ $\langle V, V\rangle \geq 0$;
– $\boxed{\textbf{\textit{définition}}}$ $\langle V, V\rangle = 0$ si et seulement si $V = 0$.

▷ ***Exemples 6.3.***

1. Les formes bilinéaires symétriques des exemples 2 et 4 de la série d'exemples 6.1 sont des produits scalaires.

1. La matrice Q est inversible, sans être orthogonale, on n'a pas $Q^t = Q^{-1}$ et les vecteurs colonnes de Q ne peuvent s'interpréter comme des vecteurs propres de la matrice A.

2. L'application a_h de l'exemple 6.1.4 n'est pas un produit scalaire, puisque $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ est de trace $-2 < 0$. Il en est de même de la forme bilinéaire a_h de l'exemple 6.1.2 puisque $a_h((1,0),(1,0)) = 0$ ou bien encore $a_h((1,-1),(1,-1)) < 0$. ◁

La $\boxed{\textit{norme}}$ d'un vecteur V est le réel positif noté $\|V\|$ égal à la racine carrée du produit scalaire (toujours positif) du vecteur par lui-même :

$$\|V\| = \sqrt{(V,V)}.$$

Proposition 6.1. *Le produit scalaire, et sa norme associée, vérifie l'inégalité de Cauchy-Schwarz*

$$|\langle V, W\rangle| \leq \|V\|\,\|W\|$$

Démonstration. — Le trinôme du second degré relativement à la variable λ

$$\|V+\lambda W\|^2 = \langle V+\lambda W, V+\lambda W\rangle = \langle V,V\rangle + 2\lambda\langle V,W\rangle + \lambda^2\langle W,W\rangle,$$

toujours positif, n'a pas de racines réelles distinctes donc un discriminant

$$4\langle V,W\rangle^2 - 4\langle V,V\rangle\langle W,W\rangle$$

négatif ou nul. □

Proposition 6.2. *La norme satisfait les conditions suivantes*

- $\boxed{\textbf{positivité}}$ $\|V\| \geq 0$;
- $\boxed{\textbf{définition}}$ $\|V\| = 0$ *si et seulement si* $V = 0$;
- $\boxed{\textbf{homogénéité}}$ $\|kV\| = |k|\|V\|$ *pour tout* k *réel ;*
- $\boxed{\textbf{inégalité triangulaire}}$ $\|V+W\| \leq \|V\| + \|W\|$.

L'inégalité triangulaire résulte de l'inégalité de Cauchy-Schwarz

$$\begin{aligned}\|V+W\|^2 &= \|V\|^2 + 2\langle V,W\rangle + \|W\|^2 \\ &\leq \|V\|^2 + 2\|V\|\|W\| + \|W\|^2 = (\|V\| + \|W\|)^2.\end{aligned}$$

Le produit scalaire permet de définir un angle $\theta \in [0,\pi]$ entre deux vecteurs V et W

$$\cos\theta = \frac{\langle V,W\rangle}{\|V\|\,\|W\|},$$

puisque le membre de droite de l'égalité précédente est dans $(-1,1)$ d'après l'inégalité de Cauchy-Schwarz.

Définition 6.1. Les deux vecteurs V, W sont *orthogonaux* si leur produit scalaire $\langle V,W\rangle$ est nul. Le vecteur V est normé si $\langle V,V\rangle = 1$.

Une base $\mathbf{b} = (b_i)$ est dite *orthonormée* si elle est constituée de vecteurs normés et orthogonaux deux à deux : $\langle b_i, b_j\rangle = \delta_{ij}$.

Les bases orthonormées, qui existent [2], ont les propriétés remarquables suivantes :

2. La preuve de l'existence d'une base orthonormée dans un espace E de dimension n est de nature très géométrique : on prend un vecteur unitaire b_1, puis un plan P_1 contenant b_1 et dans ce plan un vecteur orthogonal v à b_1 : quitte à multiplier par un scalaire, on peut supposer v unitaire, ce sera notre b_2. Si $n = 2$, alors (b_1, b_2) est une base de E, sinon on considère un espace de dimension 3 contenant le plan P_1, etc

– Le produit scalaire de deux vecteurs s'écrit simplement

$$\langle V,W\rangle = \langle\sum_{i=1}^n v_i b_i, \sum_{j=1}^n w_j b_j\rangle = \sum_{i,j=1}^n v_i w_j \langle b_i, b_j\rangle = \sum_{i,j=1}^n v_i w_j \delta_{ij} = \sum_{i=1}^n v_i w_i = \mathbf{V}^t\mathbf{W}$$

si $\mathbf{V} = (v_i), \mathbf{W} = (w_i)$ sont les vecteurs coordonnées respectifs de V et W relativement à la base orthonormée $\mathbf{b}$. Ainsi le produit scalaire de deux vecteurs coïncide avec celui défini dans la section 2.3 pour leurs vecteurs colonnes de coordonnées relativement à toute base orthonormée.

– La norme associée est $\|V\| = \sqrt{\sum_{i=1}^n v_i^2} = \sqrt{\mathbf{V}^t\mathbf{V}}$.

– La matrice de passage P d'une base orthonormée à une autre base orthonormée est une matrice orthogonale : $P^{-1} = P^t$.

Terminons en remarquant que sur un espace vectoriel E, il existe des normes (applications vérifiant les propriétés de la proposition 6.2) qui ne sont pas dérivées d'un produit scalaire, ainsi sur $\mathbb{R}^2$ les normes $\| \quad \|_1$ et $\| \quad \|_\infty$ définie par

$$\|(x,y)\|_1 = |x| + |y|, \quad \|(x,y)\|_\infty = \max(|x|,|y|), \quad (x,y) \in \mathbb{R}^2.$$

6.3. Exercices

E6.1. Dans l'espace des polynômes de degré au plus 3, calculer une base orthonormée relative au produit scalaire :

$$\langle P,Q\rangle = \int_{-1}^1 P(x)Q(x)dx, \quad P,Q \in \mathrm{Poly}_3\,.$$

E6.2. Soit P le sous-espace vectoriel de $\mathbb{R}^3$ engendré par les vecteurs $v_1 = (1,-1,2)$ et $v_2 = (1,0,1)$. Déterminer

1. une équation de P,
2. une base orthonormée de P
3. une base orthonormée de l'orthogonal $P^\perp$ de P,
4. la projection orthogonale de $(1,1,1)$ sur P.

E6.3. Soient $x_1, x_2, \ldots, x_n$ des nombres réels. Montrer que

$$\left(\sum_{i=1}^n x_i\right)^2 \le n\sum_{i=1}^n x_i^2.$$

E**6**.4. Peut-on trouver des constantes C_1, C_2, C_3 strictement positives telles que

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leq C_1 \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2,$$
$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq C_2 \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2,$$
$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \geq C_3 \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2,$$

pour tous réels $a_1, b_1, \ldots, a_n, b_n$?

E**6**.5. Dans l'espace des polynômes de degré au plus 2, muni du produit scalaire

$$\langle P, Q \rangle = \int_{-1}^{1} P(x)Q(x)dx, \quad P, Q \in \mathrm{Poly}_2,$$

on considère les bases $\mathbf{b} = (b_0 = 1, b_1 = x, b_2 = x^2)$ et $\mathbf{c} = (c_0 = 1/2, c_1 = \sqrt{3/2}x, c_2 = \sqrt{8/45}(x^2 - 1/3))$. Calculer la matrice de passage de $\mathbf{b}$ à $\mathbf{c}$ et son inverse, puis donner dans les deux bases les composantes du polynôme $P(x) = 5x^2 + 3x - 1$. Calculer la matrice de l'application *dérivation* dans les deux bases.
Dans chaque question, on vérifiera les théorèmes de changement de base.

E**6**.6. Montrer que l'application a_2 définie par

$$a_2(A, B) = \mathrm{tr}(AB^t)$$

munit l'espace des matrices carrées d'ordre n d'un produit scalaire.

E**6**.7. Réduire à une somme de carrés la forme quadratique q définie sur $\mathbb{R}^3$ par

$$q(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2ca, \quad (a, b, c) \in \mathbb{R}^3.$$

E**6**.8. L'exercice a pour objectif de démontrer une relation importante pour l'évaluation des intégrales gaussiennes dans $\mathbb{R}^n$
Démontrer que dans $\mathbb{R}$, on a

$$I(a) = \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}ax^2} = \left[\frac{2\pi}{a}\right]^{\frac{1}{2}}.$$

Généraliser le résultat dans $\mathbb{R}^n$

$$I(a_1, \ldots, a_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2)} dx_1 dx_2 \ldots dx_n = \left[\frac{(2\pi)^n}{a_1 a_2 \ldots a_n}\right]^{\frac{1}{2}}.$$

En effectuant une transformation orthogonale sur la forme quadratique en exposant, montrer le résultat général

$$I(A) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^t A x} dx_1 dx_2 \ldots dx_n = \left[\frac{(2\pi)^n}{\mathrm{d\acute{e}t}\, A}\right]^{\frac{1}{2}}.$$

où $X = (x_1 x_2 \ldots x_n)^t$ et A est une matrice symétrique à valeurs propres strictement positives.

INDEX