Table des matières

CHAPITRE I

Systèmes linéaires et matrices

1. Systèmes d'équations algébriques linéaires

Soit un système de m équations en n inconnues $x_1, \ldots, x_n$:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{cases}$$

où les coefficients a_{ij} pour $i = 1, \ldots, m$ et $j = 1, \ldots, n$, et les termes indépendants b_i pour $i = 1, \ldots, m$ sont des nombres réels ou complexes, ou plus généralement des éléments d'un corps[1] (commutatif) arbitraire K. Résoudre ce système consiste à trouver dans K des valeurs des inconnues $x_1, \ldots, x_n$ pour lesquelles ces équations sont satisfaites. Dans cette première section, on se propose de mettre au point des techniques permettant de reconnaître si un tel système admet une solution et, dans l'affirmative, de trouver *toutes* les solutions (en fonction d'un certain nombre de paramètres).

1.1. Opérations élémentaires

Les techniques que l'on utilise pour résoudre les systèmes d'équations algébriques linéaires se fondent sur trois types d'opérations sur les équations d'un système :

I) Dans un système d'équations, on peut remplacer une des équations par la somme de celle-ci et d'un multiple d'une *autre* équation du système sans modifier l'ensemble des solutions. En effet, les systèmes

$$\begin{cases} P_1(x_1, \ldots, x_n) & = & b_1 \\ & \vdots & \\ P_i(x_1, \ldots, x_n) & = & b_i \\ & \vdots & \\ P_j(x_1, \ldots, x_n) & = & b_j \\ & \vdots & \\ P_m(x_1, \ldots, x_n) & = & b_m \end{cases}$$

et

$$\begin{cases} P_1(x_1, \ldots, x_n) & = & b_1 \\ & \vdots & \\ P_i(x_1, \ldots, x_n) + \lambda P_j(x_1, \ldots, x_n) & = & b_i + \lambda b_j \\ & \vdots & \\ P_j(x_1, \ldots, x_n) & = & b_j \\ & \vdots & \\ P_m(x_1, \ldots, x_n) & = & b_m \end{cases}$$

ont les mêmes solutions.

[1] Voir l'Appendice pour une définition de la notion de *corps*.

II) Dans un système d'équations, on peut échanger deux équations sans modifier l'ensemble des solutions. En effet, il est clair que les systèmes

$$\left\{\begin{array}{rcl} P_1(x_1,\dots,x_n) & = & b_1 \\ & \vdots & \\ P_i(x_1,\dots,x_n) & = & b_i \\ & \vdots & \\ P_j(x_1,\dots,x_n) & = & b_j \\ & \vdots & \\ P_m(x_1,\dots,x_n) & = & b_m \end{array}\right. \qquad \text{et} \qquad \left\{\begin{array}{rcl} P_1(x_1,\dots,x_n) & = & b_1 \\ & \vdots & \\ P_j(x_1,\dots,x_n) & = & b_j \\ & \vdots & \\ P_i(x_1,\dots,x_n) & = & b_i \\ & \vdots & \\ P_m(x_1,\dots,x_n) & = & b_m \end{array}\right.$$

admettent les mêmes solutions.

III) Dans un système d'équations, on peut multiplier les deux membres d'une équation par un élément non nul du corps de base K sans modifier l'ensemble des solutions. En effet, les systèmes

$$\left\{\begin{array}{rcl} P_1(x_1,\dots,x_n) & = & b_1 \\ & \vdots & \\ P_i(x_1,\dots,x_n) & = & b_i \\ & \vdots & \\ P_m(x_1,\dots,x_n) & = & b_m \end{array}\right. \qquad \text{et} \qquad \left\{\begin{array}{rcl} P_1(x_1,\dots,x_n) & = & b_1 \\ & \vdots & \\ \lambda P_i(x_1,\dots,x_n) & = & \lambda b_i \\ & \vdots & \\ P_m(x_1,\dots,x_n) & = & b_m \end{array}\right.$$

admettent les mêmes solutions si $\lambda \neq 0$.

Les opérations indiquées ci-dessus sont appelées *opérations élémentaires* (sur les équations d'un système). Dans le cas d'une opération élémentaire de type III, l'élément $\lambda \in K$ par lequel on multiplie une des équations est appelé *facteur* de l'opération élémentaire.

1.2. Notation matricielle

Pour faire apprécier la portée des opérations élémentaires dans la résolution d'un système d'équations algébriques linéaires, il est commode d'adopter la notation suivante : à un système arbitraire de m équations à n inconnues

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right.$$

on associe le tableau des coefficients

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

que l'on appelle *matrice (des coefficients)* du système ; c'est un tableau à m lignes et n colonnes, chaque ligne correspondant à une équation du système. On note cette matrice $(a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ ou simplement $(a_{ij})_{i,j}$ lorsque le nombre de lignes et le nombre de colonnes sont indiqués par le contexte. Par convention, le premier indice est toujours celui des lignes et le second celui des colonnes. L'entrée a_{ij} est donc celle qui est située à l'intersection de la i-ème ligne et de la j-ème colonne.

On considère aussi le tableau

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

appelée *matrice complète* du système, obtenue en adjoignant à la matrice des coefficients du système la colonne des termes indépendants.

Comme les équations du système correspondent aux lignes de la matrice complète, les opérations élémentaires décrites précédemment correspondent à des opérations sur les *lignes* de la matrice complète ; à savoir :

I) remplacer une ligne par la somme de celle-ci et d'un multiple d'une *autre* ligne de la matrice complète ;

II) échanger deux lignes de la matrice complète ;

III) multiplier une ligne par un élément non nul du corps de base K (appelé *facteur* de l'opération élémentaire).

1.3. Échelonnement

1.1. Théorème. *Par une suite d'opérations élémentaires sur les lignes, on peut transformer toute matrice en une matrice* $U = (u_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ *possédant la propriété suivante : le nombre d'entrées nulles au début de chaque ligne augmente à chaque ligne, l'augmentation étant stricte pour les lignes qui suivent une ligne non nulle*[2] *; c'est-à-dire que pour* $i = 1, \ldots, m-1$ *et* $k \in \mathbb{N}$ *fixés, si* $u_{ij} = 0$ *pour tout* $j \leqslant k$ *alors* $u_{i+1,j} = 0$ *pour tout* $j \leqslant k+1$.

Une matrice U possédant la propriété ci-dessus est appelée *matrice à lignes échelonnées*. Si une ligne d'une telle matrice est nulle, la condition entraîne que toutes les lignes suivantes sont nulles. Dans une ligne non nulle, la première entrée non nulle est appelée *pivot* de la ligne. Par exemple, la matrice suivante est à lignes échelonnées :

$$(a_{ij})_{\substack{1\leqslant i\leqslant 5\\ 1\leqslant j\leqslant 6}} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 6 & 7 & 8 \\ 0 & 0 & 0 & 0 & 9 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

les pivots (des lignes non nulles) sont $a_{12} = 1$, $a_{24} = 6$ et $a_{35} = 9$. En revanche, la matrice suivante n'est *pas* à lignes échelonnées :

$$\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Démonstration du théorème : Soit $A = (a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ une matrice arbitraire. La stratégie de la démonstration consiste à faire apparaître des entrées nulles successivement en dessous des premières entrées non nulles de chaque ligne.

Si la première colonne est nulle, elle le reste lorsque l'on effectue des opérations élémentaires sur les lignes de la matrice. On peut donc la supprimer (mentalement) et porter son attention sur la deuxième colonne. Bien entendu, si toutes les colonnes de la matrice sont nulles — c'est-à-dire si la matrice est nulle — alors la matrice est à lignes échelonnées.

Si la première colonne n'est pas nulle, on choisit[3] un élément non nul de cette colonne, qui deviendra le premier pivot de la matrice à lignes échelonnées. Supposons avoir choisi l'entrée de la k-ème ligne : $a_{k1} \neq 0$. Par des opérations de type I sur les lignes, on peut faire apparaître des 0 à la place de toutes les autres entrées de la première colonne : en effet, si l'on remplace la ℓ-ème ligne L_ℓ (pour $\ell \neq k$) par la somme de celle-ci et de la k-ème ligne L_k multipliée par $-a_{\ell 1}a_{k1}^{-1}$, on transforme la matrice A en la matrice $A' = (a'_{ij})_{i,j}$ telle que

$$a'_{\ell j} = a_{\ell j} + a_{kj}\,(-a_{\ell 1}a_{k1}^{-1}) \quad \text{pour } j = 1, \ldots, n$$

[2] car la ligne qui suit une ligne nulle ne peut évidemment contenir plus de n entrées !

[3] Dans les calculs numériques, il y a avantage à choisir l'entrée la plus grande en valeur absolue, afin de minimiser l'effet des erreurs d'arrondi.

et

$$a'_{ij} = a_{ij} \quad \text{pour } i \neq \ell \text{ et } j = 1, \dots, n.$$

En particulier, $a'_{\ell 1} = 0$:

$$A = \begin{pmatrix} \vdots & \vdots & & \vdots \\ a_{\ell 1} & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ a_{k1} & * & \cdots & * \\ \vdots & \vdots & & \vdots \end{pmatrix} \xrightarrow{L_\ell \rightsquigarrow L_\ell - L_k \left(\frac{a_{\ell 1}}{a_{k1}}\right)} \begin{pmatrix} \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ a_{k1} & * & \cdots & * \\ \vdots & \vdots & & \vdots \end{pmatrix} = A'.$$

Après avoir ainsi opéré sur toutes les lignes autres que la k-ème, on obtient une matrice dont la première colonne ne contient qu'un seul élément non nul, à savoir a_{k1} à la k-ème ligne :

$$\begin{pmatrix} 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ a_{k1} & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}.$$

Pour faire apparaître cette entrée à la première ligne (si elle n'y est pas déjà), il suffit d'échanger la première et la k-ème ligne.

Après ces opérations, on obtient une matrice A'' dont la première colonne est conforme à ce que l'on attend d'une matrice à lignes échelonnées, puisque toutes les entrées de la première colonne sont nulles à l'exception de la première :

$$A \longrightarrow A'' = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}.$$

Pour poursuivre l'échelonnement, on n'a plus besoin de faire intervenir la première ligne ; on peut donc la supprimer mentalement, ainsi que la première colonne puisque ce qu'il en reste après avoir retiré la première ligne est constitué de 0 :

$$A'' = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} = \left(\begin{array}{c|ccc} * & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right).$$

On est ainsi ramené à une matrice B plus petite, sur les lignes de laquelle on reprend les opérations élémentaires. Bien entendu, les opérations élémentaires sur les lignes de B correspondent à des opérations élémentaires sur les lignes de A'' puisque les lignes de B deviennent des lignes de A'' lorsqu'on les fait précéder d'un 0. En poursuivant ainsi de suite, supprimant à chaque fois une colonne et une ligne (ou seulement une colonne dans le cas où la première colonne à considérer est nulle) jusqu'à ce qu'il n'en reste plus, on obtient finalement une matrice à lignes échelonnées. □

La matrice à lignes échelonnées produite par le procédé indiqué dans la démonstration n'est nullement déterminée de manière unique ; elle dépend (notamment) du choix des pivots, comme le montrent les exemples suivants (où le pivot est encadré) :

$$\begin{pmatrix} \boxed{1} & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 0 \\ 0 & \boxed{-1} & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 0 \\ 0 & \boxed{-1} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & 1 & 0 \\ \boxed{2} & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix} \to \begin{pmatrix} 0 & 1/2 & -1/2 \\ \boxed{2} & 1 & 1 \\ 0 & -1/2 & 1/2 \\ 0 & 3/2 & -3/2 \end{pmatrix} \xrightarrow{L_1 \rightsquigarrow L_1 + L_2} \begin{pmatrix} \boxed{2} & 3/2 & 1/2 \\ 2 & 1 & 1 \\ 0 & -1/2 & 1/2 \\ 0 & 3/2 & -3/2 \end{pmatrix} \xrightarrow{L_2 \rightsquigarrow L_2 - L_1}$$

$$\to \begin{pmatrix} \boxed{2} & 3/2 & 1/2 \\ 0 & -1/2 & 1/2 \\ 0 & -1/2 & 1/2 \\ 0 & \boxed{3/2} & -3/2 \end{pmatrix} \to \begin{pmatrix} \boxed{2} & 3/2 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \boxed{3/2} & -3/2 \end{pmatrix} \to \begin{pmatrix} \boxed{2} & 3/2 & 1/2 \\ 0 & \boxed{3/2} & -3/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

En revanche, on montre au §6 que le nombre de lignes non nulles de la matrice à lignes échelonnées est un invariant de la matrice donnée (c'est son *rang*).

Après avoir échelonné les lignes d'une matrice, on peut encore poursuivre les opérations élémentaires pour annuler les entrées situées *au-dessus* des pivots. En utilisant les opérations de type III qui consistent à multiplier chaque ligne non nulle par l'inverse de son pivot, on parvient finalement à une matrice à lignes échelonnées dont tous les pivots sont égaux à 1 et toutes les entrées situées au-dessus[4] d'un pivot sont nulles :

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots \\ & & & & & & & \cdots & & & & & & \end{pmatrix}.$$

Une telle matrice est dite *sous forme réduite de Gauss–Jordan.*

On a ainsi prouvé un résultat plus précis :

1.2. Corollaire. *Par une suite d'opérations élémentaires sur les lignes, on peut transformer toute matrice en une matrice sous forme réduite de Gauss–Jordan.*

On peut prouver que la matrice sous forme réduite obtenue à partir d'une matrice A donnée est déterminée de manière unique par A. Comme ce résultat n'est pas indispensable pour la suite, sa démonstration est laissée à la sagacité du lecteur[5].

1.4. Solution de systèmes d'équations algébriques linéaires

Revenons au problème initial, qui consiste à résoudre un système de m équations en n inconnues

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m. \end{cases}$$

Par des opérations élémentaires sur les équations de ce système ou, ce qui revient au même, sur les lignes de la matrice complète, on obtient un système qui a les mêmes solutions et dont la matrice complète est sous forme réduite de Gauss–Jordan. Il est facile de reconnaître si un tel système admet des solutions[6] :

– les lignes nulles de la matrice complète correspondent à des équations

$$0x_1 + \cdots + 0x_n = 0;$$

on peut donc les négliger et ne s'intéresser qu'aux lignes non nulles. Soit r le nombre de ces lignes ;

[4] Celles qui sont situées *en dessous* d'un pivot sont nulles par définition d'une matrice à lignes échelonnées.

[5] Il est toutefois conseillé de chercher l'inspiration dans le §6.

[6] En pratique, il n'est d'ailleurs pas nécessaire de poursuivre les opérations élémentaires au-delà du moment où la matrice du système est à lignes échelonnées.

– si le dernier pivot (c'est-à-dire le r-ème) est dans la dernière colonne (c'est-à-dire la $n+1$-ème), qui est la colonne des termes indépendants, alors la dernière équation non triviale est

$$0x_1 + \cdots + 0x_n = 1;$$

le système n'admet donc pas de solution ;

– si le dernier pivot n'est pas dans la dernière colonne : soient $j_1, \ldots, j_r$ les indices des colonnes de pivots, de sorte que $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$ et que la matrice complète soit du type

$$\begin{array}{c}
\begin{array}{ccccccccccccccccccccc}
 & & & j_1 & & & & & j_2 & & & & & j_3 & & & \cdots & & j_r & & \\
 & & & \downarrow & & & & & \downarrow & & & & & \downarrow & & & & & \downarrow & &
\end{array} \\
\left(\begin{array}{ccccccccccccccccccccc|c}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\
 & & & & & & & & & & & & & & \cdots & & & & & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 & * & \cdots & *
\end{array}\right)
\end{array}.$$

En faisant passer dans les membres de droite des équations correspondantes les variables d'indice $\neq j_1, \ldots, j_r$, on met le système sous la forme

$$\left\{\begin{array}{rcl}
x_{j_1} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{1j} x_j + b'_1 \\
x_{j_2} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{2j} x_j + b'_2 \\
 & \vdots & \\
x_{j_r} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{rj} x_j + b'_r.
\end{array}\right.$$

C'est un système qui donne les valeurs des variables $x_{j_1}, \ldots, x_{j_r}$ en fonction des autres. Les solutions s'obtiennent en donnant aux variables d'indice $\neq j_1, \ldots, j_r$ des valeurs arbitraires et aux variables $x_{j_1}, \ldots, x_{j_r}$ les valeurs données par le système ci-dessus.

En particulier, si $r = n$, la matrice complète sous forme réduite de Gauss–Jordan est de la forme

$$\left(\begin{array}{cccc|c}
1 & 0 & \cdots & 0 & b'_1 \\
0 & 1 & \cdots & 0 & b'_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & b'_r
\end{array}\right)$$

et la solution est unique :

$$\left\{\begin{array}{rcl}
x_1 & = & b'_1 \\
 & \vdots & \\
x_r & = & b'_r.
\end{array}\right.$$

Systèmes homogènes. Un système d'équations algébriques linéaires est dit *homogène* si les seconds membres $b_1, \ldots, b_m$ sont tous nuls. Un tel système est donc de la forme

$$\left\{\begin{array}{rcl}
a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\
 & \vdots & \\
a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0.
\end{array}\right.$$

Il admet toujours la *solution triviale* $x_1 = \cdots = x_n = 0$. Pour déterminer s'il en existe d'autres, on procède comme pour les autres systèmes d'équations algébriques linéaires. La dernière colonne de la matrice complète est nulle, et elle reste nulle lorsque l'on effectue des opérations élémentaires sur les lignes. Le système réduit est donc de la forme

$$\left\{\begin{array}{rcl}
x_{j_1} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{1j} x_j \\
x_{j_2} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{2j} x_j \\
 & \vdots & \\
x_{j_r} & = & -\sum_{j \neq j_1, \ldots, j_r} a'_{rj} x_j
\end{array}\right.$$

où r est le nombre de pivots (ou, ce qui revient au même, le nombre de lignes non nulles de la matrice sous forme réduite de Gauss–Jordan). Si $r = n$, la solution est unique (et est triviale). Si $r < n$, il y a des solutions non triviales, qui s'obtiennent en donnant aux variables d'indice $\neq j_1, \ldots, j_r$ des valeurs arbitraires (non toutes nulles) et aux variables $x_{j_1}, \ldots, x_{j_r}$ les valeurs données par le système ci-dessus.

La suite du cours a pour objectif de mettre au point et de développer le cadre théorique dans lequel s'insèrent naturellement les résultats précédents. En particulier, on verra que le nombre r d'équations non nulles du système réduit ne dépend pas de la manière dont la réduction a été effectuée et qu'il admet une interprétation géométrique susceptible de s'appliquer à des situations très diverses.

Exercices

On trouvera un choix d'exercices à la fin de chaque section. Certains de ces exercices sont notés "prérequis". Cela signifie qu'ils sont d'un niveau très élémentaire ; ne pas être capable de les résoudre doit être considéré comme une "sonnette d'alarme". À l'opposé, d'autres exercices sont notés "pour aller plus loin". Ceux-là sont plus intéressants car moins routiniers. Mais ils peuvent poser difficulté car ils mettent en jeu des objets plus abstraits, et/ou que leur résolution est plus astucieuse. Il ne faut pas s'inquiéter si l'on est incapable de les résoudre.

(1.1) Trouvez les solutions réelles du système d'équations suivant :

$$\begin{array}{rcl} x_1 - 2x_2 - 3x_3 + 4x_4 & = & 1 \\ x_1 + 2x_2 + 2x_3 & = & 6 \\ x_1 + x_2 & = & 3 \\ x_1 & = & 2. \end{array}$$

(1.2) Trouvez les solutions complexes du système d'équations suivant :

$$\begin{array}{rcl} x + iy + (1+i)z - t & = & 1 \\ x + y & = & 0 \\ (1+i)x + (i-1)z - it & = & 1. \end{array}$$

(1.3) Trouvez les solutions réelles du système d'équations suivant :

$$\begin{array}{rcl} x - y + 2z & = & 1 \\ 2x + y - z & = & 0 \\ 3x + z & = & 1 \\ 5x + y & = & 1. \end{array}$$

(1.4) Trouvez les solutions réelles du système d'équations suivant :

$$\begin{array}{rcl} x + 2y - 3z & = & 0 \\ -x + y + z & = & 0 \\ 4x - 2y + z & = & 0. \end{array}$$

(1.5) Trouvez les solutions réelles du système d'équations suivant :

$$\begin{array}{rcl} 3x_1 + 2x_2 + x_3 - x_4 - x_5 & = & 0 \\ x_1 - x_2 - x_3 - x_4 + 2x_5 & = & 0 \\ -x_1 + 2x_2 + 3x_3 + x_4 - x_5 & = & 0. \end{array}$$

(1.6) Quel est l'ensemble des polynômes P tels que $(X^2 + X + 1) \cdot P = 2X^4 + 3X^3 + 6X^2 + 4X + 3$?

(1.7) Notons N le nombre de solutions d'un système de n équations linéaires à m inconnues. On sait que $N \in \{0, 1, \infty\}$. Quelles sont les valeurs possibles de N

– si $n < m$?

– si $n = m$?

– si $n > m$?

Dans les trois cas, donner un exemple de système d'équations ayant N solutions pour chaque valeur possible de N, et si certaines valeurs de N sont impossibles, justifiez-le.

(1.8) Discutez, en fonction de la valeur du réel k, le nombre de solutions réelles du système d'équations suivant :

$$\begin{aligned} x+2y-z &= 4 \\ 2x+5y+z &= k \\ x+3y+(k^2+1)z &= 5k. \end{aligned}$$

Même question pour le nombre de solutions complexes, k pouvant prendre des valeurs complexes.

(1.9) Gérard Mambont s'est entraîné à effectuer plusieurs opérations élémentaires simultanément pour résoudre plus vite les systèmes d'équations. Il transforme le système

$$I \left\{ \begin{aligned} x_1+2x_2-x_3+x_4 &= 1 \\ x_1-x_2+x_3+x_4 &= 0 \end{aligned} \right.$$

en remplaçant la première ligne L_1 par L_1-L_2 et la ligne L_2 par L_2-L_1. Il obtient le système

$$II \left\{ \begin{aligned} 3x_2-2x_3 &= 1 \\ -3x_2+2x_3 &= -1, \end{aligned} \right.$$

dont les solutions sont tous les quadruplets $(x_1, \frac{1}{3}(1+2x_3), x_3, x_4)$ où x_1, x_3, x_4 sont arbitraires. Sa copine Jenny Croix-Guerre lui fait remarquer que $(1,1,1,0)$ est solution du système II mais non de I. Où est l'erreur ?

(1.10) (pour aller plus loin) Soit un système d'équations linéaires à coefficients réels. Prouvez que si ce système n'a pas de solutions réelles, alors il n'a pas non plus de solutions complexes.

2. Matrices

Les opérations élémentaires dont il a été question dans la première section admettent une interprétation très simple en termes de multiplication matricielle. Pour la donner, on commence par rappeler les définitions fondamentales.

Soit K un corps commutatif arbitraire. On note $K^{m\times n}$ l'ensemble des matrices $m\times n$ (ou de *genre* (m,n)) sur K, c'est-à-dire l'ensemble des tableaux rectangulaires à m lignes et n colonnes dont les entrées sont des éléments de K. L'égalité de deux matrices $A=(a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ et $B=(b_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ dans $K^{m\times n}$ revient à l'égalité de chacune de leurs entrées :

$$A=B \quad \text{si et seulement si} \quad a_{ij}=b_{ij} \text{ pour tout } i=1,\ldots,m \text{ et tout } j=1,\ldots,n.$$

Les matrices qui n'ont qu'une seule ligne (resp. colonne) sont appelées *matrices-lignes* (resp. *matrices-colonnes*). Les matrices dont le nombre de lignes est égal au nombre de colonnes sont appelées *matrices carrées*. L'*ordre* d'une matrice carrée est le nombre de ses lignes ou de ses colonnes. La *diagonale principale* d'une matrice carrée est formée des entrées dont les indices sont égaux. Une matrice *diagonale* est une matrice carrée dont toutes les entrées hors de la diagonale principale sont nulles.

Matrices particulières :

- $0_{m,n}$ est la *matrice nulle* de genre (m,n), dont toutes les entrées sont nulles.
- I_n est la *matrice unité* d'ordre n, définie par

  $$I_n=(\delta_{ij})_{1\leqslant i,j\leqslant n}$$

  où

  $$\delta_{ij}=\left\{ \begin{array}{ll} 0 & \text{si } i\neq j \\ 1 & \text{si } i=j. \end{array} \right.$$

  La fonction δ_{ij} ainsi définie est appelée *symbole de Kronecker*.

2.1. Opérations matricielles

– La *somme* de deux matrices $m \times n$ est définie par

$$(a_{ij})_{i,j} + (b_{ij})_{i,j} = (a_{ij} + b_{ij})_{i,j}.$$

– Le *produit* d'une matrice $m \times n$ par un élément de K est défini par

$$\alpha\,(a_{ij})_{i,j} = (\alpha a_{ij})_{i,j}.$$

– Le *produit* d'une matrice $m \times n$ et d'une matrice $n \times p$ est une matrice $m \times p$ définie par

$$(a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} \cdot (b_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} = (c_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant p}}$$

où

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \text{pour } i = 1, \ldots, m \text{ et } j = 1, \ldots, p.$$

Les propriétés suivantes s'établissent par des vérifications directes :

1. Propriétés de l'addition : pour $A, B, C \in K^{m\times n}$,

$$(A + B) + C = A + (B + C)$$

$$A + B = B + A$$

$$A + 0_{m,n} = A = 0_{m,n} + A$$

$$A + (-1)A = 0_{m,n} = (-1)A + A$$

2. Propriétés de la multiplication par un élément de K : pour $A, B \in K^{m\times n}$ et $\alpha, \beta \in K$,

$$\alpha\,(A + B) = \alpha\,A + \alpha\,B$$

$$(\alpha + \beta)\,A = \alpha\,A + \beta\,A$$

$$(\alpha\beta)\,A = \alpha\,(\beta\,A)$$

$$1\,A = A$$

3. Propriétés de la multiplication matricielle : pour $A \in K^{m\times n}$, $B \in K^{n\times p}$ et $C \in K^{p\times q}$,

$$(AB)C = A(BC)$$

$$A\,I_n = A = I_m\,A$$

4. Propriétés mixtes : pour $A, A' \in K^{m\times n}$, $B, B' \in K^{n\times p}$ et $\alpha \in K$,

$$(A + A')\,B = AB + A'B$$

$$A\,(B + B') = AB + AB'$$

$$(\alpha\,A)\,B = \alpha\,(A\,B) = A\,(\alpha\,B).$$

En revanche, le produit des matrices n'est pas commutatif : en général, $AB \neq BA$ (d'ailleurs, si $A \in K^{m\times n}$ et $B \in K^{n\times p}$, le produit BA n'est défini que si $p = m$). Cependant, d'après la dernière des propriétés énoncées ci-dessus, les matrices de la forme $\alpha\,I_n$ pour $\alpha \in K$, que l'on appelle *matrices scalaires*, commutent avec toutes les matrices carrées d'ordre n : pour $A \in K^{n\times n}$ et $\alpha \in K$,

$$(\alpha\,I_n)A = \alpha\,A = A\,(\alpha\,I_n).$$

Il faut remarquer de plus que les matrices n'admettent pas toutes un inverse pour la multiplication (voir la section 2.4 ci-dessous) et qu'un produit matriciel peut être nul sans qu'aucun des facteurs ne soit nul ; par exemple,

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

La multiplication matricielle permet de noter sous une forme particulièrement compacte les systèmes d'équations algébriques linéaires. En effet, le système de m équations à n inconnues

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right.$$

s'écrit simplement

$$A \cdot X = b$$

où

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

est la matrice des coefficients du système,

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

est la matrice-colonne des inconnues, et

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

est la matrice-colonne des termes indépendants.

2.2. Opérations par blocs

Si les entiers m et n sont décomposés en sommes d'entiers strictement positifs

$$m = m_1 + \cdots + m_r \qquad n = n_1 + \cdots + n_s$$

toute matrice de genre (m, n) peut être considérée comme une matrice $r \times s$ dont l'entrée d'indices i, j est une matrice de genre (m_i, n_j) que l'on appelle *bloc*. Ainsi, par exemple, pour $m = 4$ et $n = 6$, aux décompositions

$$4 = 1 + 2 + 1 \qquad 6 = 2 + 4$$

correspond la décomposition suivante d'une matrice arbitraire A de genre $(4, 6)$:

$$A = \left(\begin{array}{cc|cccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}$$

où

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \qquad A_{12} = \begin{pmatrix} a_{13} & a_{14} & a_{15} & a_{16} \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \qquad A_{22} = \begin{pmatrix} a_{23} & a_{24} & a_{25} & a_{26} \\ a_{33} & a_{34} & a_{35} & a_{36} \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} a_{41} & a_{42} \end{pmatrix} \qquad A_{32} = \begin{pmatrix} a_{43} & a_{44} & a_{45} & a_{46} \end{pmatrix}.$$

Les décompositions les plus fréquemment utilisées sont la *décomposition par lignes :*

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \vdots & \vdots & & \vdots \\ \hline a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right) = \begin{pmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{m*} \end{pmatrix}$$

où $a_{i*} = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ désigne la i-ème ligne de A, et la *décomposition par colonnes :*

$$A = \left(\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right) = \begin{pmatrix} a_{*1} & a_{*2} & \cdots & a_{*n} \end{pmatrix}$$

où $a_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ désigne la j-ème colonne de A.

2.1. Proposition. *Le résultat d'opérations effectuées sur des matrices décomposées en blocs peut s'obtenir en appliquant les règles habituelles du calcul matriciel à ces blocs comme s'ils étaient les éléments des matrices (pour autant, évidemment, que les opérations requises sur les blocs aient un sens).*

Démonstration. La proposition étant évidente pour les multiplications par des éléments de K et pour les additions matricielles, il suffit de la prouver pour les multiplications matricielles. Soient

$$A = (a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} \qquad \text{et} \qquad B = (b_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$$

deux matrices, de genres (m, n) et (n, p) respectivement, de sorte que le produit AB soit défini. Considérons des décompositions

$$m = m_1 + \cdots + m_r \qquad n = n_1 + \cdots + n_s \qquad p = p_1 + \cdots + p_t$$

et les décompositions correspondantes de A et B en blocs :

$$A = (A_{\alpha\beta})_{\substack{1\leqslant\alpha\leqslant r\\ 1\leqslant\beta\leqslant s}} \qquad B = (B_{\alpha\beta})_{\substack{1\leqslant\alpha\leqslant s\\ 1\leqslant\beta\leqslant t}}.$$

Il faut prouver :

$$AB = \left(\sum_{\gamma=1}^{s} A_{\alpha\gamma}B_{\gamma\beta}\right)_{\substack{1\leqslant\alpha\leqslant r\\ 1\leqslant\beta\leqslant t}}.$$

Or, l'élément d'indices i, j de AB est $\sum_{k=1}^{n} a_{ik}b_{kj}$, que l'on peut décomposer en somme de s termes :

$$\sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n_1} a_{ik}b_{kj} + \sum_{k=1+n_1}^{n_1+n_2} a_{ik}b_{kj} + \cdots + \sum_{k=1+n_1+\cdots+n_{s-1}}^{n} a_{ik}b_{kj}$$

et l'on reconnaît dans le membre de droite l'élément i, j de la matrice décomposée en blocs :

$$\left(\sum_{\gamma=1}^{s} A_{\alpha\gamma}B_{\gamma\beta}\right)_{\substack{1\leqslant\alpha\leqslant r\\ 1\leqslant\beta\leqslant t}}.$$

□

En particulier, dans un produit de deux matrices, on peut décomposer le facteur de droite en colonnes ou le facteur de gauche en lignes pour faire apparaître les colonnes ou les lignes du produit comme des produits matriciels : pour $A \in K^{m\times n}$ et $B \in K^{n\times p}$, on a

$$A \cdot B = A \cdot \begin{pmatrix} b_{*1} & \cdots & b_{*p} \end{pmatrix} = \begin{pmatrix} A \cdot b_{*1} & \cdots & A \cdot b_{*p} \end{pmatrix}$$

et aussi

$$A \cdot B = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{m*} \end{pmatrix} \cdot B = \begin{pmatrix} a_{1*} \cdot B \\ \vdots \\ a_{m*} \cdot B \end{pmatrix}$$

et encore
$$A \cdot B = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{m*} \end{pmatrix} \cdot \begin{pmatrix} b_{*1} & \cdots & b_{*p} \end{pmatrix} = \begin{pmatrix} a_{1*} \cdot b_{*1} & \cdots & a_{1*} \cdot b_{*p} \\ \vdots & & \vdots \\ a_{m*} \cdot b_{*1} & \cdots & a_{m*} \cdot b_{*p} \end{pmatrix}.$$

2.3. Transposition

Pour $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \in K^{m \times n}$, on pose
$$A^t = (a'_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}} \in K^{n \times m},$$
où
$$a'_{ij} = a_{ji} \qquad \text{pour } i = 1, \ldots, n \text{ et } j = 1, \ldots, m.$$
La matrice A^t est appelée *transposée* de la matrice A ; elle est obtenue en écrivant en colonnes les lignes de A (ou vice-versa).

2.2. PROPOSITION. 1. *Pour $A, B \in K^{m \times n}$ et $\alpha, \beta \in K$,*
$$(\alpha\, A + \beta\, B)^t = \alpha\, A^t + \beta\, B^t.$$

2. *Pour $A \in K^{m \times n}$ et $B \in K^{n \times p}$,*
$$(A \cdot B)^t = B^t \cdot A^t.$$

3. *Pour $A \in K^{m \times n}$,*
$$(A^t)^t = A.$$

Les démonstrations de ces propriétés sont de simples vérifications laissées au lecteur.

2.4. Inversion

Une matrice $A \in K^{m \times n}$ est dite *inversible à gauche* (resp. *à droite*) s'il existe une matrice $B \in K^{n \times m}$ telle que $B \cdot A = I_n$ (resp. $A \cdot B = I_m$). Toute matrice B satisfaisant cette condition est appelée *inverse à gauche* (resp. *inverse à droite*) de A.

Par exemple, toute matrice-ligne non nulle est inversible à droite. En effet, si $a_1, \ldots, a_n \in K$ ne sont pas tous nuls, on peut trouver (en général de plusieurs manières) $b_1, \ldots, b_n \in K$ tels que $a_1 b_1 + \cdots + a_n b_n = 1$, ce qui revient à dire
$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = (1) = I_1.$$
De même, toute matrice-colonne non nulle est inversible à gauche ; cet exemple montre que l'inverse à gauche ou à droite d'une matrice n'est généralement pas unique.

2.3. PROPOSITION. *Si une matrice est inversible à gauche et à droite, alors elle admet un unique inverse à gauche, qui est aussi l'unique inverse à droite.*

DÉMONSTRATION. Soit $A \in K^{m \times n}$ une matrice inversible à la fois à gauche et à droite. Supposons que B soit un inverse à gauche et que C soit un inverse à droite de A, de sorte que
$$BA = I_n \qquad \text{et} \qquad AC = I_m.$$
Il s'agit de prouver : $B = C$. On considère pour cela le produit BAC. D'après les facteurs que l'on multiplie en premier, on a
$$BAC = B(AC) = BI_m = B$$
ou
$$BAC = (BA)C = I_nC = C;$$
donc $B = C$. Cela prouve que tout inverse à gauche de A est égal à C et que tout inverse à droite de A est égal à B ; il y a donc un seul inverse à gauche, qui est aussi l'unique inverse à droite. □

Si une matrice A satisfait la condition de la proposition, on dit qu'elle est *inversible* et on désigne par A^{-1} son unique inverse (à gauche et à droite). On verra plus loin (Corollaire 6.6, page 39) que seules les matrices *carrées* peuvent être inversibles, ce qui n'est pas évident *a priori*. Par ailleurs, on montre dans la section suivante (voir le Théorème 2.8) que *pour une matrice carrée*, l'inversibilité à gauche est équivalente à l'inversibilité à droite (et donc aussi à l'inversibilité), et on donne un procédé explicite pour calculer l'inverse au moyen d'opérations élémentaires sur les lignes.

2.4. PROPOSITION. 1. *L'inverse de toute matrice inversible est inversible : si $A \in K^{m\times n}$ est inversible, alors $A^{-1} \in K^{n\times m}$ est inversible et*

$$(A^{-1})^{-1} = A.$$

2. *Le produit de deux matrices inversibles est inversible : si $A \in K^{m\times n}$ et $B \in K^{n\times p}$ sont inversibles, alors le produit $A \cdot B \in K^{m\times p}$ est inversible et*

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

3. *La transposée de toute matrice inversible est inversible : si $A \in K^{m\times n}$ est inversible, alors $A^t \in K^{n\times m}$ est inversible et*

$$(A^t)^{-1} = (A^{-1})^t.$$

DÉMONSTRATION. 1. Dire que A^{-1} est l'inverse de A revient à dire

$$(*) \qquad A \cdot A^{-1} = I_m \qquad \text{et} \qquad A^{-1} \cdot A = I_n.$$

Ces mêmes relations montrent que A est l'inverse de A^{-1}.
2. On a

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot I_n \cdot A^{-1} = A \cdot A^{-1} = I_m$$

et

$$(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = B^{-1} \cdot (A^{-1} \cdot A) \cdot B = B^{-1} \cdot I_n \cdot B = B^{-1} \cdot B = I_p,$$

donc $A \cdot B$ est bien inversible, d'inverse $B^{-1} \cdot A^{-1}$.
3. Cette relation s'obtient en transposant les relations $(*)$. □

2.5. Matrices élémentaires

Une *matrice élémentaire de type I* est une matrice qui ne diffère d'une matrice unité que par une seule entrée située hors de la diagonale principale. Une telle matrice est donc carrée, et de la forme

$$E = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & \lambda & & 1 & \\ & & & & 1 \end{pmatrix}$$

où l'entrée λ est non nulle et d'indices i, j avec $i \neq j$, les autres entrées hors de la diagonale principale étant nulles.

De manière équivalente, une matrice élémentaire de type I est une matrice obtenue en effectuant une opération élémentaire de type I sur les lignes ou sur les colonnes de la matrice unité. L'entrée λ étant située à l'intersection de la i-ème ligne et de la j-ème colonne, on peut en effet obtenir la matrice E ci-dessus en remplaçant la i-ème ligne de la matrice unité par la somme de cette ligne et de la j-ème ligne multipliée par λ ou, si l'on préfère, en remplaçant la j-ème colonne de la matrice unité par la somme de celle-ci et de la i-ème colonne multipliée par λ.

De la même manière, en effectuant des opérations élémentaires de type II ou III sur les lignes ou les colonnes de la matrice unité, on peut définir des matrices élémentaires de type II ou III, qui sont de la forme

$$P = \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 &&&& 1 &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& 1 &&&& 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix}$$

ou

$$M = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}$$

avec $\lambda \neq 0$. (L'élément $\lambda \in K$ est appelé *facteur* de la matrice élémentaire de type III.)

D'après la forme des matrices ainsi définies, il est clair que la transposée de toute matrice élémentaire est une matrice élémentaire du même type (et de même facteur, en ce qui concerne le type III).

2.5. PROPOSITION. *Toute matrice élémentaire est inversible, et son inverse est une matrice élémentaire du même type (et de facteur inverse, en ce qui concerne le type III).*

DÉMONSTRATION. Considérons d'abord la matrice E définie ci-dessus. Un calcul direct montre que la matrice

$$E' = \begin{pmatrix} 1 &&&& \\ & 1 &&& \\ && \ddots && \\ & -\lambda && 1 & \\ &&&& 1 \end{pmatrix},$$

où l'entrée $-\lambda$ occupe la même place que l'entrée λ dans E, est un inverse à gauche et à droite de E. On a donc $E^{-1} = E'$. Pour les matrices P et M, on vérifie de même : $P^{-1} = P$ et

$$M^{-1} = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda^{-1} &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}.$$

□

2.6. PROPOSITION. *Effectuer une opération élémentaire sur les lignes d'une matrice revient à multiplier cette matrice à gauche par une matrice élémentaire. Plus précisément, soit $A \in K^{m \times n}$ une matrice arbitraire et $A' \in K^{m \times n}$ la matrice obtenue en effectuant une certaine opération élémentaire sur les lignes de A. Soit encore $E \in K^{m \times m}$ la matrice obtenue en effectuant la* même *opération élémentaire sur les lignes de la matrice unité I_m ; alors*

$$A' = E \cdot A.$$

Démonstration. Soit

$$E = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & \lambda & & 1 & \\ & & & & 1 \end{pmatrix},$$

l'entrée λ étant située à l'intersection de la i-ème ligne et de la j-ème colonne. On calcule le produit $E \cdot A$ en décomposant A par lignes :

$$E \cdot A = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \lambda & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{m*} \end{pmatrix} = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{i*} + \lambda a_{j*} \\ \vdots \\ a_{m*} \end{pmatrix}.$$

La vérification est pareille pour les opérations élémentaires de type II ou III. □

Remarque : Le fait que la matrice E' définie dans la démonstration de la Proposition 2.5 est l'inverse de E résulte facilement aussi de la proposition précédente.

2.7. Corollaire. *Soit $A \in K^{m\times n}$ une matrice arbitraire et soit $U \in K^{m\times n}$ une matrice sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de A (voir le Corollaire 1.2, p. 5). Il existe une suite de matrices élémentaires $E_1, \dots, E_r$ telles que*

$$A = E_1 \cdot \cdots \cdot E_r \cdot U.$$

Démonstration. Vu la proposition précédente, chacune des opérations élémentaires revient à la multiplication à gauche par une matrice élémentaire ; on peut donc trouver des matrices élémentaires $F_1, \dots, F_r$ telles que

$$(\dagger) \qquad F_r \cdot \cdots \cdot F_1 \cdot A = U.$$

D'après la Proposition 2.5, chaque matrice F_i est inversible, et son inverse est aussi élémentaire. Notons $E_i = F_i^{-1}$. En multipliant les deux membres de l'équation (†) ci-dessus successivement par les matrices $F_r^{-1} = E_r, \dots, F_1^{-1} = E_1$, on obtient

$$A = F_1^{-1} \cdot \cdots \cdot F_r^{-1} \cdot U = E_1 \cdot \cdots \cdot E_r \cdot U.$$

□

Jusqu'à la fin de cette section, on considère le cas particulier des *matrices carrées*. Soit $A \in K^{n\times n}$ et soit $U \in K^{n\times n}$ une matrice sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de A. Comme U est carrée, on a l'alternative :

soit le nombre de pivots de U est n, et alors $U = I_n$;
soit le nombre de pivots de U est $< n$, et alors la dernière ligne de U est nulle.

2.8. Théorème. *Soit $A \in K^{n\times n}$ une matrice carrée arbitraire. Les conditions suivantes sont équivalentes :*

(a) *A est inversible ;*
(b) *A est inversible à gauche ;*
(c) *A est inversible à droite ;*
(d) *A est un produit de matrices élémentaires ;*
(e) *toute matrice U sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de A est la matrice unité I_n ;*
(f) *la matrice I_n peut être obtenue en effectuant des opérations élémentaires sur les lignes de A ;*

(*g*) *le système homogène* $A \cdot X = 0$ *n'admet que la solution triviale ;*

(*h*) *pour tout* $b \in K^n$*, le système* $A \cdot X = b$ *admet une et une seule solution.*

Une matrice carrée satisfaisant les conditions équivalentes ci-dessus est qualifiée de *régulière*. Les matrices non régulières sont appelées *singulières*.

Démonstration. Il suffit de démontrer les implications

$$(a) \Rightarrow (b) \Rightarrow (g) \Rightarrow (e) \Rightarrow (f) \Rightarrow (d) \Rightarrow (a)$$

pour prouver l'équivalence des conditions (a), (b), (d), (e), (f) et (g). On montrera ensuite

$$(a) \Rightarrow (h) \Rightarrow (g),$$

ce qui prouvera encore l'équivalence de (h) avec les conditions examinées précédemment, et enfin on établira

$$(a) \Rightarrow (c) \Rightarrow (a).$$

$(a) \Rightarrow (b)$: C'est clair par définition de matrice inversible.
$(b) \Rightarrow (g)$: Soit B un inverse à gauche de A. En multipliant les deux membres de $A \cdot X = 0$ à gauche par B, on trouve $X = B \cdot 0 = 0$.
$(g) \Rightarrow (e)$: Supposons que la condition (e) ne soit pas satisfaite. Soit $U \neq I_n$ une matrice sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de A ; alors le nombre de pivots de U est $< n$, donc le système homogène $A \cdot X = 0$ admet une solution non triviale, d'après ce qui a été dit dans la section 1.4.
$(e) \Rightarrow (f)$: C'est clair.
$(f) \Rightarrow (d)$: Cela résulte du Corollaire 2.7.
$(d) \Rightarrow (a)$: Soit $A = E_1 \cdot \cdots \cdot E_r$, où $E_1, \dots, E_r$ sont des matrices élémentaires. D'après la Proposition 2.5, toute matrice élémentaire est inversible. Comme tout produit de matrices inversibles est inversible (voir la Proposition 2.4, p. 13), on en déduit que A est inversible (et que $A^{-1} = E_r^{-1} \cdot \cdots \cdot E_1^{-1}$).
$(a) \Rightarrow (h)$: Si A est inversible, le système $A \cdot X = b$ admet $X = A^{-1} \cdot b$ pour unique solution.
$(h) \Rightarrow (g)$: C'est clair, puisque (g) est un cas particulier de (h) (vu que tout système homogène admet au moins la solution triviale).
$(a) \Rightarrow (c)$: Cela résulte directement de la définition de matrice inversible.
$(c) \Rightarrow (a)$: Soit B un inverse à droite de A. La relation $AB = I_n$ montre que A est un inverse à gauche de B. Or, d'après les implications $(b) \Rightarrow (g) \Rightarrow (e) \Rightarrow (f) \Rightarrow (d) \Rightarrow (a)$ démontrées ci-dessus, on sait que toute matrice carrée inversible à gauche est inversible. Dès lors B est inversible et la matrice A, qui est inverse à gauche de B, est aussi inverse à droite de B. On a donc $BA = I_n$, ce qui montre que B est un inverse à gauche de A. La matrice A est donc inversible. □

Application au calcul de l'inverse. La relation entre l'inversibilité de A et la forme de U établie dans le théorème précédent suggère une manière pratique de déterminer l'inverse de A (si elle existe) : on forme une matrice $n \times 2n$ en juxtaposant A et la matrice unité :

$$\left(\, A \;\middle|\; I_n \, \right),$$

puis on effectue sur les lignes de cette matrice les mêmes opérations élémentaires que pour réduire A à la forme de Gauss–Jordan :

$$\left(\, A \;\middle|\; I_n \, \right) \to \left(\, A_1 \;\middle|\; A'_1 \, \right) \to \cdots \to \left(\, A_r \;\middle|\; A'_r \, \right) \to \left(\, U \;\middle|\; B \, \right).$$

Chaque flèche représente une transformation par une opération élémentaire sur les lignes ou, de manière équivalente, la multiplication à gauche par une matrice élémentaire ; pour tout $i = 0, \dots, r$, il existe donc une matrice élémentaire E_i telle que

$$\left(\, A_{i+1} \;\middle|\; A'_{i+1} \, \right) = E_i \cdot \left(\, A_i \;\middle|\; A'_i \, \right) = \left(\, E_i \cdot A_i \;\middle|\; E_i \cdot A'_i \, \right).$$

(On pose $A_0 = A$, $A'_0 = I_n$, $A_{r+1} = U$ et $A'_{r+1} = B$). On a donc

$$\left(\, U \;\middle|\; B \, \right) = E_r \cdot \cdots \cdot E_0 \cdot \left(\, A \;\middle|\; I_n \, \right),$$

si bien que $B = E_r \cdot \cdots \cdot E_0$ et $U = B \cdot A$.

Si $U \neq I_n$, alors le Théorème 2.8 montre que A n'est pas inversible. En revanche, si $U = I_n$, alors B est un inverse à gauche de A. Or, le Théorème 2.8 montre que A est inversible, donc $B = A^{-1}$.

Remarque : Le rôle des opérations élémentaires *sur les lignes* a été privilégié dans cette section à cause des applications aux systèmes d'équations algébriques linéaires. Cependant, la Proposition 2.6 et le Corollaire 2.7 admettent une variante où le rôle des lignes est joué par les colonnes. La multiplication à gauche par des matrices élémentaires est alors remplacé par la multiplication *à droite*. On peut obtenir ces variantes soit par un raisonnement direct, en adaptant les arguments utilisés dans les démonstrations de la Proposition 2.6 et du Corollaire 2.7, soit en transposant leurs énoncés.

Exercices

(2.1) (prérequis) Que signifie la phrase “les deux matrices A et B commutent” ? Qu'est-ce qu'une matrice diagonale ?

(2.2) Soit un corps K, $A \in K^{n\times n}$, $\alpha \in K$ et D une matrice diagonale de la forme

$$D = \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \alpha \end{pmatrix}.$$

Démontrez que $D \cdot A = A \cdot D$.

(2.3) Calculez A^4 si $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

(2.4) On considère un corps K et trois matrices $A, B, C \in K^{n\times n}$.
Supposons que B soit inversible et que $BAB^{-1} = C$. Démontrez par induction sur k que, pour tout $k \in \mathbb{N}$, $A^k = B^{-1} \cdot C^k \cdot B$.

En généralisant l'exercice 2.3, expliquez pourquoi ce résultat est intéressant si C est une matrice diagonale.

(2.5) Prouvez que multiplier une matrice à droite par une matrice-colonne revient à effectuer une combinaison linéaire de ses colonnes ; c'est-à-dire, prouvez que

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} \cdot x_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix} \cdot x_2 + \cdots + \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix} \cdot x_m.$$

De même, prouvez que multiplier une matrice à gauche par une matrice-ligne revient à effectuer une combinaison linéaire de ses lignes :

$$(x_1 \quad \dots \quad x_n) \cdot \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} = x_1(a_{11} \quad \dots \quad a_{1m}) + \cdots + x_n(a_{n1} \quad \dots \quad a_{nm}).$$

(2.6) Soit un corps K et $A \in K^{n\times m}$.

Prouvez que $A \cdot \begin{pmatrix} b_{11} & \dots & b_{1l} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{ml} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1l} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nl} \end{pmatrix}$,

où $\begin{pmatrix} c_{11} \\ \vdots \\ c_{n1} \end{pmatrix} = A \cdot \begin{pmatrix} b_{11} \\ \vdots \\ b_{m1} \end{pmatrix}$, $\begin{pmatrix} c_{12} \\ \vdots \\ c_{n2} \end{pmatrix} = A \cdot \begin{pmatrix} b_{12} \\ \vdots \\ b_{m2} \end{pmatrix}$, $\dots$, $\begin{pmatrix} c_{1l} \\ \vdots \\ c_{nl} \end{pmatrix} = A \cdot \begin{pmatrix} b_{1l} \\ \vdots \\ b_{ml} \end{pmatrix}$.

Prouvez un résultat analogue pour les lignes.

(2.7) Soit K un corps, $A, B \in K^{n\times n}$. Prouvez que
$$(A+B)^2 = A^2 + 2AB + B^2 \quad \text{si et seulement si} \quad AB = BA.$$

(2.8) Soit K un corps, $A \in K^{n\times m}$ et $B \in K^{m\times p}$. Prouvez que $(AB)^t = B^t A^t$.

(2.9) Soit K un corps, $A \in K^{n\times n}$. On définit la trace de A comme la somme des éléments se trouvant sur sa diagonale (principale) : $\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}$. Prouvez que si $B \in K^{n\times n}$, alors $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. Déduisez-en que, si $C \in K^{n\times n}$ est inversible, alors $\mathrm{tr}(CAC^{-1}) = \mathrm{tr}(A)$.
(pour aller plus loin) Trouvez trois matrices carrées A, B, C telles que $\mathrm{tr}(ABC) \neq \mathrm{tr}(CBA)$.

(2.10) Calculez l'inverse de $\begin{pmatrix} i & 3 & -i \\ 0 & 2i & 1 \\ -i & 0 & -i \end{pmatrix}$.

(2.11) Calculez l'inverse de la matrice $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$.

(2.12) Déterminez pour quelles valeurs de $\lambda \in \mathbb{R}$ la matrice suivante est inversible :
$$\begin{pmatrix} 0 & 1 & \lambda \\ 1 & \lambda & 0 \\ \lambda & 0 & 1 \end{pmatrix}.$$
Pour ces valeurs, calculez son inverse.

(2.13) Soit M une matrice carrée. Démontrez que M est inversible si et seulement si M^t est inversible. Démontrez également que si M est inversible alors $(M^t)^{-1} = (M^{-1})^t$.

(2.14) Si M est une matrice carrée telle que $M^2 - M + I = 0$, démontrez que M est inversible et déterminez son inverse (I est une matrice unité et 0 une matrice nulle, toutes deux de même ordre que M).

(2.15) Expliquez la remarque qui suit la Proposition 2.6, à la page 14.

(2.16) (pour aller plus loin) Si M et N sont des matrices carrées réelles de même ordre, démontrez que M et N sont inversibles si et seulement si $M \cdot N$ est inversible. Calculez l'inverse de M et l'inverse de N en fonction de M, N et $(M \cdot N)^{-1}$.

(2.17) (pour aller plus loin) Développez la remarque de la page 17 : démontrez que l'on peut déterminer si une matrice est inversible et, le cas échéant, calculer son inverse, en faisant des opérations élémentaires sur les *colonnes* plutôt que sur les lignes.

Pourrait-on aussi faire des opérations élémentaires *à la fois* sur les lignes *et* sur les colonnes ?

(2.18) (pour aller plus loin) Est-il possible qu'une matrice carrée réelle ait un inverse complexe mais pas d'inverse réel ?

(2.19) (pour aller plus loin) À partir d'une matrice quelconque $A \in K^{m\times n}$, on forme une matrice $A' \in K^{n\times m}$ en faisant des opérations élémentaires sur les lignes de A et en transposant le résultat obtenu.

- Peut-on obtenir la matrice A' en faisant des opérations élémentaires sur les lignes de la matrice transposée A^t ?
- Peut-on obtenir la matrice A' en faisant des opérations élémentaires sur les colonnes de A^t ?

CHAPITRE II

Espaces vectoriels

3. Notions fondamentales

3.1. La notion d'espace vectoriel

Définition. Soit K un corps (commutatif) et soit E un ensemble muni de deux lois de composition, l'une interne et notée additivement :

$$\begin{aligned} + \quad : \quad & E \times E \to E \\ & (x, y) \mapsto x + y \end{aligned}$$

l'autre externe et notée multiplicativement :

$$\begin{aligned} . \quad : \quad & E \times K \to E \\ & (x, \alpha) \mapsto x\alpha. \end{aligned}$$

On dit que E est un *espace vectoriel* sur K si les conditions suivantes sont satisfaites :

A) E est un groupe commutatif pour $+$.

B) Pour $x, y \in E$ et $\alpha, \beta \in K$,

$$\begin{aligned} x(\alpha + \beta) &= x\alpha + x\beta \\ (x + y)\alpha &= x\alpha + y\alpha \\ x(\alpha\beta) &= (x\alpha)\beta \\ x.1 &= x. \end{aligned}$$

Les éléments de E sont alors appelés *vecteurs* et ceux de K sont appelés *scalaires*. Dans les applications, les scalaires sont le plus souvent des nombres réels ($K = \mathbb{R}$) ou, plus généralement, des nombres complexes ($K = \mathbb{C}$). On parle alors d'espace vectoriel *réel* ou *complexe*, respectivement.

Le résultat de la multiplication d'un vecteur x et d'un scalaire α est ici noté conventionnellement $x\alpha$. Il est cependant indifférent de noter αx pour $x\alpha$. La notation αx est plus habituelle dans certains exemples mais la notation $x\alpha$ conduit à des formules plus élégantes, notamment lors de changements de base.

Soient $v_1, \dots, v_n$ un nombre fini de vecteurs d'un espace vectoriel E. Une *combinaison linéaire* (ou *combili*) de $v_1, \dots, v_n$ est un vecteur $x \in E$ qui admet une décomposition comme somme de multiples de $v_1, \dots, v_n$:

$$x = v_1\alpha_1 + \cdots + v_n\alpha_n$$

pour certains coefficients $\alpha_1, \dots, \alpha_n \in K$.

Exemples.

A) 1. Pour tout corps K, l'ensemble K^n des n-uples d'éléments de K est un espace vectoriel sur K, l'addition vectorielle et la multiplication scalaire étant définies respectivement par :

$$\begin{aligned} (x_1, \dots, x_n) + (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n) \\ (x_1, \dots, x_n)\alpha &= (x_1\alpha, \dots, x_n\alpha). \end{aligned}$$

2. L'ensemble $K^{m \times n}$ des matrices de genre (m, n) sur le corps K est un espace vectoriel sur K, l'addition vectorielle et la multiplication scalaire étant définies comme dans la section 2.1.

B) Pour tout ensemble X et tout espace vectoriel E sur un corps K, l'ensemble $\mathcal{F}(X;E)$ de toutes les fonctions de X vers E est un espace vectoriel sur K, l'addition vectorielle et la multiplication scalaire étant définies respectivement par :

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (\alpha f)(x) &= \alpha f(x).\end{aligned}$$

C) Pour tout corps K, l'ensemble $K[X]$ des polynômes en une indéterminée X à coefficients dans K est un espace vectoriel sur K, l'addition des polynômes et le produit d'un polynôme par un élément de K étant définis de la manière usuelle.

Plus généralement, l'ensemble $K[X_1,\ldots,X_n]$ des polynômes en n indéterminées $X_1, \ldots, X_n$ à coefficients dans K est un espace vectoriel sur K pour les opérations usuelles.

D) L'ensemble E_O des vecteurs de l'espace usuel ayant pour origine un point O arbitrairement choisi est un espace vectoriel réel, l'addition étant définie par la règle du parallélogramme et la multiplication scalaire étant définie de manière évidente. Le point O étant fixé, on identifie souvent E_O à l'ensemble des points de l'espace usuel en faisant correspondre à tout vecteur d'origine O son extrémité.

3.2. La notion de sous-espace vectoriel

Définition. Une partie V d'un espace vectoriel E sur un corps K est un *sous-espace vectoriel* si les conditions suivantes sont satisfaites :

1. $0 \in V$.
2. Pour $x, y \in V$, on a $x + y \in V$.
3. Pour $x \in V$ et $\alpha \in K$, on a $x\alpha \in V$.

De manière équivalente, un sous-espace vectoriel de E est une partie non vide de E qui est stable par combinaisons linéaires : pour $v_1, \ldots, v_n \in V$ et pour $\alpha_1, \ldots, \alpha_n \in K$,

$$v_1\alpha_1 + \cdots + v_n\alpha_n \in V.$$

Une vérification directe montre que tout sous-espace vectoriel est lui-même un espace vectoriel.

Exemples.

A) L'ensemble des solutions d'un système de m équations linéaires homogènes en n indéterminées est un sous-espace vectoriel de K^n.

B)
   1. L'ensemble $C\big(]a,b[;\mathbb{R}\big)$ des fonctions continues sur un intervalle $]a,b[$ à valeurs dans $\mathbb{R}$ est un sous-espace vectoriel de l'espace $\mathcal{F}\big(]a,b[;\mathbb{R}\big)$ de toutes les fonctions de $]a,b[$ dans $\mathbb{R}$.
   2. Pour tout entier $i \geqslant 0$, l'ensemble $C^i\big(]a,b[;\mathbb{R}\big)$ des fonctions dérivables i fois sur l'intervalle $]a,b[$ et dont la dérivée i-ème est continue est un sous-espace vectoriel de l'espace $\mathcal{F}\big(]a,b[;\mathbb{R}\big)$ de toutes les fonctions de $]a,b[$ dans $\mathbb{R}$.
   3. L'ensemble des solutions d'une équation différentielle linéaire homogène est un sous-espace vectoriel de l'espace $\mathcal{F}\big(]a,b[;\mathbb{R}\big)$ de toutes les fonctions de $]a,b[$ dans $\mathbb{R}$.

C) Pour tout entier d, l'ensemble $K[X]_{\leqslant d}$ des polynômes de degré au plus d (y compris le polynôme nul) est un sous-espace vectoriel de l'espace vectoriel $K[X]$ de tous les polynômes à coefficients dans K.

D) Dans l'espace usuel E_O, tout plan et toute droite passant par O est un sous-espace vectoriel. Les droites et les plans ne passant pas par O ne sont pas des sous-espaces vectoriels.

Opérations sur les sous-espaces vectoriels. A. Intersection

Si $V_1, \ldots, V_n$ sont des sous-espaces vectoriels d'un espace vectoriel E, l'intersection $V_1 \cap \ldots \cap V_n$ est un sous-espace vectoriel de E.

B. Somme

Si $V_1, \ldots, V_n$ sont des sous-espaces vectoriels d'un espace vectoriel E, on définit la *somme* de $V_1, \ldots, V_n$ par :

$$V_1 + \cdots + V_n = \{v_1 + \cdots + v_n \mid v_1 \in V_1, \ldots, v_n \in V_n\}.$$

Une vérification directe montre que cet ensemble est un sous-espace vectoriel de E.

On dit que la somme $V_1 + \cdots + V_n$ est *directe* si la condition suivante est satisfaite :

$$\text{Pour } v_1 \in V_1, \ \ldots, v_n \in V_n, \quad v_1 + \cdots + v_n = 0 \Rightarrow v_1 = \cdots = v_n = 0.$$

La somme $V_1 + \cdots + V_n$ est alors notée $V_1 \oplus \cdots \oplus V_n$.

Lorsque la condition ci-dessus est satisfaite, tout vecteur $x \in V_1 \oplus \cdots \oplus V_n$ s'écrit de manière unique comme somme de vecteurs de $V_1, \ldots, V_n$. Supposons en effet

$$x = v_1 + \cdots + v_n = v_1' + \cdots + v_n',$$

où $v_1, v_1' \in V_1, \ \ldots, v_n, v_n' \in V_n$. La dernière égalité donne, en rassemblant tous les termes dans un membre :

$$(v_1 - v_1') + \cdots + (v_n - v_n') = 0.$$

Comme $v_i - v_i' \in V_i$ pour tout i, la condition de somme directe entraîne alors :

$$v_1 - v_1' = \cdots = v_n - v_n' = 0,$$

c'est-à-dire : $v_i = v_i'$ pour tout i.

Lorsqu'il n'y a que deux termes, la condition de somme directe se simplifie : la somme $V + W$ est directe si et seulement si $V \cap W = \{0\}$. (On dit alors que V et W sont des sous-espaces *disjoints*). En effet, supposons d'abord $V \cap W = \{0\}$. Toute relation

$$v + w = 0 \qquad v \in V, \, w \in W$$

entraîne $v = -w$, donc $v \in V \cap W = \{0\}$ et par conséquent $v = w = 0$, ce qui prouve que la somme $V + W$ est directe. Réciproquement, si la somme est directe et si $x \in V \cap W$, alors dans l'équation

$$x + (-x) = 0$$

on peut considérer le premier terme comme un élément de V et le second comme un élément de W ; la condition de somme directe entraîne alors : $x = -x = 0$, ce qui prouve : $V \cap W = \{0\}$.

Exemples :

1. Dans l'espace usuel E_O, la somme de deux droites distinctes passant par O est directe ; la somme de trois droites par O est directe si et seulement si les trois droites ne sont pas dans un même plan.
2. Une matrice carrée $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathbb{R}^{n \times n}$ est dite *symétrique* (resp. *anti-symétrique*) si elle est égale (resp. opposée) à sa transposée : $A = A^t$ (resp. $A = -A^t$). On dit qu'elle est *triangulaire supérieure* (resp. *triangulaire inférieure*, resp. *diagonale*) si $a_{ij} = 0$ pour $i > j$ (resp. pour $i < j$, resp. pour $i \neq j$).

   L'ensemble S des matrices symétriques et l'ensemble T des matrices anti-symétriques de genre (n, n) sont des sous-espaces vectoriels de $\mathbb{R}^{n \times n}$ dont la somme est directe : on a

   $$S \oplus T = \mathbb{R}^{n \times n}.$$

   L'ensemble U des matrices triangulaires supérieures, l'ensemble L des matrices triangulaires inférieures et l'ensemble D des matrices diagonales de genre (n, n) sont également des sous-espaces vectoriels de $\mathbb{R}^{n \times n}$. On a

   $$\mathbb{R}^{n \times n} = L + U,$$

   mais la somme n'est pas directe car $L \cap U = D \neq \{0\}$.

Sous-espace vectoriel engendré. Pour tout vecteur v d'un espace vectoriel E, l'ensemble des multiples scalaires de v :

$$vK = \{v\alpha \mid \alpha \in K\}$$

est un sous-espace vectoriel de E. C'est même le plus petit sous-espace vectoriel de E contenant v, puisque tout sous-espace qui contient v doit aussi contenir ses multiples ; c'est pourquoi on l'appelle *sous-espace vectoriel engendré par* v.

Plus généralement, pour $v_1, \ldots, v_n \in E$, l'ensemble C des combinaisons linéaires de $v_1, \ldots, v_n$:

$$C = \{v_1\alpha_1 + \cdots + v_n\alpha_n \mid \alpha_1, \ldots, \alpha_n \in K\}$$

est le plus petit sous-espace vectoriel de E contenant $v_1, \ldots, v_n$, puisque tout sous-espace contenant $v_1, \ldots, v_n$ contient aussi les combinaisons linéaires de ces vecteurs. Le sous-espace C est appelé *sous-espace vectoriel engendré par* $v_1, \ldots, v_n$ et noté $\mathrm{sev}\langle v_1, \ldots, v_n\rangle$. Vu la définition de somme de sous-espaces vectoriels, on a

$$\mathrm{sev}\langle v_1, \ldots, v_n\rangle = v_1K + \cdots + v_nK.$$

Comme $\mathrm{sev}\langle v_1, \ldots, v_n\rangle$ est le plus petit sous-espace vectoriel contenant $v_1, \ldots, v_n$, on a pour tout sous-espace vectoriel $V \subset E$:

$$\mathrm{sev}\langle v_1, \ldots, v_n\rangle \subset V \qquad \text{si et seulement si} \qquad v_1, \ldots, v_n \in V.$$

3.1. Proposition. *Si $v_1, \ldots, v_n$ et $w_1, \ldots, w_m$ sont deux familles de vecteurs de E, on a*

$$\mathrm{sev}\langle v_1, \ldots, v_n\rangle \subset \mathrm{sev}\langle w_1, \ldots, w_m\rangle$$

si et seulement si $v_1, \ldots, v_n$ sont combinaisons linéaires de $w_1, \ldots, w_m$, et par conséquent

$$\mathrm{sev}\langle v_1, \ldots, v_n\rangle = \mathrm{sev}\langle w_1, \ldots, w_m\rangle$$

si et seulement si $v_1, \ldots, v_n$ sont combinaisons linéaires de $w_1, \ldots, w_m$ et $w_1, \ldots, w_m$ combinaisons linéaires de $v_1, \ldots, v_n$.

Démonstration. Cette proposition découle directement de l'observation précédente ; en effet, on a

$$\mathrm{sev}\langle v_1, \ldots, v_n\rangle \subset \mathrm{sev}\langle w_1, \ldots, w_m\rangle$$

si et seulement si $v_1, \ldots, v_n \in \mathrm{sev}\langle w_1, \ldots, w_m\rangle$, ce qui revient à dire que $v_1, \ldots, v_n$ sont combinaisons linéaires de $w_1, \ldots, w_m$. $\square$

Exercices

(3.1) (prérequis) Décrivez la structure habituelle d'espace vectoriel réel de $\mathbb{R}^n$, $\mathbb{R}[X]$, $\mathbb{R}^{n\times m}$, $\mathcal{F}(\mathbb{R};\mathbb{R})$ (l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$).

Décrivez la structure habituelle d'espace vectoriel de $\mathbb{C}^n$, d'une part sur $\mathbb{R}$ et d'autre part sur $\mathbb{C}$.

(3.2) (pour aller plus loin) Voici trois définitions d'opérations $\oplus\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ et $\odot\colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2$. Dans quel(s) cas la structure $\langle \mathbb{R}^2, \oplus, \odot\rangle$ est-elle un espace vectoriel sur $\mathbb{R}$?

– $(x_1, y_1) \oplus (x_2, y_2) = \Big((\sqrt[3]{x_1} + \sqrt[3]{x_2})^3 , (\sqrt[5]{y_1} + \sqrt[5]{y_2})^5\Big)$
  $(x, y) \odot \alpha = (x\alpha^3, y\alpha^5)$;
– $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, 0)$
  $(x, y) \odot \alpha = (x\alpha, 0)$;
– $(x_1, y_1) \oplus (x_2, y_2) = (y_1 + y_2, x_1 + x_2)$
  $(x, y) \odot \alpha = (y\alpha, x\alpha)$.

(3.3) $\mathbb{C}^2$, avec les opérations usuelles, est à la fois un espace vectoriel sur $\mathbb{C}$ et un espace vectoriel sur $\mathbb{R}$. Tout espace vectoriel complexe peut-il être considéré comme un espace vectoriel réel ?

(3.4) (pour aller plus loin) Soit $\langle E, +_E, \cdot_E \rangle$, un espace vectoriel sur un corps K. Considérons un ensemble A tel qu'il existe une bijection $f : A \to E$.

Si $a, b \in A$ et $\alpha \in K$, on définit

$$\begin{aligned} a \oplus b &= f^{-1}\Big(f(a) +_E f(b)\Big) \\ a \odot \alpha &= f^{-1}\Big(f(a) \cdot_E \alpha\Big). \end{aligned}$$

Démontrez que $\langle A, \oplus, \odot \rangle$ est un espace vectoriel sur K.

(3.5) Soient $(E_1, +_{E_1}, \cdot_{E_1})$ et $(E_2, +_{E_2}, \cdot_{E_2})$ deux espaces vectoriels sur un corps K. On définit des opérations $\oplus : (E_1 \times E_2) \times (E_1 \times E_2) \to (E_1 \times E_2)$ et $\odot : (E_1 \times E_2) \times K \to (E_1 \times E_2)$ de la manière suivante :

- $(x_1, x_2) \oplus (y_1, y_2) = (x_1 +_{E_1} y_1, x_2 +_{E_2} y_2)$;
- $(x_1, x_2) \odot \alpha = (x_1 \cdot_{E_1} \alpha, x_2 \cdot_{E_2} \alpha)$.

Démontrez que $(E_1 \times E_2, \oplus, \odot)$ est un espace vectoriel sur K.

(3.6) Démontrez qu'un sous-espace vectoriel est un espace vectoriel.

(3.7) Soit E un espace vectoriel sur un corps K et $V \subseteq E$. Démontrez que

- V est un sous-espace vectoriel de E
  ssi $V \neq \emptyset$ et $(\forall x, y \in V)(\forall \alpha \in K)(x + y \in V \text{ et } x\alpha \in V)$;
- V est un sous-espace vectoriel de E
  ssi $0 \in V$ et $(\forall x, y \in V)(\forall \alpha, \beta \in K)(x\alpha + y\beta \in V)$.

(3.8) Soit E un espace vectoriel sur un corps K et $V \subseteq E$. Parmi les affirmations suivantes, lesquelles sont vraies ? Justifiez : quand une affirmation est vraie, prouvez-la, et quand elle fausse, trouvez un contre-exemple (le plus simple possible !) montrant qu'elle est fausse.

1. Si V n'est *pas* un sous-espace vectoriel de E, alors $0 \notin V$.
2. Si $0 \notin V$, alors V n'est *pas* un sous-espace vectoriel de E.
3. Si $V \neq \emptyset$ et que V n'est *pas* un sous-espace vectoriel de E, alors pour tous $x, y \in V$ et pour tous $\alpha, \beta \in K$, $x\alpha + y\beta \notin V$.
4. Si V n'est *pas* un sous-espace vectoriel de E, alors on peut trouver $x, y \in V$ et $\alpha, \beta \in K$ tels que $x\alpha + y\beta \notin V$.
5. Si on peut trouver $x, y \in V$ et $\alpha, \beta \in K$ tels que $x\alpha + y\beta \notin V$, alors V n'est *pas* un sous-espace vectoriel de E.

(3.9) On donne ci-dessous quelques exemples d'espaces vectoriels E sur un corps K, ainsi que des vecteurs $v, v_1, \ldots, v_n \in E$. Déterminez dans chaque cas s'il est possible d'écrire v comme combili de $v_1, \ldots, v_n$, c'est-à-dire si $v = v_1\alpha_1 + \cdots + v_n\alpha_n$, pour certains $\alpha_1, \ldots, \alpha_n \in K$.

1. $E = \mathbb{R}^3$, $K = \mathbb{R}$, $v = (2, 4, 6)$, $v_1 = (1, 2, 3), v_2 = (4, 5, 6), v_3 = (7, 8, 9)$.
2. $E = \mathbb{R}[X]$, $K = \mathbb{R}$, $v = 2 + 3X + 4X^2$, $v_1 = 1 + X, v_2 = X + X^2, v_3 = X^2 + X^3$.
3. $E = \mathbb{R}^{2\times 2}$, $K = \mathbb{R}$, $v = \begin{pmatrix} 2 & 3 \\ 4 & 0 \end{pmatrix}$,
   $v_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.
4. $E = \mathbb{C}^2$, $K = \mathbb{C}$, $v = (1 + 2i, 3 - 4i)$, $v_1 = (i, 0), v_2 = (0, i)$.
5. $E = \mathbb{C}^2$, $K = \mathbb{R}$, $v = (1 + 2i, 3 - 4i)$, $v_1 = (i, 0), v_2 = (0, i)$.

(3.10) Parmi les ensembles suivants, lesquels sont des sous-espaces vectoriels de $\mathbb{R}^3$?

- $A = \{(0, y, 0) \mid y \in \mathbb{R}\}$.
- $B = \{(1, y, 0) \mid y \in \mathbb{R}\}$.
- $C = \{(x, y, x + y) \mid x, y \in \mathbb{R}\}$.
- $D = \{(x, y, x + 1) \mid x, y \in \mathbb{R}\}$.

- $E = \{(x, x^2, x^3) \mid x \in \mathbb{R}\}$.
- $F = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 0\}$.
- $G = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = 2y + 3 = z - 1\}$.
- $H = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 0\}$.
- $I = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$.
- $J = \{(x + y, -z, u + t) \in \mathbb{R}^3 \mid x - u = z + t = 2y - t = 0\}$.

(3.11) Parmi les ensembles suivants, lesquels sont des sous-espaces vectoriels de l'espace $\mathcal{F}(\mathbb{R}; \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$?

- $A = \{f : \mathbb{R} \to \mathbb{R} \mid \forall x \in \mathbb{R}\ f(x) \geqslant 0\}$.
- $B = \{f : \mathbb{R} \to \mathbb{R} \mid f$ est croissante ou f est décroissante$\}$,
  où f est croissante ssi $(\forall x, y \in \mathbb{R})(x \leqslant y \Rightarrow f(x) \leqslant f(y))$ et
  f est décroissante ssi $(\forall x, y \in \mathbb{R})(x \leqslant y \Rightarrow f(x) \geqslant f(y))$.
- $C = \{f : \mathbb{R} \to \mathbb{R} \mid f(0) = 0\}$.
- $D = \{f : \mathbb{R} \to \mathbb{R} \mid f(a) = b\}$ (discutez en fonction des valeurs de a et de b).

(3.12) Parmi les ensembles suivants, lesquels sont des sous-espaces vectoriels de l'espace vectoriel complexe $\mathbb{C}[X]$?

- $A = \{P \in \mathbb{C}[X]_{\leqslant 2} \mid P'(0) = 0\}$.
- $B = \{P \in \mathbb{C}[X] \mid P(0) = 2P(1)\}$.

(3.13) (prérequis) Quand une matrice est-elle (1) symétrique ; (2) antisymétrique ?

(3.14) Parmi les ensembles suivants, lesquels sont des sous-espaces vectoriels de $\mathbb{R}^{2\times 2}$?

- $A = \{M \in \mathbb{R}^{2\times 2} \mid M$ est symétrique$\}$.
- $B = \{M \in \mathbb{R}^{2\times 2} \mid M$ est inversible$\}$.
- $C = \{M \in \mathbb{R}^{2\times 2} \mid M_{1,1} = a\}$ (discutez en fonction de la valeur de a).

(3.15) Soit
$$S = \{(x, y, z) \in \mathbb{R}^3 \mid \begin{array}{l} x + y + z = 0 \\ x + 52y + 37z = 0 \\ 31x + 1287y + 389z = 0\}. \end{array}$$
S est-il un sous-espace vectoriel de $\mathbb{R}^3$?
En est-il de même pour l'ensemble des solutions (1) de n'importe quel système homogène ? (2) de n'importe quel système non-homogène ?

(3.16) (pour aller plus loin) Soit $\mathcal{F} = \{f \mid f$ est une fonction $\mathbb{R} \to \mathbb{R}_+\}$ (où $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$).
On définit
$$f \oplus g : \mathbb{R} \to \mathbb{R}^+ : x \mapsto f(x).g(x)$$
$$f \odot \alpha : \mathbb{R} \to \mathbb{R}^+ : x \mapsto (f(x))^\alpha$$
où $f, g \in \mathcal{F}$ et $\alpha \in \mathbb{R}$. Démontrez que ces opérations font de $\mathcal{F}$ un espace vectoriel sur $\mathbb{R}$. $\mathcal{F}$ est-il un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}; \mathbb{R})$ (muni des opérations usuelles) ?

(3.17) On sait que $\mathbb{C}[X]$ est à la fois un espace vectoriel sur $\mathbb{C}$ et un espace vectoriel sur $\mathbb{R}$.
$\mathbb{R}[X]$ est-il un sous-espace vectoriel réel de $\mathbb{C}[X]$? Est-il aussi un sous-espace vectoriel complexe de $\mathbb{C}[X]$?

(3.18) (pour aller plus loin) Démontrez que $V = \{(a + bi, -b - ai) \mid a, b \in \mathbb{R}\}$ est un sous-espace vectoriel de $\mathbb{C}^2$ considéré comme espace vectoriel réel, mais que ce n'est pas un sous-espace vectoriel de $\mathbb{C}^2$ considéré comme espace vectoriel complexe.

(3.19) Démontrez que l'intersection de deux sous-espaces vectoriels est encore un sous-espace vectoriel.

(3.20) La réunion de deux sous-espaces vectoriels est-elle toujours un sous-espace vectoriel ? Si oui, prouvez-le. Si non, la réunion de deux sous-espaces vectoriels n'est-elle jamais un sous-espace vectoriel, ou existe-t-il des sous-espaces vectoriels dont la réunion soit un sous-espace vectoriel et d'autres dont la réunion n'en soit pas un ?

(3.21) Expliquez l'exemple 1 de la page 21 : montrez que la somme de trois droites passant par O est directe ssi les trois droites ne sont pas dans un même plan.

(3.22) (pour aller plus loin) Montrez par induction sur k qu'une somme $V_1 + \cdots + V_k$ de sous-espaces vectoriels est directe ssi $(V_1 + \cdots + V_{i-1}) \cap V_i = \{0\}$, pour tous les $i \in \{2, \ldots, k\}$.

(3.23) (prérequis) Qu'est-ce que le sous-espace vectoriel engendré par un ensemble de vecteurs $v_1, \ldots, v_n$? Démontrez que c'est le plus petit sous-espace vectoriel contenant $v_1, \ldots, v_n$.

4. Bases et dimension

4.1. Bases

Soit $(e_1, \ldots, e_n)$ une suite de vecteurs d'un espace vectoriel E sur un corps K.

On dit que la suite $(e_1, \ldots, e_n)$ est une *suite génératrice* de E si le sous-espace vectoriel engendré par les vecteurs de cette suite est E tout entier :

$$\mathrm{sev}\langle e_1, \ldots, e_n\rangle = E,$$

ce qui revient à dire que tout vecteur $x \in E$ est combinaison linéaire de $e_1, \ldots, e_n$:

$$x = e_1\alpha_1 + \cdots + e_n\alpha_n \qquad \text{pour certains scalaires } \alpha_1, \ldots, \alpha_n \in K.$$

On dit que la suite $(e_1, \ldots, e_n)$ est *libre* (ou que les vecteurs $e_1, \ldots, e_n$ sont *linéairement indépendants*) s'il n'est pas possible de trouver des scalaires $\alpha_1, \ldots, \alpha_n$ non tous nuls tels que $e_1\alpha_1 + \cdots + e_n\alpha_n = 0$; en d'autres termes, la seule combinaison linéaire nulle de cette suite est celle dont tous les coefficients sont nuls :

$$e_1\alpha_1 + \cdots + e_n\alpha_n = 0 \quad \text{entraîne} \quad \alpha_1 = \cdots = \alpha_n = 0.$$

Par convention, la suite vide () est aussi considérée comme libre, et on convient que le sous-espace vectoriel engendré par cette suite est $\{0\}$ (qui est en effet le plus petit de tous les sous-espaces).

Enfin, on dit que la suite $(e_1, \ldots, e_n)$ est *une base* de E si elle est à la fois libre et génératrice. Par exemple, la suite vide () est une base de l'espace nul $\{0\}$, par convention.

Nous limiterons notre discussion de la notion de base et de dimension au cas où l'espace E admet une suite génératrice finie[1]. Un tel espace est dit *finiment engendré*.

4.1. Théorème. *Tout espace vectoriel finiment engendré admet une base.*

Démonstration. L'espace nul admet comme base la suite vide, comme observé ci-dessus. On peut donc se limiter dans la suite de la démonstration à considérer des espaces vectoriels $E \neq \{0\}$. Soit $(e_1, \ldots, e_n)$ une suite génératrice de l'espace E. Si l'un des vecteurs de cette suite (par exemple e_n) est combinaison linéaire des autres, la suite obtenue en retirant ce vecteur est également génératrice : l'égalité

$$\mathrm{sev}\langle e_1, \ldots, e_{n-1}\rangle = \mathrm{sev}\langle e_1, \ldots, e_n\rangle (= E)$$

découle en effet de la Proposition 3.1 (p. 22). À partir de la suite génératrice donnée, on peut donc obtenir une nouvelle suite génératrice plus courte en supprimant tous les vecteurs qui sont combinaisons linéaires des autres. Pour la commodité des notations, supposons que l'on ait éliminé les derniers vecteurs de la suite donnée, de sorte que la nouvelle suite génératrice soit $(e_1, \ldots, e_m)$ pour un certain $m \leqslant n$. On va prouver que cette dernière suite est une base. Comme elle est génératrice, il suffit de prouver que cette suite est libre. Supposons avoir une relation de dépendance linéaire

$$e_1\alpha_1 + \cdots + e_m\alpha_m = 0,$$

avec $\alpha_1, \ldots, \alpha_m$ non tous nuls. Supposons par exemple $\alpha_m \neq 0$. En divisant la relation précédente par α_m et en isolant e_m, on obtient alors

$$e_m = e_1(-\alpha_1\alpha_m^{-1}) + \cdots + e_{m-1}(-\alpha_{m-1}\alpha_m^{-1}).$$

[1] De toute manière, si l'espace ne contient pas de suite génératrice finie, il ne contient pas non plus de base finie puisque les bases sont des suites génératrices particulières. Les bases infinies ne présentent pas le même caractère d'utilité que les bases finies, et ne seront pas utilisées dans ce cours. Voir l'Appendice pour quelques remarques sur les espaces non finiment engendrés.

C'est une contradiction, puisque, par construction, aucun des vecteurs $e_1, \ldots, e_m$ n'est combinaison linéaire des autres. □

Cette démonstration montre de plus que de toute suite génératrice d'un espace vectoriel on peut extraire une base.

Si $e = (e_1, \ldots, e_n)$ est une base de $E(\neq \{0\})$, alors tout vecteur de E s'écrit d'une et d'une seule manière comme combinaison linéaire de e. En effet, tout vecteur $x \in E$ s'écrit comme combinaison linéaire de cette suite car elle est génératrice ; de plus, l'expression est unique, car si

$$x = e_1\alpha_1 + \cdots + e_n\alpha_n = e_1\beta_1 + \cdots + e_n\beta_n,$$

alors, en faisant passer le troisième membre dans le deuxième, on trouve

$$e_1(\alpha_1 - \beta_1) + \cdots + e_n(\alpha_n - \beta_n) = 0,$$

d'où $\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0$ car la suite e est libre. L'unique colonne de scalaires $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ telle que $x = e_1x_1 + \cdots + e_nx_n$ est appelée suite des *coordonnées*[2] de x par rapport à la base e. On la note ${}_e(x)$. L'application

$${}_e\gamma\colon\ E \to K^n$$

définie par

$${}_e\gamma(x) = {}_e(x)$$

est bijective, puisqu'elle admet une application inverse $K^n \to E$, à savoir l'application qui envoie $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$ sur $e_1\alpha_1 + \cdots + e_n\alpha_n \in E$. De plus, on a pour $x, y \in E$ et $\alpha, \beta \in K$:

$${}_e\gamma(x\,\alpha + y\,\beta) = {}_e\gamma(x)\,\alpha + {}_e\gamma(y)\,\beta,$$

car si $x = e_1x_1 + \cdots + e_nx_n$ et $y = e_1y_1 + \cdots + e_ny_n$, alors

$$x\,\alpha + y\,\beta = e_1(x_1\alpha + y_1\beta) + \cdots + e_n(x_n\alpha + y_n\beta).$$

On traduit ces propriétés en disant que ${}_e\gamma$ est un *isomorphisme* d'espaces vectoriels de E vers K^n, ce qui signifie que ${}_e\gamma$ permet d'identifier E et K^n en tant qu'espaces vectoriels. Par exemple, dans $E = K^{2\times 2}$, la suite

$$e = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

est une base de E ; l'isomorphisme correspondant ${}_e\gamma\colon\ K^{2\times 2} \to K^4$ est défini par

$${}_e\gamma\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix};$$

il permet d'identifier $K^{2\times 2}$ et K^4 *en tant qu'espaces vectoriels*, mais on dispose en plus d'un produit sur $K^{2\times 2}$, à savoir le produit matriciel, alors que l'on n'a pas défini de produit sur K^4.

Les applications entre espaces vectoriels seront étudiées de manière détaillée dans le chapitre suivant.

[2]Les coordonnées d'un vecteur sont donc toujours des *matrices-colonnes*. Par abus de notation, on note K^n pour $K^{n\times 1}$.

4.2. Dimension

La *dimension* d'un espace vectoriel finiment engendré est le nombre d'éléments d'une base quelconque. Pour justifier cette définition, il faut prouver que toutes les bases d'un espace vectoriel finiment engendré ont le même nombre d'éléments ; c'est ce que l'on se propose de faire dans cette section.

Le principal outil technique est la propriété suivante, connue sous le nom de *lemme d'échange :*

4.2. Lemme. *Soient $e_1, \ldots, e_r, x, y$ des vecteurs d'un espace vectoriel E. Supposons que $y \in \operatorname{sev}\langle e_1, \ldots, e_r, x\rangle$ et que le coefficient de x dans une expression de y comme combinaison linéaire de $e_1, \ldots, e_r, x$ soit non nul :*

$$(*) \qquad y = e_1\alpha_1 + \cdots + e_r\alpha_r + x\beta \qquad \text{avec } \beta \neq 0;$$

alors $x \in \operatorname{sev}\langle e_1, \ldots, e_r, y\rangle$ et

$$\operatorname{sev}\langle e_1, \ldots, e_r, x\rangle = \operatorname{sev}\langle e_1, \ldots, e_r, y\rangle.$$

Démonstration. En divisant les deux membres de la relation $(*)$ par β et en isolant x dans un membre, on obtient

$$x = e_1(-\alpha_1\beta^{-1}) + \cdots + e_r(-\alpha_r\beta^{-1}) + y\beta^{-1},$$

ce qui prouve la première affirmation. La seconde en découle par la Proposition 3.1. □

4.3. Proposition. *Soit $g = (g_1, \ldots, g_n)$ une suite génératrice d'un espace vectoriel E et $\ell = (\ell_1, \ldots, \ell_m)$ une suite libre de E. On peut remplacer m vecteurs de la suite g par les vecteurs de ℓ sans que la suite cesse d'être génératrice. En particulier, $n \geqslant m$.*

Démonstration. Le principe de la démonstration est le suivant : à l'aide du lemme précédent, on échange successivement m vecteurs de la suite g contre les vecteurs de la suite ℓ sans changer le sous-espace engendré, c'est-à-dire en conservant une suite génératrice.

Considérons d'abord ℓ_1. Comme la suite g est génératrice, on a

$$\ell_1 = g_1\alpha_1 + \cdots + g_n\alpha_n$$

pour certains coefficients $\alpha_1, \ldots, \alpha_n \in K$. Si $\alpha_1 = \cdots = \alpha_n = 0$, alors $\ell_1 = 0$ et la suite ℓ n'est pas libre, contrairement à l'hypothèse. L'un des coefficients est donc non nul ; quitte à changer la numérotation de $g_1, \ldots, g_n$, on peut supposer $\alpha_1 \neq 0$. D'après le lemme d'échange on a alors

$$\operatorname{sev}\langle \ell_1, g_2, \ldots, g_n\rangle = \operatorname{sev}\langle g_1, \ldots, g_n\rangle,$$

donc la suite $(\ell_1, g_2, \ldots, g_n)$ est aussi génératrice.

On répète le même raisonnement successivement pour $\ell_2, \ell_3, \ldots$. Supposons par exemple avoir établi que la suite $(\ell_1, \ldots, \ell_i, g_{i+1}, \ldots, g_n)$ est génératrice. Alors ℓ_{i+1} est combili de cette suite :

$$\ell_{i+1} = \ell_1\alpha_1 + \cdots + \ell_i\alpha_i + g_{i+1}\beta_{i+1} + \cdots + g_n\beta_n$$

pour certains $\alpha_1, \ldots, \alpha_i, \beta_{i+1}, \ldots, \beta_n \in K$. Si les coefficients $\beta_{i+1}, \ldots, \beta_n$ sont tous nuls, alors la suite ℓ n'est pas libre car

$$\ell_1\alpha_1 + \cdots + \ell_i\alpha_i + \ell_{i+1}(-1) + \ell_{i+2}0 + \cdots + \ell_m 0 = 0.$$

(Le coefficient de ℓ_{i+1} est non nul !) Quitte à changer la numérotation de $g_{i+1}, \ldots, g_n$, on peut donc supposer $\beta_{i+1} \neq 0$. D'après le lemme d'échange, on peut alors, dans la suite génératrice $(\ell_1, \ldots, \ell_i, g_{i+1}, \ldots, g_n)$, échanger g_{i+1} contre ℓ_{i+1} pour conclure que la suite $(\ell_1, \ldots, \ell_{i+1}, g_{i+2}, \ldots, g_n)$ est génératrice.

Si $m > n$, alors après avoir appliqué n fois le raisonnement précédent on voit que la suite $(\ell_1, \ldots, \ell_n)$ est génératrice. Le vecteur ℓ_{n+1} est donc combinaison linéaire de cette suite, ce qui contredit comme précédemment l'hypothèse que la suite ℓ est libre. On a donc $m \leqslant n$ et après m itérations les arguments ci-dessus montrent que la suite $(\ell_1, \ldots, \ell_m, g_{m+1}, \ldots, g_n)$ est génératrice (mais la numérotation des vecteurs de la suite g peut avoir changé !) □

4.4. Théorème. *Toutes les bases d'un espace vectoriel finiment engendré ont le même nombre d'éléments.*

Démonstration. Soient $(e_1, \dots, e_n)$ et $(f_1, \dots, f_m)$ deux bases d'un espace vectoriel E. Comme $(e_1, \dots, e_n)$ est une suite génératrice et $(f_1, \dots, f_m)$ une suite libre, la proposition précédente montre que l'on a $m \leqslant n$. Par ailleurs, on peut échanger les rôles des suites $(e_1, \dots, e_n)$ et $(f_1, \dots, f_m)$: comme $(f_1, \dots, f_m)$ est génératrice et que $(e_1, \dots, e_n)$ est libre, la proposition précédente montre que $n \leqslant m$. Ainsi, $n = m$ et toutes les bases de E ont donc le même nombre d'éléments. □

Si un espace E est finiment engendré, on appelle *dimension* de E le nombre d'éléments d'une quelconque de ses bases. Ce nombre entier est noté $\dim E$. Dans le cas contraire, on dit qu'il est de dimension infinie et l'on note $\dim E = \infty$.

Notons encore les conséquences suivantes des résultats précédents :

4.5. Corollaire. *Dans un espace vectoriel E de dimension finie n,*

1. *de toute suite génératrice on peut extraire une base ;*
2. *toute suite génératrice de n éléments est une base ;*
3. *toute suite libre peut être prolongée en base ;*
4. *toute suite libre de n éléments est une base.*

Démonstration. La première propriété a été observée dans la démonstration du Théorème 4.1. La deuxième propriété s'en déduit, car la base extraite de la suite génératrice doit avoir le même nombre d'éléments que celle-ci. La troisième propriété résulte de la Proposition 4.3 et de la propriété précédente : si un espace E admet une suite libre $(x_1, \dots, x_m)$ et une base $(g_1, \dots, g_n)$, alors la Proposition 4.3 montre que l'on peut prolonger $(x_1, \dots, x_m)$ en une suite génératrice de n éléments. Cette suite génératrice est une base, d'après la propriété précédente. Enfin, la quatrième propriété se déduit de la précédente car la base prolongeant la suite libre donnée doit avoir le même nombre d'éléments que celle-ci. □

Exemples :

A) L'espace vectoriel K^n sur un corps K admet une base $(c_1, \dots, c_n)$ appelée *base canonique* (ou *standard*), dont le i-ème vecteur est
$$c_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i}}{1}, 0, \dots, 0).$$
On a donc $\dim K^n = n$.

B) L'espace vectoriel $K[X]_{\leqslant n}$ des polynômes de degré au plus n admet pour base la suite $(1, X, X^2, \dots, X^n)$. On a donc $\dim K[X]_{\leqslant n} = n + 1$.

En revanche, l'espace $K[X]$ de tous les polynômes est de dimension infinie. En effet, si $(P_1, \dots, P_m)$ est une suite finie de polynômes, on peut trouver un entier d strictement plus grand que les degrés de $P_1, \dots, P_m$. Un polynôme de degré d n'est donc pas combinaison linéaire de $(P_1, \dots, P_m)$. L'espace $K[X]$ n'est donc pas finiment engendré, d'où $\dim K[X] = \infty$.

C) Dans l'espace usuel E_O (où l'on a choisi arbitrairement un point de référence O), une suite de vecteurs $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est une base si et seulement si ces vecteurs ne sont pas contenus dans un même plan. On a donc $\dim E_O = 3$.

D) La suite vide () est la seule base de l'espace vectoriel $\{0\}$. On a donc $\dim\{0\} = 0$.

Exercices

(4.1) (prérequis) Soit E un espace vectoriel sur un corps K et $e_1, \dots, e_n \in E$. Ecrivez des formules mathématiques qui signifient :

– la suite $(e_1, \dots, e_n)$ est libre ;
– la suite $(e_1, \dots, e_n)$ n'est pas libre ;
– la suite $(e_1, \dots, e_n)$ est génératrice.

(4.2) Soit E un espace vectoriel sur un corps K et $e_1, \ldots, e_n \in E$ ($n \geqslant 2$). Montrez que la suite $(e_1, \ldots, e_n)$ est libre ssi pour chaque $i \in \{1, \ldots, n\}$, e_i n'est pas combili de $e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n$.

(4.3) Soit E un espace vectoriel sur un corps K et $e_1, \ldots, e_n \in E$.
Prouvez que $(e_1, \ldots, e_n)$ est une base ssi tout élément de E s'écrit de manière unique comme combili de $e_1, \ldots, e_n$.

(4.4) Soit E un espace vectoriel sur un corps K, e une base de E et $x, y, v, v_1, \ldots, v_n \in E$. Démontrez les affirmations suivantes (la fonction ${}_e\gamma$ est définie à la page 26) :

– Si v est combili de x et de y, alors ${}_e\gamma(v)$ est combili de ${}_e\gamma(x)$ et de ${}_e\gamma(y)$.

– Si $(v_1, \ldots, v_n)$ est une suite libre, alors $\Big({}_e\gamma(v_1), \ldots, {}_e\gamma(v_n)\Big)$ est libre.

– Si $(v_1, \ldots, v_n)$ est une base de E, alors $\Big({}_e\gamma(v_1), \ldots, {}_e\gamma(v_n)\Big)$ est une base de K^n.

Démontrez également les trois réciproques.

(4.5) (prérequis) Quelles sont les bases usuelles de $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$ et $\mathbb{R}[X]_{\leqslant n}$, munis des structures habituelles d'espace vectoriel sur $\mathbb{R}$?

(4.6) Donnez une base de $\mathbb{C}$ muni de la structure habituelle d'espace vectoriel sur $\mathbb{R}$.

(4.7) Si E et F sont des espaces vectoriels de dimension finie sur un corps K, donnez une base de l'espace vectoriel $E \times F$ (voir l'exercice 3.5, page 23, pour la définition de $E \times F$).

(4.8) Donnez une base des sous-espaces vectoriels suivants :

– $A = \{P \in \mathbb{R}[X]_{\leqslant 4} \mid P(2) = P(3) = P'(-1)\}$ (sous-espace réel de $\mathbb{R}[X]_{\leqslant 4}$).

– $B = \left\{M \in \mathbb{R}^{2\times 2} \mid M \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot M\right\}$ (sous-espace réel de $\mathbb{R}^{2\times 2}$).

(4.9) La suite $b = \Big((1,0,2), (2,5,4), (-2,1,0)\Big)$ est une base de $\mathbb{R}^3$ (vous pouvez le vérifier). Trouvez les coordonnées de $(17, 26, 18)$ dans b. Trouvez les coordonnées dans b d'un élément (x, y, z) quelconque de $\mathbb{R}^3$.

(4.10) En utilisant les affirmations de l'exercice 4.4 et les calculs de l'exercice 4.9,

– montrez que $b = \Big(1 + 2X^2\,,\; 2 + 5X + 4X^2\,,\; -2 + X\Big)$ est une base de $\mathbb{R}[X]_{\leqslant 2}$;

– trouvez les coordonnées dans b d'un polynôme quelconque $a_0 + a_1X + a_2X^2$ de $\mathbb{R}[X]_{\leqslant 2}$.

(4.11) (pour aller plus loin) Quel est l'ensemble des matrices de $\mathbb{R}^{2\times 2}$ qui commutent avec toutes les matrices de $\mathbb{R}^{2\times 2}$? Montrez que cet ensemble est un sous-espace vectoriel de $\mathbb{R}^{2\times 2}$. Trouvez-en une base.

5. Bases de sous-espaces

Dans cette section, on se propose de donner des indications générales sur les dimensions de sous-espaces d'un espace vectoriel fixé et de mettre au point diverses techniques permettant de trouver des bases de sous-espaces.

5.1. Sous-espaces donnés par une suite génératrice

Le résultat technique suivant est souvent utile pour prouver l'indépendance linéaire d'une suite de vecteurs :

5.1. Lemme. *Soit V un sous-espace vectoriel d'un espace vectoriel E sur un corps K et soit $(v_1, \ldots, v_m)$ une suite libre de V. Si u est un vecteur de E qui n'est pas dans V :*

$$u \in E \smallsetminus V,$$

alors la suite $(v_1, \ldots, v_m, u)$ est une suite libre de E.

DÉMONSTRATION. Soit

$$(*) \qquad v_1\alpha_1 + \cdots + v_m\alpha_m + u\beta = 0$$

une relation de dépendance linéaire. Il faut prouver $\alpha_1 = \cdots = \alpha_m = \beta = 0$. Si $\beta \neq 0$, alors on peut diviser les deux membres de l'équation ci-dessus par β et isoler u dans un membre :

$$u = v_1(-\alpha_1/\beta) + \cdots + v_m(-\alpha_m/\beta).$$

Dès lors u est une combinaison linéaire de $v_1, \ldots, v_m$ et il est donc dans V contrairement à l'hypothèse. On doit donc avoir $\beta = 0$; alors la relation $(*)$ ci-dessus entraîne $\alpha_1 = \cdots = \alpha_m = 0$ puisque $v_1, \ldots, v_m$ sont linéairement indépendants. □

5.2. PROPOSITION. *Soit V un sous-espace vectoriel d'un espace vectoriel E de dimension n (finie).*

1. *V admet une suite génératrice dont le nombre de vecteurs est au plus n. En particulier,*

$$\dim V \leqslant n.$$

2. *Si $\dim V = n$, alors $V = E$.*
3. *Toute base de V peut être prolongée en base de E.*

DÉMONSTRATION. 1. Supposons au contraire que V n'admette pas de suite génératrice de moins de $n+1$ vecteurs. Pour toute suite $(v_1, \ldots, v_m)$ de vecteurs de V, on a donc l'inclusion *stricte*

$$\mathrm{sev}\langle v_1, \ldots, v_m\rangle \subsetneq V \qquad \text{tant que } m \leqslant n.$$

On construit alors dans V une suite libre de $n+1$ vecteurs, de la manière suivante : Soit v_1 un vecteur non nul de V et soit $v_2 \in V \smallsetminus \mathrm{sev}\langle v_1\rangle$. Le lemme précédent (appliqué au sous-espace $\mathrm{sev}\langle v_1\rangle \subsetneq V$) montre que la suite (v_1, v_2) est libre. Si $n \geqslant 2$, choisissons encore $v_3 \in V \smallsetminus \mathrm{sev}\langle v_1, v_2\rangle$, puis (si $n \geqslant 3$) $v_4 \in V \smallsetminus \mathrm{sev}\langle v_1, v_2, v_3\rangle$ et ainsi de suite. Le lemme précédent montre successivement que les suites (v_1, v_2, v_3), (v_1, v_2, v_3, v_4), etc. sont libres. On peut de la sorte construire une suite libre $(v_1, \ldots, v_{n+1})$ de V — donc aussi de E. Cela contredit la Proposition 4.3 (p. 27) ; il est donc absurde de supposer que V n'admet pas de suite génératrice d'au plus n vecteurs.[3]
2. Si $\dim V = n$, alors toute base de V est une suite libre de E dont le nombre d'éléments est égal à la dimension de E ; c'est donc une base de E, d'après le Corollaire 4.5 (p. 28). L'espace engendré par cette suite est donc $V = E$.
3. Cela résulte directement du Corollaire 4.5, car les bases de V sont des suites libres de E. □

Les opérations élémentaires dont il a été question dans la section 1.1 livrent un algorithme permettant de déterminer une base d'un sous-espace vectoriel dont on connaît une suite génératrice. Remarquons d'abord que ces opérations peuvent aussi être utilisées sur les suites de vecteurs : une opération élémentaire de type I consiste à remplacer un des vecteurs de la suite par la somme de ce même vecteur et d'un multiple d'un autre vecteur de la suite ; une opération élémentaire de type II consiste à échanger deux vecteurs de la suite, et une opération élémentaire de type III à multiplier un vecteur par un scalaire non nul, appelé *facteur* de l'opération élémentaire.

5.3. PROPOSITION. *On ne modifie pas le sous-espace vectoriel engendré par une suite de vecteurs quand on effectue sur celle-ci une ou plusieurs opérations élémentaires.*

DÉMONSTRATION. Cela est clair pour les opérations de type II ou III. Pour les opérations élémentaires de type I, il s'agit de prouver :

$$\mathrm{sev}\langle x_1, \ldots, x_n\rangle = \mathrm{sev}\langle x_1, \ldots, x_{i-1}, x_i + x_j\lambda, x_{i+1}, \ldots, x_n\rangle$$

(où $i \neq j$ et λ est un scalaire arbitraire). Or, la forme des vecteurs de la suite de droite montre que chacun de ceux-ci est combinaison linéaire de ceux de gauche. Inversement, montrons que chaque vecteur de gauche est combinaison linéaire de ceux de droite : il suffit de le vérifier pour x_i, puisque tous les autres apparaîssent aussi à droite ; pour x_i, on a

$$x_i = (x_i + x_j\lambda) - x_j\lambda.$$

[3]Cela ne veut pas dire que *toutes* les suites génératrices de V ont au plus n vecteurs. Au contraire, on peut former des suites génératrices dont le nombre de vecteurs est arbitrairement grand en ajoutant des vecteurs nuls à une suite génératrice quelconque.

L'égalité des sous-espaces engendrés résulte à présent de la Proposition 3.1 (p. 22). □

Supposons à présent que les vecteurs donnés soient dans l'espace K^n : soit

$$v_1, \dots, v_m \in K^n.$$

Chacun des vecteurs v_i est donc un n-uple :

$$v_i = (v_{i1}, \dots, v_{in}).$$

On dit que la suite $(v_1, \dots, v_m)$ est *échelonnée* si le nombre de composantes nulles au début de chaque v_i augmente avec l'indice i, l'augmentation étant stricte tant que $v_i \neq 0$:

$$\text{si } v_{ij} = 0 \text{ pour } j \leqslant k, \text{ alors } v_{i+1,j} = 0 \text{ pour } j \leqslant k+1.$$

En d'autres termes, la suite $(v_1, \dots, v_m)$ est échelonnée si et seulement si la matrice de genre (m, n)

$$\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

dont les lignes sont $v_1, \dots, v_m$ est une matrice à lignes échelonnées.

5.4. Proposition. *Toute suite échelonnée de vecteurs non nuls de K^n est libre.*

Démonstration. Soit $(v_1, \dots, v_m)$ une suite échelonnée de vecteurs non nuls de K^n. Supposons avoir trouvé $\alpha_1, \dots, \alpha_m \in K$ tels que

$$(\dagger) \qquad v_1\alpha_1 + \cdots + v_m\alpha_m = 0.$$

Il faut prouver : $\alpha_1 = \cdots = \alpha_m = 0$.

Soit j_1 le plus petit indice tel que $v_{1j_1} \neq 0$. Comme la suite $(v_1, \dots, v_m)$ est échelonnée, on a $v_{ij_1} = 0$ pour tout $i > 1$. Dès lors, la composante d'indice j_1 des deux membres de l'égalité (†) est

$$v_{1j_1}\alpha_1 + 0\alpha_2 + \cdots + 0\alpha_m = 0.$$

On en déduit $\alpha_1 = 0$. L'égalité (†) peut alors se réécrire

$$v_2\alpha_2 + \cdots + v_m\alpha_m = 0.$$

Soit j_2 le plus petit indice tel que $v_{2j_2} \neq 0$. En comparant les composantes d'indice j_2 dans les deux membres de l'égalité précédente, on obtient de même

$$\alpha_2 = 0,$$

et on continue ainsi de suite. Le raisonnement s'applique pour tout $i = 1, \dots, m$, puisque chacun des vecteurs $v_1, \dots, v_m$ possède une composante non nulle. □

Soit $(v_1, \dots, v_m)$ une suite finie quelconque de vecteurs de K^n. Par des opérations élémentaires sur les lignes de la matrice

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in K^{m\times n}$$

dont les lignes sont $v_1, \dots, v_m$, on peut obtenir une matrice à lignes échelonnées, d'après le Théorème 1.1. Soit

$$A' = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \in K^{m\times n}$$

une telle matrice. La Proposition 5.3 montre que le sous-espace de K^n engendré par les lignes de A' est le même que celui qui est engendré par les lignes de A :

$$\text{sev}\langle v_1, \dots, v_m\rangle = \text{sev}\langle w_1, \dots, w_m\rangle.$$

Comme la suite $(w_1, \dots, w_m)$ est échelonnée, les vecteurs non nuls de cette suite sont situés au début ; soit $(w_1, \dots, w_r)$ la sous-suite des vecteurs non nuls. Cette suite est libre d'après la Proposition 5.4, et l'on a

$$\mathrm{sev}\langle w_1, \dots, w_m\rangle = \mathrm{sev}\langle w_1, \dots, w_r\rangle$$

puisque $w_{r+1} = \dots = w_m = 0$. Dès lors, $(w_1, \dots, w_r)$ est une base de $\mathrm{sev}\langle v_1, \dots, v_m\rangle$.

Le procédé précédent s'applique également aux sous-espaces V d'un espace arbitraire E de dimension finie, moyennant le choix d'une base $e = (e_1, \dots, e_n)$ de E. Une telle base détermine en effet un isomorphisme ${}_e\gamma\colon E \to K^n$ qui permet d'identifier V à un sous-espace de K^n.

5.2. Sous-espaces donnés par des équations

Comme indiqué dans la section 3.2, l'ensemble des solutions d'un système homogène d'équations algébriques linéaires en n indéterminées est un sous-espace vectoriel de K^n. La manière de résoudre de tels systèmes a par ailleurs été détaillée dans la section 1.4. Le but de la présente section est de compléter cette information en indiquant comment déterminer une base d'un sous-espace (et donc sa dimension) à partir des solutions trouvées par la méthode de la section 1.4.

Soit un système homogène arbitraire de m équations en n inconnues

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{array}\right.$$

dont on note V l'ensemble des solutions. L'ensemble V est donc un sous-espace vectoriel de K^n :

$$V = \left\{ (x_1, \dots, x_n) \in K^n \;\middle|\; \sum_{j=1}^{n} a_{ij}x_j = 0 \;\text{ pour } i = 1, \dots, m \right\}.$$

Le système étant homogène, la dernière colonne de la matrice complète est nulle. Il suffit donc de considérer la matrice des coefficients. Suivant la méthode de la section 1.4, on effectue des opérations élémentaires sur les lignes de cette matrice pour obtenir une matrice sous forme réduite de Gauss–Jordan qui est la matrice des coefficients d'un système dont l'ensemble des solutions est encore V. On peut négliger les lignes nulles de cette matrice, puisqu'elles correspondent à des équations triviales

$$0\,x_1 + \cdots + 0\,x_n = 0.$$

Supposons qu'il reste r lignes non nulles ; soient $j_1, \dots, j_r$ les indices des colonnes de pivots, de sorte que $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$ et que la matrice des coefficients du système soit du type

$$\begin{array}{c} \begin{array}{ccccccccccccccccccc} & & & j_1 & & & & j_2 & & & & j_3 & & & \cdots & & j_r & & \\ & & & \downarrow & & & & \downarrow & & & & \downarrow & & & & & \downarrow & & \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & \cdots & * & 0 & * & \cdots & * \\ & & & & & & & & & & & & & & \cdots & & & & & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 & * & \cdots & * \end{pmatrix} \end{array}.$$

Pour la commodité des notations, on suppose de plus $j_k = k$ pour $k = 1, \dots, r$. Cette hypothèse ne restreint pas la généralité, car on peut toujours se ramener à cette situation en permutant les composantes des n-uples, ce qui revient à renuméroter les inconnues $x_1, \dots, x_n$. La matrice des coefficients du système est alors de la forme

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & a_{1,r+1} & a_{1,r+2} & \cdots & a_{1n} \\ 0 & 1 & \cdots & 0 & a_{2,r+1} & a_{2,r+2} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & a_{r,r+1} & a_{r,r+2} & \cdots & a_{rn} \end{pmatrix}$$

et le système peut se mettre sous la forme

$$\begin{cases} x_1 &= -(a_{1,r+1}x_{r+1} + a_{1,r+2}x_{r+2} + \cdots + a_{1n}x_n) \\ x_2 &= -(a_{2,r+1}x_{r+1} + a_{2,r+2}x_{r+2} + \cdots + a_{2n}x_n) \\ &\vdots \\ x_r &= -(a_{r,r+1}x_{r+1} + a_{r,r+2}x_{r+2} + \cdots + a_{rn}x_n) \end{cases}$$

Ses solutions sont

$$V = \left\{ \left(-\sum_{j=r+1}^{n} a_{1,j}x_j, -\sum_{j=r+1}^{n} a_{2,j}x_j, \ldots, -\sum_{j=r+1}^{n} a_{rj}x_j, x_{r+1}, \ldots, x_n \right) \middle| \; x_{r+1}, \ldots, x_n \in K \right\}.$$

Considérons les solutions particulières suivantes obtenues en donnant à l'une des variables $x_{r+1}, \ldots, x_n$ la valeur 1 et aux autres la valeur 0 :

$$\begin{aligned} e_{r+1} &= (-a_{1,r+1}, -a_{2,r+1}, \ldots, -a_{r,r+1}, 1, 0, \ldots, 0) \\ e_{r+2} &= (-a_{1,r+2}, -a_{2,r+2}, \ldots, -a_{r,r+2}, 0, 1, \ldots, 0) \\ &\vdots \\ e_n &= (-a_{1n}, -a_{2n}, \ldots, -a_{rn}, 0, 0, \ldots, 1); \end{aligned}$$

alors

$$\left(-\sum_{j=r+1}^{n} a_{1,j}x_j \,, -\sum_{j=r+1}^{n} a_{2,j}x_j \,, \ldots, -\sum_{j=r+1}^{n} a_{rj}x_j \,, x_{r+1}, \ldots, x_n \right) =$$

$$e_{r+1}x_{r+1} + e_{r+2}x_{r+2} + \cdots + e_nx_n,$$

ce qui montre que

$$V = \operatorname{sev}\langle e_{r+1}, \ldots, e_n \rangle.$$

Par ailleurs, pour $\alpha_{r+1}, \ldots, \alpha_n \in K$,

$$e_{r+1}\alpha_{r+1} + \cdots + e_n\alpha_n = (*, \ldots, *, \alpha_{r+1}, \ldots, \alpha_n),$$

donc le membre de gauche ne peut être nul que si $\alpha_{r+1} = \cdots = \alpha_n = 0$. Cela prouve que la suite $(e_{r+1}, \ldots, e_n)$ est libre ; c'est donc une base de V, et par conséquent

$$\dim V = n - r.$$

On a ainsi prouvé :

5.5. Proposition. *La dimension de l'espace des solutions d'un système d'équations linéaires homogènes en n inconnues est $n - r$ où r est le nombre de lignes non nulles d'une matrice sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de la matrice des coefficients du système.*

On pourrait voir que l'entier r qui intervient dans cet énoncé est aussi le nombre de lignes non nulles de toute matrice à lignes échelonnées obtenue en effectuant des opérations élémentaires sur les lignes de la matrice des coefficients. (Voir aussi la Proposition 6.9, p. 41.)

5.3. Dimension d'une somme de sous-espaces

5.6. Proposition. *Si $V_1, \ldots, V_n$ sont des sous-espaces d'un espace vectoriel E, alors une suite génératrice de $V_1 + \cdots + V_n$ s'obtient en juxtaposant des suites génératrices de $V_1, \ldots, V_n$. Si de plus la somme $V_1 + \cdots + V_n$ est directe, alors une base de $V_1 \oplus \cdots \oplus V_n$ s'obtient en juxtaposant des bases de $V_1, \ldots, V_n$; par conséquent,*

$$\dim(V_1 \oplus \cdots \oplus V_n) = \dim V_1 + \cdots + \dim V_n.$$

Démonstration. Soient

$$\begin{gathered}(e_1, \dots, e_r) \text{ une suite génératrice de } V_1,\\ (f_1, \dots, f_s) \text{ une suite génératrice de } V_2,\\ \dots\\ (g_1, \dots, g_t) \text{ une suite génératrice de } V_n.\end{gathered}$$

Il faut prouver que la suite

$$u = (e_1, \dots, e_r, f_1, \dots, f_s, \dots, g_1, \dots, g_t)$$

engendre $V_1 + \cdots + V_n$.

Tout vecteur de cette somme s'écrit

$$x = v_1 + \cdots + v_n$$

où $v_1 \in V_1$, $v_2 \in V_2$, ..., $v_n \in V_n$. En décomposant chaque v_i comme combinaison linéaire de la suite génératrice donnée de V_i, on obtient

$$\begin{gathered}v_1 = e_1\alpha_1 + \cdots + e_r\alpha_r,\\ v_2 = f_1\beta_1 + \cdots + f_s\beta_s,\\ \dots\\ v_n = g_1\gamma_1 + \cdots + g_t\gamma_t;\end{gathered}$$

d'où, en additionnant,

$$x = e_1\alpha_1 + \cdots + e_r\alpha_r + f_1\beta_1 + \cdots + f_s\beta_s + \cdots + g_1\gamma_1 + \cdots + g_t\gamma_t.$$

Cela montre que la suite u engendre $V_1 + \cdots + V_n$ et prouve la première partie de l'énoncé.

Pour prouver la seconde partie, supposons que la somme $V_1 + \cdots + V_n$ soit directe et que les suites génératrices choisies dans $V_1, \dots, V_n$ soient des bases, et montrons que la suite u est une base de $V_1 \oplus \cdots \oplus V_n$. Il suffit de prouver qu'elle est libre, puisque l'on vient de prouver qu'elle engendre $V_1 \oplus \cdots \oplus V_n$. Supposons

$$(\ddagger) \qquad e_1\alpha_1 + \cdots + e_r\alpha_r + f_1\beta_1 + \cdots + f_s\beta_s + \cdots + g_1\gamma_1 + \cdots + g_t\gamma_t = 0.$$

En rassemblant les termes, on fait apparaître des vecteurs de $V_1, \dots, V_n$ dont la somme est nulle :

$$\underbrace{(e_1\alpha_1 + \cdots + e_r\alpha_r)}_{\in V_1} + \underbrace{(f_1\beta_1 + \cdots + f_s\beta_s)}_{\in V_2} + \cdots + \underbrace{(g_1\gamma_1 + \cdots + g_t\gamma_t)}_{\in V_n} = 0.$$

Comme la somme de $V_1, \dots, V_n$ est directe, cette relation entraîne

$$\begin{gathered}e_1\alpha_1 + \cdots + e_r\alpha_r = 0,\\ f_1\beta_1 + \cdots + f_s\beta_s = 0,\\ \dots\\ g_1\gamma_1 + \cdots + g_t\gamma_t = 0.\end{gathered}$$

Comme par ailleurs les suites $(e_1, \dots, e_r)$, $(f_1, \dots, f_s)$, ..., $(g_1, \dots, g_t)$ sont libres (puisque ce sont des bases de V_1, V_2, ..., V_n respectivement), on en déduit

$$\begin{gathered}\alpha_1 = \cdots = \alpha_r = 0,\\ \beta_1 = \cdots = \beta_s = 0,\\ \dots\\ \gamma_1 = \cdots = \gamma_t = 0.\end{gathered}$$

Tous les coefficients de la relation de dépendance linéaire (‡) sont donc nuls, et la suite u est libre. C'est donc bien une base de $V_1 \oplus \cdots \oplus V_n$. □

Lorsque la somme $V_1 + \cdots + V_n$ n'est pas directe, il est plus difficile d'en trouver une base. Voici cependant une indication utile concernant les sommes de deux sous-espaces :

5.7. Proposition. *Si V et W sont deux sous-espaces de dimension finie d'un espace vectoriel E, alors $V + W$ est de dimension finie et*

$$\dim(V + W) + \dim(V \cap W) = \dim V + \dim W.$$

DÉMONSTRATION. Soit $(z_1, \dots, z_r)$ une base de $V \cap W$. On prolonge cette suite d'une part en base de V, soit $(z_1, \dots, z_r, v_1, \dots, v_s)$, d'autre part en base $(z_1, \dots, z_r, w_1, \dots, w_t)$ de W. On a donc $\dim(V \cap W) = r$, $\dim V = r + s$ et $\dim W = r + t$. On se propose de prouver que la suite $(z_1, \dots, z_r, v_1, \dots, v_s, w_1, \dots, w_t)$ est une base de $V + W$. Il en résultera que $V + W$ est de dimension finie et

$$\dim(V + W) = r + s + t = (r + s) + (r + t) - r = \dim V + \dim W - \dim(V \cap W),$$

comme annoncé.

Montrons d'abord que la suite ci-dessus engendre $V + W$. D'après la Proposition 5.6, la suite $(z_1, \dots, z_r, v_1, \dots, v_s, z_1, \dots, z_r, w_1, \dots, w_t)$, obtenue en juxtaposant des bases de V et de W, engendre $V + W$. Or, on a évidemment

$$\operatorname{sev}\langle z_1, \dots, z_r, v_1, \dots, v_s, z_1, \dots, z_r, w_1, \dots, w_t\rangle = \operatorname{sev}\langle z_1, \dots, z_r, v_1, \dots, v_s, w_1, \dots, w_t\rangle,$$

donc la suite $(z_1, \dots, z_r, v_1, \dots, v_s, w_1, \dots, w_t)$ engendre $V + W$.

Montrons pour terminer que cette suite est libre. Supposons avoir trouvé une relation de dépendance linéaire

$$z_1\alpha_1 + \cdots + z_r\alpha_r + v_1\beta_1 + \cdots + v_s\beta_s + w_1\gamma_1 + \cdots + w_t\gamma_t = 0.$$

Il s'agit de prouver que tous les coefficients α_i, β_j, γ_k sont nuls. Pour cela, on fait passer dans le second membre les derniers termes de l'égalité précédente :

$$z_1\alpha_1 + \cdots + z_r\alpha_r + v_1\beta_1 + \cdots + v_s\beta_s = -(w_1\gamma_1 + \cdots + w_t\gamma_t).$$

Comme les vecteurs z_i sont dans $V \cap W$ et les v_j dans V, le premier membre de cette dernière équation est dans V ; cependant, le second membre est dans W puisque les vecteurs w_k sont dans W. Dès lors, les deux membres sont dans $V \cap W$, et l'on peut écrire le second membre comme combinaison linéaire de la base $(z_1, \dots, z_r)$ de $V \cap W$:

$$-(w_1\gamma_1 + \cdots + w_t\gamma_t) = z_1\delta_1 + \cdots + z_r\delta_r$$

pour certains scalaires $\delta_1, \dots, \delta_r$. En rassemblant tous les termes dans un membre, on obtient

$$z_1\delta_1 + \cdots + z_r\delta_r + w_1\gamma_1 + \cdots + w_t\gamma_t = 0,$$

et l'on en tire $\delta_1 = \cdots = \delta_r = \gamma_1 = \cdots = \gamma_t = 0$, car $(z_1, \dots, z_r, w_1, \dots, w_t)$ est une suite libre (c'est même une base de W). En revenant à la relation de dépendance linéaire dont on est parti, et en tenant compte de ce que les coefficients γ_k sont nuls, on a

$$z_1\alpha_1 + \cdots + z_r\alpha_r + v_1\beta_1 + \cdots + v_s\beta_s = 0.$$

Comme $(z_1, \dots, z_r, v_1, \dots, v_s)$ est une suite libre (c'est même une base de V), on en déduit : $\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_s = 0$. □

Exercices

(5.1) Soit $V = \operatorname{sev}\langle(1, -2, 0, 3), (2, 3, 0, 1), (2, -1, 2, 1)\rangle \subseteq \mathbb{R}^4$.
Trouvez une base de V.

(5.2) Soit $e_1 = (1, 2, 3, 4)$, $e_2 = (2, 5, 10, 13)$, $e_3 = (1, 4, 11, 18)$ et $e_4 = (0, 1, 4, 7)$. Considérons $V = \operatorname{sev}\langle e_1, e_2, e_3, e_4\rangle \subseteq \mathbb{R}^4$. Sachant que $\dim V = 3$, trouvez une base de V dont les éléments soient parmi e_1, e_2, e_3, e_4. A priori, suffit-il de prendre 3 éléments e_i au hasard sans effectuer aucune vérification ?

(5.3) Considérons un espace vectoriel E de dimension n sur un corps K et V un sous-espace vectoriel de E, de dimension $k < n$. Soient $b = (e_1, \dots, e_k)$ et $b' = (e'_1, \dots, e'_k)$ deux bases de V. On prolonge b en une base $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ de E. La suite $(e'_1, \dots, e'_k, e_{k+1}, \dots, e_n)$ obtenue en prolongeant b' au moyen des mêmes éléments est-elle également une base de E ?

(5.4) Soit E un espace vectoriel. Considérons trois vecteurs e_1, e_2, e_3 tels que
$\dim \operatorname{sev}\langle e_1, e_2, e_3\rangle = 2$.

- Est-il vrai qu'il existe $i, j \in \{1, 2, 3\}$ tels que (e_i, e_j) soit une base de $\operatorname{sev}\langle e_1, e_2, e_3\rangle$?

– Est-il vrai que, quels que soient $i, j \in \{1, 2, 3\}$, (e_i, e_j) est une base de $\mathrm{sev}\langle e_1, e_2, e_3\rangle$?

(5.5) On considère les matrices

$$M = \begin{pmatrix} 1 & i \\ 3 & -i \end{pmatrix}, \qquad N = \begin{pmatrix} i & 0 \\ 2 & i \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 \\ 1 & -2i \end{pmatrix}.$$

Trouvez une base de $V = \mathrm{sev}\langle M, N, L\rangle \subseteq \mathbb{C}^{2\times 2}$ et complétez-la en une base de $\mathbb{C}^{2\times 2}$.

(5.6) Soit E, un espace vectoriel et $V, W \subseteq E$, des sous-espaces vectoriels. Démontrez que si $\dim V + \dim W > \dim E$, alors $V \cap W \neq \{0\}$.

(5.7) On considère V et W, deux sous-espaces vectoriels d'un espace vectoriel E. Soit $(e_1, \ldots, e_n)$, une base de V et $(f_1, \ldots, f_m)$, une base de W.

– Démontrez que $(e_1, \ldots, e_n, f_1, \ldots, f_m)$ est une suite génératrice de $V + W$.

– Démontrez que si $V \cap W = \{0\}$,
alors $(e_1, \ldots, e_n, f_1, \ldots, f_m)$ est une base de $V \oplus W$.

– Démontrez que si $(e_1, \ldots, e_n, f_1, \ldots, f_m)$ est une base de $V + W$,
alors $V \cap W = \{0\}$.

– Trouvez un contre-exemple montrant que $(e_1, \ldots, e_n, f_1, \ldots, f_m)$ n'est pas nécessairement une base de $V + W$.

– Une suite dont les éléments sont ceux de l'ensemble $\{e_1, \ldots, e_n\} \cap \{f_1, \ldots, f_m\}$ est-elle une base de $V \cap W$? Si non, une telle suite peut-elle être prolongée en une base de $V \cap W$?

(5.8) Soient
$$\begin{aligned} V_1 &= \mathrm{sev}\langle (1,2,3), (4,5,6)\rangle \subseteq \mathbb{R}^3 \\ V_2 &= \mathrm{sev}\langle (0,3,6), (0,9,12)\rangle \subseteq \mathbb{R}^3 \\ V_3 &= \mathrm{sev}\langle (0,3,6), (10,11,12)\rangle \subseteq \mathbb{R}^3. \end{aligned}$$
Quelles sont les dimensions de V_1, V_2, V_3 ? V_1 est-il égal à V_2 ? V_1 est-il égal à V_3 ? Quelle est la dimension de $V_1 \cap V_2$? Donnez une base de $V_1 + V_3$ et de $V_2 \cap V_3$. (Essayez de faire le moins de calculs possible, en utilisant des arguments de dimension.)

(5.9) Donnez une base de $V_1 \cap V_2$ et de $V_1 + V_2$, où V_1 et V_2 sont les deux sous-espaces complexes de $\mathbb{C}[X]_{\leqslant 3}$ suivants :

$$\begin{aligned} V_1 &= \{P \in \mathbb{C}[X]_{\leqslant 3} \mid P(0) = 2P(1)\} \\ V_2 &= \{P \in \mathbb{C}[X]_{\leqslant 3} \mid P(0) = P(1)\}. \end{aligned}$$

(5.10) Donnez une base de $V_1 \cap V_2$ et de $V_1 + V_2$, où V_1 et V_2 sont les deux sous-espaces réels de $\mathbb{R}^{2\times 2}$ suivants :

$$\begin{aligned} V_1 &= \{M \in \mathbb{R}^{2\times 2} \mid M \text{ est antisymétrique}\} \\ V_2 &= \mathrm{sev}\left\langle \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 4 & 4 \end{pmatrix}, \begin{pmatrix} 3 & 5 \\ 10 & 12 \end{pmatrix} \right\rangle. \end{aligned}$$

CHAPITRE III

Rang et déterminant

6. Rang de matrices

6.1. Espace des lignes et espace des colonnes

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in K^{m \times n}$ une matrice quelconque de genre (m, n) sur un corps K, que l'on décompose par lignes :

$$A = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{m*} \end{pmatrix}$$

et par colonnes :

$$A = \begin{pmatrix} a_{*1} & \cdots & a_{*n} \end{pmatrix}.$$

Chacune des lignes a_{i*} est un n-uple : $a_{i*} \in K^n$, et on peut considérer le sous-espace de K^n engendré par ces lignes. Soit

$$\mathcal{L}(A) = \mathrm{sev}\langle a_{1*}, \ldots, a_{m*} \rangle \subset K^n$$

ce sous-espace, que l'on appelle *espace des lignes* de A. De même, on peut considérer l'*espace des colonnes* de A, défini par

$$\mathcal{C}(A) = \mathrm{sev}\langle a_{*1}, \ldots, a_{*n} \rangle \subset K^m.$$

Le résultat principal de cette section (Théorème 6.3) indique que ces sous-espaces ont la même dimension. Nous commençons par quelques observations préliminaires qui seront aussi utilisées pour établir un critère d'inversibilité dans la section suivante.

6.1. Proposition. *Soient A, B, C des matrices sur un corps K telles que*

$$A = BC;$$

alors $\mathcal{C}(A) \subset \mathcal{C}(B)$ et $\mathcal{L}(A) \subset \mathcal{L}(C)$.

Démonstration. Soient $B \in K^{m \times n}$ et $C \in K^{n \times p}$, de sorte que $A \in K^{m \times p}$. Si l'on effectue le produit BC en décomposant B par colonnes, on obtient la matrice A décomposée par colonnes :

$$\begin{pmatrix} a_{*1} & \cdots & a_{*p} \end{pmatrix} = \begin{pmatrix} b_{*1} & \cdots & b_{*n} \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{np} \end{pmatrix}.$$

Cette égalité donne

$$a_{*j} = \sum_{i=1}^{n} b_{*i} c_{ij} \qquad \text{pour } j = 1, \ldots, p,$$

ce qui montre que les colonnes de A sont des combinaisons linéaires des colonnes de B et par conséquent $\mathcal{C}(A) \subset \mathcal{C}(B)$.

Si l'on effectue le produit BC en décomposant C par lignes, on obtient de même

$$\begin{pmatrix} a_{1*} \\ \vdots \\ a_{m*} \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \begin{pmatrix} c_{1*} \\ \vdots \\ c_{n*} \end{pmatrix}$$

d'où

$$a_{i*} = \sum_{j=1}^{n} b_{ij} c_{j*} \qquad \text{pour } i = 1, \dots, m,$$

ce qui montre que les lignes $a_{1*}, \dots, a_{m*}$ de A sont des combinaisons linéaires des lignes $c_{1*}, \dots, c_{n*}$ de C dans l'espace vectoriel K^p, donc $\mathcal{L}(A) \subset \mathcal{L}(C)$. □

Cette proposition admet une sorte de réciproque :

6.2. Proposition. *Soient A, B_1, C_2 des matrices sur un corps K. On suppose que A est de genre (m,p), que B_1 est de genre (m,n) et que C_2 est de genre (n,p) (pour certains entiers m,n,p).*

1. *Si $\mathcal{C}(A) \subset \mathcal{C}(B_1)$, alors il existe une matrice $C_1 \in K^{n\times p}$ telle que*
$$A = B_1 C_1.$$
2. *Si $\mathcal{L}(A) \subset \mathcal{L}(C_2)$, alors il existe une matrice $B_2 \in K^{m\times n}$ telle que*
$$A = B_2 C_2.$$

Démonstration. Il suffit de parcourir les mêmes étapes que dans la démonstration précédente en sens inverse : si $\mathcal{C}(A) \subset \mathcal{C}(B_1)$, alors les colonnes $a_{*1}, \dots, a_{*p}$ de A sont combinaisons linéaires des colonnes $b_{*1}, \dots, b_{*n}$ de B_1 (dans l'espace vectoriel K^m), donc on a des relations du type

$$a_{*j} = \sum_{i=1}^{n} b_{*i} c_{ij} \qquad \text{pour } j = 1, \dots, p$$

pour certains scalaires c_{ij} $(i = 1, \dots, n\,;\ j = 1, \dots, p)$. En posant

$$C_1 = (c_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$$

on a donc

$$\begin{pmatrix} a_{*1} & \cdots & a_{*p} \end{pmatrix} = \begin{pmatrix} b_{*1} & \cdots & b_{*n} \end{pmatrix} C_1$$

c'est-à-dire

$$A = B_1 C_1.$$

La deuxième partie se démontre de manière analogue. □

6.3. Théorème. *Soit A une matrice de genre (m,n) sur un corps K. La dimension du sous-espace vectoriel de K^m engendré par les n colonnes de A est égale à la dimension du sous-espace vectoriel de K^n engendré par les m lignes de A :*

$$\dim \mathcal{C}(A) = \dim \mathcal{L}(A).$$

Démonstration. Soit $\ell = \dim \mathcal{L}(A)$ et $c = \dim \mathcal{C}(A)$. Il s'agit de prouver : $\ell = c$. Il suffit pour cela de prouver : $\ell \geqslant c$ et $\ell \leqslant c$.

Considérons d'abord une matrice $B_1 \in K^{m\times c}$ dont les c colonnes forment une base de $\mathcal{C}(A)$; alors $\mathcal{C}(A) = \mathcal{C}(B_1)$, donc la proposition précédente livre une matrice $C_1 \in K^{c\times n}$ telle que

$$A = B_1 C_1.$$

La Proposition 6.1 montre alors que $\mathcal{L}(A) \subset \mathcal{L}(C_1)$, donc $\ell \leqslant \dim \mathcal{L}(C_1)$. Comme la matrice C_1 n'a que c lignes, on a forcément $\dim \mathcal{L}(C_1) \leqslant c$ et en combinant les deux inégalités précédentes on obtient

$$\ell \leqslant c.$$

Pour démontrer l'inégalité réciproque, on considère une matrice $C_2 \in K^{\ell\times n}$ dont les ℓ lignes forment une base de $\mathcal{L}(A)$. Comme précédemment, on a alors

$$A = B_2 C_2$$

pour une certaine matrice $B_2 \in K^{m\times\ell}$, d'où $\mathcal{C}(A) \subset \mathcal{C}(B_2)$. Comme B_2 n'a que ℓ colonnes, on en déduit

$$c \leqslant \ell.$$

□

On appelle *rang* d'une matrice $A \in K^{m\times n}$ la dimension du sous-espace vectoriel $\mathcal{C}(A)$ de K^m engendré par les n colonnes de A ou, ce qui revient au même d'après le théorème précédent, la dimension du sous-espace vectoriel $\mathcal{L}(A)$ de K^n engendré par les m lignes de A. D'après cette définition, on a clairement

$$\operatorname{rang} A \leqslant \min(m, n)$$

et

$$\operatorname{rang} A = \operatorname{rang} A^t.$$

Comme les opérations élémentaires sur une suite de vecteurs ne modifient pas le sous-espace engendré (Proposition 5.3, p. 30), les opérations élémentaires sur les lignes ou les colonnes d'une matrice ne modifient pas son rang. Par ailleurs, si une matrice est à lignes (resp. colonnes) échelonnées, alors ses lignes (resp. colonnes) non nulles sont linéairement indépendantes, d'après la Proposition 5.4 (p. 31), et forment donc une base de l'espace des lignes (resp. colonnes). Le rang d'une telle matrice est donc le nombre de ses lignes (resp. colonnes) non nulles. En pratique, il suffit donc d'échelonner les lignes (ou les colonnes) d'une matrice par des opérations élémentaires pour déterminer son rang.

6.4. Proposition. *Soient A, B, C des matrices sur un corps K telles que $A = BC$; alors*

$$\operatorname{rang} A \leqslant \min(\operatorname{rang} B, \operatorname{rang} C).$$

Démonstration. La Proposition 6.1 indique que $\mathcal{C}(A) \subset \mathcal{C}(B)$, donc

$$\operatorname{rang} A \leqslant \operatorname{rang} B$$

puisque $\operatorname{rang} A = \dim \mathcal{C}(A)$ et $\operatorname{rang} B = \dim \mathcal{C}(B)$.

De même, la Proposition 6.1 indique que $\mathcal{L}(A) \subset \mathcal{L}(C)$, donc

$$\operatorname{rang} A \leqslant \operatorname{rang} C.$$

□

6.2. Rang et inversibilité

6.5. Proposition. *Une matrice $A \in K^{m\times n}$ est inversible à gauche (resp. à droite) si et seulement si* $\operatorname{rang} A = n$ *(resp.* $\operatorname{rang} A = m$*).*

Démonstration. Supposons d'abord que A soit inversible à gauche. Il existe alors une matrice $B \in K^{n\times m}$ telle que

$$BA = I_n.$$

Il résulte alors de la Proposition 6.1 (p. 37) que $\mathcal{L}(I_n) \subset \mathcal{L}(A)$. Comme les lignes de I_n forment la base canonique de K^n, l'espace engendré par les lignes de A est K^n tout entier, donc la dimension de cet espace est n :

$$\operatorname{rang} A = n.$$

(De manière équivalente, on peut utiliser la Proposition 6.4 pour obtenir $\operatorname{rang} I_n \leqslant \operatorname{rang} A$, d'où $\operatorname{rang} A = n$ puisque $\operatorname{rang} I_n = n$ et $\operatorname{rang} A \leqslant n$.) Réciproquement, si $\operatorname{rang} A = n$, alors les lignes de A forment une suite génératrice de K^n, donc $\mathcal{L}(I_n) \subset \mathcal{L}(A)$. D'après la Proposition 6.2 (p. 38), on peut alors trouver une matrice $B \in K^{n\times m}$ telle que

$$BA = I_n.$$

Pour prouver que A est inversible à droite si et seulement si $\operatorname{rang} A = m$, on utilise les mêmes arguments que ci-dessus, mais on raisonne sur les colonnes au lieu des lignes et avec K^m au lieu de K^n. On peut aussi déduire la condition d'inversibilité à droite de la condition d'inversibilité à gauche, en utilisant la propriété $(AB)^t = B^t A^t$ pour montrer d'abord qu'une matrice est inversible à droite si et seulement si sa transposée est inversible à gauche. Les détails de cette démonstration sont laissés au lecteur. □

6.6. Corollaire. *Toute matrice inversible est carrée. De plus, pour toute matrice carrée A d'ordre n, les conditions équivalentes du Théorème 2.8 sont équivalentes à*

$$\operatorname{rang} A = n;$$

c'est-à-dire qu'une matrice carrée est régulière si et seulement si son rang est égal à son ordre.

Démonstration. Si une matrice $A \in K^{m\times n}$ est inversible, alors elle est inversible à gauche et à droite, et la proposition précédente montre que

$$\operatorname{rang} A = m \qquad \text{et} \qquad \operatorname{rang} A = n,$$

d'où $m = n$ et A est donc carrée. De plus, une matrice carrée d'ordre n est inversible si et seulement si son rang est n. □

6.7. Corollaire. *Le rang d'une matrice $A \in K^{m\times n}$ est l'ordre de la plus grande sous-matrice (carrée) régulière de A.*

Démonstration. Soit $r = \operatorname{rang} A$; on peut donc trouver r lignes de A linéairement indépendantes. La sous-matrice A' de A formée de ces r lignes est alors de genre (r, n) et de rang r. D'après la définition du rang d'une matrice, on peut trouver r colonnes de A' linéairement indépendantes ; la sous-matrice A'' de A' formée de ces r colonnes est une sous-matrice carrée d'ordre r de A qui est régulière puisque son rang est r ; cela prouve que la plus grande sous-matrice régulière de A est d'ordre au moins égal au rang de A.

Inversement, si A'' est une sous-matrice régulière d'ordre s de A, alors les lignes et les colonnes de A'' sont linéairement indépendantes. Soit A' la sous-matrice de genre (s, n) de A qui contient A''. (La matrice A' est donc obtenue à partir de A'' en lui adjoignant les colonnes de A qui lui manquent). Comme la matrice A' contient s colonnes linéairement indépendantes (à savoir celles de A''), son rang est s. Ses s lignes sont donc linéairement indépendantes et comme les lignes de A' sont des lignes de A, la matrice A est de rang au moins égal à s. Cela montre que le rang de A est au moins égal à l'ordre de la plus grande sous-matrice régulière, et achève la démonstration. □

6.3. Rang et systèmes d'équations linéaires

Soit

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right.$$

un système de m équations linéaires en n inconnues, que l'on peut aussi mettre sous la forme

$$A \cdot X = b$$

où

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in K^{m\times n}$$

est la matrice des coefficients du système, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est la colonne des indéterminées et $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ est la colonne des termes indépendants. Rappelons encore que la *matrice complète* du système est

$$(A|b) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \in K^{m\times(n+1)}.$$

6.8. Proposition. *Les conditions suivantes sont équivalentes :*

(a) *le système $A \cdot X = b$ admet une solution ;*

(b) $\operatorname{rang}(A|b) = \operatorname{rang} A$ *;*

(c) $b \in \mathcal{C}(A)$.

En supposant ces conditions satisfaites, la solution du système $A \cdot X = b$ est unique si et seulement si $\operatorname{rang} A = n$*, si et seulement si les colonnes de A sont linéairement indépendantes.*

Démonstration. (a) $\iff$ (b) : Soit $(A'|b')$ une matrice sous forme réduite de Gauss–Jordan obtenue par des opérations élémentaires sur les lignes de $(A|b)$. On a vu dans la section 1.4 que la condition nécessaire et suffisante pour que le système $A \cdot X = b$ admette une solution est que la matrice $(A'|b')$ ne contienne pas de pivot dans la dernière colonne. Cela revient à dire que le nombre de lignes non nulles de $(A'|b')$ doit être le même que celui de A'. Comme les matrices A' et $(A'|b')$ sont à lignes échelonnées, le nombre de lignes non nulles de ces matrices est leur rang. Par ailleurs, on a

$$(*) \qquad \operatorname{rang}(A'|b') = \operatorname{rang}(A|b) \qquad \text{et} \qquad \operatorname{rang} A' = \operatorname{rang} A$$

puisque les opérations élémentaires sur les lignes d'une matrice ne modifient pas son rang. Dès lors, une condition nécessaire et suffisante pour que le système $A \cdot X = b$ admette une solution est :

$$\operatorname{rang}(A|b) = \operatorname{rang} A.$$

(a) $\iff$ (c) : En décomposant A par colonnes : $A = \begin{pmatrix} a_{*1} & \cdots & a_{*n} \end{pmatrix}$, le produit $A \cdot X$ prend la forme

$$A \cdot X = a_{*1}x_1 + \cdots + a_{*n}x_n.$$

Dès lors, il existe $x_1, \ldots, x_n \in K$ tels que $A \cdot X = b$ si et seulement si b est combinaison linéaire de $a_{*1}, \ldots, a_{*n}$, c'est-à-dire

$$b \in \operatorname{sev}\langle a_{*1}, \ldots, a_{*n} \rangle = \mathcal{C}(A).$$

Bien que cela ne soit pas indispensable du point de vue logique, on peut aussi démontrer facilement l'équivalence (b) $\iff$ (c) : en effet, vu l'inclusion $\mathcal{C}(A) \subset \mathcal{C}(A|b)$, la condition (b) est équivalente à : $\mathcal{C}(A) = \mathcal{C}(A|b)$. La Proposition 3.1 montre que cette égalité a lieu si et seulement si $b \in \mathcal{C}(A)$.

Supposons à présent les conditions (a), (b), (c) satisfaites. Comme au début de la démonstration, soit $(A'|b')$ une matrice sous forme réduite de Gauss–Jordan obtenue par des opérations élémentaires sur les lignes de $(A|b)$. D'après la section 1.4, la solution du système $A \cdot X = b$ est unique si et seulement si la matrice $(A'|b')$ est du type

$$\left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & * \\ 0 & 1 & \cdots & 0 & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{array}\right)$$

ce qui revient à dire :

$$\operatorname{rang}(A'|b') = \operatorname{rang} A' = n.$$

Dès lors, vu les égalités $(*)$, la solution du système est unique si et seulement si $\operatorname{rang}(A) = n$. Comme n est le nombre de colonnes de A, cette égalité vaut si et seulement si les colonnes de A forment une base de $\mathcal{C}(A)$, c'est-à-dire si et seulement si les colonnes de A sont linéairement indépendantes. □

Systèmes homogènes.

6.9. Proposition. *Soit $A \in K^{m \times n}$. L'ensemble des solutions du système homogène $A \cdot X = 0$ est un sous-espace vectoriel de K^n de dimension $n - \operatorname{rang} A$.*

Démonstration. La Proposition 5.5 (p. 33) montre que la dimension de l'espace des solutions de $A \cdot X = 0$ est $n - r$, où r est le nombre de lignes non nulles d'une matrice U sous forme réduite de Gauss–Jordan obtenue en effectuant des opérations élémentaires sur les lignes de A. Comme la matrice U est à lignes échelonnées, on a $r = \operatorname{rang} U$, et comme les opérations élémentaires sur les lignes d'une matrice ne modifient pas son rang, on a $\operatorname{rang} A = \operatorname{rang} U$. □

Exercices

(6.1) Calculez le rang de la matrice suivante :

$$A = \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \\ 2 & 8 & 6 \\ 3 & 5 & 2 \end{pmatrix}.$$

(6.2) Calculez le rang de la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 10 & 13 \\ 1 & 4 & 11 & 14 \\ 0 & 1 & 4 & 9 \end{pmatrix}.$$

Trouvez une sous-matrice de A, carrée et régulière, et dont l'ordre soit égal au rang de A.

Mêmes questions pour les matrices

$$B = \begin{pmatrix} 1 & i & 3 & -i \\ i & 0 & 2 & i \\ 1 & 0 & 1 & -2i \end{pmatrix} \qquad \text{et} \qquad C = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

(6.3) Calculez le rang de la matrice suivante, en fonction de la valeur de $m \in \mathbb{R}$:

$$A = \begin{pmatrix} 2 & m & 2 \\ m-2 & -3/2 & -1 \\ 2m & 2(m+1) & m+1 \end{pmatrix}.$$

(6.4) (pour aller plus loin) Soient $A \in K^{m\times n}$ et $B \in K^{n\times k}$. Montrez que si $A \cdot B = 0$, alors $\operatorname{rang} A + \operatorname{rang} B \leqslant n$. (*Indication :* inspirez-vous de la proposition 6.9, page 41.)

7. Déterminant des matrices carrées

À chaque matrice carrée A sur un corps K, on se propose d'associer un élément de K, appelé *déterminant* de A et noté dét A. Il est facile de définir le déterminant des matrices d'ordre 1 ou 2 : on pose

$$\text{dét}\,(a) = a \qquad \text{pour } a \in K$$

et

$$\text{dét} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \qquad \text{pour } a, b, c, d \in K;$$

mais les formules explicites qui permettent de définir le déterminant d'une matrice d'ordre n deviennent rapidement très compliquées lorsque n augmente. Cependant, le déterminant possède un certain nombre de propriétés remarquables qui sont indépendantes de l'ordre des matrices considérées, et qui font son importance. Un petit nombre (en fait, trois) de ces propriétés permettent même de *caractériser* le déterminant. Dès lors, nous procèderons comme suit :

1. nous définirons le déterminant par ses propriétés caractéristiques ;
2. nous montrerons *l'unicité* de la fonction satisfaisant ces propriétés (en supposant qu'il en existe une) ;
3. nous établirons *l'existence* d'une telle fonction.

7.1. Propriétés

Soit n un entier $\geqslant 1$ et K un corps quelconque. Dans cette section, les matrices carrées d'ordre n sont systématiquement décomposées par lignes : pour $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, on considère A comme un n-uple de lignes :

$$A = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{n*} \end{pmatrix}$$

où $a_{i*} = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$ pour $i = 1, \ldots, n$.

Le *déterminant* est l'unique application

$$\text{dét}: \; K^{n\times n} \to K$$

possédant les propriétés suivantes :

(dét.1): Pour tout $i = 1, \ldots, n$ et pour $a_{1*}, \ldots, a_{i-1*}, a_{i+1*}, \ldots, a_{n*}, x, y \in K^n$ et $\alpha, \beta \in K$,

$$\text{dét} \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i-1*} \\ \alpha x + \beta y \\ a_{i+1*} \\ \vdots \\ a_{n*} \end{pmatrix} = \alpha \, \text{dét} \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i-1*} \\ x \\ a_{i+1*} \\ \vdots \\ a_{n*} \end{pmatrix} + \beta \, \text{dét} \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i-1*} \\ y \\ a_{i+1*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

On exprime cette propriété en disant que dét est une *fonction linéaire* (ou qu'il *dépend de manière linéaire*) des lignes de la matrice.

(dét.2): dét est une fonction alternée des lignes, c'est-à-dire que le déterminant d'une matrice ayant deux lignes identiques est nul : si $a_{i*} = a_{j*}$ pour certains indices $i \neq j$, alors

$$\text{dét} \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = 0.$$

(dét.3):

$$\text{dét}\, I_n = 1.$$

Ni l'existence, ni l'unicité d'une telle application ne sont évidentes pour $n \geqslant 2$. Afin d'en préparer la démonstration, on peut déjà observer que les propriétés ci-dessus suffisent à déterminer le comportement du déterminant lorsque la matrice est transformée par des opérations élémentaires sur les lignes.

7.1. Proposition. *Supposons qu'il existe une application $\delta\colon \; K^{n\times n} \to K$ possédant les propriétés (dét.1) et (dét.2) ci-dessus.*

(I) *La valeur de δ ne change pas lorsqu'on remplace une ligne par la somme de celle-ci et d'un multiple d'une autre ligne :*

$$\delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} + \lambda a_{j*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = \delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

(II) *La valeur de δ change de signe lorsqu'on échange deux lignes :*

$$\delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = -\delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

(III) *La valeur de δ est multipliée par $\lambda \in K$ lorsqu'on multiplie une ligne par λ :*

$$\delta \begin{pmatrix} a_{1*} \\ \vdots \\ \lambda a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix} = \lambda\, \delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

Démonstration. (I) : En utilisant la linéarité de δ suivant la i-ème ligne, on trouve

$$\delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} + \lambda a_{j*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = \delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} + \lambda\, \delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

Le deuxième terme du second membre est nul vu que la ligne a_{j*} y apparaît deux fois.
(II) : Comme δ est une fonction alternée des lignes, la matrice dont la i-ème et la j-ème lignes sont égales à $a_{i*} + a_{j*}$ annule δ :

$$\delta \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} + a_{j*} \\ \vdots \\ a_{i*} + a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = 0.$$

On peut utiliser la linéarité de δ suivant la i-ème et la j-ème ligne pour développer le premier membre :

$$\delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*}+a_{j*} \\ \vdots \\ a_{i*}+a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} = \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix} + \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} + \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix} + \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

Dans le membre de droite, les deux termes extrêmes sont nuls car δ est une fonction alternée. En faisant disparaître les termes nuls, il reste

$$0 = \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix} + \delta\begin{pmatrix} a_{1*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix},$$

ce qui établit la propriété (II).
Enfin, la propriété (III) résulte directement de la linéarité de δ suivant la i-ème ligne. □

On peut à présent démontrer l'unicité du déterminant :

7.2. Proposition. *Si $\delta, \delta' \colon K^{n\times n} \to K$ sont deux applications satisfaisant les propriétés (dét.1), (dét.2) et (dét.3), alors $\delta = \delta'$.*

Démonstration. Soit $A \in K^{n\times n}$ et soit $U \in K^{n\times n}$ une matrice sous forme réduite de Gauss–Jordan obtenue par des opérations élémentaires sur les lignes de A. Supposons avoir fait, pour passer de A à U, un nombre r d'opérations élémentaires de type II et un nombre s d'opérations élémentaires de type III, de facteurs $\lambda_1, \ldots, \lambda_s$ (et un certain nombre d'opérations élémentaires de type I). D'après la Proposition 7.1, on a alors

$$(*) \qquad \delta(U) = (-1)^r\,\lambda_1\ldots\lambda_s\,\delta(A) \qquad \text{et} \qquad \delta'(U) = (-1)^r\,\lambda_1\ldots\lambda_s\,\delta'(A).$$

Or, la matrice U est une matrice carrée sous forme réduite de Gauss–Jordan. Comme observé précédemment (juste avant le Théorème 2.8), on a alors l'alternative

soit $U = I_n$,

soit la dernière ligne de U est nulle.

Dans le premier cas, on a $\delta(U) = 1 = \delta'(U)$ d'après la propriété (dét.3), d'où, vu les équations $(*)$,

$$\delta(A) = (-1)^r(\lambda_1\ldots\lambda_s)^{-1} = \delta'(A).$$

Dans le second cas, on a $\delta(U) = 0 = \delta'(U)$ d'après la Proposition 7.1 (utiliser la partie *(III)* avec $\lambda = 0$). Les équations $(*)$ donnent alors

$$\delta(A) = 0 = \delta'(A).$$

Dans tous les cas, on a bien $\delta(A) = \delta'(A)$. □

La proposition précédente montre que les conditions (dét.1), (dét.2) et (dét.3) suffisent à caractériser de manière unique le déterminant. Cependant, à ce stade, il n'est pas clair que le déterminant existe, c'est-à-dire qu'il y a bien une fonction satisfaisant les conditions (dét.1), (dét.2) et (dét.3) (sauf pour $n = 1$ ou 2, où l'on a donné une formule explicite). Nous différons la démonstration de l'existence du déterminant à la section suivante. Nous la supposons acquise pour la suite de la présente section, où nous nous proposons de tirer des conséquences de l'unicité prouvée ci-dessus.

Pour commencer, remarquons que la démonstration précédente donne une manière explicite de calculer dét A. Comme le rang d'une matrice ne change pas lorsque l'on effectue des opérations élémentaires, on a $\operatorname{rang} A = \operatorname{rang} U$, donc le cas $U = I_n$ correspond à : $\operatorname{rang} A = n$ et le cas où la dernière ligne de U est nulle est celui où $\operatorname{rang} A < n$. Avec les notations précédentes, on a donc

$$\text{dét}\, A = \begin{cases} (-1)^r(\lambda_1 \dots \lambda_s)^{-1} (\neq 0) & \text{si } \operatorname{rang} A = n \\ 0 & \text{si } \operatorname{rang} A < n. \end{cases}$$

Cette formule ne peut cependant pas servir de *définition* au déterminant de A, car les opérations élémentaires qui permettent de transformer A en $U (= I_n)$ ne sont pas déterminées de manière unique.

7.3. Proposition. *Pour $A, B \in K^{n \times n}$,*

1. *$\text{dét}\, A \neq 0$ si et seulement si A est régulière.*
2. *$\text{dét}\,(AB) = \text{dét}\, A\, \text{dét}\, B$.*
3. *$\text{dét}\, A^t = \text{dét}\, A$.*

Démonstration. La première partie découle des remarques ci-dessus sur la manière de calculer le déterminant.
(2). Si $\operatorname{rang} A < n$, alors la Proposition 6.4 montre que $\operatorname{rang}(AB) < n$. Dans ce cas, on a donc

$$\text{dét}\, A = \text{dét}\,(AB) = 0,$$

et la relation $\text{dét}\,(AB) = \text{dét}\, A\, \text{dét}\, B$ est satisfaite.

Supposons $\operatorname{rang} A = n$. La matrice réduite de Gauss–Jordan U obtenue par des opérations élémentaires sur les lignes de A est alors I_n. Soient $E_1, \dots, E_\ell$ les matrices élémentaires correspondantes, parmi lesquelles on compte r matrices élémentaires de type II et s matrices élémentaires de type III, de facteurs $\lambda_1, \dots, \lambda_s$. On a, d'après la Proposition 2.6,

$$E_\ell \dots E_1\, A = U = I_n$$

donc

$$A = E_1^{-1} \dots E_\ell^{-1} \tag{†}$$

et

$$AB = E_1^{-1} \dots E_\ell^{-1} B.$$

Comme l'inverse d'une matrice élémentaire est une matrice élémentaire de même type (mais de facteur inverse, en ce qui concerne le type III), l'égalité précédente montre que la matrice AB est obtenue en effectuant les opérations élémentaires qui correspondent à $E_\ell^{-1}, \dots, E_1^{-1}$ sur les lignes de B. Parmi ces opérations élémentaires, on compte r opérations de type II et s opérations de type III, de facteurs $\lambda_1^{-1}, \dots, \lambda_s^{-1}$ (et un certain nombre d'opérations élémentaires de type I). Comme l'effet d'une opération élémentaire sur le déterminant a été déterminé dans la Proposition 7.1, on en déduit

$$\text{dét}\,(AB) = (-1)^r \lambda_1^{-1} \dots \lambda_s^{-1}\, \text{dét}\, B.$$

Or, les observations précédentes sur le calcul de dét A donnent

$$\text{dét}\, A = (-1)^r \lambda_1^{-1} \dots \lambda_s^{-1}.$$

La relation (2) est donc démontrée.
(3). Si $\operatorname{rang} A < n$, alors $\operatorname{rang} A^t < n$ puisque $\operatorname{rang} A^t = \operatorname{rang} A$. Dans ce cas, on a

$$\text{dét}\, A = 0 = \text{dét}\, A^t.$$

Si $\operatorname{rang} A = n$, alors d'après le Théorème 2.8 ou l'équation (†) ci-dessus, A est un produit de matrices élémentaires. Soit

$$A = F_1 \dots F_\ell$$

où $F_1, \dots, F_\ell$ sont des matrices élémentaires. En transposant les deux membres de cette égalité, on trouve

$$A^t = F_\ell^t \dots F_1^t.$$

La propriété de multiplicativité du déterminant (c'est-à-dire la deuxième partie de l'énoncé) démontrée ci-dessus donne alors

$$\text{dét}\, A = \text{dét}\, F_1 \dots \text{dét}\, F_\ell \qquad \text{et} \qquad \text{dét}\, A^t = \text{dét}\, F_\ell^t \dots \text{dét}\, F_1^t.$$

Il suffit donc d'établir : dét F_i = dét F_i^t pour tout i. Rappelons pour cela que toute matrice élémentaire est obtenue en effectuant une opération élémentaire sur les lignes de la matrice unité. Comme la matrice unité est de déterminant 1, on déduit de la Proposition 7.1 (p. 43)

$$\text{dét}\, F_i = \begin{cases} 1 & \text{si } F_i \text{ est de type I,} \\ -1 & \text{si } F_i \text{ est de type II,} \\ \lambda & \text{si } F_i \text{ est de type III, de facteur } \lambda. \end{cases}$$

Or, la transposée d'une matrice élémentaire est une matrice élémentaire de même type (et de même facteur, en ce qui concerne le type III), donc dét F_i = dét F_i^t pour tout i. □

Comme le déterminant d'une matrice est égal au déterminant de sa transposée, toute propriété du déterminant établie pour les lignes des matrices vaut également pour les colonnes ; en particulier, les propriétés (dét.1) et (dét.2) sont également vraies pour les colonnes.

7.4. Corollaire. *Pour $A, B \in K^{n\times n}$ avec A régulière,*

1. *dét$(A^{-1}) = ($dét$A)^{-1}$.*
2. *dét$(ABA^{-1}) =$ détB.*

Démonstration. D'après la multiplicativité du déterminant démontrée ci-dessus, on a

$$\text{dét}\, A\, \text{dét}\,(A^{-1}) = \text{dét}\,(AA^{-1}) = \text{dét}\, I_n = 1,$$

ce qui prouve la première partie. De même, on a

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

Comme les déterminants sont des éléments de K, leur produit est commutatif, donc on peut simplifier le premier facteur du second membre avec le dernier. □

7.2. Existence du déterminant

Développement en mineurs. Démontrer l'existence du déterminant consiste à prouver l'existence pour tout entier $n \geqslant 1$ d'une fonction de $K^{n\times n}$ vers K satisfaisant les conditions (dét.1), (dét.2) et (dét.3). Dans cette section, on définit une telle fonction[1] par induction sur n : on suppose que le déterminant des matrices d'ordre $n-1$ est définie et on l'utilise pour définir le déterminant des matrices d'ordre n. Comme pour les matrices d'ordre 1 le déterminant est défini :

$$\text{dét}\,(a) = a,$$

la définition ci-dessous donne successivement le déterminant des matrices d'ordre 2,3,4, etc.

Fixons un entier $n \geqslant 2$ et supposons avoir défini le déterminant des matrices d'ordre $n-1$. Soit $A \in K^{n\times n}$ une matrice carrée d'ordre n. Pour $i, j = 1, \ldots, n$, on note A_{ij} la matrice d'ordre $n-1$ obtenue en supprimant la i-ème ligne et la j-ème colonne de A.

Pour $k = 1, \ldots, n$, on pose

$$\delta_k(A) = \sum_{\ell=1}^{n} (-1)^{\ell+k} a_{\ell k}\, \text{dét}\, A_{\ell k}.$$

7.5. Proposition. *Pour $k = 1, \ldots, n$, les fonctions $\delta_k \colon K^{n\times n} \to K$ satisfont les conditions (dét.1), (dét.2) et (dét.3). Ces fonctions sont donc toutes égales et définissent le déterminant des matrices d'ordre n.*

Démonstration. (dét.1) : Montrons que δ_k est linéaire suivant la i-ème ligne de A. Considérons pour cela tous les termes

$$(-1)^{\ell+k} a_{\ell k}\, \text{dét}\, A_{\ell k}.$$

Si $i \neq \ell$, alors $a_{\ell k}$ ne dépend pas de la i-ème ligne, et dét $A_{\ell k}$ est linéaire suivant cette ligne. Si $i = \ell$, alors $a_{\ell k}$ dépend linéairement de la i-ème ligne et dét $A_{\ell k}$ n'en dépend pas. Dans tous les cas le produit $(-1)^{\ell+k} a_{\ell k}$ dét $A_{\ell k}$ dépend linéairement de la i-ème ligne de A ; il en est donc de même de δ_k.

[1] En fait, on en définit n. D'après l'unicité du déterminant démontrée dans la section précédente, toutes ces fonctions sont égales.

(dét.2) : Montrons d'abord que $\delta_k(A) = 0$ si A contient deux lignes *consécutives* identiques, soit $a_{i*} = a_{i+1*}$. Les matrices $A_{\ell k}$ contiennent alors deux lignes identiques pour $\ell \neq i, i+1$, donc les termes d'indice $\ell \neq i, i+1$ sont nuls. Il reste

$$\delta_k(A) = (-1)^{i+k} a_{ik} \operatorname{dét} A_{ik} + (-1)^{i+k+1} a_{i+1,k} \operatorname{dét} A_{i+1,k}.$$

Comme la i-ème et la $(i+1)$-ème lignes de A sont identiques, on a $a_{ik} = a_{i+1,k}$ et $A_{ik} = A_{i+1,k}$, donc les deux termes du second membre s'annulent et $\delta_k(A) = 0$.

Le même raisonnement que dans la Proposition 7.1 montre alors que δ_k change de signe quand on permute deux lignes *consécutives* de A.

Supposons enfin que A contienne deux lignes identiques : $a_{i*} = a_{m*}$ pour certains indices $i \neq m$. Par des échanges de lignes consécutives, on peut amener les lignes identiques à être consécutives ; alors δ_k s'annule, comme on l'a vu ci-dessus. Or, des échanges de lignes consécutives ne changent au plus que le signe de δ_k, donc $\delta_k(A) = 0$.

(dét.3) : Comme toutes les entrées de la k-ème colonne de I_n sont nulles sauf la k-ème qui est égale à 1, seul le terme d'indice k de la somme $\delta_k(I_n)$ subsiste :

$$\delta_k(I_n) = (-1)^{k+k}\, 1 \operatorname{dét} I_{n-1} = 1.$$

□

7.6. Corollaire. *Soit* $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in K^{n \times n}$.

1. Développement du déterminant suivant la j-ème colonne :
   pour tout $j = 1, \ldots, n$,
$$\operatorname{dét} A = \sum_{\ell=1}^{n} (-1)^{\ell+j} a_{\ell j} \operatorname{dét} A_{\ell j}.$$
2. Développement du déterminant suivant la i-ème ligne :
   pour tout $i = 1, \ldots, n$,
$$\operatorname{dét} A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \operatorname{dét} A_{ik}.$$
3. *Pour* $i, j = 1, \ldots, n$ *avec* $i \neq j$,
$$0 = \sum_{k=1}^{n} (-1)^{j+k} a_{ik} \operatorname{dét} A_{jk}.$$
4. *Pour* $i, j = 1, \ldots, n$ *avec* $i \neq j$,
$$0 = \sum_{k=1}^{n} (-1)^{k+i} a_{kj} \operatorname{dét} A_{ki}.$$

Démonstration. (1). Le second membre est $\delta_j(A)$. La relation (1) est donc une conséquence directe de la proposition précédente.

(2). Le second membre est $\delta_i(A^t)$. La relation (2) résulte donc de la relation (1) et de la propriété $\operatorname{dét} A = \operatorname{dét} A^t$.

(3). Soit $A' \in K^{n \times n}$ la matrice obtenue en remplaçant dans A la j-ème ligne par la i-ème :

$$\text{si } A = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{j*} \\ \vdots \\ a_{n*} \end{pmatrix}, \quad \text{alors} \quad A' = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{i*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

Comme la matrice A' a deux lignes identiques, son déterminant est nul. Cependant, en appliquant la formule de développement du déterminant suivant la j-ème ligne on trouve

$$\text{dét}\, A' = \sum_{k=1}^{n} (-1)^{j+k} a'_{jk} \,\text{dét}\, A'_{jk}.$$

Or, $a'_{jk} = a_{ik}$ et $A'_{jk} = A_{jk}$, d'où la formule (3).
La formule (4) s'établit en appliquant la formule (3) à la matrice transposée A^t ou en développant suivant la i-ème colonne le déterminant de la matrice obtenue en remplaçant la i-ème colonne de A par la j-ème. □

L'expression $(-1)^{i+j}\text{dét}\, A_{ij}$ qui apparaît dans les développements du déterminant est appelée *cofacteur d'indices i, j* de la matrice A. La matrice

$$\text{cof}\, A = \big((-1)^{i+j}\text{dét}\, A_{ij}\big)_{1\leqslant i,j\leqslant n}$$

est appelée *matrice des cofacteurs de A*.

7.7. Corollaire. *Pour toute matrice $A \in K^{n\times n}$,*

$$A \cdot (\text{cof}\, A)^t = (\text{cof}\, A)^t \cdot A = \textit{dét}\, A \cdot I_n.$$

Démonstration. L'élément d'indices i, j de la matrice produit $A \cdot (\text{cof}\, A)^t$ est

$$\sum_{k=1}^{n} a_{ik}\, (-1)^{j+k}\text{dét}\, A_{jk}.$$

Si $i \neq j$ cette somme est nulle d'après la troisième formule ci-dessus ; si $i = j$ elle vaut dét A d'après le développement du déterminant de A suivant la i-ème ligne. Dès lors, le produit $A \cdot (\text{cof}\, A)^t$ est une matrice diagonale, tous les éléments diagonaux étant égaux à dét A :

$$A \cdot (\text{cof}\, A)^t = \text{dét}\, A \cdot I_n.$$

On raisonne de même pour établir l'égalité $(\text{cof}\, A)^t \cdot A = \text{dét}\, A \cdot I_n$, en utilisant la quatrième formule du corollaire précédent et le développement de dét A suivant la j-ème colonne. □

D'après ce corollaire, l'inverse de toute matrice carrée régulière est donné par

$$A^{-1} = (\text{dét}\, A)^{-1} \cdot (\text{cof}\, A)^t.$$

Étant donné le nombre d'opérations nécessaires pour calculer le déterminant d'une matrice, cette formule n'est cependant utile que d'un point de vue théorique, ou pour calculer l'inverse d'une matrice d'ordre petit ou ayant des propriétés particulières.

On peut aussi l'utiliser pour obtenir une formule donnant la solution d'un système de n équations en n inconnues, lorsque la matrice des coefficients est régulière :

7.8. Corollaire (Cramer). *Soit $A \cdot X = b$ un système de n équations algébriques linéaires en n inconnues. Pour $i = 1, \ldots, n$, on désigne par B_i la matrice carrée d'ordre n obtenue en remplaçant la i-ème colonne de A par la colonne b des termes indépendants. Si A est régulière, alors la valeur de la i-ème inconnue x_i est donnée par*

$$x_i = \frac{\textit{dét}\, B_i}{\textit{dét}\, A}.$$

Démonstration. En développant dét B_i suivant la i-ème colonne, on trouve

$$\begin{aligned}\text{dét}\, B_i &= \sum_{k=1}^{n} (-1)^{i+k} b_k \,\text{dét}\, (B_i)_{ki}\\ &= \sum_{k=1}^{n} (-1)^{i+k} b_k \,\text{dét}\, A_{ki}.\end{aligned}$$

Ce calcul montre que dét B_i est aussi la i-ème ligne de la matrice-colonne $(\text{cof}\, A)^t \cdot b$. Or, si A est régulière, on tire de $A \cdot X = b$

$$X = A^{-1} \cdot b = (\text{dét}\, A)^{-1} \cdot (\text{cof}\, A)^t \cdot b.$$

□

Formule explicite. En utilisant systématiquement la propriété de linéarité du déterminant suivant chaque ligne, on se propose de donner une formule explicite pour le déterminant d'une matrice carrée d'ordre n.

Soit $A = (a_{ij})_{1\leqslant i,j\leqslant n} \in K^{n\times n}$ une matrice carrée d'ordre n, que l'on décompose par lignes :

$$A = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

Décomposons la première ligne comme combinaison linéaire de la base canonique $(c_1, \ldots, c_n)$ de K^n :

$$a_{1*} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \end{pmatrix} = a_{11}c_1 + \cdots + a_{1n}c_n.$$

Comme le déterminant est linéaire suivant la première ligne, on en déduit

$$\begin{aligned} \text{dét}\, A &= a_{11}\,\text{dét}\begin{pmatrix} c_1 \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix} + a_{12}\,\text{dét}\begin{pmatrix} c_2 \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix} + \cdots + a_{1n}\,\text{dét}\begin{pmatrix} c_n \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix} \\ &= \sum_{i_1=1}^{n} a_{1i_1}\,\text{dét}\begin{pmatrix} c_{i_1} \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix}. \end{aligned}$$

Décomposons ensuite la deuxième ligne :

$$a_{2*} = a_{21}c_1 + \cdots + a_{2n}c_n = \sum_{i_2=1}^{n} a_{2i_2}c_{i_2}.$$

Comme le déterminant est linéaire suivant la deuxième ligne, on obtient

$$\text{dét}\begin{pmatrix} c_{i_1} \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix} = \sum_{i_2=1}^{n} a_{2i_2}\,\text{dét}\begin{pmatrix} c_{i_1} \\ c_{i_2} \\ a_{3*} \\ \vdots \\ a_{n*} \end{pmatrix},$$

d'où

$$\text{dét}\, A = \sum_{i_1,i_2=1}^{n} a_{1i_1}a_{2i_2}\,\text{dét}\begin{pmatrix} c_{i_1} \\ c_{i_2} \\ a_{3*} \\ \vdots \\ a_{n*} \end{pmatrix}.$$

En décomposant de même chaque ligne de A, on obtient

$$(\ddagger) \qquad \text{dét}\, A = \sum_{i_1,\ldots,i_n=1}^{n} a_{1i_1}\ldots a_{ni_n}\,\text{dét}\begin{pmatrix} c_{i_1} \\ \vdots \\ c_{i_n} \end{pmatrix}.$$

Si deux des indices $i_1, \ldots, i_n$ prennent la même valeur, alors la suite $c_{i_1}, \ldots, c_{i_n}$ contient deux fois le même vecteur, donc

$$\text{dét}\begin{pmatrix} c_{i_1} \\ \vdots \\ c_{i_n} \end{pmatrix} = 0.$$

Si tous les indices $i_1, \ldots, i_n$ prennent des valeurs différentes, alors

$$\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$$

et la suite $i = (i_1, \dots, i_n)$ est une permutation de $(1, \dots, n)$. Par un certain nombre s d'opérations de type II (échanges de deux vecteurs), les vecteurs de la suite $(c_{i_1}, \dots, c_{i_n})$ peuvent être rangés dans l'ordre naturel $(c_1, \dots, c_n)$. D'après la Proposition 7.1, on a alors

$$\text{dét} \begin{pmatrix} c_{i_1} \\ \vdots \\ c_{i_n} \end{pmatrix} = (-1)^s \,\text{dét} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

et comme

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n,$$

on en déduit

$$\text{dét} \begin{pmatrix} c_{i_1} \\ \vdots \\ c_{i_n} \end{pmatrix} = (-1)^s.$$

Le nombre $(-1)^s$ est appelé *signature* de la permutation $i = (i_1, \dots, i_n)$; on le note $\operatorname{sgn} i$. On a donc $\operatorname{sgn} i = \pm 1$, et plus précisément $\operatorname{sgn} i = 1$ (resp. $\operatorname{sgn} i = -1$) si la suite $(i_1, \dots, i_n)$ peut être rangée dans l'ordre naturel par un nombre *pair* (resp. *impair*) d'échanges de deux éléments. Par exemple, $\operatorname{sgn}(1, \dots, n) = 1$ puisque $(1, \dots, n)$ est dans l'ordre naturel.

En revenant à la relation (‡), on obtient la formule explicite suivante pour le déterminant de A :

$$\text{dét}\, A = \sum_i \operatorname{sgn} i\, a_{1i_1} \dots a_{ni_n}$$

où $i = (i_1, \dots, i_n)$ parcourt l'ensemble des permutations de $(1, \dots, n)$.

Cette formule est surtout utilisable pour $n = 2$ ou 3. Pour $n = 2$, on a $\operatorname{sgn}(1, 2) = 1$ et $\operatorname{sgn}(2, 1) = -1$ car il suffit d'échanger 1 et 2 pour revenir à l'ordre naturel. Dès lors,

$$\text{dét}\, A = a_{11}a_{22} - a_{12}a_{21}.$$

Pour $n = 3$ on a

$$\operatorname{sgn}(2, 1, 3) = \operatorname{sgn}(3, 2, 1) = \operatorname{sgn}(1, 3, 2) = -1$$

et

$$\operatorname{sgn}(2, 3, 1) = 1$$

car pour passer de $(2, 3, 1)$ à $(1, 2, 3)$ il suffit de deux échanges : par exemple, on échange d'abord 1 et 2, puis 2 et 3. De même,

$$\operatorname{sgn}(3, 1, 2) = 1;$$

donc

$$\begin{aligned} \text{dét}\, A &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}. \end{aligned}$$

Exercices

(7.1) (prérequis) Démontrez que le déterminant d'une matrice est nul si au moins une de ses lignes est nulle.

(7.2) Calculez le déterminant de

$$A = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1+i & 1+2i \\ 1-i & 0 & 2-3i \\ 1-2i & 2+3i & 0 \end{pmatrix}.$$

Quel est le déterminant de $A \cdot B$ et celui de $-A$? Après avoir calculé ces deux déterminants, vous y prendriez-vous de la même manière pour calculer le déterminant de $A+A$ et le déterminant de $A+B$?

(7.3) Calculez le déterminant des deux matrices suivantes, en effectuant le moins de calculs possible :

$$A = \begin{pmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{pmatrix}.$$

(7.4) Démontrez que le déterminant de toute matrice complexe antisymétrique d'ordre impair est nul.

(7.5) On considère une matrice $A \in K^{n \times n}$ telle que $a_{ij} = 0$ si $j \leqslant n - i$. Cela signifie que A est de la forme

$$\begin{pmatrix} 0 & \dots & & 0 & a_{1,n} \\ 0 & \dots & 0 & a_{2,n-1} & a_{2,n} \\ \vdots & & & \vdots & \\ 0 & a_{n-1,2} & & \dots & a_{n-1,n} \\ a_{n,1} & & \dots & & a_{n,n} \end{pmatrix}.$$

Quel est le déterminant de A ?

CHAPITRE IV

Applications linéaires

8. Notions fondamentales

8.1. Définition

Soient E et F deux espaces vectoriels sur un même corps K. Une application $A\colon E \to F$ est dite *linéaire* si les conditions suivantes sont satisfaites :

1. $A(x+y) = A(x) + A(y)$ pour $x, y \in E$.
2. $A(x\alpha) = A(x)\alpha$ pour $x \in E$ et $\alpha \in K$.

De manière équivalente, l'application A est linéaire si

$$A(x\alpha + y\beta) = A(x)\alpha + A(y)\beta$$

pour $x, y \in E$ et $\alpha, \beta \in K$. Plus généralement, on a alors $A(\sum_{i\in I} x_i\alpha_i) = \sum_{i\in I} A(x_i)\alpha_i$ pour tout ensemble fini d'indices I.

Une vérification directe montre que l'ensemble des applications linéaires de E dans F est un sous-espace vectoriel de l'espace de toutes les applications de E dans F et que, de plus, la composée de deux applications linéaires est linéaire.

Une terminologie particulière s'attache à certains types d'applications linéaires : les applications linéaires $A\colon E \to E$ sont appelées *opérateurs linéaires* sur E, ou parfois aussi *endomorphismes* de E. Les applications linéaires injectives (resp. surjectives) sont parfois appelées *monomorphismes* (resp. *épimorphismes*) et les applications linéaires bijectives sont appelées *isomorphismes* (d'espaces vectoriels). Enfin, les endomorphismes bijectifs sont appelés *automorphismes*.

La notion suivante est étroitement liée à celle d'application linéaire : une *équation linéaire* est une équation de la forme

$$A(X) = b$$

où $A\colon E \to F$ est une application linéaire entre deux espaces vectoriels E et F, et $b \in F$. Résoudre cette équation consiste à trouver tous les vecteurs $x \in E$ qui sont envoyés sur b par A. Une telle équation satisfait le *principe de superposition des solutions :* si $x_1, x_2 \in E$ sont solutions de $A(X) = b_1$ et $A(X) = b_2$ respectivement, alors pour $\alpha_1, \alpha_2 \in K$ le vecteur $\alpha_1x_1 + \alpha_2x_2 \in E$ est solution de $A(X) = \alpha_1b_1 + \alpha_2b_2$. Ce principe de superposition des solutions n'est en effet qu'une reformulation de l'hypothèse que A est linéaire.

8.2. Exemples

A) Toute matrice A de genre (m,n) sur un corps K définit une application linéaire $\tilde{A}\colon K^n \to K^m$ par $\tilde{A}(x) = A \cdot x$. Pour $b \in K^m$, le système d'équations algébriques linéaires

$$A \cdot X = b$$

peut alors aussi s'écrire

$$\tilde{A}(X) = b;$$

c'est donc un cas particulier d'équation linéaire.

B) La dérivation est une application linéaire $D\colon C^i\big(]a,b[;\mathbb{R}\big) \to C^{i-1}\big(]a,b[;\mathbb{R}\big)$.

Plus généralement, si $a_1, \ldots, a_n \in C\big(]a,b[;\mathbb{R}\big)$ sont des applications continues sur un intervalle $]a,b[\subset \mathbb{R}$, on peut définir une application linéaire

$$A\colon C^n\big(]a,b[;\mathbb{R}\big) \to C\big(]a,b[;\mathbb{R}\big)$$

par

$$A = D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n I,$$

c'est-à-dire

$$A(y) = y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y$$

pour $y \in C^n\big(]a,b[;\mathbb{R}\big)$. Pour toute fonction $f \in C\big(]a,b[;\mathbb{R}\big)$, l'équation différentielle ordinaire

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = f$$

est donc un exemple d'équation linéaire.

C) L'évaluation des polynômes en un point $a \in K$ est une application linéaire $v_a\colon K[X] \to K$ définie par $v_a(P) = P(a)$ pour $P \in K[X]$.

Plus généralement, étant donnés n éléments $a_1, \ldots, a_n \in K$, on peut définir une application linéaire $A\colon K[X] \to K^n$ par

$$A(P) = (P(a_1), \ldots, P(a_n)).$$

Résoudre l'équation linéaire $A(P) = (b_1, \ldots, b_n)$ revient à déterminer tous les polynômes $P \in K[X]$ qui prennent en chaque a_i la valeur b_i.

D) Dans l'espace usuel E_O, les transformations de symétrie par rapport à O ou par rapport à un plan passant par O, les projections sur un plan passant par O parallèlement à une direction donnée, les rotations dont l'axe passe par O sont des transformations linéaires. Par contre, les transformations de symétrie par rapport à un point autre que O, les rotations dont l'axe ne passe pas par O ne sont pas linéaires.

En relation avec le premier exemple ci-dessus, on peut déterminer *toutes* les applications linéaires de K^n vers K^m :

8.1. Proposition. *Toute application linéaire $L\colon K^n \to K^m$ est de la forme $L = \tilde{A}$ pour une certaine matrice $A \in K^{m\times n}$.*

Démonstration. À toute application linéaire $L\colon K^n \to K^m$, on associe la matrice $A \in K^{m\times n}$ dont les colonnes sont les images des vecteurs de la base canonique de K^n :

$$A = \big(L(c_1), \ldots, L(c_n)\big) \in K^{m\times n}.$$

Si $L(c_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$, alors pour tout $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c_1 x_1 + \cdots + c_n x_n \in K^n$, on a

$$L(x) = L(c_1)x_1 + \cdots + L(c_n)x_n = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = A \cdot x.$$

Dès lors, l'application linéaire $\tilde{A}$ associée à la matrice A est bien la même que L. □

8.3. Noyau et image

À toute application linéaire $A\colon E \to F$, on fait correspondre un sous-espace vectoriel de E appelé *noyau* de A, défini par

$$\operatorname{Ker} A = \{x \in E \mid A(x) = 0\}$$

et un sous-espace vectoriel de F appelé *image* de A, défini par

$$\operatorname{Im} A = \{A(x) \in F \mid x \in E\}$$

ou, de manière équivalente, par

$$\operatorname{Im} A = \{y \in F \mid \exists x \in E \text{ tel que } A(x) = y\}.$$

Le fait que ces ensembles sont bien des sous-espaces vectoriels de E et de F respectivement résulte d'une simple vérification. (Pour voir que $0 \in \operatorname{Ker} A$, c'est-à-dire que $A(0) = 0$, il suffit de remplacer α par 0 dans la relation $A(x\alpha) = A(x)\alpha$.)

8.2. Proposition. *Une application* $A\colon E \to F$ *est injective si et seulement si* $\operatorname{Ker} A = \{0\}$ *; elle est surjective si et seulement si* $\operatorname{Im} A = F$.

Démonstration. Si A est injective, alors seul le vecteur nul de E peut être envoyé sur le vecteur nul de F, donc $\operatorname{Ker} A = \{0\}$. Réciproquement, supposons que $\operatorname{Ker} A = \{0\}$ et que $x_1, x_2 \in E$ ont la même image dans F :

$$A(x_1) = A(x_2);$$

alors $A(x_1 - x_2) = 0$, donc $x_1 - x_2 \in \operatorname{Ker} A$. Comme $\operatorname{Ker} A = \{0\}$, cela entraîne $x_1 - x_2 = 0$, donc $x_1 = x_2$. La deuxième partie de l'énoncé résulte directement des définitions. □

Le noyau et l'image ont une interprétation naturelle en termes d'équations linéaires. En effet, $\operatorname{Ker} A$ est l'ensemble des solutions de l'équation linéaire homogène $A(x) = 0$, tandis que $\operatorname{Im} A$ est l'ensemble des seconds membres $b \in F$ pour lesquels l'équation $A(X) = b$ admet une solution. En particulier, l'équation $A(X) = b$ admet une solution si et seulement si $b \in \operatorname{Im} A$. La structure de l'ensemble des solutions est donnée par la proposition suivante :

8.3. Proposition. *Soit* $A\colon E \to F$ *une application linéaire et* $b \in F$*. Supposons que l'équation linéaire* $A(X) = b$ *admette* $u \in E$ *comme solution. L'ensemble des solutions de cette équation est alors l'ensemble*

$$u + \operatorname{Ker} A = \{u + v \mid v \in \operatorname{Ker} A\}.$$

Démonstration. Pour tout $v \in \operatorname{Ker} A$, on a $A(u+v) = A(u) + A(v) = b + 0$, donc tout élément de $u + \operatorname{Ker} A$ est solution de $A(X) = b$. Inversement, si s est solution de $A(X) = b$, alors

$$A(s) = b = A(u),$$

donc $s - u \in \operatorname{Ker} A$ et $s = u + (s - u) \in u + \operatorname{Ker} A$. □

Étant donné une solution particulière $u \in E$ de l'équation $A(X) = b$, on obtient donc la solution générale de cette équation sous la forme

$$u + t_1\alpha_1 + \cdots + t_n\alpha_n,$$

où $t_1, \ldots, t_n$ est une base de $\operatorname{Ker} A$ (supposé de dimension finie) et $\alpha_1, \ldots, \alpha_n$ sont des scalaires arbitraires. En particulier, si A est injective, alors toute équation $A(X) = b$ admet au plus une solution. (Elle en admet exactement une lorsque $b \in \operatorname{Im} A$, elle n'en admet pas si $b \notin \operatorname{Im} A$.)

Pour faire le lien entre le point de vue abstrait des applications linéaires et les résultats des premiers chapitres, considérons le cas particulier où l'application linéaire est associée à une matrice : soit $A \in K^{m\times n}$ et $\tilde{A}\colon K^n \to K^m$ l'application linéaire définie par

$$\tilde{A}(x) = A \cdot x.$$

Par définition,

$$\operatorname{Ker} \tilde{A} = \{x \in K^n \mid A \cdot x = 0\},$$

donc le noyau de $\tilde{A}$ est l'espace des solutions du système d'équations linéaires homogènes

$$A \cdot X = 0.$$

De même, par définition,

$$\operatorname{Im} \tilde{A} = \{b \in K^m \mid \exists x \in K^n \text{ tel que } A \cdot x = b\},$$

donc l'image de $\tilde{A}$ est l'ensemble des seconds membres $b \in K^m$ pour lesquels le système d'équations

$$A \cdot X = b$$

admet une solution. En effectuant le produit $A \cdot X$ après avoir décomposé A par colonnes, on obtient

$$A \cdot X = \begin{pmatrix} a_{*1} & \cdots & a_{*n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_{*1}x_1 + \cdots + a_{*n}x_n,$$

donc

$$\operatorname{Im} \tilde{A} = \operatorname{sev}\langle a_{*1}, \ldots, a_{*n}\rangle = \mathcal{C}(A)$$

est l'espace des colonnes de A. (Voir la Proposition 6.8, p. 40). On a donc

$$\operatorname{rang} A = \dim \operatorname{Im} \tilde{A}.$$

La Proposition 6.9 établit une relation entre la dimension de $\operatorname{Ker} \tilde{A}$ et celle de $\operatorname{Im} \tilde{A}$. Notre prochain objectif est d'obtenir une relation du même type pour des applications linéaires générales.

8.4. Rang et nullité

Le *rang* d'une application linéaire $A\colon\ E \to F$ est la dimension de l'image de A :

$$\text{rang } A = \text{dim Im } A.$$

La *nullité* de A est la dimension du noyau de A :

$$\text{null } A = \text{dim Ker } A.$$

8.4. Théorème.

$$\operatorname{rang} A + \operatorname{null} A = \dim E.$$

Démonstration. Quoique l'énoncé soit vrai également pour les espaces de dimension infinie, nous supposerons dans cette démonstration que la dimension de E est finie.

Soit $(e_1, \ldots, e_m)$ une base de Ker A, que l'on prolonge en base de E, soit $(e_1, \ldots, e_n)$. On a donc null $A = m$ et $\dim E = n$. Pour établir le théorème, on va prouver que $\big(A(e_{m+1}), \ldots, A(e_n)\big)$ est une base de Im A. Il en résultera rang $A = n - m = \dim E - \operatorname{null} A$, comme annoncé.

Montrons d'abord que la suite $\big(A(e_{m+1}), \ldots, A(e_n)\big)$ engendre $\operatorname{Im} A$. Comme cette famille est constituée d'éléments de Im A, il suffit de prouver que tout vecteur de Im A en est une combinaison linéaire. Soit $y \in \operatorname{Im} A$; on a donc $y = A(x)$ pour un certain vecteur $x \in E$. Exprimons x comme combinaison linéaire de la base de E ci-dessus :

$$x = e_1\alpha_1 + \cdots + e_n\alpha_n.$$

En appliquant A aux deux membres de cette égalité, et en tenant compte de ce que $A(e_1) = \cdots = A(e_m) = 0$, puisque $e_1, \ldots, e_m$ sont dans le noyau de A, on obtient

$$y = A(x) = A(e_{m+1})\alpha_{m+1} + \cdots + A(e_n)\alpha_n,$$

ce qui montre que tout vecteur de Im A est bien combinaison linéaire de la suite $\big(A(e_{m+1}), \ldots, A(e_n)\big)$.

Montrons enfin que cette suite est libre. Soit

$$A(e_{m+1})\beta_{m+1} + \cdots + A(e_n)\beta_n = 0$$

une relation de dépendance linéaire entre les vecteurs de la suite. Il s'agit de prouver que tous les coefficients β_i sont nuls. Vu la linéarité de A, la relation ci-dessus peut aussi s'écrire

$$A(e_{m+1}\beta_{m+1} + \cdots + e_n\beta_n) = 0;$$

elle montre alors que le vecteur $e_{m+1}\beta_{m+1} + \cdots + e_n\beta_n$ est dans le noyau de A. On peut donc exprimer ce vecteur comme combinaison linéaire de la base $(e_1, \ldots, e_m)$ de Ker A :

$$e_{m+1}\beta_{m+1} + \cdots + e_n\beta_n = e_1\gamma_1 + \cdots + e_m\gamma_m$$

pour certains scalaires $\gamma_1, \ldots, \gamma_m$. En rassemblant tous les termes dans un membre, on trouve

$$e_1(-\gamma_1) + \cdots + e_m(-\gamma_m) + e_{m+1}\beta_{m+1} + \cdots + e_n\beta_n = 0.$$

Comme $(e_1, \ldots, e_n)$ est une suite libre (c'est même une base de E), on tire de la relation précédente

$$\gamma_1 = \cdots = \gamma_m = \beta_{m+1} = \cdots = \beta_n = 0.$$

□

8.5. Corollaire. *Soit* $A\colon E \to F$ *une application linéaire. On suppose* $\dim E = n$ *et* $\dim F = m$ *(avec* m, n *finis). Les conditions* (1) *à* (3) *sont équivalentes et, de même, les conditions* (1') *à* (3') *sont équivalentes :*

(1) *A est injective*	(1') *A est surjective*
(2) $\operatorname{null} A = 0$	(2') $\operatorname{null} A = n - m$
(3) $\operatorname{rang} A = n$	(3') $\operatorname{rang} A = m$

En particulier, si $n = m$, *alors toutes les conditions ci-dessus sont équivalentes, et sont encore équivalentes à :* A *est bijective.*

Démonstration. Du théorème précédent, on déduit tout de suite la relation

$$\operatorname{rang} A + \operatorname{null} A = n,$$

qui établit l'équivalence de (2) et (3) d'une part et celle de (2') et (3') d'autre part.

D'après la Proposition 8.2, la condition (1) est équivalente à $\operatorname{Ker} A = \{0\}$, et cette dernière condition est équivalente à (2). De même, la condition (1') est équivalente à $\operatorname{Im} A = F$. Cette dernière condition est encore équivalente à $\operatorname{rang} A = \dim F$, c'est-à-dire à (3'), par la Proposition 5.2 (p. 30). □

Exercices

(8.1) (prérequis) Soient E et F deux espaces vectoriels sur un corps K et $A\colon E \to F$ une application linéaire. Prouvez que $A(0) = 0$.

(8.2) (prérequis) Soient E et F deux espaces vectoriels sur un corps K et $A : E \to F$ une application linéaire. Si $x_1, \ldots, x_n \in E$ et $\alpha_1, \ldots, \alpha_n \in K$, prouvez par induction sur $n \geqslant 1$ que $A\left(\sum_{i=1}^n x_i\alpha_i\right) = \sum_{i=1}^n A(x_i)\alpha_i$.

(8.3) Soit une application linéaire $A : \mathbb{R}^3 \to \mathbb{R}^4$ telle que $A(1,1,1) = (1,1,1,3)$, $A(1,1,0) = (1,1,1,2)$ et $A(1,0,0) = (1,1,1,1)$. Que vaut $A(3,2,1)$? Que vaut $A(x,y,z)$ pour un élément (x,y,z) quelconque de $\mathbb{R}^3$?

(8.4) Soit une application linéaire $B\colon \mathbb{R}^3 \to \mathbb{C}[X]_{\leqslant 2}$ telle que $B(1,0,0) = (1+i) + (2-i)X + iX^2$, $B(0,1,0) = i + (1+i)X + X^2$ et $B(0,0,1) = X + (1+i)X^2$. Que vaut $B(a,b,c)$ pour un élément (a,b,c) quelconque de $\mathbb{R}^3$?

(8.5) Soient E et F deux espaces vectoriels sur un corps K. Prouvez que $L(E,F)$, l'ensemble des application linéaires $E \to F$ est un sous-espace vectoriel de l'espace $\mathcal{F}(E;F)$ de toutes les fonctions $E \to F$.

(8.6) Soit une application linéaire $C\colon \mathbb{R}^2 \to L(\mathbb{R}^2, \mathbb{R})$ telle que

$$\Big(C(1,0)\Big)(1,0) = 2 \qquad \Big(C(1,0)\Big)(0,1) = 10$$
$$\Big(C(0,1)\Big)(1,1) = 6 \qquad \Big(C(0,1)\Big)(0,1) = 5.$$

Décrire l'image par C d'un élément (x,y) quelconque de $\mathbb{R}^2$.

(8.7) Démontrez que la composée de deux applications linéaires est encore une application linéaire.

(8.8) (pour aller plus loin) On considère un espace vectoriel E et une application linéaire $A : E \to E$. On suppose que pour un certain $n \in \mathbb{N}$, $A^n = 0$ tandis que $A^{n-1} \neq 0$ ($A^n = \underbrace{A \circ A \circ \ldots \circ A}_{n \text{ fois}}$).

Démontrez que $A^0 = I, A, \ldots, A^{n-1}$ sont linéairement indépendants dans l'espace $L(E,E)$.

(8.9) Vérifiez que l'opérateur de dérivation $C^i\big(]a,b[;\mathbb{R}\big) \to C^{i-1}\big(]a,b[;\mathbb{R}\big)$ est une application linéaire.

(8.10) Soient un corps K et $a_1, \ldots, a_n \in K$. Démontrez que la fonction

$$A\colon K[X] \to K^n : P \mapsto \big(P(a_1), \ldots, P(a_n)\big)$$

est une application linéaire. (Vous pouvez donner une démonstration directe, ou utiliser une généralisation de l'exercice 8.17).

(8.11) Soient E, F, G, H, quatre espaces vectoriels sur un corps K. Démontrez que si $A : E \to F$ et $B : G \to H$ sont des applications linéaires, alors l'application

$$C : E \times G \to F \times H : (e, g) \mapsto (A(e), B(g))$$

est une application linéaire (pour la définition de $\times$, voir exercice 3.5).

(8.12)

- Montrez que l'opérateur de conjugaison $\mathbb{C} \to \mathbb{C} : z \mapsto \bar{z}$ est une application linéaire si $\mathbb{C}$ est considéré comme un espace vectoriel sur $\mathbb{R}$, mais qu'elle ne l'est plus si $\mathbb{C}$ est considéré comme un espace vectoriel sur $\mathbb{C}$.
- Montrez que l'application $A : \mathbb{C}^2 \to \mathbb{C}^2 : (z_1, z_2) \mapsto (\bar{z}_1 + iz_2, \bar{z}_1 - z_2)$ est linéaire si l'on considère $\mathbb{C}^2$ comme espace vectoriel sur $\mathbb{R}$, mais qu'elle ne l'est plus si l'on considère $\mathbb{C}^2$ comme espace vectoriel sur $\mathbb{C}$.
- Si $F : \mathbb{C}^n \to \mathbb{C}^m$ est une application linéaire lorsque l'on considère $\mathbb{C}^n$ et $\mathbb{C}^m$ comme espaces vectoriels sur $\mathbb{C}$, alors F sera-t-elle encore linéaire si l'on considère $\mathbb{C}^n$ et $\mathbb{C}^m$ comme espaces vectoriels sur $\mathbb{R}$? Et si l'on considère $\mathbb{C}^n$ comme espace vectoriel sur $\mathbb{C}$ et $\mathbb{C}^m$ comme espace vectoriel sur $\mathbb{R}$?

(8.13) (prérequis) Soient A et B, deux ensembles quelconques et $f : A \to B$, une fonction ; on suppose que le domaine de f est A (tout point de A a une image). Écrire des formules mathématiques qui signifient que (1) f est injective, (2) f est surjective, (3) f est bijective, (4) f n'est pas injective, (5) f n'est pas surjective.

Qu'est-ce que l'image de f ?

(8.14) (pour aller plus loin) Soient A et B, deux ensembles quelconques, et $f : A \to B$, une fonction.

Démontrez que si f est injective alors il existe une fonction $g : B \to A$ telle que $g \circ f = \mathrm{id}_A$ (où id_A est la fonction identité $A \to A : x \mapsto x$).

Démontrez que si f est surjective alors il existe une fonction $g : B \to A$ telle que $f \circ g = \mathrm{id}_B$.

(8.15) Voici quelques fonctions entre des espaces vectoriels réels. Montrez que ces fonctions sont linéaires. Calculez-en le noyau et l'image.

- $A : \mathcal{F}(\mathbb{R}; \mathbb{R}) \to \mathbb{R} : f \mapsto f(a)$, où a est un réel fixé ;
- $B : \mathbb{R}^4 \to \mathbb{R}^2 : (x_1, x_2, x_3, x_4) \mapsto (x_1, x_3)$;
- $C : C\big([0, 1]; \mathbb{R}\big) \to \mathbb{R} : f \mapsto \displaystyle\int_0^1 f(x)dx$;
- $D : C^\infty\big([-1, 1]; \mathbb{R}\big) \to C^\infty\big([-1, 1]; \mathbb{R}\big) : f \mapsto g$, où $(\forall x \in [-1, 1])(g(x) = xf'(x))$;
- $E : \mathcal{F}(\mathbb{R}; \mathbb{R}) \to \mathcal{F}(\mathbb{R}; \mathbb{R}) : f \mapsto g$, où $(\forall x \in \mathbb{R})(g(x) = \frac{1}{2}(f(x) - f(-x)))$;
- $F : \mathbb{R}^4 \to C^\infty(\mathbb{R}; \mathbb{R}) : (x_1, x_2, x_3, x_4) \mapsto f$,
  où $(\forall t \in \mathbb{R})(f(t) = x_1 e^{t+1} + x_2 e^t t + x_3(t + 1)e^t + x_4 e^{t+2})$.

(8.16) Les applications suivantes, entre espaces vectoriels réels, sont-elles linéaires ? Lorsque c'est le cas, donnez-en le noyau, l'image et le rang.

- $A : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}[X]_{\leqslant 3} : P \mapsto Q$, où $(\forall x \in \mathbb{R})(Q(x) = x.P(x))$;
- $B : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}[X]_{\leqslant 2} : P \mapsto Q$, où $(\forall x \in \mathbb{R})(Q(x) = P(x + 1))$;
- $C : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}[X]_{\leqslant 4} : P \mapsto Q$, où $(\forall x \in \mathbb{R})(Q(x) = x^2.P(x) - P'(x))$;
- $D : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x + 1, y + 1)$;
- $E : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (3(x + 1) - 2(y + 2) + 1, (y - 3) + 3(x + 2) - 3)$;
- $F : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x^2, y^2)$;
- $G : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (e^x, e^y)$;
- $H : \Big(\mathcal{F}(\mathbb{R}; \mathbb{R})\Big)^2 \mapsto \mathcal{F}(\mathbb{R}; \mathbb{R}) : (f, g) \mapsto f.g$;
- $I : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}^3 : P \mapsto (P(1), P(2), P(3))$;

– $J : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}^3 : P \mapsto (P(1), P(1), P(0))$;
– $K = L \circ D$, où $L : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x + y - 2, 2x + y - 3)$.

(8.17) Soient E, F, G, des espaces vectoriels sur un corps K, et $A : E \to F$ et $B : E \to G$ des applications linéaires. Montrez que l'application $C : E \to F \times G : x \mapsto (A(x), B(x))$ est linéaire. Déterminez $\operatorname{Ker} C$ et $\operatorname{Im} C$.

(8.18) Soient un espace vectoriel E et deux sous-espaces vectoriels V et W. Si $E = V \oplus W$, tout $x \in E$ s'écrit de manière unique comme somme d'un vecteur $p_V(x) \in V$ et d'un vecteur $p_W(x) \in W$.

– Montrez que les applications $p_V : E \to V$ et $p_W : E \to W$ sont linéaires et surjectives.
– Montrez que $\operatorname{Im} p_V = \operatorname{Ker} p_W = V$ et que $\operatorname{Im} p_W = \operatorname{Ker} p_V = W$.

(8.19) (pour aller plus loin) Appliquez l'exercice 8.18 pour démontrer la propriété suivante (réciproque de l'exercice 8.17). Soient E, F, G, trois espaces vectoriels sur un corps K et $C : E \to F \times G$, une application linéaire. Si $x \in E$, notons $C(x) = (a_x, b_x)$. Démontrez que les fonctions $A : E \to F : x \mapsto a_x$ et $B : E \to G : x \mapsto b_x$ sont des applications linéaires.

(8.20) Avec les notations de la proposition 8.3, page 55, l'ensemble $u + \operatorname{Ker} A$ est-il un sous-espace vectoriel de E ?

(8.21) Donnez une autre preuve pour l'exercice 6.4, page 42, en considérant A et B comme des matrices représentant des applications linéaires.

(8.22) Soient deux espaces vectoriels E et F sur un corps K, et une application linéaire $A : E \to F$. Démontrez les affirmations suivantes :

– A est injective ssi l'image par A de n'importe quelle suite libre est encore une suite libre.
– A est surjective ssi l'image par A de n'importe quelle suite génératrice de E est une suite génératrice de F.

9. Construction d'applications linéaires

9.1. Principe de construction

Soient E et F deux espaces vectoriels sur un même corps K. On suppose que E est de dimension finie.

9.1. THÉORÈME. *Soit $(e_1, \ldots, e_n)$ une base de E et soit $(v_1, \ldots, v_n)$ une suite quelconque de vecteurs de F. Il existe une et une seule application linéaire $A\colon E \to F$ telle que $A(e_i) = v_i$ pour tout $i = 1, \ldots, n$. Cette application satisfait*

$$\operatorname{Im} A = \operatorname{sev}\langle v_1, \ldots, v_n\rangle$$

et

$$\operatorname{Ker} A = \{e_1x_1 + \cdots + e_nx_n \mid v_1x_1 + \cdots + v_nx_n = 0\}.$$

De plus, l'application A est injective si et seulement si $(v_1, \ldots, v_n)$ est une suite libre ; elle est surjective si et seulement si $(v_1, \ldots, v_n)$ est une suite génératrice et, par conséquent, elle est bijective si et seulement si la suite $(v_1, \ldots, v_n)$ est une base de F.

DÉMONSTRATION. *Existence de A :* Tout $x \in E$ s'écrit d'une et d'une seule manière comme combinaison linéaire de la base $(e_1, \ldots, e_n)$; soit

$$x = e_1x_1 + \cdots + e_nx_n.$$

On définit une application $A\colon E \to F$ en posant

$$A(x) = v_1x_1 + \cdots + v_nx_n.$$

Il est clair que pour l'application A ainsi définie, on a bien $A(e_i) = v_i$ pour tout $i = 1, \ldots, n$. Il suffit donc de prouver que A est linéaire. Soient

$$x = e_1x_1 + \cdots + e_nx_n \quad \text{et} \quad y = e_1y_1 + \cdots + e_ny_n$$

des vecteurs de E. On a alors, par définition de A,

$$A(x) = v_1x_1 + \cdots + v_nx_n \quad \text{et} \quad A(y) = v_1y_1 + \cdots + v_ny_n$$

et, par ailleurs, pour $\alpha, \beta \in K$,

$$x\alpha + y\beta = e_1(x_1\alpha + y_1\beta) + \cdots + e_n(x_n\alpha + y_n\beta),$$

d'où, toujours par définition de A,

$$A(x\alpha + y\beta) = v_1(x_1\alpha + y_1\beta) + \cdots + v_n(x_n\alpha + y_n\beta).$$

Comme

$$\begin{aligned} v_1(x_1\alpha + y_1\beta) + \cdots + v_n(x_n\alpha + y_n\beta) &= (v_1x_1 + \cdots + v_nx_n)\alpha + (v_1y_1 + \cdots + v_ny_n)\beta \\ &= A(x)\alpha + A(y)\beta, \end{aligned}$$

on a bien $A(x\alpha+y\beta) = A(x)\alpha+A(y)\beta$, comme annoncé. Les descriptions de $\operatorname{Im} A$ et de $\operatorname{Ker} A$ résultent directement de la définition de A.

Unicité de A : Si $B \colon E \to F$ est une application linéaire telle que $B(e_i) = v_i$ pour tout $i = 1, \ldots, n$, alors pour tout $x = e_1x_1 + \cdots + e_nx_n \in E$ on a

$$B(x) = B(e_1)x_1 + \cdots + B(e_n)x_n = v_1x_1 + \cdots + v_nx_n.$$

En comparant avec la définition de l'application A ci-dessus, on voit que $B(x) = A(x)$ pour tout $x \in E$, donc $B = A$.

Injectivité de A : D'après la description de $\operatorname{Ker} A$, on voit tout de suite que A est injective si la suite $(v_1, \ldots, v_n)$ est libre, puisqu'alors la relation $v_1x_1 + \cdots + v_nx_n = 0$ entraîne $x_1 = \cdots = x_n = 0$. Inversement, si la suite $(v_1, \ldots, v_n)$ n'est pas libre, soit

$$v_1\alpha_1 + \cdots + v_n\alpha_n = 0,$$

les scalaires $\alpha_1, \ldots, \alpha_n$ n'étant pas tous nuls ; alors $e_1\alpha_1 + \cdots + e_n\alpha_n$ est un vecteur non nul du noyau de A, donc A n'est pas injective.

Surjectivité de A : Comme $\operatorname{Im} A = \operatorname{sev}\langle v_1, \ldots, v_n\rangle$, on a $\operatorname{Im} A = F$ si et seulement si $(v_1, \ldots, v_n)$ est une suite génératrice de F. □

9.2. Inversibilité des applications linéaires

Le principe de construction de la section précédente permet en particulier de construire des inverses à gauche ou à droite de certaines applications linéaires.

Une application linéaire $A \colon E \to F$ est *inversible à gauche* (resp. *inversible à droite*) s'il existe une application linéaire $B \colon F \to E$, appelée *inverse de A à gauche* (resp. *inverse de A à droite*), telle que $B \circ A = I_E$ (resp. $A \circ B = I_F$). L'application A est *inversible* s'il existe une application linéaire $B \colon F \to E$, appelée *inverse de A*, telle que $B \circ A = I_E$ et $A \circ B = I_F$.

Bien entendu, si une application linéaire est inversible, elle est à la fois inversible à gauche et inversible à droite. Réciproquement, si une application linéaire A est inversible à gauche et à droite, alors elle est inversible, et elle admet un unique inverse à gauche, qui est aussi l'unique inverse à droite et donc l'unique inverse. En effet, si l'on a $B \circ A = I_E$ et $A \circ B' = I_F$, alors on peut calculer la composée $B \circ A \circ B'$ de deux manières différentes :

$$(B \circ A) \circ B' = I_E \circ B' = B' \quad \text{et} \quad B \circ (A \circ B') = B \circ I_F = B$$

donc $B = B'$. L'inverse d'une application A (inversible !) est notée A^{-1}.

9.2. Proposition. *Pour qu'une application linéaire soit inversible à gauche (resp. à droite), il faut et il suffit qu'elle soit injective (resp. surjective) ; pour qu'une application linéaire soit inversible, il faut et il suffit qu'elle soit bijective.*

Démonstration. Pour la démonstration, nous supposerons que les espaces de départ et d'arrivée sont de dimension finie, quoique l'énoncé soit vrai sans cette hypothèse. Soit $A \colon E \to F$ une application linéaire, soit $(e_1, \ldots, e_n)$ une base de E et soit $v_i = A(e_i)$ pour tout $i = 1, \ldots, n$.

Si A est inversible à gauche, soit B un inverse à gauche de A. Si $x, x' \in E$ sont tels que $A(x) = A(x')$, alors en appliquant B aux deux membres, on trouve $B \circ A(x) = B \circ A(x')$ et comme $B \circ A = I_E$ on en déduit $x = x'$, donc A est injective.

Réciproquement, si A est injective, alors d'après le théorème précédent la suite $(v_1, \ldots, v_n)$ est libre. Soit $(v_1, \ldots, v_m)$ $(m \geqslant n)$ une base de F qui prolonge cette suite libre. Le Théorème 9.1 montre qu'il existe une application linéaire $B\colon\ F \to E$ telle que $B(v_i) = e_i$ pour $i = 1, \ldots, n$ et $B(v_i) = 0$ pour $i = n+1, \ldots, m$. On a alors

$$B \circ A(e_i) = e_i = I_E(e_i) \quad \text{pour tout } i = 1, \ldots, n.$$

Comme une application linéaire est déterminée de manière unique par les images des vecteurs d'une base, d'après le Théorème 9.1, on a donc $B \circ A = I_E$.

Si A est inversible à droite, soit B un inverse de A à droite. Pour tout $y \in F$, on peut trouver $x \in E$ tel que $A(x) = y$: il suffit en effet de choisir $x = B(y)$, car alors

$$A(x) = A \circ B(y) = I_F(y) = y;$$

cela montre que A est surjective.

Réciproquement, si A est surjective, alors le Théorème 9.1 montre que la suite $(v_1, \ldots, v_n)$ est génératrice ; on peut donc en extraire une base de F. Pour la commodité des notations, supposons qu'une telle base est formée des premiers vecteurs de la suite, soit $(v_1, \ldots, v_m)$ (avec $m \leqslant n$). Le Théorème 9.1 montre qu'il existe une application linéaire $B\colon\ F \to E$ telle que $B(v_i) = e_i$ pour tout $i = 1, \ldots, m$. On a alors

$$A \circ B(v_i) = v_i = I_F(v_i) \quad \text{pour tout } i = 1, \ldots, m,$$

d'où on conclut comme précédemment que $A \circ B = I_F$. □

Remarquons que la construction des applications inverses à gauche ou à droite ci-dessus fait intervenir le choix d'une base de F qui est extraite ou qui prolonge la suite $(v_1, \ldots, v_n)$. Ce choix n'étant généralement pas unique, l'inverse à gauche ou à droite d'une application linéaire n'est pas unique, en général.

Nous pouvons à présent compléter le Corollaire 8.5 (p. 57) de la manière suivante :

9.3. Corollaire. *Soit $A\colon\ E \to F$ une application linéaire. On suppose* $\dim E = n$ *et* $\dim F = m$ *(avec m, n finis). Les conditions* (1) *à* (4) *sont équivalentes et, de même, les conditions* (1') *à* (4') *sont équivalentes :*

(1) *A est injective*	(1') *A est surjective*
(2) *A est inversible à gauche*	(2') *A est inversible à droite*
(3) $\operatorname{null} A = 0$	(3') $\operatorname{null} A = n - m$
(4) $\operatorname{rang} A = n$	(4') $\operatorname{rang} A = m$

En particulier, si $n = m$, alors toutes les conditions ci-dessus sont équivalentes entre elles, et sont encore équivalentes à : A est inversible et à : A est bijective.

9.4. Corollaire. *Soient E, F deux espaces vectoriels de dimension finie. S'il existe un isomorphisme*

$$A\colon\ E \to F,$$

alors $\dim E = \dim F$. *Réciproquement, si* $\dim E = \dim F$, *alors il existe un isomorphisme $A\colon\ E \to F$.*

Démonstration. Soit $\dim E = n$ et $\dim F = m$, comme dans l'énoncé précédent. S'il existe un isomorphisme $A\colon\ E \to F$, alors de l'injectivité de A il découle : $\operatorname{rang} A = n$ et de la surjectivité de A on tire $\operatorname{rang} A = m$, donc $n = m$.

Réciproquement, si $n = m$, soit $e = (e_1, \ldots, e_n)$ et $f = (f_1, \ldots, f_n)$ des bases de E et F respectivement. Le principe de construction (Théorème 9.1) montre l'existence d'une application $A\colon\ E \to F$ telle que $A(e_i) = f_i$ pour tout $i = 1, \ldots, n$. Comme f est une base, le même théorème établit la bijectivité de A, donc A est un isomorphisme. □

9.3. Représentation matricielle des applications linéaires

Soient E et F deux espaces vectoriels de dimension finie sur un même corps K. Soit $e = (e_1, \dots, e_n)$ une base de E et $f = (f_1, \dots, f_m)$ une base de F. À toute application linéaire $A\colon E \to F$, on associe une matrice ${}_f(A)_e \in K^{m\times n}$ par

$$ {}_f(A)_e \in K^{m\times n} = (a_{i,j})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} \qquad \text{si } A(e_j) = \sum_{i=1}^m f_i a_{i,j}. $$

La j-ème colonne de la matrice ${}_f(A)_e$ est donc formée des coordonnées par rapport à la base f de $A(e_j)$, l'image par A du j-ème vecteur de la base e. Comme une application linéaire est déterminée de manière unique par les images des vecteurs d'une base, l'application linéaire A est déterminée de manière unique par la matrice ${}_f(A)_e$ (les bases e et f étant fixées).

Cette matrice permet de calculer les coordonnées de l'image d'un vecteur quelconque de E. En effet, si $x = \sum_{j=1}^n e_j x_j$, alors en appliquant A on obtient $A(x) = \sum_{j=1}^n A(e_j)x_j$, d'où, en substituant $A(e_j) = \sum_{i=1}^m f_i a_{i,j}$,

$$ A(x) = \sum_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} f_i a_{i,j} x_j $$

ou encore

$$ A(x) = \sum_{i=1}^m f_i \left(\sum_{j=1}^n a_{i,j} x_j \right). $$

La i-ème coordonnée de $A(x)$ par rapport à la base f est donc $\sum_{j=1}^n a_{i,j}x_j$. En posant $y = A(x) \in F$, on a donc

$$ y_i = \sum_{j=1}^n a_{i,j} x_j \quad \text{pour } i = 1, \dots, m. $$

Cette formule peut aussi s'exprimer par un produit matriciel :

$$ \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. $$

En désignant par ${}_e(x)$ (resp. ${}_f(y)$) la matrice-colonne des coordonnées de x par rapport à la base e (resp. de y par rapport à la base f), on a donc

$$ \boxed{{}_f(y) = {}_f(A)_e \cdot {}_e(x) \quad \text{où } \ y = A(x).} $$

9.5. Proposition. 1. *Pour tout scalaire α,*

$$ {}_f(\alpha A)_e = \alpha\, {}_f(A)_e. $$

2. *Si A et B sont deux applications linéaires de E vers F,*

$$ {}_f(A+B)_e = {}_f(A)_e + {}_f(B)_e. $$

3. *Soit $g = (g_1, \dots, g_\ell)$ une base d'un troisième espace vectoriel G. Si $A\colon E \to F$ et $B\colon F \to G$ sont des applications linéaires,*

$$ {}_g(B \circ A)_e = {}_g(B)_f \cdot {}_f(A)_e. $$

Démonstration. Vérifions seulement la dernière propriété (les deux autres se vérifient de manière semblable, quoique plus simple). Soient

$$ {}_f(A)_e = (a_{i,j})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} \qquad \text{et} \quad {}_g(B)_f = (b_{i,j})_{\substack{1\leqslant i\leqslant \ell\\ 1\leqslant j\leqslant m}} $$

les matrices de A et de B respectivement, de sorte que

$$ (*) \qquad A(e_j) = \sum_{i=1}^m f_i a_{i,j} \quad \text{et} \quad B(f_j) = \sum_{i=1}^\ell g_i b_{i,j}. $$

En appliquant B aux deux membres de la première égalité, et en remplaçant l'indice de sommation i par k, on trouve

$$B \circ A(e_j) = \sum_{k=1}^{m} B(f_k) a_{k,j}.$$

En substituant dans le second membre l'expression de $B(f_k)$ que donne la deuxième égalité $(*)$, on obtient

$$B \circ A(e_j) = \sum_{\substack{1 \leqslant i \leqslant \ell \\ 1 \leqslant k \leqslant m}} g_i b_{i,k} a_{k,j} = \sum_{i=1}^{\ell} g_i \left(\sum_{k=1}^{m} b_{i,k} a_{k,j} \right).$$

Cette dernière égalité montre que la matrice de $B \circ A$ par rapport aux bases e et g est

$${}_g(B \circ A)_e = \left(\sum_{k=1}^{m} b_{i,k} a_{k,j} \right).$$

On reconnaît dans le deuxième membre l'expression de la matrice produit des matrices ${}_g(B)_f = (b_{i,j})$ et ${}_f(A)_e = (a_{i,j})$, et la démonstration est donc achevée. □

Variante : On peut aussi démontrer la troisième propriété ci-dessus en utilisant la multiplication par blocs : comme les colonnes de la matrice ${}_f(A)_e$ sont les coordonnées de $A(e_1), \ldots, A(e_n)$ par rapport à la base f, la décomposition de ${}_f(A)_e$ par colonnes est

$${}_f(A)_e = \begin{pmatrix} {}_f(A(e_1)) & \cdots & {}_f(A(e_n)) \end{pmatrix}.$$

Dès lors, le produit de ${}_g(B)_f$ et ${}_f(A)_e$ peut se calculer par blocs :

$${}_g(B)_f \cdot {}_f(A)_e = \begin{pmatrix} {}_g(B)_f \cdot {}_f(A(e_1)) & \cdots & {}_g(B)_f \cdot {}_f(A(e_n)) \end{pmatrix}.$$

Or, la formule encadrée ci-dessus montre que pour tout $i = 1, \ldots, n$,

$${}_g(B)_f \cdot {}_f(A(e_i)) = {}_g(B \circ A(e_i)).$$

Dès lors, les colonnes de la matrice produit ${}_g(B)_f \cdot {}_f(A)_e$ sont les coordonnées de $B \circ A(e_1), \ldots, B \circ A(e_n)$ par rapport à la base g ; cette matrice est donc ${}_g(B \circ A)_e$. □

Montrons pour terminer que le calcul du rang (ou, ce qui revient au même, de la nullité) d'une application linéaire peut se faire par l'intermédiaire de la matrice de l'application par rapport à n'importe quelles bases :

9.6. Proposition. *Soient e et f des bases d'espaces vectoriels E et F respectivement. Pour toute application linéaire $A\colon E \to F$,*

$$\operatorname{rang} A = \operatorname{rang} {}_f(A)_e.$$

Démonstration. Comme $e = (e_1, \ldots, e_n)$ est une base de E, les éléments de l'image de A sont de la forme

$$A(x) = A(e_1 x_1 + \cdots + e_n x_n) = A(e_1)x_1 + \cdots + A(e_n)x_n,$$

donc

$$\operatorname{Im} A = \operatorname{sev} \langle A(e_1), \ldots, A(e_n) \rangle.$$

(Cela résulte aussi du Théorème 9.1). Soit ${}_f\gamma\colon F \to K^m$ l'isomorphisme qui envoie tout vecteur de F sur ses coordonnées par rapport à la base f. On a

$${}_f\gamma(\operatorname{Im} A) = \operatorname{sev} \langle {}_f(A(e_1)), \ldots, {}_f(A(e_n)) \rangle;$$

or, les coordonnées de $A(e_1), \ldots, A(e_n)$ par rapport à la base f sont les colonnes de ${}_f(A)_e$, donc

$${}_f\gamma(\operatorname{Im} A) = \mathcal{C}\big({}_f(A)_e\big).$$

Dès lors, ${}_f\gamma$ établit un isomorphisme entre $\operatorname{Im} A$ et l'espace des colonnes de ${}_f(A)_e$. Les dimensions de ces deux espaces sont donc égales, d'après le Corollaire 9.4 :

$$\dim \operatorname{Im} A = \dim \mathcal{C}\big({}_f(A)_e\big),$$

ce qui établit l'égalité du rang de A et du rang de la matrice ${}_f(A)_e$ (vu comme dimension de l'espace engendré par les colonnes). □

9.4. Changements de base

Il est important de remarquer que ce qui précède s'applique également dans le cas où $F = E$, et où $A = I_E$ est l'identité sur E. Si les bases e et f sont les mêmes, alors un simple calcul montre que l'on obtient la matrice unité :

$$_e(I_E)_e = I_n;$$

cependant, si $f \neq e$, alors la matrice $_f(I_E)_e$ n'est pas la matrice unité ; c'est la matrice dont la j-ème colonne est formée des coordonnées du vecteur e_j $(= I_E(e_j))$ par rapport à la base f. Cette matrice permet de contrôler les changements de base, de la manière suivante : si $_f(x)$ et $_e(x)$ désignent les matrices-colonnes des coordonnées d'un même vecteur $x \in E$ par rapport aux bases f et e respectivement, alors la formule encadrée plus haut (p. 62) donne (dans le cas particulier où $A = I_E$)

$$\boxed{_f(x) = {_f(I_E)_e} \cdot {_e(x)} \quad \text{pour tout } x \in E.}$$

Par ailleurs, si $A\colon E \to F$ est une application linéaire et si e, e' désignent deux bases de E et f, f' deux bases de F, alors les matrices de A par rapport aux bases e et f d'une part, e' et f' d'autre part sont liées par la formule

$$\boxed{_{f'}(A)_{e'} = {_{f'}(I_F)_f} \cdot {_f(A)_e} \cdot {_e(I_E)_{e'}}.}$$

Cette formule se déduit en appliquant la Proposition 9.5(3) au cas où l'application composée est $A = I_F \circ A \circ I_E$.

Remarquons pour terminer que toute matrice de changement de base $_e(I_E)_{e'}$ est régulière :

9.7. Proposition. *Soient e et e' deux bases d'un espace vectoriel E de dimension n (finie). La matrice de changement de base $_e(I_E)_{e'}$ est régulière, et*

$$_e(I_E)_{e'}^{-1} = {_{e'}(I_E)_e}.$$

De plus, pour toute matrice régulière $P \in K^{n\times n}$ il existe des bases f et f' de E telles que

$$P = {_e(I_E)_f} \qquad \text{et} \qquad P = {_{f'}(I_E)_{e'}}.$$

Démonstration. La Proposition 9.5(3) donne

$$_e(I_E)_{e'} \cdot {_{e'}(I_E)_e} = {_e(I_E)_e} = I_n \quad \text{et} \quad {_{e'}(I_E)_e} \cdot {_e(I_E)_{e'}} = {_{e'}(I_E)_{e'}} = I_n.$$

Cela établit la première partie de l'énoncé.

Soit $P = (p_{ij})_{1\leqslant i,j\leqslant n} \in K^{n\times n}$ une matrice régulière. On définit une suite $f = (f_1, \dots, f_n)$ par

$$(\dagger) \qquad f_j = \sum_{i=1}^n e_i p_{ij} \qquad \text{pour } j = 1, \dots, n.$$

Montrons que la suite f est libre : si $f_1\alpha_1 + \cdots + f_n\alpha_n = 0$ pour certains scalaires $\alpha_1, \dots, \alpha_n \in K$, alors

$$\sum_{j=1}^n \left(\sum_{i=1}^n e_i p_{ij}\right)\alpha_j = 0,$$

ce que l'on peut réécrire

$$\sum_{i=1}^n e_i \left(\sum_{j=1}^n p_{ij}\alpha_j\right) = 0.$$

Comme la suite e est libre (puisque c'est une base de E), cette relation entraîne

$$\sum_{j=1}^n p_{ij}\alpha_j = 0 \qquad \text{pour tout } i = 1, \dots, n,$$

donc

$$P \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = 0.$$

Comme P est régulière, on en déduit $\alpha_1 = \cdots = \alpha_n = 0$. La suite f est donc libre ; comme $\dim E = n$, c'est une base de E. Par ailleurs, la relation (†) montre que l'on a

$$P = {}_e(I_E)_f.$$

De la même manière, on peut trouver une base f' de E telle que

$$P^{-1} = {}_{e'}(I_E)_{f'}.$$

En inversant les deux membres, on obtient

$$P = {}_{f'}(I_E)_{e'}.$$

□

Déterminant d'un opérateur linéaire. Les formules de changement de base rendent possible la définition du *déterminant d'un opérateur linéaire* sur un espace de dimension finie.

Soit $A\colon E \to E$ un opérateur linéaire sur un espace de dimension finie et soient e, e' deux bases de E. D'après la formule de changement de base et la proposition précédente on a alors

$$\begin{aligned} {}_{e'}(A)_{e'} &= {}_{e'}(I_E)_e \cdot {}_e(A)_e \cdot {}_e(I_E)_{e'} \\ &= {}_{e'}(I_E)_e \cdot {}_e(A)_e \cdot {}_{e'}(I_E)_e^{-1}. \end{aligned}$$

D'après le corollaire 7.4, on a donc

$$\text{dét}\, {}_{e'}(A)_{e'} = \text{dét}\, {}_e(A)_e,$$

ce qui montre que le déterminant de la matrice de A par rapport à une base de E ne dépend pas du choix de la base. On peut donc poser

$$\text{dét}\, A = \text{dét}\, {}_e(A)_e,$$

où e est une base quelconque de E.

9.8. Proposition. *Soit E un espace vectoriel de dimension finie.*

1. *Pour $A, B\colon E \to E$ deux opérateurs linéaires sur E,*

$$\text{dét}\, A \circ B = \text{dét}\, A\ \text{dét}\, B.$$

2.

$$\text{dét}\, I_E = 1.$$

3. *Un opérateur linéaire $A\colon E \to E$ est inversible si et seulement si $\text{dét}\, A \neq 0$.*

Démonstration. Soit $e = (e_1, \dots , e_n)$ une base de E.
1. Comme la composée des applications linéaires correspond au produit des matrices qui les représentent (Proposition 9.5(3)), on a

$${}_e(A \circ B)_e = {}_e(A)_e \cdot {}_e(B)_e.$$

Vu la multiplicativité du déterminant, la propriété s'en déduit en prenant le déterminant des deux membres.
2. Cela résulte des égalités ${}_e(I_E)_e = I_n$ et $\text{dét}\, I_n = 1$.
3. D'après le Corollaire 9.3, l'opérateur A est bijectif si et seulement si $\text{rang}\, A = n$. Or, $\text{rang}\, A = \text{rang}\, {}_e(A)_e$ par la Proposition 9.6, et la Proposition 7.3 (p. 46) montre que $\text{rang}\, {}_e(A)_e = n$ si et seulement si $\text{dét}\, {}_e(A)_e \neq 0$. □

Exercices

(9.1) Avec les notations du théorème 9.1, page 59, que peut-on dire, d'une part, de l'existence de A et, d'autre part, de son unicité, si

- la suite $(e_1, \dots, e_n)$ est libre mais peut-être pas génératrice ;
- la suite $(e_1, \dots, e_n)$ est génératrice mais peut-être pas libre.

(9.2) Soit $L : E \to F$ une application linéaire entre deux espaces vectoriels E et F. Peut-on déterminer l'image de n'importe quel $x \in E$ si l'on connaît

1. $L(e_1), \dots, L(e_n)$, où $(e_1, \dots, e_n)$ est une base ?
2. $L(e_1), \dots, L(e_n)$, où $(e_1, \dots, e_n)$ est une suite génératrice ?
3. $L(e_1), \dots, L(e_n)$, où $(e_1, \dots, e_n)$ est une suite libre ?

(9.3)

- Existe-t-il une application linéaire $A : \mathbb{R}^3 \to \mathbb{R}^4$ telle que $A(1,1,3) = (1,1,1,1)$, $A(1,1,2) = (1,1,1,0)$ et $A(1,1,1) = (1,1,0,0)$?
- Existe-t-il une application linéaire $A : \mathbb{R}^3 \to \mathbb{R}^4$ telle que $A(1,1,3) = (1,1,1,1)$ et $A(1,1,2) = (1,1,1,0)$?
- Existe-t-il une application linéaire $A : \mathbb{R}^3 \to \mathbb{R}^4$ telle que $A(1,1,3) = (1,1,1,1)$, $A(1,1,2) = (1,1,1,0)$ et $A(1,0,0) = (1,1,0,0)$?

(9.4) $L(\mathbb{R}^2, \mathbb{R}^3)$, l'ensemble des applications linéaires $\mathbb{R}^2 \to \mathbb{R}^3$ est un espace vectoriel. Donnez-en une base.

(9.5) Dans la deuxième partie de la démonstration de la proposition 9.2, aurait-on pu choisir une autre valeur que 0 pour les $B(v_i)$ tels que $i \in \{n+1, \dots, m\}$?

(9.6) Les applications A, B et C des exercices 8.3, 8.4 et 8.6 (page 57) sont-elles inversibles à gauche ? Sont-elle inversibles à droite ? Définissez explicitement de tels inverses, lorsqu'ils existent.

(9.7) (pour aller plus loin) Soient deux espaces vectoriels E et F de dimension finie sur un corps K, et une application linéaire $A : E \to F$. Démontrez les affirmations suivantes :

- Si $\dim E > 0$ et que A possède un unique inverse à gauche, alors elle est bijective.
- Si $\dim F > 0$ et que A possède un unique inverse à droite, alors elle est bijective.

(9.8) (prérequis) Soient E et F deux espaces vectoriels de dimension finie sur un même corps K, e une base de E et f une base de F. Si $A : E \to F$ est une application linéaire, comment construit-on la matrice ${}_f(A)_e$ qui représente A dans les bases e et f ? Pourquoi dit-on que ${}_f(A)_e$ représente A ? Justifiez la construction de ${}_f(A)_e$.

(9.9) Soit $A : \mathbb{R}^3 \to \mathbb{R}[X]_{\leqslant 1}$, l'application linéaire définie par sa matrice

$$
{}_u(A)_c = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \end{pmatrix}
$$

par rapport à la base usuelle u de $\mathbb{R}[X]_{\leqslant 1}$ et à la base canonique c de $\mathbb{R}^3$.

Que vaut $A(x, y, z)$, pour un élément quelconque $(x, y, z) \in \mathbb{R}^3$?

(9.10) Soit $A : \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}^2$, l'application linéaire définie par sa matrice

$$
{}_f(A)_e = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \end{pmatrix}
$$

par rapport à la base $e = (1, 1+X, X+X^2)$ de $\mathbb{R}[X]_{\leqslant 2}$ et à la base $f = ((2,0),(0,-1))$ de $\mathbb{R}^2$.

Que vaut $A(a_0 + a_1X + a_2X^2)$, pour un élément quelconque $a_0 + a_1X + a_2X^2 \in \mathbb{R}[X]_{\leqslant 2}$?

(9.11) Soit $A : \mathbb{C}^3 \to \mathbb{C}^3$, l'application linéaire définie par sa matrice

$$_e(A)_f = \begin{pmatrix} 2i & i & i \\ -i & 0 & 0 \\ i & 1 & 0 \end{pmatrix}$$

par rapport aux bases $e = \Big((1,0,0),(1,1,0),(1,1,1)\Big)$ et $f = \Big((i,1,1),(1,1,0),(i,0,0)\Big)$ de $\mathbb{C}^3$ (comme espace vectoriel complexe).

Que vaut $A(z_1, z_2, z_3)$, pour un élément quelconque $(z_1, z_2, z_3) \in \mathbb{C}^3$?

(9.12) Voici quelques applications linéaires entre des espaces vectoriels réels. Donnez-en les représentations matricielles par rapport aux bases usuelles de ces espaces vectoriels.

- $A\colon\ \mathbb{R}^3 \to \mathbb{R}^2 : (x,y,z) \mapsto (3x - 2y + z, x + 3y)$;
- $B\colon\ \mathbb{R}^2 \to \mathbb{R}^3 : (x,y) \mapsto (x - y, x + y, 2x)$;
- $C\colon\ \mathbb{R}^3 \to \mathbb{R} : (x,y,z) \mapsto x - y + z$;
- $D\colon\ \mathbb{R} \to \mathbb{R}^3 : x \mapsto (3x, x, -2x)$;
- $E\colon\ \mathbb{C} \to \mathbb{C} : z \mapsto \bar{z}$;
- $F\colon\ \mathbb{R}[X]_{\leqslant 4} \to \mathbb{R}[X]_{\leqslant 3} : P \mapsto P'$;
- $G\colon\ \mathbb{R}[X]_{\leqslant 4} \to \mathbb{R}[X]_{\leqslant 4} : P \mapsto P'$;
- $H\colon\ \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a+b & b+c \\ c+d & d+a \end{pmatrix}$;
- $I\colon\ \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}[X]_{\leqslant 4} : P \mapsto X^2.P - P'$;
- $J\colon\ \mathbb{R}[X]_{\leqslant 3} \to \mathbb{R} : P \mapsto \int_0^2 P(x)dx$;
- $K\colon\ \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}[X]_{\leqslant 2} : P \mapsto Q$, où $(\forall x \in \mathbb{R})(Q(x) = P(x+1))$.
- $L\colon\ \mathbb{R}[X]_{\leqslant 2} \to \mathbb{R}^3 : P \mapsto (P(a_1), P(a_2), P(a_3))$, où a_1, a_2 et a_3 sont trois éléments quelconques de $\mathbb{R}$.

(9.13) Soit E un espace vectoriel réel de dimension n, e une base de E, et c la base canonique de $\mathbb{R}^n$. Quelle est la matrice $_c({_e\gamma})_e$ qui représente $_e\gamma\colon\ E \to \mathbb{R}^n : x \mapsto {_e(x)}$ par rapport à la base e et à la base c ?

(9.14) Donnez la matrice qui représente, dans la base canonique de $\mathbb{R}^3$, l'application $B \circ A$, où A et B sont définies dans l'exercice 9.12. Calculez cette matrice de deux manières différentes.

(9.15) Soit $A\colon\ \mathbb{C}^2 \to \mathbb{C}^2 : (z_1, z_2) \mapsto (z_1 + iz_2, z_1 - iz_2)$.

- Calculez $_c(A)_c$, où c est la base canonique de $\mathbb{C}^2$, considéré comme espace vectoriel complexe.
- Calculez $_u(A)_u$, où u est la base usuelle de $\mathbb{C}^2$, considéré comme espace vectoriel réel.

(9.16) Soit $A\colon\ \mathbb{R}^{2\times 2} \to \mathbb{R}^2$, l'application linéaire définie par

$$A\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) = (2,3) \qquad A\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = (1,4)$$

$$A\left(\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\right) = (0,0) \qquad A\left(\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}\right) = (2,2).$$

- Calculez $_c(A)_u$, où c est la base canonique de $\mathbb{R}^2$ et u la base usuelle de $\mathbb{R}^{2\times 2}$.
- Que vaut $A(M)$, où $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est une matrice quelconque de $\mathbb{R}^{2\times 2}$?

(9.17) Soit E un espace vectoriel et $e = (e_1, e_2, e_3, e_4)$ une base de E. On considère l'application linéaire $A\colon\ E \to E$ définie par

$$A(e_1) = e_1 + e_2, \quad A(e_2) = 2e_2 + e_3, \quad A(e_3) = 3e_3 + e_4, \quad A(e_4) = 4e_4.$$

- Montrez que $A(e) = \Big(A(e_1), A(e_2), A(e_3), A(e_4)\Big)$ est une base de E.
- Calculez $_e(A)_e$ et $_{A(e)}(A)_{A(e)}$. Pourquoi ces deux matrices sont-elles identiques ?

(9.18) (pour aller plus loin) Soit $A\colon \mathbb{R}[X]_{\leqslant 3} \to \mathbb{R}^{2\times 2}$, l'application linéaire définie par

$$A(X^2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A(X^3) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\operatorname{Ker} A = \{a_0 + a_1X + a_2X^2 + a_3X^3 \mid a_0 - a_1 - a_2 = 0 \text{ et } a_0 - a_3 = 0\}.$$

Calculez ${}_u(A)_{u'}$, où u est la base usuelle de $\mathbb{R}^{2\times 2}$ et u' celle de $\mathbb{R}[X]_{\leqslant 3}$.

(9.19) (pour aller plus loin) Soit E un espace vectoriel de dimension finie et $A\colon E \to E$ une application linéaire telle que $\operatorname{Im} f \subseteq \operatorname{Ker} f$.

- Montrez que $\operatorname{rang} f \leqslant \frac{1}{2}\dim E$.
- Montrez qu'il existe une base e de E telle que ${}_e(A)_e$ soit une matrice de la forme

$$\begin{pmatrix} 0 & \dots & 0 & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{lk} & \dots & a_{ln} \\ 0 & & & \dots & & 0 \\ \vdots & & & & & \vdots \\ 0 & & & \dots & & 0 \end{pmatrix}.$$

(9.20) Dans $\mathbb{R}^2$, on considère la base canonique c et la base $e = \big((1,1),(1,-1)\big)$. Calculez la matrice de changement de coordonnées ${}_e(I_{\mathbb{R}^2})_c$. Calculez ${}_e(A)_e$, où $A\colon \mathbb{R}^2 \to \mathbb{R}^2 : (x,y) \mapsto (4x-2y, -2x+4y)$.

(9.21) Dans $\mathbb{R}^3$, on considère la base canonique c, ainsi que les bases f et g suivantes :

$$f = \big((1,1,1),(1,1,0),(1,0,0)\big)$$
$$g = \big((1,-1,0),(2,-1,-2),(1,-1,-2)\big).$$

- Calculez les matrices de changements de coordonnées ${}_f(I_{\mathbb{R}^3})_c$ et ${}_f(I_{\mathbb{R}^3})_g$.
- Si $A\colon \mathbb{R}^3 \to \mathbb{R}^3 : (x,y,z) \mapsto (3x-2y+z, x+3y, x-y+z)$, calculez ${}_c(A)_f$ et ${}_f(A)_g$.

(9.22) Soient f et g les deux bases suivantes de $\mathbb{C}^{2\times 2}$, considéré comme espace vectoriel complexe :

$$f = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right)$$
$$g = \left(\begin{pmatrix} i & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Calculez la matrice de changement de coordonnées ${}_f(I_{\mathbb{C}^{2\times 2}})_g$.

CHAPITRE V

Espaces euclidiens

Dans certains espaces vectoriels, tels que l'espace usuel, il est possible de calculer la longueur d'un vecteur ou l'angle entre deux vecteurs au moyen d'un produit scalaire. Cette notion, déjà abordée dans le cours d'analyse MATH1140[1], fait l'objet du présent chapitre. Insistons cependant sur le fait que la notion centrale est ici celle de produit scalaire, alors que dans le cours d'analyse la notion fondamentale est celle de norme, et qu'il existe sur des espaces vectoriels des normes qui ne se déduisent *pas* d'un produit scalaire (comme par exemple les normes $|\cdot|_1$ et $|\cdot|_\infty$ sur $\mathbb{R}^n$).

10. Produits scalaires

10.1. Définitions

Une *forme bilinéaire symétrique* sur un espace vectoriel E sur le corps $\mathbb{R}$ est une application

$$(\cdot|\cdot)\colon\ E \times E \to \mathbb{R}$$

satisfaisant les conditions suivantes :

1. Bilinéarité : pour $x, y, z \in E$ et $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned}(x\alpha + y\beta|z) &= \alpha\,(x|z) + \beta\,(y|z).\\ (z|x\alpha + y\beta) &= (z|x)\,\alpha + (z|y)\,\beta.\end{aligned}$$

2. Symétrie : pour $x, y \in E$,

$$(y|x) = (x|y).$$

Un *produit scalaire* est une forme bilinéaire symétrique qui possède de plus la propriété suivante :

3. Définie positivité : pour $x \in E$, $x \neq 0$,

$$(x|x) > 0.$$

La bilinéarité entraîne par ailleurs

$$(0|0) = 0.$$

Un espace muni d'un produit scalaire est appelé *espace euclidien*.

D'autres notations sont parfois utilisées pour le produit scalaire $(x|y)$, comme par exemple

$$x \cdot y \quad \text{ou} \quad (x, y) \quad \text{ou} \quad \langle x, y\rangle \quad \text{ou} \quad \langle x|y\rangle.$$

Exemples.

A. Le *produit scalaire usuel* sur $\mathbb{R}^n$ est défini par

$$\Big((x_1, \ldots, x_n)\,|\,(y_1, \ldots, y_n)\Big) = x_1y_1 + \cdots + x_ny_n.$$

(Voir [An, p.24]). D'autres formes bilinéaires symétriques peuvent être définies de la manière suivante : soit $A \in \mathbb{R}^n$ une matrice symétrique, c'est-à dire telle que $A^t = A$. On note les éléments de $\mathbb{R}^n$ en colonnes et on pose, pour $x, y \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$,

$$(x|y)_A = x^t \cdot A \cdot y.$$

La bilinéarité de $(\cdot|\cdot)_A$ résulte des propriétés du produit matriciel et l'hypothèse que A est symétrique assure

$$(y|x)_A = (x|y)_A \qquad \text{pour } x, y \in \mathbb{R}^n.$$

[1]Les références à l'ouvrage "Analyse — Fondements, techniques, évolution (2e édition)" de Jean Mawhin (De Boeck-Université, Paris-Bruxelles, 1997) seront indiquées par [An].

Pour certains choix de A, la forme $(\cdot|\cdot)_A$ est en outre définie positive et définit dès lors un produit scalaire sur $\mathbb{R}^n$, différent du produit scalaire usuel. Des critères de définie positivité sont établis dans le chapitre VII.

B. Dans l'espace usuel E_O (où l'on a choisi un point de référence O), on note $\| \vec{x} \|$ la longueur d'un vecteur $\vec{x}$. Le produit scalaire canonique est défini par

$$(\vec{x}|\vec{y}) = \| \vec{x} \| \, \| \vec{y} \| \, \cos\theta = \tfrac{1}{2}(\| \vec{x} + \vec{y} \|^2 - \| \vec{x} \|^2 - \| \vec{y} \|^2),$$

où θ est l'angle entre les vecteurs $\vec{x}$ et $\vec{y}$.

C. Soient $a_1, \dots, a_n \in \mathbb{R}$ des scalaires fixés. On peut définir une forme bilinéaire symétrique sur l'espace $\mathbb{R}[X]$ en posant

$$(P|Q) = \sum_{i=1}^{n} P(a_i)Q(a_i) \qquad \text{pour } P, Q \in \mathbb{R}[X].$$

Nous laissons au lecteur le soin de montrer que si $a_1, \dots, a_n$ sont distincts deux à deux, la restriction de cette forme au sous-espace $\mathbb{R}[X]_{\leqslant n-1}$ des polynômes de degré au plus $n-1$ est un produit scalaire.

D. Sur l'ensemble $C\big([0,1];\mathbb{R}\big)$ des fonctions continues sur l'intervalle $[0,1]$, on peut définir un produit scalaire par

$$(f|g) = \int_0^1 f(x)g(x)\,dx.$$

Normes et distances. Dans un espace euclidien E, on définit la *norme* d'un vecteur x par

$$\| x \| = \sqrt{(x|x)} \in \mathbb{R}.$$

10.1. Proposition. *L'application $\| \cdot \| \colon E \to \mathbb{R}$ satisfait les conditions qui définissent une norme (abstraite) au sens de l'analyse [An, p.20], à savoir :*

(i) $\| x \| \geqslant 0$ *pour tout* $x \in E$.

(ii) $\| x \| = 0$ *entraîne* $x = 0$.

(iii) $\| \alpha x \| = |\alpha| \; \| x \|$ *pour* $x \in E$ *et* $\alpha \in \mathbb{R}$.

(iv) $\| x + y \| \leqslant \| x \| + \| y \|$ *pour* $x, y \in E$.

(Comparer à la Proposition de [An, p.22]).

Démonstration. La propriété (i) est claire, la propriété (ii) découle directement de la définie positivité du produit scalaire et la propriété (iii) résulte de sa linéarité. Pour établir la propriété (iv) (connue sous le nom d'*inégalité triangulaire* ou *de Minkowski*), nous utiliserons le lemme suivant :

10.2. Lemme (Inégalité de Cauchy-Schwarz). *Pour $x, y \in E$,*

$$|(x|y)| \leqslant \| x \| \, \| y \| \,.$$

Démonstration. Considérons le vecteur $tx + y \in E$, où $t \in \mathbb{R}$ est un paramètre variable. D'après la définie positivité du produit scalaire, on doit avoir

$$(*) \qquad (tx + y|tx + y) \geqslant 0 \qquad \text{quel que soit } t.$$

Or, la bilinéarité du produit scalaire permet de réécrire le membre de gauche sous la forme d'un trinôme du second degré en t :

$$(tx + y|tx + y) = t^2(x|x) + 2t(x|y) + (y|y).$$

L'inégalité $(*)$ montre que ce trinôme en t ne peut pas avoir deux racines distinctes, donc son discriminant est négatif ou nul. Cela donne

$$(x|y)^2 - (x|x)\,(y|y) \leqslant 0.$$

□

(De manière équivalente, on peut considérer le vecteur $z = (x|x)y - (x|y)x$ et calculer

$$(z|z) = (x|x)\,[(x|x)\,(y|y) - (x|y)^2].$$

L'inégalité de Cauchy-Schwarz est claire si $x = 0$ et résulte de $(z|z) \geqslant 0$ si $x \neq 0$.)

On peut à présent compléter la démonstration de la Proposition 10.1 en démontrant la propriété (iv) : par définition, on a

$$\| x+y \|^2 = (x+y|x+y) = (x|x) + 2(x|y) + (y|y).$$

Or, par l'inégalité de Cauchy-Schwarz,

$$(x|y) \leqslant |(x|y)| \leqslant \| x \| \, \| y \|,$$

donc

$$\| x+y \|^2 \leqslant \| x \|^2 + 2 \, \| x \| \, \| y \| + \| y \|^2 = (\| x \| + \| y \|)^2.$$

□

Vu la Proposition 10.1, on peut définir une *distance* (vectorielle) entre deux vecteurs $x, y \in E$ en posant

$$d(x,y) = \| x - y \| \in \mathbb{R}.$$

(Voir [An, Proposition, p.678]).

10.2. Bases orthonormées

Dans l'exemple B de la section précédente, nous voyons que si deux vecteurs $\vec{x}$ et $\vec{y}$ sont orthogonaux au sens géométrique du terme, alors leur produit scalaire est nul. Nous allons généraliser cette notion à des espaces vectoriels plus abstraits par la définition suivante : deux vecteurs x, y d'un espace euclidien E sont dits *orthogonaux* si $(x|y) = 0$. Comme $(y|x) = (x|y)$, cette relation est symétrique. Une base $e = (e_1, \ldots, e_n)$ d'un espace euclidien E est dite *orthogonale* si les vecteurs de base sont orthogonaux deux à deux[2] :

$$(e_i|e_j) = 0 \qquad \text{si } i \neq j.$$

Elle est dite *orthonormée* si de plus

$$\| e_i \| = 1 \qquad \text{pour tout } i = 1, \ldots, n.$$

Si $e = (e_1, \ldots, e_n)$ est une base orthogonale, alors la suite $e' = (e'_1, \ldots, e'_n)$ définie par

$$e'_i = \frac{e_i}{\| e_i \|}$$

est une base orthonormée.

Le lemme suivant montre que les produits scalaires s'expriment de manière particulièrement simple en fonction des coordonnées par rapport à une base orthonormée.

10.3. Lemme. *Soit e une base orthonormée d'un espace euclidien E. Pour $x, y \in E$,*

$$(x|y) = {}_e(x)^t \cdot {}_e(y).$$

Démonstration. Soit $x = \sum_{i=1}^n e_i x_i$ et $y = \sum_{i=1}^n e_j y_j$, de sorte que

$${}_e(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad {}_e(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

En utilisant la bilinéarité du produit scalaire, on trouve

$$(x|y) = \left(\sum_{i=1}^n e_i x_i \,\middle|\, \sum_{j=1}^n e_j y_j \right) = \sum_{i,j=1}^n x_i (e_i|e_j) y_j.$$

[2]En particulier, toute base réduite à un élément est considérée comme orthogonale.

Or, comme la base e est orthonormée, on a $(e_i|e_j) = \delta_{ij}$ (symbole de Kronecker), donc

$$(x|y) = \sum_{i=1}^{n} x_i y_i = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

$\square$

Le principal résultat de cette section est le suivant :

10.4. Théorème. *Tout espace euclidien de dimension finie non nulle admet une base orthonormée.*

Démonstration. Comme on peut toujours normer les vecteurs d'une base en les divisant par leur norme, il suffit de prouver l'existence d'une base orthogonale. Si l'espace est de dimension 1, le théorème est clair puisque toute base contient un seul vecteur et est donc orthogonale. On peut donc supposer $\dim E \geqslant 2$ et raisonner par induction, c'est-à-dire supposer de plus que l'énoncé est vrai pour les espaces euclidiens de dimension $< \dim E$.

Soit $(e_1, \ldots, e_n)$ une base quelconque de E. Pour $i = 2, \ldots, n$, posons

$$e_i' = e_i - e_1 \frac{(e_1|e_i)}{(e_1|e_1)}.$$

Ces vecteurs sont définis de manière à être orthogonaux à e_1. En effet, pour $i = 2, \ldots, n$, on a

$$(e_1|e_i') = (e_1|e_i) - (e_1|e_1)\frac{(e_1|e_i)}{(e_1|e_1)} = 0.$$

Par ailleurs, la suite $(e_1, e_2', \ldots, e_n')$ est obtenue à partir de la suite $(e_1, \ldots, e_n)$ par des opérations élémentaires (de type I) ; elle engendre donc aussi l'espace E et est par conséquent une base de E.

Par hypothèse d'induction, le sous-espace $V = \operatorname{sev}\langle e_2', \ldots, e_n'\rangle$ admet une base orthogonale $(e_2'', \ldots, e_n'')$, puisqu'il est de dimension $n-1$. Tout vecteur de E est somme d'un multiple de e_1 et d'un vecteur de V, car $(e_1, e_2', \ldots, e_n')$ est une base de E. Dès lors la suite $(e_1, e_2'', \ldots, e_n'')$ engendre E ; c'est donc une base de E. On va prouver que cette base est orthogonale.

Comme les vecteurs e_i'' sont combilis de $e_2', \ldots, e_n'$, ils sont orthogonaux à e_1 ; en effet, si $e_i'' = \sum_{j=2}^n e_j'\alpha_j$, alors

$$(e_1|e_i'') = \sum_{j=2}^{n} (e_1|e_j')\alpha_j = 0$$

puisque $(e_1|e_j') = 0$ pour $j = 2, \ldots, n$. (Voir aussi le Lemme 11.1 ci-dessous.) De plus, les vecteurs $e_2'', \ldots, e_n''$ sont orthogonaux deux à deux par construction, donc $(e_1, e_2'', \ldots, e_n'')$ est une base orthogonale de E. $\square$

Remarquons pour terminer que si e est une base orthonormée d'un espace euclidien E de dimension n, alors l'isomorphisme ${}_e\gamma\colon E \to \mathbb{R}^n$ qui à tout vecteur de E associe ses coordonnées par rapport à e est une *isométrie* d'espaces euclidiens ($\mathbb{R}^n$ étant muni du produit scalaire usuel), dans le sens que le produit scalaire de deux vecteurs $x, y \in E$ est égal au produit scalaire usuel de leurs coordonnées ${}_e(x)$, ${}_e(y)$. En effet, comme la base e est orthonormée, le Lemme 10.3 donne

$$(x|y) = {}_e(x)^t \cdot {}_e(y) \qquad \text{pour } x, y \in E.$$

Exercices

(10.1) Prouvez que les applications suivantes sont des produits scalaires.

– $(\cdot|\cdot)_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : \Big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\Big) \mapsto \sum_{k=1}^n k x_k y_k$.

– $(\cdot|\cdot)_2 : \mathbb{R}[X]_{\leqslant n-1} \times \mathbb{R}[X]_{\leqslant n-1} \to \mathbb{R} : (P, Q) \mapsto \sum_{i=1}^n P(a_i)Q(a_i)$, où $a_1, \ldots, a_n$ sont des réels fixés, deux à deux distincts.

– $(\cdot|\cdot)_3 : C\big([0,1];\mathbb{R}\big) \times C\big([0,1];\mathbb{R}\big) \to \mathbb{R} : (f, g) \mapsto \displaystyle\int_0^1 f(x)g(x)dx$.

(10.2) Prouvez que $(\cdot|\cdot) : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R} : (A,B) \mapsto \mathrm{tr}(A^t.B)$ est un produit scalaire ($\mathrm{tr}(M) = \sum_{i=1}^n M_{ii}$).

(10.3) Soit S le sous-espace vectoriel de $\mathbb{R}^{n\times n}$ formé des matrices symétriques.
Prouvez que $(\cdot|\cdot) : S \times S \to \mathbb{R} : (A,B) \mapsto \mathrm{tr}(A.B)$ est un produit scalaire.
Peut-on déduire ceci de l'exercice 10.2 ?

(10.4) Obtient-on encore un produit scalaire si l'on remplace A^t par A dans la définition de $(\cdot|\cdot)$ dans l'exercice 10.2 ?

(10.5) On considère l'application

$$(\cdot|\cdot) : C(\mathbb{R};\mathbb{R}) \times C(\mathbb{R};\mathbb{R}) \to \mathbb{R} : (f,g) \mapsto f(\pi)g(\pi) + f(\pi/2)g(\pi/2) + f(0)g(0).$$

– Montrez que $(\cdot|\cdot)$ n'est *pas* un produit scalaire.
– Soit V le sous-espace vectoriel de $C(\mathbb{R};\mathbb{R})$ engendré par la fonction constante 1, la fonction sin et la fonction cos. Montrez que la restriction $(\cdot|\cdot)_{|V}$ de $(\cdot|\cdot)$ à V est un produit scalaire.

(10.6) Dans $\mathbb{R}^{2\times 2}$ muni du produit scalaire $(\cdot|\cdot)$ défini par $(A|B) = \mathrm{tr}(A^t.B)$, quel est l'orthogonal du sous-espace formé des matrices symétriques ?

(10.7) Prouvez que

$$(\cdot|\cdot) : \mathbb{R}[X]_{\leqslant n} \times \mathbb{R}[X]_{\leqslant n} \to \mathbb{R} : (P,Q) \mapsto \sum_{i=0}^n \frac{1}{i!} P^{(i)}(0).Q^{(i)}(0)$$

est un produit scalaire ($P^{(i)}$ et $Q^{(i)}$ désignent respectivement les dérivées d'ordre i de P et de Q).

Pour $n = 3$, calculez l'orthogonal du sous-espace vectoriel formé des polynômes P tels que $P(0) = 0$ et $P''(0) = 0$.

(10.8) Dans $\mathbb{R}[X]_{\leqslant 2}$, muni du produit scalaire défini par

$$(P|Q) = P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1),$$

quel est l'orthogonal du sous-espace $V = \{P \in \mathbb{R}[X]_{\leqslant 2} \mid P(1) = 0\}$?

11. Projections orthogonales

Dans cette section, nous ne nous intéressons qu'à des espaces euclidiens de dimension finie. Nous allons définir la projection orthogonale d'un vecteur sur un sous-espace et montrer que cette notion permet de construire des bases orthonormées à partir d'une base quelconque et de résoudre des problèmes d'approximation.

11.1. Définition

Soit E un espace euclidien de dimension finie. Pour toute partie $V \subset E$, l'orthogonal $V^\perp$ est l'ensemble des vecteurs orthogonaux à tous les vecteurs de V :

$$V^\perp = \{x \in E \mid (x|v) = 0 \text{ pour tout } v \in V\}.$$

Cet ensemble est un sous-espace vectoriel de E.

11.1. Lemme. *Si V est un sous-espace vectoriel de base $(e_1, \dots, e_m)$, alors*

$$V^\perp = \{x \in E \mid (e_1|x) = \cdots = (e_m|x) = 0\}.$$

Démonstration. L'inclusion de $V^\perp$ dans le membre de droite est claire, car tout vecteur $x \in V^\perp$ est orthogonal à $e_1, \dots, e_m$ puisque ces vecteurs sont dans V. Réciproquement, si $x \in E$ est orthogonal à $e_1, \dots, e_m$ et si $v \in V$, soit

$$v = e_1 v_1 + \cdots + e_m v_m,$$

alors

$$(v|x) = v_1\,(e_1|x) + \cdots + v_m\,(e_m|x) = 0,$$

donc $x \in V^\perp$. □

La projection orthogonale d'un vecteur sur un sous-espace est définie par la proposition suivante :

11.2. PROPOSITION. *Soit V un sous-espace d'un espace euclidien E. Pour tout $x \in E$, il existe un et un seul vecteur $p \in V$ tel que $x - p \in V^\perp$. De plus, si $(e_1, \dots, e_r)$ est une base orthonormée de V, alors*

$$p = \sum_{i=1}^{r} e_i(e_i|x).$$

Le vecteur p est appelé *projection orthogonale* du vecteur x sur le sous-espace V.

DÉMONSTRATION. Le vecteur p défini par la formule ci-dessus est dans V puisqu'il est combinaison linéaire de la base $(e_1, \dots, e_r)$ de V. Pour prouver que $x - p \in V^\perp$, il suffit, d'après le Lemme 11.1, de montrer que $(e_j|x-p) = 0$ pour $j = 1, \dots, r$. Or,

$$(e_j|x-p) = \left(e_j \,\middle|\, x - \sum_{i=1}^{r} e_i\,(e_i|x)\right) = (e_j|x) - \sum_{i=1}^{r}(e_j|e_i)(e_i|x).$$

Comme la base $(e_1, \dots, e_r)$ est orthogonale, tous les termes de la somme du membre de droite sont nuls, sauf peut-être celui où l'indice i est égal à j :

$$(e_j|x) - \sum_{i=1}^{r}(e_j|e_i)(e_i|x) = (e_j|x) - (e_j|e_j)(e_j|x) = 0.$$

On a donc

$$(e_j|x-p) = 0 \qquad \text{pour } j = 1, \dots, r,$$

ce qui montre que $x-p \in V^\perp$ et achève de démontrer l'existence d'un vecteur p possédant les propriétés de l'énoncé.

Pour prouver l'unicité de ce vecteur p, supposons avoir trouvé deux vecteurs $p, p' \in V$ tels que $x - p, x - p' \in V^\perp$. Alors $p - p' \in V$, mais

$$p - p' = (x - p') - (x - p) \in V^\perp,$$

donc $p - p' \in V \cap V^\perp$. En particulier, le vecteur $p - p'$ est orthogonal à lui-même :

$$(p - p'|p - p') = \| p - p' \|^2 = 0,$$

donc $\| p - p' \| = 0$ et $p = p'$. □

Il est maintenant facile de déterminer la dimension de l'orthogonal d'un sous-espace :

11.3. COROLLAIRE. *Pour tout sous-espace vectoriel V d'un espace euclidien E de dimension finie,*

$$E = V \oplus V^\perp,$$

donc

$$\dim V^\perp = \dim E - \dim V.$$

De plus,

$$(V^\perp)^\perp = V.$$

DÉMONSTRATION. Pour tout vecteur $x \in E$, la projection orthogonale p de x sur V satisfait : $p \in V$ et $x - p \in V^\perp$. Par conséquent,

$$x = p + (x - p)$$

est une décomposition de x en somme d'un vecteur de V et d'un vecteur de $V^\perp$. Cela prouve : $E = V + V^\perp$. Par ailleurs, $V \cap V^\perp = \{0\}$ puisque les vecteurs de cette intersection sont orthogonaux à eux-mêmes. On a donc bien

$$E = V \oplus V^\perp,$$

et la formule pour $\dim V^\perp$ s'en déduit.

On peut appliquer cette formule deux fois de suite pour trouver la dimension de $(V^\perp)^\perp$:

$$\dim(V^\perp)^\perp = \dim E - \dim V^\perp = \dim E - (\dim E - \dim V) = \dim V.$$

Or, par définition de l'orthogonal d'un sous-espace, on a immédiatement

$$V \subseteq (V^{\perp})^{\perp}.$$

Comme ces deux sous-espaces ont la même dimension, il sont égaux. □

11.2. Procédé d'orthogonalisation de Gram–Schmidt

La technique des projections orthogonales permet de construire systématiquement une base orthogonale à partir d'une base quelconque.

Soit $e = (e_1, \ldots, e_n)$ une base d'un espace euclidien E et soit

$$V_k = \text{sev}\langle e_1, \ldots, e_k\rangle \qquad \text{pour } k = 1, \ldots, n,$$

de sorte que $V_n = E$. On peut projeter tout vecteur de E orthogonalement sur tout sous-espace V_k. Pour $k = 2, \ldots, n$, soit $p_k \in V_{k-1}$ la projection orthogonale de e_k sur V_{k-1} et soit

$$u_k = e_k - p_k.$$

On pose aussi $u_1 = e_1$.

11.4. Théorème. *La suite $u = (u_1, u_2, \ldots, u_n)$ est une base orthogonale de E.*

Démonstration. On prouve, par récurrence sur k, que la suite

$$u(k) = (u_1, u_2, \ldots, u_k)$$

est une base orthogonale de V_k pour $k = 1, \ldots, n$. Comme $V_n = E$, cela établira l'énoncé.

Pour $k = 1$, on a $u(1) = (e_1)$; c'est une base de V_1, par définition de ce sous-espace, et elle est orthogonale car elle n'a qu'un seul élément.

Supposons ensuite avoir prouvé que $u(k-1)$ est une base orthogonale de V_{k-1}. Comme p_k est la projection orthogonale de e_k sur V_{k-1}, on a

$$u_k = e_k - p_k \in V_{k-1}^{\perp}.$$

Si $u_k \in V_{k-1}$, alors $u_k = 0$ puisque seul le vecteur nul est orthogonal à lui-même. Alors $e_k = p_k \in V_{k-1} = \text{sev}\langle e_1, \ldots, e_{k-1}\rangle$, ce qui est en contradiction avec l'indépendance linéaire de la suite $(e_1, \ldots, e_k)$. Dès lors, on a

$$u_k \notin V_{k-1}$$

et il résulte alors du Lemme 5.1 (p. 29) que la suite $u(k)$ est libre. Elle est contenue dans V_k car $p_k \in V_{k-1} \subset V_k$ et $e_k \in V_k$. Par ailleurs, $\dim V_k = k$ puisque $V_k = \text{sev}\langle e_1, \ldots, e_k\rangle$; donc la suite $u(k)$ est une base de V_k, par le Corollaire 4.5 (p. 28). Comme $u(k-1)$ est une base orthogonale de V_{k-1} et que $u_k \in V_{k-1}^{\perp}$, cette base est orthogonale. □

Un aspect important du théorème précédent est qu'il peut être mis en œuvre en pratique pour obtenir une base orthonormée d'un espace à partir d'une base quelconque. En effet, pour obtenir le dernier vecteur de la base orthogonale $u(k)$ de V_k, il suffit de calculer la projection orthogonale p_k de e_k sur V_{k-1}. Or, au moment de calculer cette projection, on connaît déjà une base orthogonale de V_{k-1}, à savoir $u(k-1)$. Quitte à normer cette base, on peut donc calculer cette projection à l'aide de la formule de la Proposition 11.2. Explicitement, on définit successivement

$$\begin{aligned}
u_1 &= e_1,\\
q_1 &= \frac{u_1}{\|u_1\|},\\
u_2 &= e_2 - q_1\,(q_1|e_2),\\
q_2 &= \frac{u_2}{\|u_2\|},\\
u_3 &= e_3 - q_1\,(q_1|e_3) - q_2\,(q_2|e_3),\\
q_3 &= \frac{u_3}{\|u_3\|},\\
&\ \vdots\\
u_k &= e_k - q_1\,(q_1|e_k) - q_2\,(q_2|e_k) - \cdots - q_{k-1}\,(q_{k-1}|e_k),\\
q_k &= \frac{u_k}{\|u_k\|}.
\end{aligned}$$

En isolant dans le membre de droite le vecteur e_k, on obtient une expression de ce vecteur comme combinaison linéaire de $q_1, \ldots, q_{k-1}, u_k$:

$$\begin{aligned}
e_1 &= u_1\\
&= q_1\,\|u_1\|,\\
e_2 &= q_1(q_1|e_2) + u_2\\
&= q_1(q_1|e_2) + q_2\,\|u_2\|,\\
e_3 &= q_1(q_1|e_3) + q_2(q_2|e_3) + u_3\\
&= q_1(q_1|e_3) + q_2(q_2|e_3) + q_3\,\|u_3\|,\\
&\ \vdots\\
e_k &= q_1(q_1|e_k) + \cdots + q_{k-1}(q_{k-1}|e_k) + u_k\\
&= q_1(q_1|e_k) + \cdots + q_{k-1}(q_{k-1}|e_k) + q_k\,\|u_k\|\,.
\end{aligned}$$

La matrice de changement de base ${}_q(I_E)_e$ s'en déduit :

11.5. Corollaire. *Soit q la base orthonormée obtenue en appliquant le procédé de Gram–Schmidt à une base e d'un espace euclidien E. La matrice de changement de base ${}_q(I_E)_e$ est triangulaire supérieure avec des entrées diagonales non nulles.*

Démonstration. Les calculs précédents montrent que la matrice ${}_q(I_E)_e$ est

$$ {}_q(I_E)_e = \begin{pmatrix} \|u_1\| & (q_1|e_2) & (q_1|e_3) & \cdots & (q_1|e_n)\\ 0 & \|u_2\| & (q_2|e_3) & \cdots & (q_2|e_n)\\ 0 & 0 & \|u_3\| & \cdots & (q_3|e_n)\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & \|u_n\| \end{pmatrix}. $$

□

11.3. Problèmes d'approximation

Les projections orthogonales satisfont la propriété d'approximation suivante :

11.6. Proposition. *Soit E un espace euclidien et $V \subset E$ un sous-espace. Pour tout $x \in E$, la projection orthogonale p de x sur V est un vecteur de V dont la distance à x est minimale. De plus, pour tout vecteur $v \in V$ autre que p, on a*

$$d(x, v) > d(x, p).$$

Démonstration. Pour montrer la première partie de l'énoncé, il s'agit de prouver :

$$d(x,p) \leqslant d(x,y) \qquad \text{pour tout } y \in V,$$

ou, de manière équivalente,

$$\| x-p \|^2 \leqslant \| x-y \|^2 \qquad \text{pour tout } y \in V.$$

On décompose pour cela : $(x-y) = (x-p) + (p-y)$, et on calcule

$$(x-y|x-y) = (x-p|x-p) + (x-p|p-y) + (p-y|x-p) + (p-y|p-y),$$

c'est-à-dire

$$\| x-y \|^2 = \| x-p \|^2 + (x-p|p-y) + (p-y|x-p) + \| p-y \|^2 .$$

Comme $x-p \in V^\perp$ et $p-y \in V$, on a

$$(x-p|p-y) = (p-y|x-p) = 0,$$

et l'égalité précédente entraîne

$$\| x-y \|^2 = \| x-p \|^2 + \| p-y \|^2 \geqslant \| x-p \|^2 .$$

On voit d'ailleurs que l'égalité $\| x-y \| = \| x-p \|$ (avec $y \in V$) ne peut être satisfaite que si $\| p-y \| = 0$, c'est-à-dire si $y = p$. □

Un exemple typique d'application de cette propriété d'approximation concerne les systèmes d'équations linéaires : soit $A \in \mathbb{R}^{m\times n}$ et $b \in \mathbb{R}^m$. Pour que le système

$$A \cdot X = b$$

admette une solution, il faut et il suffit que b soit dans le sous-espace $\mathcal{C}(A) \subset \mathbb{R}^m$ engendré par les colonnes de A, puisqu'en décomposant A par colonnes l'équation précédente s'écrit

$$a_{*1}x_1 + \cdots + a_{*n}x_n = b.$$

Si $b \notin \mathcal{C}(A)$, on peut cependant envisager de trouver une solution approchée du système $A \cdot X = b$ en remplaçant b par le vecteur $b' \in \mathcal{C}(A)$ qui est le plus proche de b (au sens de la distance euclidienne usuelle sur $\mathbb{R}^m$, induite par le produit scalaire usuel). Toute solution du système d'équations linéaires

$$A \cdot X = b'$$

où b' est la projection orthogonale de b sur $\mathcal{C}(A)$ est appelée *solution approchée de $A \cdot X = b$ au sens des moindres carrés.*

Le procédé de Gram–Schmidt permet de trouver une base orthonormée de $\mathcal{C}(A)$, à l'aide de laquelle la projection orthogonale b' est facile à calculer. Dans le cas où les colonnes de A sont linéairement indépendantes, les calculs peuvent s'organiser de manière particulièrement simple, et prennent la forme d'une factorisation matricielle :

11.7. Proposition. *Pour toute matrice $A \in \mathbb{R}^{m\times n}$ de rang n, il existe une matrice $Q \in \mathbb{R}^{m\times n}$ telle que $Q^t \cdot Q = I_n$ et une matrice $R \in \mathbb{R}^{n\times n}$ triangulaire supérieure dont toutes les entrées diagonales sont non nulles telles que*

$$A = Q \cdot R.$$

De plus, pour tout $y \in \mathbb{R}^m$, le vecteur $Q \cdot Q^t \cdot y \in \mathbb{R}^m$ est la projection orthogonale de y sur $\mathcal{C}(A)$.

La condition $Q^t \cdot Q = I_n$ entraîne que $Q^t = Q^{-1}$ si $m = n$; la matrice Q n'est cependant pas inversible à droite si $m > n$.

DÉMONSTRATION. Soit $a = (a_{*1}, \dots, a_{*n})$ la suite des colonnes de A. Comme $\operatorname{rang} A = n$, la suite a est libre ; c'est donc une base de $\mathcal{C}(A)$. En appliquant le procédé de Gram–Schmidt à la suite a, on obtient une base orthogonale de $\mathcal{C}(A)$, que l'on norme pour former une base orthonormée $q = (q_1, \dots, q_n)$. Le Corollaire 11.5 montre que la matrice de changement de base ${}_q(I_{\mathcal{C}(A)})_a$ est triangulaire ; on la note $R = (r_{ij})_{1 \leqslant i,j \leqslant n}$. On a donc

$$\begin{array}{rcl} a_{*1} & = & q_1 r_{11} \\ a_{*2} & = & q_1 r_{12} + q_2 r_{22} \\ & \vdots & \\ a_{*n} & = & q_1 r_{1n} + q_2 r_{2n} + \cdots + q_n r_{nn} \end{array}$$

et les entrées diagonales r_{ii} sont non nulles. Si l'on note

$$Q = \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix} \in \mathbb{R}^{m \times n}$$

la matrice dont les colonnes sont $q_1, \dots, q_n$, les égalités ci-dessus montrent que l'on a

$$\begin{pmatrix} a_{*1} & \cdots & a_{*n} \end{pmatrix} = \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

c'est-à-dire

$$A = Q \cdot R.$$

Comme $(q_1, \dots, q_n)$ est une base orthonormée de $\mathcal{C}(A)$ pour le produit scalaire usuel de $\mathbb{R}^m$, on a

$$q_i^t \cdot q_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j, \end{cases}$$

donc

$$Q^t \cdot Q = \begin{pmatrix} q_1^t \\ \vdots \\ q_n^t \end{pmatrix} \cdot \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix} = \begin{pmatrix} q_1^t \cdot q_1 & \cdots & q_1^t \cdot q_n \\ \vdots & & \vdots \\ q_n^t \cdot q_1 & \cdots & q_n^t \cdot q_n \end{pmatrix} = I_n.$$

Par ailleurs, comme $(q_1, \dots, q_n)$ est une base orthonormée de $\mathcal{C}(A)$, la Proposition 11.2 montre que la projection orthogonale sur $\mathcal{C}(A)$ d'un vecteur $y \in \mathbb{R}^m$ est donnée par

$$\begin{array}{rcl} p & = & q_1 \cdot (q_1^t \cdot y) + \cdots + q_n \cdot (q_n^t \cdot y) \\ & = & \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix} \cdot \begin{pmatrix} q_1^t \\ \vdots \\ q_n^t \end{pmatrix} \cdot y \\ & = & Q \cdot Q^t \cdot y. \end{array}$$

□

11.8. COROLLAIRE. *Soit $A \in \mathbb{R}^{m \times n}$ une matrice de rang n, et soit $A = Q \cdot R$ la factorisation matricielle obtenue dans la proposition précédente. Pour tout $b \in \mathbb{R}^m$, le système $A \cdot X = b$ admet une et une seule solution exacte (si $b \in \mathcal{C}(A)$) ou approchée au sens des moindres carrés (si $b \notin \mathcal{C}(A)$), qui est aussi l'unique solution du système*

$$R \cdot X = Q^t \cdot b.$$

DÉMONSTRATION. Soit b' la projection orthogonale de b sur $\mathcal{C}(A)$, de sorte que $b' = b$ si $b \in \mathcal{C}(A)$. D'après la proposition précédente, on a $b' = Q \cdot Q^t \cdot b$; toute solution (exacte ou approchée au sens des moindres carrés) de $A \cdot X = b$ est donc solution (exacte) de

$$Q \cdot R \cdot X = Q \cdot Q^t \cdot b.$$

En multipliant les deux membres par la gauche par Q^t, et en tenant compte de ce que $Q^t \cdot Q = I_n$, on voit que l'équation précédente est équivalente à

$$R \cdot X = Q^t \cdot b.$$

Comme la matrice R est régulière, ce système d'équations admet une et une seule solution. $\square$

Lorsque l'on dispose de la factorisation $A = Q \cdot R$, le corollaire précédent donne une manière particulièrement simple de résoudre toute équation $A \cdot X = b$ (de manière exacte ou approchée), puisque la matrice R est triangulaire avec des entrées diagonales non nulles. Le système $R \cdot X = Q^t \cdot b$ est donc échelonné.

Exercices

(11.1) (prérequis) Soit E un espace euclidien et V un sous-espace de E. Si $v \in E$ et que p_v est la projection de v sur V, démontrez que, alors, $v - p_v$ est la projection de v sur $V^\perp$.

(11.2) On considère $\mathbb{R}^3$ muni du produit scalaire usuel.

- Soit $V = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = 0\}$, où a, b, c sont trois réels fixés. Calculez $V^\perp$.
- Soit $W = \operatorname{sev}\langle(k, l, m)\rangle$, où (k, l, m) est un élément fixé de $\mathbb{R}^3$. Calculez $W^\perp$.

(11.3) Si E est un espace euclidien et V et W des sous-espaces de E, est-il vrai que

$$\begin{aligned} (V + W)^\perp &= V^\perp \cap W^\perp? \\ (V \cap W)^\perp &= V^\perp + W^\perp? \end{aligned}$$

(11.4) On considère $\mathbb{R}^3$ muni du produit scalaire usuel.

- En appliquant le procédé de Gram–Schmidt à la base $\big((1, 0, 0), (1, 1, 0), (1, 1, 1)\big)$, trouvez une base orthonormée de $\mathbb{R}^3$.
- Même question en partant de la base $\big((1, 1, 0), (2, 1, 0), (1, 1, 1)\big)$.
- Trouvez les coordonnées d'un élément quelconque de $\mathbb{R}^3$ dans les deux bases orthonormées ainsi obtenues.

(11.5) Construire une base orthonormée de

$$V = \{M \in \mathbb{R}^{2\times 2} \mid M \text{ est symétrique}\},$$

pour le produit scalaire $(\cdot|\cdot)$ défini par $(A|B) = \operatorname{tr}(A.B)$.

(11.6) Dans $C^2\big([-\pi, \pi]; \mathbb{R}\big)$, on définit un produit scalaire $(\cdot|\cdot)$ par

$$(f|g) = \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Donnez une base orthonormée de $V = \operatorname{sev}\langle 1, \sin x, \cos x, \sin 5x, \cos 10x\rangle$.

(11.7) Dans $\mathbb{R}^3$ muni du produit scalaire usuel, déterminez la projection orthogonale du vecteur $(-1, 0, 8)$ sur le sous-espace engendré par les vecteurs $(1, 0, 1)$ et $(0, 1, 1)$.

(11.8) Dans l'espace vectoriel réel $C\big([-1, 1]; \mathbb{R}\big)$, avec le produit scalaire $(\cdot|\cdot)$ défini par

$$(f|g) = \int_{-1}^{1} f(x)g(x)dx,$$

trouvez le polynôme du premier degré P le plus proche de $f(x) = e^x$. L'ayant trouvé, donnez sa distance à $f(x)$.

(11.9) Dans l'espace vectoriel réel $C\big([0, 2\pi]; \mathbb{R}\big)$, avec le produit scalaire $(\cdot|\cdot)$ défini par

$$(f|g) = \int_{0}^{2\pi} f(x)g(x)dx,$$

trouvez la projection orthogonale de la fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto x$ sur le sous-espace V engendré par les fonctions 1, cos et sin.

(11.10) On considère $\mathbb{R}[X]_{\leqslant 2}$ muni du produit scalaire défini dans l'exercice 10.8, ainsi que le sous-espace V défini dans ce même exercice.

– Trouvez une base orthonormée de $V^{\perp}$.

– Calculez les projections orthogonales du polynôme X sur V et sur $V^{\perp}$.

(*Conseil :* relire l'exercice 11.1, page 79.)

(11.11) Soient u et v deux vecteurs non nuls fixés dans un espace euclidien E. Trouvez le vecteur le plus court qui soit de la forme $u + v\lambda$ ($\lambda \in \mathbb{R}$) et montrez qu'il est orthogonal à v.

(11.12) Mettez les matrices $A_i \in \mathbb{R}^{m\times n}$ suivantes sous la forme d'un produit $Q \cdot R$ où $Q \in \mathbb{R}^{m\times n}$ est telle que $Q^t \cdot Q = I_n$ et $R \in \mathbb{R}^{n\times n}$ est triangulaire supérieure avec toutes ses entrées diagonales non nulles.

$$A_1 = \begin{pmatrix} 1 & -2 \\ \sqrt{3} & 2\sqrt{3} \end{pmatrix} \qquad A_2 = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad A_5 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \\ 0 & 1. \end{pmatrix} \qquad A_6 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

(11.13) Résolvez les systèmes d'équations linéaires suivants au sens des moindres carrés.

$$\text{(i)} \begin{array}{rcl} x-y & = & 1 \\ x & = & 0 \\ y & = & 2 \end{array} \quad \text{(ii)} \begin{array}{rcl} x+y & = & 3 \\ x & = & 1 \\ x-y & = & -1 \end{array} \quad \text{(iii)} \begin{array}{rcl} 3x & = & 2 \\ 4x & = & 11 \end{array} \quad \text{(iv)} \begin{array}{rcl} 2x+y & = & -1 \\ x+y & = & 1 \\ -x+y & = & 3 \\ y & = & -1. \end{array}$$

(11.14) Trouvez la solution, au sens des moindres carrés, du système

$$\left\{ \begin{array}{rcl} x & = & b_1 \\ x & = & b_2 \\ & \vdots & \\ x & = & b_n. \end{array} \right.$$

Calculez ensuite la distance entre le second membre du système original et celui du système utilisé pour calculer la solution au sens des moindres carrés.

CHAPITRE VI

Opérateurs linéaires

Un opérateur linéaire est une application linéaire d'un espace vectoriel dans lui-même. On peut donc considérer un opérateur linéaire sur un espace comme une transformation de cet espace, et envisager de trouver les vecteurs fixes sous cette transformation, ou les vecteurs transformés en leurs opposés, ou d'une manière générale les vecteurs dont l'image est un multiple d'eux-mêmes, que l'on appelle *vecteurs propres* de l'opérateur.

La méthode permettant de déterminer les vecteurs propres d'un opérateur linéaire sur un espace de dimension finie, ainsi que les formes particulières que peut prendre la matrice d'un tel opérateur, sont indiquées dans la première section. Dans la deuxième section, on considère les opérateurs linéaires sur des espaces euclidiens, où le produit scalaire permet d'associer à chaque opérateur linéaire un opérateur adjoint. L'aboutissement de cette section est le *théorème spectral*, qui donne une condition nécessaire et suffisante pour qu'il existe une base orthonormée constituée de vecteurs propres.

Toutes les notions concernant les opérateurs linéaires admettent un analogue matriciel, obtenu par la correspondance qui à chaque matrice *carrée* $A \in K^{n\times n}$ associe l'opérateur linéaire $\tilde{A}$ sur K^n qui envoie $x \in K^n$ sur $\tilde{A}(x) = A\cdot x$, le n-uple x étant considéré comme une colonne : $x \in K^n = K^{n\times 1}$. Dans tout le chapitre, on passe constamment des opérateurs linéaires aux matrices carrées et vice-versa, suivant que les propriétés en vue sont plus faciles à établir pour les matrices ou pour les opérateurs linéaires.

12. Propriétés caractéristiques

12.1. Valeurs propres et vecteurs propres

Soit $A\colon E \to E$ un opérateur linéaire sur un espace vectoriel E sur un corps arbitraire K. Un *vecteur propre* de A est un vecteur $v \in E$ *non nul* tel que

$$A(v) = v\lambda \qquad \text{pour un certain } \lambda \in K.$$

Le scalaire λ est appelé *valeur propre* de A associée à v ; on dit aussi que v est un vecteur propre de A de valeur propre λ.

L'équation $A(v) = v\lambda$ peut aussi s'écrire

$$(\lambda I_E - A)(v) = 0,$$

donc les vecteurs propres de valeur propre λ sont les vecteurs non nuls de $\operatorname{Ker}(\lambda I_E - A)$.

Pour tout $\lambda \in K$, on appelle *espace propre de valeur propre* λ le sous-espace défini par

$$E(\lambda) = \operatorname{Ker}(\lambda I_E - A).$$

Avec cette définition, on peut dire que λ est valeur propre de A si et seulement si $E(\lambda) \neq \{0\}$ ou, ce qui revient au même, si et seulement si l'opérateur $\lambda I_E - A$ n'est pas injectif.

Supposons à partir d'ici que l'espace E est de dimension finie non nulle. D'après la Proposition 9.8 (p. 65), la condition que $\lambda I_E - A$ n'est pas injectif est encore équivalente à : dét $(\lambda I_E - A) = 0$.

On appelle *polynôme caractéristique* d'un opérateur linéaire A sur un espace E de dimension finie le polynôme

$$\mathrm{Pc}_A(X) = \text{dét}\,(X\, I_E - A).$$

La formule explicite du déterminant donnée dans la section 7.2 (p. 51) montre que le membre de droite définit bien un polynôme en X ; de plus, ce polynôme est unitaire, c'est-à-dire que le coefficient du terme de plus haut degré est 1, et son degré est égal à la dimension de E :

$$\deg \mathrm{Pc}_A = \dim E.$$

12.1. Proposition. *Les valeurs propres d'un opérateur linéaire sur un espace vectoriel de dimension finie sont les racines du polynôme caractéristique de cet opérateur. Le nombre de valeurs propres d'un opérateur linéaire est donc au plus égal à la dimension de l'espace.*

Démonstration. On a vu ci-dessus que $\lambda \in K$ est valeur propre de A si et seulement si dét $(\lambda I_E - A) = 0$. □

Pour $K = \mathbb{C}$, le théorème fondamental de l'algèbre affirme que tout polynôme non constant se décompose en produit de polynômes du premier degré, et admet par conséquent au moins une racine. Tout opérateur linéaire sur un espace complexe admet donc au moins une valeur propre, d'où aussi un vecteur propre. En revanche, si $K = \mathbb{R}$, certains opérateurs linéaires peuvent ne pas admettre de valeur propre : par exemple, dans le plan usuel, une rotation d'angle $\neq 0, \pi$ autour du point de référence O n'a pas de vecteur propre (ni par conséquent de valeur propre).

En utilisant d'une part le polynôme caractéristique, d'autre part les espaces propres, on peut associer à chaque valeur propre λ d'un opérateur linéaire A sur un espace E de dimension finie deux types de muliplicités : on appelle *multiplicité algébrique* de λ et on note $m_a(\lambda)$ l'exposant de $X - \lambda$ dans la décomposition de Pc_A en produit de facteurs irréductibles ; en d'autres termes, $m_a(\lambda)$ est l'exposant de la plus haute puissance de $X - \lambda$ qui divise Pc_A :

$$m_a(\lambda) = m \qquad \text{si } \mathrm{Pc}_A(X) = (X-\lambda)^m Q(X)$$

où $Q(X)$ est un polynôme tel que $Q(\lambda) \neq 0$.

On appelle *multiplicité géométrique* de λ et on note $m_g(\lambda)$ la dimension de l'espace propre $E(\lambda)$:

$$m_g(\lambda) = \dim E(\lambda).$$

Comme $E(\lambda) = \mathrm{Ker}(\lambda I_E - A)$, on a aussi, d'après le Théorème 8.4 (p. 56),

$$m_g(\lambda) = \dim E - \mathrm{rang}(\lambda I_E - A).$$

12.2. Proposition. *Pour toute valeur propre λ d'un opérateur linéaire A sur un espace E de dimension finie,*

$$m_a(\lambda) \geqslant m_g(\lambda) \geqslant 1.$$

Démonstration. Soit $(e_1, \ldots, e_r)$ une base de $E(\lambda)$, que l'on prolonge en une base $e = (e_1, \ldots, e_r, \ldots, e_n)$ de E. On a donc posé : $m_g(\lambda) = r$, $\dim E = n$.

Comme $e_1, \ldots, e_r$ sont des vecteurs propres de A de valeur propre λ, on a $A(e_i) = e_i\lambda$ pour $i = 1, \ldots, r$, donc la matrice de A par rapport à la base e est de la forme

$${}_e(A)_e = \left(\begin{array}{c|c} \lambda I_r & U \\ \hline 0_{n-r,r} & V \end{array}\right)$$

où $U \in K^{r\times(n-r)}$ et $V \in K^{(n-r)\times(n-r)}$. En développant le déterminant suivant la première colonne et en raisonnant par induction sur r, on obtient

$$\mathrm{Pc}_A(X) = (X-\lambda)^r \,\text{dét}\,(X\, I_{n-r} - V).$$

Cela montre que Pc_A est divisible par $(X-\lambda)^r$, donc $r \leqslant m_a(\lambda)$, ce qui achève la démonstration puisque $r = m_g(\lambda)$. □

Traduction matricielle. Les notions définies ci-dessus se transposent immédiatement aux matrices par l'intermédiaire de l'opérateur linéaire associé. Pour $A \in K^{n\times n}$, l'opérateur associé $\tilde{A}\colon K^n \to K^n$ est défini par $\tilde{A}(x) = A \cdot x$ pour $x \in K^n$. (On abuse des notations en notant les éléments de K^n en colonnes : $K^n = K^{n\times 1}$).

Un *vecteur propre* de $A \in K^{n\times n}$ est donc un vecteur non nul $x \in K^n$ tel que

$$A \cdot x = x\,\lambda \qquad \text{pour un certain } \lambda \in K.$$

Le scalaire λ est appelé *valeur propre* de A associée au vecteur propre x. Le *polynôme caractéristique* de $A \in K^{n\times n}$ est le polynôme caractéristique de l'opérateur $\tilde{A}$. Comme la matrice de $\tilde{A}$ par rapport à la base canonique de K^n est A, on a donc

$$\mathrm{Pc}_A(X) = \mathrm{dét}\,(X\,I_n - A).$$

Ce polynôme est unitaire de degré n et ses racines sont les valeurs propres de A.

12.3. Proposition. *Soient $A, P \in K^{n\times n}$. Si la matrice P est régulière, les matrices A et $P^{-1}\cdot A\cdot P$ ont le même polynôme caractéristique :*

$$\mathrm{Pc}_A = \mathrm{Pc}_{P^{-1}\cdot A\cdot P}\,.$$

Démonstration. D'après la multiplicativité du déterminant, on a

$$\begin{aligned}\mathrm{dét}\,(X\,I_n - P^{-1}\cdot A\cdot P) &= \mathrm{dét}\,[P^{-1}\cdot(X\,I_n - A)\cdot P]\\ &= (\mathrm{dét}\,P)^{-1}\mathrm{dét}\,(X\,I_n - A)\,\mathrm{dét}\,P\\ &= \mathrm{dét}\,(X\,I_n - A).\end{aligned}$$

□

Remarques : 1. Il y a aussi une relation entre les vecteurs propres de A et ceux de $P^{-1}\cdot A\cdot P$: si x est un vecteur propre de A de valeur propre λ, alors

$$A\cdot x = x\,\lambda;$$

en multipliant les deux membres par P^{-1}, on obtient

$$P^{-1}\cdot A\cdot x = P^{-1}\cdot x\,\lambda,$$

c'est-à-dire

$$(P^{-1}\cdot A\cdot P)\cdot(P^{-1}\cdot x) = (P^{-1}\cdot x)\,\lambda,$$

donc $P^{-1}\cdot x$ est un vecteur propre de $P^{-1}\cdot A\cdot P$ de valeur propre λ. Inversement, si x est un vecteur propre de $P^{-1}\cdot A\cdot P$ de valeur propre λ, alors $P\cdot x$ est un vecteur propre de A de valeur propre λ.
2. D'un point de vue heuristique, il est clair *a priori* que les valeurs propres de A et de $P^{-1}\cdot A\cdot P$ sont les mêmes, puisque ces deux matrices représentent le même opérateur linéaire par rapport à des bases différentes.

12.2. Triangularisation

12.4. Proposition. *Pour une matrice carrée $A \in K^{n\times n}$, les conditions suivantes sont équivalentes :*

(a) *Le polynôme caractéristique se décompose en produit de facteurs linéaires, c'est-à-dire qu'il n'admet pas de facteur irréductible de degré > 1.*

(b) *Il existe une matrice régulière $P \in K^{n\times n}$ telle que la matrice $P^{-1}\cdot A\cdot P$ soit triangulaire supérieure.*

Si ces conditions sont satisfaites, les valeurs propres de A sont les entrées diagonales de la matrice $P^{-1}\cdot A\cdot P$.

Démonstration. (b) ⇒ (a) : Soit

$$P^{-1}\cdot A\cdot P = \begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}.$$

D'après la Proposition 12.3, on a $\mathrm{Pc}_A = \mathrm{Pc}_{P^{-1}\cdot A\cdot P}$, et ce polynôme est

$$\begin{aligned}\mathrm{dét}\,(X\,I_n - P^{-1}\cdot A\cdot P) &= \mathrm{dét}\begin{pmatrix}X-\lambda_1 & & * \\ & \ddots & \\ 0 & & X-\lambda_n\end{pmatrix}\\ &= (X-\lambda_1)\ldots(X-\lambda_n).\end{aligned}$$

Il se décompose donc en produit de facteurs linéaires. On voit de plus que les valeurs propres de A sont les entrées diagonales de $P^{-1} \cdot A \cdot P$.
(a) $\Rightarrow$ (b) : On raisonne par induction sur n. Si $n = 1$, on peut choisir $P = I_1$ puisque toute matrice d'ordre 1 est triangulaire supérieure. Pour la suite de la démonstration, on peut supposer que l'énoncé est vrai pour les matrices d'ordre $n - 1$. Comme Pc_A se décompose en produit de facteurs linéaires, ce polynôme admet (au moins) une racine. Soit $\lambda \in K$ une valeur propre de A et soit $v \in K^n$ un vecteur propre correspondant. On a $v \neq 0$, donc (v) est une suite libre de K^n, et l'on peut trouver des vecteurs $v_2, \ldots, v_n \in K^n$ tels que la suite $(v, v_2, \ldots, v_n)$ soit une base de K^n. Soit

$$P_1 = \begin{pmatrix} v & v_2 & \cdots & v_n \end{pmatrix} \in K^{n \times n}$$

la matrice dont les colonnes sont $v, v_2, \ldots, v_n$. Comme la suite $(v, v_2, \ldots, v_n)$ est une base de K^n, l'espace engendré par les colonnes de P_1 est K^n c'est-à-dire que $\mathcal{C}(P_1) = K^n$, donc $\operatorname{rang} P_1 = n$ et la matrice P_1 est inversible. Calculons le produit $P_1^{-1} \cdot A \cdot P$ en décomposant P_1 par colonnes :

$$\begin{aligned} P_1^{-1} \cdot A \cdot P_1 &= \begin{pmatrix} P_1^{-1} \cdot A \cdot v & P_1^{-1} \cdot A \cdot v_2 & \cdots & P_1^{-1} \cdot A \cdot v_n \end{pmatrix} \\ &= \begin{pmatrix} P_1^{-1} \cdot v\,\lambda & P_1^{-1} \cdot A \cdot v_2 & \cdots & P_1^{-1} \cdot A \cdot v_n \end{pmatrix}. \end{aligned}$$

Comme v est la première colonne de P_1, on a

$$v\,\lambda = P_1 \cdot \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

donc $P_1^{-1} \cdot v\,\lambda = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. La matrice $P_1^{-1} \cdot A \cdot P_1$ a donc la forme suivante

$$P_1^{-1} \cdot A \cdot P_1 = \left(\begin{array}{c|c} \lambda & u \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right)$$

pour une certaine ligne $u \in K^{1 \times (n-1)}$ et une certaine matrice carrée $B \in K^{(n-1) \times (n-1)}$. D'après la Proposition 12.3, on a $\mathrm{Pc}_A = \mathrm{Pc}_{P_1^{-1} \cdot A \cdot P_1}$; donc

$$\mathrm{Pc}_A(X) = \mathrm{dét} \left(\begin{array}{c|c} X - \lambda & -u \\ \hline 0 & \\ \vdots & X\, I_{n-1} - B \\ 0 & \end{array}\right).$$

En développant le déterminant suivant la première colonne, on obtient

$$\mathrm{Pc}_A(X) = (X - \lambda) \cdot \mathrm{Pc}_B(X),$$

donc tout facteur irréductible de Pc_B divise aussi Pc_A. D'après l'hypothèse sur Pc_A, le polynôme caractéristique Pc_B n'admet donc pas de facteur irréductible de degré > 1. On peut dès lors utiliser l'hypothèse d'induction pour trouver une matrice régulière Q d'ordre $n - 1$ telle que $Q^{-1} \cdot B \cdot Q$ soit triangulaire supérieure. On pose alors

$$P_2 = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right) \in K^{n \times n}.$$

Un calcul direct donne

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right) \cdot \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q \cdot Q^{-1} & \\ 0 & & & \end{array}\right) = I_n,$$

donc la matrice P_2 est régulière et

$$P_2^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right).$$

On a

$$\begin{aligned} P_2^{-1} \cdot P_1^{-1} \cdot A \cdot P_1 \cdot P_2 = & \\ = \quad & \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right) \cdot \left(\begin{array}{c|c} \lambda & u \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \cdot \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right) \\ = \quad & \left(\begin{array}{c|c} \lambda & u \cdot Q \\ \hline 0 & \\ \vdots & Q^{-1} \cdot B \cdot Q \\ 0 & \end{array}\right). \end{aligned}$$

Or, la matrice $Q^{-1} \cdot B \cdot Q$ est triangulaire supérieure ; il en est donc de même de $P_2^{-1} \cdot P_1^{-1} \cdot A \cdot P_1 \cdot P_2$, et la condition (b) est satisfaite avec $P = P_1 \cdot P_2$. □

12.5. Corollaire. *Pour toute matrice $A \in \mathbb{C}^{n\times n}$, il existe une matrice régulière $P \in \mathbb{C}^{n\times n}$ telle que $P^{-1} \cdot A \cdot P$ est triangulaire supérieure.*

Démonstration. Le “théorème fondamental de l'algèbre” indique que tout polynôme irréductible de $\mathbb{C}[X]$ est de degré 1. Par conséquent, la condition (a) (et donc aussi la condition (b)) de la proposition précédente est satisfaite pour toute matrice carrée complexe. □

En utilisant la réduction à la forme triangulaire, on peut donner une preuve élémentaire d'un théorème fondamental sur les polynômes matriciels. Si

$$f(X) = a_0 + a_1 X + \cdots + a_m X^m \in K[X]$$

et $A \in K^{n\times n}$, on note

$$f(A) = a_0 I_n + a_1 A + \cdots + a_m A^m \in K^{n\times n}.$$

Remarquons que pour toute matrice régulière $P \in K^{n\times n}$ et tout entier $k > 0$

$$\begin{aligned} (P^{-1} \cdot A \cdot P)^k &= \underbrace{(P^{-1} \cdot A \cdot P) \cdot (P^{-1} \cdot A \cdot P) \cdot \cdots \cdot (P^{-1} \cdot A \cdot P)}_{k} \\ &= P^{-1} \cdot \underbrace{A \cdot \cdots \cdot A}_{k} \cdot P \\ &= P^{-1} \cdot A^k \cdot P, \end{aligned}$$

donc

$$\begin{aligned} f(P^{-1} \cdot A \cdot P) &= a_0 I_n + a_1 P^{-1} \cdot A \cdot P + \cdots + a_m (P^{-1} \cdot A \cdot P)^m \\ &= P^{-1} \cdot (a_0 I_n + a_1 A + \cdots + a_m A^m) \cdot P \\ &= P^{-1} \cdot f(A) \cdot P. \end{aligned}$$

12.6. Théorème (Cayley–Hamilton). *Pour toute matrice carrée $A \in \mathbb{C}^{n\times n}$,*

$$\mathrm{Pc}_A(A) = 0.$$

Démonstration. D'après le Corollaire 12.5, on peut trouver une matrice régulière $P \in \mathbb{C}^{n\times n}$ telle que $P^{-1} \cdot A \cdot P$ soit triangulaire supérieure. Or,

$$\mathrm{Pc}_A(P^{-1} \cdot A \cdot P) = P^{-1} \cdot \mathrm{Pc}_A(A) \cdot P,$$

et $\mathrm{Pc}_A = \mathrm{Pc}_{P^{-1}\cdot A\cdot P}$ d'après la Proposition 12.3. Il suffit donc de prouver :

$$\mathrm{Pc}_{P^{-1}\cdot A\cdot P}(P^{-1} \cdot A \cdot P) = 0,$$

c'est-à-dire qu'il suffit de prouver le théorème pour les matrices triangulaires supérieures.

Pour cela, on raisonne par induction sur l'ordre n des matrices. Si $T = (a) \in \mathbb{C}^{1\times 1}$, alors $\mathrm{Pc}_T(X) = X - a$ et

$$\mathrm{Pc}_T(T) = (a - a) = 0 \in \mathbb{C}^{1\times 1},$$

ce qui prouve le théorème pour les matrices d'ordre 1.

Supposons le théorème vrai pour les matrices triangulaires d'ordre $n-1$ et soit $T \in \mathbb{C}^{n\times n}$ une matrice triangulaire supérieure d'ordre n. On décompose

$$T = \left(\begin{array}{c|c} \lambda & u \\ \hline 0 & \\ \vdots & T' \\ 0 & \end{array}\right)$$

où $u \in \mathbb{C}^{1\times(n-1)}$ et T' est une matrice triangulaire supérieure d'ordre $n-1$. Comme dans la démonstration de la Proposition 12.4, on a alors

$$\mathrm{Pc}_T(X) = (X - \lambda) \cdot \mathrm{Pc}_{T'}(X),$$

d'où

$$\mathrm{Pc}_T(T) = (T - \lambda\, I_n) \cdot \mathrm{Pc}_{T'}(T).$$

Or,

$$T - \lambda\, I_n = \left(\begin{array}{c|c} 0 & u \\ \hline 0 & \\ \vdots & T' - \lambda\, I_{n-1} \\ 0 & \end{array}\right)$$

et

$$\mathrm{Pc}_{T'}(T) = \left(\begin{array}{c|c} \mathrm{Pc}_{T'}(\lambda) & u' \\ \hline 0 & \\ \vdots & \mathrm{Pc}_{T'}(T') \\ 0 & \end{array}\right)$$

pour une certaine ligne $u' \in \mathbb{C}^{1\times(n-1)}$. L'hypothèse d'induction indique que $\mathrm{Pc}_{T'}(T') = 0$, donc les $n-1$ dernières lignes de la matrice $\mathrm{Pc}_{T'}(T)$ sont nulles. Dès lors,

$$\mathrm{Pc}_T(T) = \left(\begin{array}{c|c} 0 & u \\ \hline 0 & \\ \vdots & T' - \lambda\, I_{n-1} \\ 0 & \end{array}\right) \cdot \left(\begin{array}{c|c} \mathrm{Pc}_{T'}(\lambda) & u' \\ \hline 0 & \\ \vdots & 0 \\ 0 & \end{array}\right) = 0.$$

□

12.3. Diagonalisation

Dans ce numéro, A désigne un opérateur linéaire sur un espace E de dimension finie sur un corps K. On dit que l'opérateur A est *diagonalisable* s'il existe une base e de E par rapport à laquelle la matrice ${}_e(A)_e$ est diagonale. L'objectif est d'obtenir un critère permettant de reconnaître les opérateurs diagonalisables.

12.7. LEMME. *Soient $\lambda_1, \ldots, \lambda_s \in K$ des scalaires distincts deux à deux. La somme des espaces propres $E(\lambda_1) + \cdots + E(\lambda_s)$ est directe.*

DÉMONSTRATION. On raisonne par induction sur s. Si $s = 1$, il n'y a rien à prouver ; on peut donc supposer que toute somme d'au plus $s - 1$ espaces propres est directe. Soit

$$(*) \qquad x_1 + \cdots + x_s = 0 \qquad \text{avec } x_i \in E(\lambda_i) \text{ pour } i = 1, \ldots, s.$$

En appliquant A aux deux membres de cette égalité, on obtient

$$x_1\lambda_1 + \cdots + x_s\lambda_s = 0,$$

puisque $A(x_i) = x_i\lambda_i$. Par ailleurs, en multipliant par λ_s les deux membres de l'égalité $(*)$, on obtient

$$x_1\lambda_s + \cdots + x_s\lambda_s = 0.$$

En soustrayant membre à membre les deux dernières égalités, il vient

$$x_1(\lambda_1 - \lambda_s) + \cdots + x_{s-1}(\lambda_{s-1} - \lambda_s) = 0.$$

Comme $x_i(\lambda_i - \lambda_s) \in E(\lambda_i)$, on en déduit par l'hypothèse d'induction

$$x_i(\lambda_i - \lambda_s) = 0 \qquad \text{pour } i = 1, \ldots, s - 1,$$

d'où $x_i = 0$ pour $i = 1, \ldots, s-1$ car $\lambda_i - \lambda_s \neq 0$. En revenant à l'égalité $(*)$, on obtient aussi $x_s = 0$. $\square$

12.8. PROPOSITION. *Pour un opérateur linéaire A sur un espace E de dimension finie, les conditions suivantes sont équivalentes :*

(a) *A est diagonalisable.*

(b) *Il existe une base de E constituée de vecteurs propres de A.*

(c) *Si $\lambda_1, \ldots, \lambda_r$ sont les différentes valeurs propres de A,*

$$E = E(\lambda_1) \oplus \cdots \oplus E(\lambda_r).$$

(d) *Tous les facteurs irréductibles du polynôme caractéristique* Pc_A *sont du premier degré, et pour toute valeur propre λ de A,*

$$m_a(\lambda) = m_g(\lambda).$$

DÉMONSTRATION. On va prouver :

$$\text{(a)} \Leftrightarrow \text{(b)} \Rightarrow \text{(d)} \Rightarrow \text{(c)} \Rightarrow \text{(b)}.$$

La démonstration montrera de plus que pour une base $e = (e_1, \ldots, e_n)$ donnée, la matrice ${}_e(A)_e$ est diagonale si et seulement si les vecteurs $e_1, \ldots, e_n$ sont propres. La i-ème entrée diagonale α_i est alors la valeur propre associée à e_i.

(a) $\iff$ (b) : Supposons que $e = (e_1, \ldots, e_n)$ soit une base de E telle que ${}_e(A)_e$ soit diagonale :

$${}_e(A)_e = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

Comme les colonnes de cette matrice sont les coordonnées de $A(e_1), \ldots, A(e_n)$ par rapport à la base e, on a pour $i = 1, \ldots, n$

$$A(e_i) = e_1 0 + \cdots + e_i\alpha_i + \cdots + e_n 0 = e_i\alpha_i.$$

Cette relation montre que e_i est vecteur propre de A de valeur propre α_i.

Réciproquement, si $e = (e_1, \dots, e_n)$ est une base de E dont les éléments sont des vecteurs propres de A, soit $A(e_i) = e_i\alpha_i$, alors en plaçant en colonnes les coordonnées par rapport à la base e des vecteurs $A(e_1), \dots, A(e_n)$ on trouve une matrice diagonale :

$$_e(A)_e = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}.$$

(b) $\Rightarrow$ (d) : Soit e une base de A constituée de vecteurs propres de A. Soient $\lambda_1, \dots, \lambda_r \in K$ les différentes valeurs propres des vecteurs de la base e. Quitte à changer la numérotation des vecteurs de base, on peut rassembler les vecteurs propres de même valeur propre et supposer

$$e = (e_{11}, \dots, e_{1m_1}, e_{21}, \dots, e_{2m_2}, \dots, e_{r1}, \dots, e_{rm_r})$$

où pour tout $i = 1, \dots, r$ les vecteurs $e_{i1}, \dots, e_{im_i}$ sont des vecteurs propres de valeur propre λ_i :

$$A(e_{ij}) = e_{ij}\lambda_i \qquad \text{pour } j = 1, \dots, m_i.$$

La matrice $_e(A)_e$ est alors de la forme

$$_e(A)_e = \begin{pmatrix} \lambda_1 I_{m_1} & 0_{m_1,m_2} & \cdots & 0_{m_1,m_r} \\ 0_{m_2,m_1} & \lambda_2 I_{m_2} & \cdots & 0_{m_2,m_r} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{m_r,m_1} & 0_{m_r,m_2} & \cdots & \lambda_r I_{m_r} \end{pmatrix}.$$

Son polynôme caractéristique est facile à calculer :

$$\mathrm{Pc}_A(X) = \text{dét}\,(X\,I_n - {}_e(A)_e) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}.$$

Cette égalité prouve que tout facteur irréductible de Pc_A est de degré 1 et que A n'a pas d'autre valeur propre que $\lambda_1, \dots, \lambda_r$. On voit de plus

$$m_a(\lambda_i) = m_i \qquad \text{pour tout } i = 1, \dots, r.$$

Comme $e_{i1}, \dots, e_{im_i}$ sont linéairement indépendants et sont des vecteurs propres de valeur propre λ_i, la dimension de l'espace propre $E(\lambda_i)$ est au moins égale à m_i :

$$m_g(\lambda_i) \geqslant m_i = m_a(\lambda_i) \qquad \text{pour tout } i = 1, \dots, r.$$

Or, la Proposition 12.2 montre que la multiplicité géométrique d'une valeur propre ne peut dépasser la multiplicité algébrique, donc

$$m_a(\lambda_i) = m_g(\lambda_i) \qquad \text{pour tout } i = 1, \dots, r,$$

ce qui établit la condition (d).

(d) $\Rightarrow$ (c) : D'après le Lemme 12.7, la somme $E(\lambda_1) + \cdots + E(\lambda_r)$ est directe. On a donc

$$\dim(E(\lambda_1) \oplus \cdots \oplus E(\lambda_r)) = \dim E(\lambda_1) + \cdots + \dim E(\lambda_r) = m_g(\lambda_1) + \cdots + m_g(\lambda_r),$$

et il suffit pour établir (c) de prouver :

$$m_g(\lambda_1) + \cdots + m_g(\lambda_r) = \dim E.$$

D'après *(d)*, on a

$$\mathrm{Pc}_A(X) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

et $m_g(\lambda_i) = m_a(\lambda_i) = m_i$ pour tout $i = 1, \dots, r$. Comme le degré du polynôme caractéristique est égal à la dimension de l'espace, on a

$$\dim E = m_1 + \cdots + m_r = m_g(\lambda_1) + \cdots + m_g(\lambda_r).$$

(c) $\Rightarrow$ (b) : Cela résulte immédiatement de la Proposition 5.6, p. 33. □

12.9. Corollaire. *Si le polynôme caractéristique d'un opérateur linéaire se décompose en produit de facteurs linéaires distincts, alors l'opérateur est diagonalisable.*

Démonstration. L'hypothèse signifie que tous les facteurs irréductibles du polynôme caractéristique sont de degré 1 et que de plus $m_a(\lambda) = 1$ pour toute valeur propre λ. La suite de relations

$$1 = m_a(\lambda) \geqslant m_g(\lambda) \geqslant 1$$

qui se déduit de la Proposition 12.2 indique qu'alors

$$m_a(\lambda) = m_g(\lambda) = 1$$

pour toute valeur propre λ, donc la condition (d) de la proposition précédente est satisfaite, et l'opérateur est diagonalisable. □

Traduction matricielle. Le résultat suivant indique comment traduire en termes matriciels l'existence d'une base de vecteurs propres :

12.10. Lemme. *Soient $A, P \in K^{n\times n}$. On suppose que P est régulière. Pour que la matrice $P^{-1} \cdot A \cdot P$ soit diagonale, il faut et il suffit que les colonnes de P forment une base de K^n constituée de vecteurs propres de A. Les entrées diagonales de $P^{-1} \cdot A \cdot P$ sont alors les valeurs propres de A.*

Démonstration. Remarquons d'abord que P est régulière si et seulement si son rang est n, ce qui revient à dire que les n colonnes de P forment une suite génératrice et donc une base de K^n. Soit

$$P = \begin{pmatrix} p_{*1} & \cdots & p_{*n} \end{pmatrix}$$

la décomposition de P en colonnes. La condition

$$A \cdot p_{*i} = p_{*i}\,\lambda_i \qquad \text{pour } i = 1, \ldots, n$$

(avec $\lambda_i \in K$) est équivalente à

$$A \cdot P = \begin{pmatrix} p_{*1}\lambda_1 & \cdots & p_{*n}\lambda_n \end{pmatrix}.$$

Or, le membre de droite de cette dernière équation est le produit matriciel $P \cdot \Lambda$ où

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Comme l'égalité $A \cdot P = P \cdot \Lambda$ revient à

$$P^{-1} \cdot A \cdot P = \Lambda,$$

le lemme est démontré. □

Voici la traduction matricielle du critère permettant de reconnaître les opérateurs diagonalisables (Proposition 12.8) :

12.11. Proposition. *Pour une matrice carrée $A \in K^{n\times n}$, les conditions suivantes sont équivalentes :*

(a) *Il existe une matrice régulière $P \in K^{n\times n}$ telle que $P^{-1} \cdot A \cdot P$ soit une matrice diagonale.*

(b) *Il existe une base de K^n constituée de vecteurs propres de A.*

(c) *Tous les facteurs irréductibles du polynôme caractéristique Pc_A sont du premier degré, et pour toute valeur propre λ de A,*
$$m_a(\lambda) = m_g(\lambda).$$

Démonstration. (a) $\Leftrightarrow$ (b) : Cela résulte du lemme précédent.
(b) $\Leftrightarrow$ (c) : Soit c la base canonique de K^n et soit $\tilde{A}\colon K^n \to K^n$ l'opérateur linéaire associé à A. On a donc ${}_c(\tilde{A})_c = A$. L'équivalence de (b) et (c) résulte de la Proposition 12.8, car, par définition, l'opérateur $\tilde{A}$ et la matrice A ont le même polynôme caractéristique, les mêmes valeurs propres et les mêmes vecteurs propres. □

12.12. Corollaire. *Si le polynôme caractéristique d'une matrice $A \in K^{n\times n}$ se décompose en produit de facteurs linéaires distincts, alors il existe une matrice régulière $P \in K^{n\times n}$ telle que la matrice $P^{-1} \cdot A \cdot P$ soit diagonale.*

Démonstration. Comme dans la démonstration du Corollaire 12.9, l'hypothèse sur le polynôme caractéristique entraîne

$$m_a(\lambda) = m_g(\lambda) = 1$$

pour toute valeur propre de A ; le corollaire découle alors de l'implication (c) ⇒ (a) de la proposition précédente. □

Exercices

(12.1) (prérequis) Sans effectuer aucun calcul,

(a) Citez toutes les valeurs propres des matrices réelles suivantes :

$$A = \begin{pmatrix} 5 & 7 & 14 & -8 \\ 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 7 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}.$$

Citez également un vecteur propre de ces matrices et déterminez si elles sont diagonalisables.

(b) Citez une valeur propre et un vecteur propre des matrices complexes suivantes :

$$A = \begin{pmatrix} 4i & 9 & 0 \\ 0 & -2 & 8+4i \\ 0 & i & 7 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 & 0 \\ -i & 1+i & 0 \\ 2 & 2i & 2+3i \end{pmatrix}.$$

(12.2) Soit A un opérateur linéaire sur un espace de dimension n (finie).

(a) Montrez que le terme constant du polynôme caractéristique de Pc_A est $(-1)^n \operatorname{dét} A$.

(b) Montrez que 0 est valeur propre de A si et seulement si A est non inversible.

(c) Montrez que A et A^t ont les mêmes valeurs propres.

(12.3) Trouvez les valeurs propres et vecteurs propres des opérateurs linéaires suivants sur $\mathbb{R}^3$. Déterminez lesquels sont diagonalisables.

(a) $A(x, y, z) = (2x + y + z, 2x + 3y + 4z, -x - y - 2z)$.

(b) $A(x, y, z) = (2x - y + z, 3y - z, 2x + y + 3z)$.

(c) $A(x, y, z) = (2x + y + z, 2x + 3y + 2z, 3x + 3y + 4z)$.

(12.4) Pour chacune des matrices A suivantes, trouvez les valeurs propres réelles, une base de chaque sous-espace propre, et, lorsque cela est possible, une matrice réelle P telle que $P^{-1} \cdot A \cdot P$ soit diagonale.

$$\text{(a)} \begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix} \quad \text{(b)} \ \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -6 & 17 & -17 & 7 \end{pmatrix}$$

$$\text{(d)} \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(e)} \begin{pmatrix} 2-i & 0 & i \\ 0 & 1+i & 0 \\ i & 0 & 2-i \end{pmatrix}$$

(12.5) Cherchez les valeurs propres et les vecteurs propres des matrices A suivantes. Si possible, trouvez une matrice P telle que $P^{-1} \cdot A \cdot P$ soit diagonale.

$$\text{(a)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(d)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{(e)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(f)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{(g)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(12.6) Pour chacune des matrices A suivantes, trouvez une matrice régulière P telle que $P^{-1} \cdot A \cdot P$ soit triangulaire.

$$\text{(a)} \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{pmatrix}$$

(12.7) Soit

$$A_\alpha = \begin{pmatrix} 2 & \alpha & 1 \\ 0 & 2 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

où α est un paramètre réel.

1. En discutant sur la valeur de α, trouvez les valeurs propres de A_α et leur multiplicité algébrique, les vecteurs propres de A_α et une base des sous-espaces propres.
2. Trouvez l'ensemble des valeurs de α pour lesquelles A_α est diagonalisable.

13. Théorie spectrale

Dans cette section, on s'intéresse à des opérateurs linéaires sur des espaces euclidiens de dimension finie non nulle. Le corps des scalaires est donc $\mathbb{R}$ et les espaces considérés sont munis d'un produit scalaire.

13.1. Opérateur adjoint

13.1. Proposition. *Soit $A\colon E \to F$ une application linéaire entre deux espaces euclidiens de dimension finie, dont les produits scalaires sont notés respectivement $(\cdot|\cdot)_E$ et $(\cdot|\cdot)_F$. Il existe une et une seule application linéaire $A^*\colon F \to E$ telle que*

$$(*) \qquad (A(x)|y)_F = (x|A^*(y))_E \qquad \textit{pour tout } x \in E \textit{ et tout } y \in F.$$

De plus, si e est une base orthonormée de E et f une base orthonormée de F, alors

$$_e(A^*)_f = {}_f(A)_e^t.$$

L'application A^* est appelée *adjointe* de l'application linéaire A.

Démonstration. *Existence :* Fixons une base orthonormée e de E et une base orthonormée f de F, et montrons que l'application A^* dont la matrice par rapport aux bases e et f est la transposée de celle de A :

$$_e(A^*)_f = {}_f(A)_e^t$$

satisfait la propriété $(*)$.

Soient $x \in E$ et $y \in F$, de coordonnées $_e(x)$ et $_f(y)$. Les coordonnées de $A(x)$ et de $A^*(y)$ sont alors

$$_f(A(x)) = {}_f(A)_e \cdot {}_e(x)$$

et

$$_e(A^*(y)) = {}_e(A^*)_f \cdot {}_f(y) = {}_f(A)_e^t \cdot {}_f(y).$$

Comme les bases e et f sont orthonormées, on a d'après le Lemme 10.3, p. 71,

$$\begin{aligned} (A(x)|y)_F &= [{}_f(A)_e \cdot {}_e(x)]^t \cdot {}_f(y) \\ &= {}_e(x)^t \cdot {}_f(A)_e^t \cdot {}_f(y) \end{aligned}$$

et, de même,

$$\begin{aligned} (x|A^*(y))_E &= {}_e(x)^t \cdot {}_e(A^*(y)) \\ &= {}_e(x)^t \cdot {}_f(A)_e^t \cdot {}_f(y). \end{aligned}$$

La propriété $(*)$ est donc établie.

Unicité : Supposons que $B, C \colon F \to E$ soient deux applications linéaires satisfaisant la propriété $(*)$. Alors pour tout $x \in E$ et tout $y \in F$,

$$(x|B(y))_E = (A(x)|y)_F = (x|C(y))_E,$$

donc la différence entre le membre de gauche et le membre de droite est nulle :

$$(x|B(y) - C(y))_E = 0 \qquad \text{pour tout } x \in E \text{ et tout } y \in F.$$

En particulier, pour $x = B(y) - C(y)$ on trouve

$$\| B(y) - C(y) \|^2 = 0,$$

donc

$$B(y) - C(y) = 0 \qquad \text{pour tout } y \in F,$$

ce qui signifie : $B = C$. □

Comme, en termes de matrices par rapport à des bases orthonormées, l'adjonction des applications linéaires se ramène à la transposition des matrices, on peut établir pour les applications adjointes des propriétés analogues à celles des matrices transposées :

Propriétés.

1. Pour $A, B \colon E \to F$ et $\alpha, \beta \in \mathbb{R}$,
$$(\alpha A + \beta B)^* = \alpha A^* + \beta B^*.$$
2. Pour $A \colon E \to F$ et $B \colon F \to G$,
$$(B \circ A)^* = A^* \circ B^*.$$
3. Pour $A \colon E \to F$,
$$(A^*)^* = A.$$

Ces propriétés se démontrent soit en comparant les matrices par rapport à des bases orthonormées soit en revenant à la propriété caractéristique $(*)$ des applications adjointes. Montrons par exemple la propriété 3 : si e et f sont des bases orthonormées de E et F respectivement, on a

$$_f((A^*)^*)_e = {}_e(A^*)^t_f = ({}_f(A)^t_e)^t = {}_f(A)_e;$$

donc $(A^*)^*$ et A ont la même matrice par rapport aux bases e et f, et par conséquent $(A^*)^* = A$. On peut aussi raisonner comme suit : par définition, $(A^*)^*$ est l'unique application linéaire $(A^*)^* \colon E \to F$ telle que pour tout $x \in E$ et tout $y \in F$,

$$(A^*(y)|x)_E = (y|(A^*)^*(x))_F.$$

Or, l'application A possède cette propriété puisque, d'après $(*)$,

$$(y|A(x))_F = (A(x)|y)_F = (x|A^*(y))_E = (A^*(y)|x)_E.$$

On doit donc avoir $(A^*)^* = A$. □

13.2. Théorème spectral

Le théorème spectral indique que les opérateurs pour lesquels il existe une base *orthonormée* de vecteurs propres sont exactement les opérateurs qui sont leur propre adjoint. Nous commençons par montrer que le polynôme caractéristique d'un opérateur auto-adjoint n'a pas de facteur irréductible de degré 2.

Pour toute matrice complexe $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \in \mathbb{C}^{m \times n}$, on note $\overline{A}$ la matrice dont les entrées sont les nombres complexes conjugués des entrées de A :

$$\overline{A} = (\overline{a_{ij}})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \in \mathbb{C}^{m \times n}.$$

En particulier, si A est une matrice réelle, alors $\overline{A} = A$.

13.2. LEMME. *Soit $A \in \mathbb{C}^{n\times n}$. Si $\overline{A} = A^t$, alors toutes les valeurs propres de A sont réelles ; en particulier, toutes les valeurs propres dans $\mathbb{C}$ d'une matrice symétrique réelle sont réelles, donc son polynôme caractéristique n'a pas de facteur irréductible de degré 2 dans $\mathbb{R}[X]$.*

DÉMONSTRATION. Soit $x \in \mathbb{C}^n = \mathbb{C}^{n\times 1}$ un vecteur propre de A de valeur propre λ :

$$A \cdot x = x\,\lambda.$$

En multipliant à gauche par $\overline{x}^t$, on obtient

$$\overline{x}^t \cdot A \cdot x = \overline{x}^t \cdot x\,\lambda.$$

En transposant et en conjuguant les deux membres, on obtient par ailleurs

$$\overline{x}^t \cdot \overline{A}^t \cdot x = \overline{x}^t \cdot x\,\overline{\lambda},$$

donc, comme $\overline{A}^t = A$,

$$(\dagger) \qquad \overline{x}^t \cdot x\,\lambda = \overline{x}^t \cdot x\,\overline{\lambda}.$$

Or, si $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, alors

$$\overline{x}^t \cdot x = \overline{x_1}x_1 + \cdots + \overline{x_n}x_n = |x_1|^2 + \cdots + |x_n|^2,$$

en notant $|z|$ le module d'un nombre complexe z. Comme $x \neq 0$ puisque x est un vecteur propre, on a $\overline{x}^t \cdot x \neq 0$, donc l'égalité (†) entraîne

$$\lambda = \overline{\lambda}.$$

□

13.3. THÉORÈME SPECTRAL. *Pour un opérateur linéaire A sur un espace euclidien E de dimension finie, les conditions suivantes sont équivalentes :*

(a) *Il existe une base orthonormée e de E telle que la matrice ${}_e(A)_e$ soit diagonale.*

(b) *Il existe une base orthonormée e de E constituée de vecteurs propres de A.*

(c) *Si $\lambda_1, \ldots, \lambda_r$ sont les différentes valeurs propres de A, l'espace E se décompose en somme directe orthogonale :*

$$E = E(\lambda_1) \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} E(\lambda_r),$$

c'est-à-dire que

$$E = E(\lambda_1) \oplus \cdots \oplus E(\lambda_r)$$

et que de plus pour $i \neq j$ tout vecteur de $E(\lambda_i)$ est orthogonal à tout vecteur de $E(\lambda_j)$.

(d) *L'opérateur A est auto-adjoint :*

$$A^* = A.$$

DÉMONSTRATION. (a) $\iff$ (b) : Les arguments utilisés dans la démonstration de l'équivalence (a) $\iff$ (b) de la Proposition 12.8 s'appliquent ici. Ils montrent qu'une base e de E est constituée de vecteurs propres de A si et seulement si la matrice ${}_e(A)_e$ est diagonale.
(a) $\Rightarrow$ (d) : Comme e est une base orthonormée, on a ${}_e(A^*)_e = {}_e(A)_e^t$. Comme de plus la matrice ${}_e(A)_e$ est diagonale, on en déduit

$${}_e(A^*)_e = {}_e(A)_e,$$

donc $A^* = A$ puisqu'un opérateur linéaire est déterminé de manière unique par sa matrice par rapport à la base e.
(d) $\Rightarrow$ (b) : On raisonne par induction sur la dimension de E. Si $\dim E = 1$, il n'y a rien à prouver puisque tous les vecteurs non nuls de E sont propres. Supposons $\dim E = n > 1$. Vu le Lemme 13.2, on peut trouver dans E un vecteur propre de A. Quitte à le diviser par sa norme, on peut supposer que ce vecteur est normé. Notons-le e_1, et considérons son orthogonal dans E :

$$V = \mathrm{sev}\langle e_1\rangle^\perp.$$

C'est un sous-espace de dimension $n-1$ (voir le Corollaire 11.3, p. 74). Pour appliquer l'hypothèse d'induction, on va prouver que la restriction de A à V :

$$A|_V \colon V \to E \quad x \mapsto A(x)$$

est en fait un opérateur linéaire sur V, c'est-à-dire que $A(V) \subset V$.

Soit $A(e_1) = e_1\lambda_1$. Comme A est auto-adjoint, on a pour $x \in E$

$$\big(e_1|A(x)\big) = \big(A(e_1)|x\big) = \lambda_1(e_1|x).$$

Dès lors, $\big(e_1|A(x)\big) = 0$ si $(e_1|x) = 0$, c'est-à-dire si $x \in V$. On a donc bien $A(V) \subset V$.

Dire que A est auto-adjoint revient à dire que $\big(A(x)|y\big) = \big(x|A(y)\big)$ pour x, $y \in E$. Si cette condition est satisfaite pour x, $y \in E$, elle l'est *a fortiori* pour x, $y \in V$, donc $A|_V$ est un opérateur auto-adjoint sur V. Par hypothèse d'induction, il existe une base orthonormée $(e_2, \ldots, e_n)$ de V constituée de vecteurs propres de $A|_V$, donc de A. Comme $E = \mathrm{sev}\langle e_1\rangle \oplus V$ (voir le Corollaire 11.3, p. 74) et que e_1 est orthogonal à tout vecteur de V, la suite $(e_1, e_2, \ldots, e_n)$ est une base orthonormée de E constituée de vecteurs propres de A.

(b) $\Rightarrow$ (c) : Vu la Proposition 12.8, la condition (b) entraîne

$$E = E(\lambda_1) \oplus \cdots \oplus E(\lambda_r)$$

où $\lambda_1, \ldots, \lambda_r$ sont les différentes valeurs propres de A. Il ne reste donc plus qu'à prouver

$$(x|y) = 0 \qquad \text{si } x \in E(\lambda_i) \text{ et } y \in E(\lambda_j) \text{ avec } i \neq j.$$

Calculons pour cela $(A(x)|y)$. Comme $x \in E(\lambda_i)$, on a $A(x) = x\lambda_i$, donc

$$(A(x)|y) = \lambda_i(x|y). \tag{‡}$$

Par ailleurs, on a $(A(x)|y) = (x|A^*(y))$. Comme on a déjà prouvé que la condition (b) entraîne $A^* = A$, et comme $A(y) = y\lambda_j$ puisque $y \in E(\lambda_j)$, on en déduit

$$(A(x)|y) = (x|y)\,\lambda_j. \tag{§}$$

Si $i \neq j$, alors $\lambda_i \neq \lambda_j$ et les égalités (‡) et (§) entraînent $(x|y) = 0$.

(c) $\Rightarrow$ (b) : En juxtaposant des bases orthonormées de $E(\lambda_1), \ldots, E(\lambda_r)$, on obtient une base orthonormée de E constituée de vecteurs propres de A. □

Traduction matricielle. Pour traduire le théorème 13.3 en termes de matrices, on passe, comme dans la section 12.3, par l'intermédiaire de l'opérateur linéaire associé à une matrice carrée.

Soit $A \in \mathbb{R}^{n\times n}$ et soit $\tilde{A} \colon \mathbb{R}^n \to \mathbb{R}^n$ l'opérateur linéaire associé. On munit $\mathbb{R}^n$ du produit scalaire usuel, pour lequel la base canonique c est orthonormée. Comme A est la matrice de $\tilde{A}$ par rapport à la base canonique, on a, par la Proposition 13.1,

$$_c(\tilde{A}^*)_c = A^t, \tag{¶}$$

donc $\tilde{A}$ est auto-adjoint si et seulement si $A = A^t$.

Pour traduire en termes matriciels les changements de bases orthonormées de $\mathbb{R}^n$, on utilise le résultat suivant :

13.4. Lemme. *Les colonnes d'une matrice carrée $Q \in \mathbb{R}^{n\times n}$ forment une base orthonormée de $\mathbb{R}^n$ si et seulement si la matrice Q est inversible et $Q^{-1} = Q^t$.*

Les matrices qui possèdent les propriétés de l'énoncé sont appelées *matrices orthogonales*.[1]

Démonstration. On sait que Q est inversible si et seulement si $\operatorname{rang} Q = n$ (voir le Corollaire 6.6). Comme le rang d'une matrice est la dimension de l'espace engendré par ses colonnes, cette condition est réalisée si et seulement si les colonnes de Q forment une base de $\mathbb{R}^n$. Pour établir que

[1] Pour qu'une matrice de $\mathbb{R}^{n\times n}$ soit orthogonale, il ne suffit donc pas que ses colonnes forment une base *orthogonale* de $\mathbb{R}^n$!

cette base est orthonormée si et seulement si $Q^t = Q^{-1}$, on calcule le produit Q^tQ en décomposant Q par colonnes. Soit $Q = \begin{pmatrix} q_{*1} & \cdots & q_{*n} \end{pmatrix}$; alors

$$Q^t \cdot Q = \begin{pmatrix} q_{*1}^t \\ \vdots \\ q_{*n}^t \end{pmatrix} \cdot \begin{pmatrix} q_{*1} & \cdots & q_{*n} \end{pmatrix} = \begin{pmatrix} q_{*1}^t \cdot q_{*1} & \cdots & q_{*1}^t \cdot q_{*n} \\ \vdots & & \vdots \\ q_{*n}^t \cdot q_{*1} & \cdots & q_{*n}^t \cdot q_{*n} \end{pmatrix}.$$

On a donc $Q^tQ = I_n$ si et seulement si $q_{*i}^t \cdot q_{*j} = \delta_{ij}$, ce qui revient à dire que la base $(q_{*1}, \ldots, q_{*n})$ est orthonormée, puisque $q_{*i}^t \cdot q_{*j} = (q_{*i}|q_{*j})$. Comme la matrice Q est carrée, la condition $Q^tQ = I_n$ est équivalente à $Q^t = Q^{-1}$ par le Théorème 2.8. □

Voici à présent la version matricielle du théorème spectral :

13.5. Théorème spectral. *Pour une matrice réelle $A \in \mathbb{R}^{n\times n}$, les conditions suivantes sont équivalentes :*

(a) *Il existe une matrice orthogonale $Q \in \mathbb{R}^{n\times n}$ telle que $Q^{-1} \cdot A \cdot Q$ soit une matrice diagonale.*

(b) *Il existe une base orthonormée de $\mathbb{R}^n$ pour le produit scalaire usuel constituée de vecteurs propres de A.*

(c) *La matrice A est symétrique :*

$$A^t = A.$$

De plus, les matrices orthogonales Q satisfaisant la condition (a) *sont les matrices dont les colonnes forment une base orthonormée de $\mathbb{R}^n$ constituée de vecteurs propres de A et les entrées diagonales de $Q^{-1} \cdot A \cdot Q$ sont les valeurs propres de A.*

Démonstration. (a) $\iff$ (b) : Cela résulte des Lemmes 12.10 et 13.4.
(a) $\iff$ (c) : Cela découle du Théorème 13.3, appliqué à l'opérateur $\tilde{A}\colon \mathbb{R}^n \to \mathbb{R}^n$ associé à A. En effet, cet opérateur est auto-adjoint si et seulement si A est symétrique, vu la relation (¶). □

Remarque : On peut aussi donner une démonstration directe de l'équivalence (b) $\iff$ (c). Vu le Lemme 13.2, le polynôme caractéristique $\mathrm{Pc}_A(X) \in \mathbb{R}[X]$ se décompose en produit de facteurs réels de degré 1. La Proposition 12.4 livre alors une matrice régulière $P \in \mathbb{R}^{n\times n}$ telle que $P^{-1} \cdot A \cdot P$ soit triangulaire supérieure ; soit

$$(\|) \qquad P^{-1} \cdot A \cdot P = T.$$

D'après la Proposition 11.7 (p. 77), on peut factoriser

$$P = Q \cdot R$$

où $Q \in \mathbb{R}^{n\times n}$ est une matrice orthogonale et $R \in \mathbb{R}^{n\times n}$ est une matrice triangulaire supérieure dont les entrées diagonales sont non nulles. De l'équation (‖) on tire alors

$$Q^{-1} \cdot A \cdot Q = R \cdot T \cdot R^{-1}.$$

Or, l'inverse d'une matrice triangulaire supérieure est une matrice triangulaire supérieure et tout produit de matrices triangulaires supérieures est une matrice triangulaire supérieure, donc $Q^{-1} \cdot A \cdot Q$ est une matrice triangulaire supérieure.

Par ailleurs, comme A est symétrique, un calcul montre que $Q^{-1} \cdot A \cdot Q$ est aussi symétrique. Dès lors, $Q^{-1} \cdot A \cdot Q$ est diagonale puisqu'elle est à la fois triangulaire supérieure et symétrique.

Pour établir la réciproque, il suffit de remarquer que les matrices diagonales sont symétriques et que si $Q^{-1} \cdot A \cdot Q$ est symétrique pour une matrice Q orthogonale, alors A est symétrique.

Exercices

(13.1) Pour chacune des matrices A suivantes, trouvez une matrice orthogonale P telle que $P^{-1} \cdot A \cdot P$ soit diagonale :

$$\text{(a)} \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} \qquad \text{(b)} \begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix}$$

(13.2) Soit A un opérateur sur un espace E de dimension finie. Démontrez : $\operatorname{Ker} A^* = (\operatorname{Im} A)^\perp$ et $\operatorname{Im} A^* = (\operatorname{Ker} A)^\perp$.

(13.3) Soit $Q \in \mathbb{R}^{n\times n}$. Montrer que Q est orthogonale si et seulement si Q^t est orthogonale. En déduire qu'une matrice est orthogonale si et seulement si ses lignes forment une base orthonormée de vecteurs propres de $\mathbb{R}^n$ (pour le produit scalaire usuel).

(13.4) Soit $P\colon E \to E$ un opérateur linéaire sur un espace euclidien de dimension finie. Montrer que P est une projection orthogonale sur un sous-espace de E si et seulement si P est auto-adjoint et $P^2 = P$.

(13.5) Soient P, P' des opérateurs linéaires sur un espace euclidien E de dimension finie. On suppose que P est la projection orthogonale sur un sous-espace $V \subset E$ et que P' est la projection orthogonale sur un sous-espace $V' \subset E$. Montrer que V et V' sont orthogonaux si et seulement si $P \circ P' = P' \circ P = 0$.

(13.6) (pour aller plus loin) En utilisant les deux exercices précédents, montrer que l'on peut exprimer le théorème spectral sous la forme suivante : pour tout opérateur auto-adjoint A sur un espace euclidien E de dimension finie, il existe des opérateurs $P_1, \ldots, P_r$ et des nombres réels $\lambda_1, \ldots, \lambda_r$ distincts deux à deux tels que $P_i \circ P_j = 0$ pour $i \neq j$, $P_i^2 = P_i = P_i^*$, $P_1 + \cdots + P_r = I_E$ et $A = \lambda_1 P_1 + \cdots + \lambda_R P_r$.

CHAPITRE VII

Formes quadratiques

Une *forme quadratique* réelle en n indéterminées $x_1, \ldots, x_n$ est un polynôme homogène de degré 2 à coefficients dans $\mathbb{R}$. Un tel polynôme s'écrit sous la forme

$$q(x_1, \ldots, x_n) = \sum_{1 \leqslant i \leqslant j \leqslant n} a_{ij} x_i x_j.$$

Deux formes quadratiques $q(x_1, \ldots, x_n)$ et $q'(x'_1, \ldots, x'_n)$ sont déclarées *équivalentes* si l'on peut passer de l'une à l'autre par un changement linéaire de variables, c'est-à-dire par un changement de variables du type

$$x'_i = \sum_{j=1}^{n} p_{ij} x_j \qquad x_i = \sum_{j=1}^{n} p'_{ij} x'_j$$

pour certains coefficients $p_{ij}, p'_{ij} \in \mathbb{R}$.

Le premier objectif de ce chapitre est de classifier les formes quadratiques réelles à un changement linéaire de variables près. Dans la deuxième partie, on développe un point de vue plus géométrique en montrant que les formes quadratiques correspondent à des produits scalaires généralisés.

14. Classification des formes quadratiques réelles

À toute forme quadratique

$$q(x_1, \ldots, x_n) = \sum_{1 \leqslant i \leqslant j \leqslant n} a_{ij} x_i x_j \in \mathbb{R}[x_1, \ldots, x_n]$$

on associe la matrice symétrique

$$A = \begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \cdots & \frac{a_{2n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \cdots & a_{nn} \end{pmatrix},$$

qui est la seule matrice symétrique telle que

$$q(x_1, \ldots, x_n) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Cette dernière égalité se note simplement

$$q(X) = X^t \cdot A \cdot X \qquad \text{où } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Pour préciser quelle relation sur les matrices de coefficients correspond à l'équivalence des formes quadratiques, considérons une autre forme quadratique

$$q'(x'_1, \ldots, x'_n) = \sum_{1 \leqslant i \leqslant j \leqslant n} a'_{ij} x'_i x'_j,$$

de matrice symétrique associée A', de sorte que

$$q(X') = {X'}^t \cdot A' \cdot X'.$$

Un changement linéaire de variables s'exprime sous la forme

$$X' = P \cdot X, \qquad X = P' \cdot X'$$

où P, $P' \in \mathbb{R}^{n\times n}$. Ces relations indiquent que $P \cdot P' = P' \cdot P = I_n$, donc P et P' sont deux matrices inverses l'une de l'autre. La condition que le changement de variables transforme la forme $q'(X')$ en $q(X)$ s'exprime comme suit :

$$q'(P \cdot X) = q(X)$$

c'est-à-dire

$$(P \cdot X)^t \cdot A' \cdot P \cdot X = X^t \cdot A \cdot X.$$

Comme une forme quadratique détermine de manière unique la matrice symétrique associée, l'égalité

$$X^t \cdot P^t \cdot A' \cdot P \cdot X = X^t \cdot A \cdot X$$

ne peut avoir lieu que si $P^t \cdot A' \cdot P = A$. L'équivalence des formes quadratiques revient donc à l'équivalence des matrices symétriques associées sous la relation suivante : les matrices symétriques A et A' sont *congruentes* s'il existe une matrice régulière P telle que

$$A = P^t \cdot A' \cdot P.$$

On peut dès lors utiliser les résultats des chapitres précédents sur les matrices symétriques pour obtenir des informations concernant les formes quadratiques.

14.1. Diagonalisation

14.1. Proposition. *Toute forme quadratique réelle en n indéterminées est équivalente à une forme diagonale*

$$a_1x_1^2 + \cdots + a_nx_n^2$$

pour certains scalaires $a_1, \ldots, a_n \in \mathbb{R}$.

Démonstration. Soit $q(X) = X^t \cdot A \cdot X$ une forme quadratique en n indéterminées, de matrice symétrique associée A. D'après le théorème spectral 13.5, on peut trouver une matrice orthogonale Q telle que $Q^{-1} \cdot A \cdot Q$ soit diagonale ; soit

$$Q^{-1} \cdot A \cdot Q = \Lambda,$$

les entrées diagonales de Λ étant les valeurs propres $\lambda_1, \ldots, \lambda_n$ de A. Comme $Q^t = Q^{-1}$, le changement de variables

$$X' = Q^t \cdot X \qquad X = Q \cdot X'$$

transforme la forme quadratique $q(X)$ en

$$q'(X') = q(Q \cdot X') = X'^t \cdot Q^t \cdot A \cdot Q \cdot X' = X'^t \cdot \Lambda \cdot X';$$

donc $q(X)$ est équivalente à la forme

$$q'(X') = \lambda_1 x'^2_1 + \cdots + \lambda_n x'^2_n.$$

□

Techniques de diagonalisation. Plusieurs méthodes permettent de réduire une forme quadratique réelle $q(X) = X^t \cdot A \cdot X$ en n indéterminées, de matrice symétrique associée $A \in \mathbb{R}^{n\times n}$, à la forme diagonale.

1. La méthode la moins économique est celle utilisée dans la démonstration de la Proposition 14.1. Elle consiste à trouver une base orthonormée de $\mathbb{R}^n$ constituée de vecteurs propres de A. Si Q est la matrice dont les colonnes sont les vecteurs de cette base orthonormée, alors le changement de variables $X = Q \cdot X'$ réduit la forme quadratique $q(X)$ à la forme diagonale

$$q'(X') = \lambda_1 x'^2_1 + \cdots + \lambda_n x'^2_n$$

où $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A.

2. On peut utiliser les opérations élémentaires sur la matrice A, à condition de faire suivre chaque opération élémentaire sur les lignes par la même opération élémentaire sur les colonnes. Cela revient en effet à multiplier la matrice A à gauche par une matrice élémentaire E et à droite par sa transposée E^t. On obtient ainsi une nouvelle matrice $A' = E \cdot A \cdot E^t$ qui est congruente à la matrice A ; la forme quadratique correspondante est donc équivalente à la forme quadratique donnée.

3. Un procédé particulièrement économique est la méthode de *complétion des carrés*. Il s'agit d'une suite de changements de variables par lesquels les termes rectangles $x_i x_j$ (avec $i \neq j$) sont successivement éliminés.

Supposons d'abord que les coefficients des termes carrés x_i^2 ne soient pas tous nuls. Soit par exemple $a_{11}^2 \neq 0$. Écrivons alors la forme quadratique donnée en rassemblant tous les termes qui ne contiennent pas x_1 :

$$q(x_1, \ldots, x_n) = a_{11}x_1^2 + (a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n) + q_1(x_2, \ldots, x_n).$$

Comme on a supposé $a_{11} \neq 0$, on peut poser

$$x_1' = x_1 + \frac{a_{12}}{2a_{11}}x_2 + \cdots + \frac{a_{1n}}{2a_{11}}x_n,$$

de sorte que

$${x_1'}^2 = x_1^2 + \frac{1}{a_{11}}(a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n) + \left(\frac{a_{12}}{2a_{11}}x_2 + \cdots + \frac{a_{1n}}{2a_{11}}x_n\right)^2;$$

alors

$$q(x_1, \ldots, x_n) = a_{11}{x_1'}^2 + q_1(x_2, \ldots, x_n) - \frac{1}{4a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n)^2.$$

En posant $q_2(x_2, \ldots, x_n) = q_1(x_2, \ldots, x_n) - \frac{1}{4a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n)^2$, on a donc

$$q(x_1, \ldots, x_n) = a_{11}{x_1'}^2 + q_2(x_2, \ldots, x_n),$$

et l'on peut envisager de répéter l'opération précédente sur la forme quadratique $q_2(x_2, \ldots, x_n)$, dont le nombre de variables est $n - 1$.

Cela n'est possible *a priori* que si les coefficients des termes carrés de q_2 ne sont pas tous nuls. Cependant, il est possible de modifier la procédure ci-dessus pour faire face à la situation où les coefficients des termes carrés sont tous nuls.

Supposons par exemple $a_{11} = a_{22} = 0$ et $a_{12} \neq 0$. Le changement de variables

$$\begin{aligned} x_1 &= x_1' + x_2' \\ x_2 &= x_1' - x_2' \end{aligned}$$

fait alors apparaître les termes carrés

$$a_{12}{x_1'}^2 - a_{12}{x_2'}^2$$

et ramène à la forme précédente.

On peut formaliser cette méthode tout en explicitant la condition pour qu'il ne soit pas nécessaire de recourir à la modification ci-dessus :

14.2. Proposition. *Soit* $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathbb{R}^{n \times n}$ *une matrice symétrique. Pour* $k = 1, \ldots, n$, *on considère la sous-matrice symétrique*

$$A_k = (a_{ij})_{1 \leqslant i,j \leqslant k}$$

et on pose

$$\delta_k = \textit{dét}\, A_k \qquad \textit{pour } k = 1, \ldots, n.$$

Soit aussi

$$\delta_0 = 1.$$

Si $\delta_k \neq 0$ pour $k = 1, \ldots, n$, alors il existe une matrice triangulaire supérieure $R \in \mathbb{R}^{n\times n}$ dont toutes les entrées diagonales sont égales à 1 telle que

$$R^t \cdot A \cdot R = \begin{pmatrix} \delta_1/\delta_0 & & & 0 \\ & \delta_2/\delta_1 & & \\ & & \ddots & \\ 0 & & & \delta_n/\delta_{n-1} \end{pmatrix}.$$

En particulier, la matrice R est régulière et le changement de variables $X = R \cdot X'$ transforme la forme quadratique $q(X) = X^t \cdot A \cdot X$ en

$$q'(X') = \frac{\delta_1}{\delta_0}\, {x'}_1^2 + \cdots + \frac{\delta_n}{\delta_{n-1}}\, {x'}_n^2.$$

Démonstration. On raisonne par induction sur n. Si $n = 1$, il suffit de prendre $R = (1)$. Supposons donc la propriété établie pour les matrices d'ordre $n-1$. On peut l'appliquer à A_{n-1} pour trouver une matrice triangulaire R_{n-1} dont toutes les entrées diagonales sont égales à 1 et telle que

$$R_{n-1}^t \cdot A_{n-1} \cdot R_{n-1} = \begin{pmatrix} \delta_1/\delta_0 & & 0 \\ & \ddots & \\ 0 & & \delta_{n-1}/\delta_{n-2} \end{pmatrix}.$$

Posons

$$R = \left(\begin{array}{c|c} R_{n-1} & v \\ \hline 0 & 1 \end{array}\right)$$

où $v \in \mathbb{R}^{n-1}$ est un vecteur que l'on se propose de déterminer de telle sorte que les conditions requises soient satisfaites. Quel que soit v, la matrice R est triangulaire supérieure avec des 1 sur la diagonale. Pour déterminer sous quelle condition la matrice $R^t \cdot A \cdot R$ est diagonale, décomposons la matrice A de manière semblable à R :

$$A = \left(\begin{array}{c|c} A_{n-1} & u \\ \hline u^t & a_{nn} \end{array}\right).$$

Un calcul par blocs donne

$$(*) \qquad R^t \cdot A \cdot R = \left(\begin{array}{c|c} R_{n-1}^t \cdot A_{n-1} \cdot R_{n-1} & R_{n-1}^t \cdot (A_{n-1} \cdot v + u) \\ \hline (v^t \cdot A_{n-1} + u^t) \cdot R_{n-1} & x \end{array}\right)$$

où $x = v^t \cdot A_{n-1} \cdot v + v^t \cdot u + u^t \cdot v + a_{nn}$. Comme par hypothèse la matrice A_{n-1} est régulière, le système linéaire

$$A_{n-1} \cdot X + u = 0$$

admet une solution. Si l'on choisit pour v cette solution, alors le coin supérieur droit de $R^t \cdot A \cdot R$ dans la décomposition $(*)$ est nul. Le coin inférieur gauche l'est aussi, puisque la matrice $R^t \cdot A \cdot R$ est symétrique. Comme la matrice $R_{n-1}^t \cdot A_{n-1} \cdot R_{n-1}$ est diagonale, il en est de même de $R^t \cdot A \cdot R$; soit

$$(\dagger) \qquad R^t \cdot A \cdot R = \begin{pmatrix} \delta_1/\delta_0 & & & 0 \\ & \ddots & & \\ & & \delta_{n-1}/\delta_{n-2} & \\ 0 & & & x \end{pmatrix}.$$

Pour achever la démonstration, il suffit de prouver que $x = \delta_n/\delta_{n-1}$.

Comparons pour cela les déterminants des deux membres de $(\dagger)$. Comme R est une matrice triangulaire avec des 1 sur la diagonale, on a $\operatorname{dét} R = 1$, donc

$$\operatorname{dét} A = \operatorname{dét}(R^t \cdot A \cdot R) = \frac{\delta_1}{\delta_0} \cdot \frac{\delta_2}{\delta_1} \cdot \cdots \cdot \frac{\delta_{n-1}}{\delta_{n-2}} \cdot x.$$

Comme $\operatorname{dét} A = \delta_n$ et $\delta_0 = 1$, cette dernière relation peut s'écrire $\delta_n = \delta_{n-1} \cdot x$, donc

$$x = \frac{\delta_n}{\delta_{n-1}}.$$

□

14.2. Invariants

L'objectif de cette section est de dégager des invariants de formes quadratiques réelles qui permettent de classifier celles-ci à équivalence près. Le premier de ces invariants utilise une notion connue :

Rang.

14.3. PROPOSITION. *Le rang de la matrice symétrique associée à une forme quadratique ne varie pas par changement de variables.*

DÉMONSTRATION. Il suffit de prouver que des matrices congruentes ont le même rang. Soit

$$A' = P^t \cdot A \cdot P$$

où P est une matrice régulière. Comme le rang d'un produit de matrices est majoré par le rang de chacun des facteurs, on a

$$\operatorname{rang} A' \leqslant \operatorname{rang} A.$$

Cependant, comme P est régulière, on a aussi

$$(P^t)^{-1} \cdot A' \cdot P^{-1} = A$$

et le même argument donne

$$\operatorname{rang} A' \geqslant \operatorname{rang} A.$$

On a donc bien

$$\operatorname{rang} A' = \operatorname{rang} A.$$

□

Le rang de la matrice symétrique associée à une forme quadratique est donc un invariant de la forme à équivalence près. En abusant légèrement de la terminologie, on l'appelle *rang* de la forme quadratique ; si A est la matrice symétrique telle que $q(X) = X^t \cdot A \cdot X$, on pose donc

$$\operatorname{rang} q = \operatorname{rang} A.$$

Indices. Deux autres invariants peuvent être définis à partir d'une diagonalisation quelconque :

14.4. THÉORÈME (Loi d'inertie de Sylvester). *Soit q une forme quadratique réelle en n indéterminées et soit*

$$d(x_1, \dots, x_n) = a_1 x_1^2 + \cdots + a_n x_n^2$$

une forme quadratique diagonale équivalente à q. Soit encore r (resp. s, resp. t) le nombre de coefficients $a_i > 0$ (resp. $a_i < 0$, resp. $a_i = 0$), de sorte que

$$r + s + t = n.$$

Les entiers r, s et t ne dépendent que de la forme quadratique q et non du choix de la forme diagonale d équivalente à q, et

$$\operatorname{rang} q = r + s.$$

DÉMONSTRATION. Supposons que

$$d'(x'_1, \dots, x'_n) = a'_1 {x'_1}^2 + \cdots + a'_n {x'_n}^2$$

soit une autre forme quadratique diagonale équivalente à q. Quitte à changer la numérotation des indéterminées, on peut supposer

$$\begin{aligned} a_i &> 0 && \text{pour } i = 1, \dots, r \\ a_i &< 0 && \text{pour } i = r+1, \dots, r+s \\ a_i &= 0 && \text{pour } i = r+s+1, \dots, r+s+t = n. \end{aligned}$$

De même, supposons

$$\begin{aligned} a'_i &> 0 && \text{pour } i = 1, \dots, r' \\ a'_i &< 0 && \text{pour } i = r'+1, \dots, r'+s' \\ a'_i &= 0 && \text{pour } i = r'+s'+1, \dots, r'+s'+t' = n. \end{aligned}$$

Il s'agit de prouver

$$r = r' \qquad s = s' \qquad t = t'$$

et

$$r + s = \operatorname{rang} q = r' + s'. \tag{‡}$$

Soit A la matrice symétrique associée à q :

$$q(X) = X^t \cdot A \cdot X$$

et soient D et D' les matrices diagonales dont les entrées diagonales sont respectivement $a_1, \dots, a_n$ et $a'_1, \dots, a'_n$:

$$D = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \qquad D' = \begin{pmatrix} a'_1 & & 0 \\ & \ddots & \\ 0 & & a'_n \end{pmatrix}.$$

Par hypothèse, ces matrices sont congruentes à A ; il existe donc des matrices régulières (de changement de variables) P et P' telles que

$$P^t \cdot A \cdot P = D \qquad {P'}^t \cdot A \cdot P' = D'.$$

D'après la Proposition 14.3, on a

$$\operatorname{rang} D = \operatorname{rang} A = \operatorname{rang} D',$$

ce qui établit la relation (‡). On en déduit

$$t = n - (r + s) = n - (r' + s') = t'.$$

Pour conclure la démonstration, il suffit à présent de prouver $r = r'$ car alors

$$s = (\operatorname{rang} q) - r = (\operatorname{rang} q) - r' = s'.$$

Soit $V \subset \mathbb{R}^n$ le sous-espace engendré par les r premières colonnes de P. Comme P est régulière, ses colonnes sont linéairement indépendantes, donc

$$\dim V = r.$$

Par ailleurs, les vecteurs de V s'écrivent tous sous la forme

$$v = P \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

pour certains $x_1, \dots, x_r \in \mathbb{R}$, et

$$\begin{aligned} q(v) &= v^t \cdot A \cdot v = \begin{pmatrix} x_1 & \cdots & x_r & 0 & \cdots & 0 \end{pmatrix} \cdot D \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\ &= a_1 x_1^2 + \cdots + a_r x_r^2. \end{aligned}$$

Comme $a_1, \dots, a_r > 0$, on voit que $q(v) \geqslant 0$ pour tout $v \in V$, et l'égalité $q(v) = 0$ n'a lieu que si $x_1 = \cdots = x_r = 0$, c'est-à-dire si $v = 0$.

Considérons par ailleurs le sous-espace $W \subset \mathbb{R}^n$ engendré par les $s' + t'$ dernières colonnes de P'. On a

$$\dim W = s' + t' = n - r'$$

et tout vecteur de W s'écrit

$$w = P' \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x'_{r+1} \\ \vdots \\ x'_n \end{pmatrix}$$

pour certains $x'_{r+1}, \dots, x'_n \in \mathbb{R}$. Pour un tel w, on a

$$q(w) = a'_{r+1} {x'_{r+1}}^2 + \cdots + a'_n {x'_n}^2$$

donc

$$q(w) \leqslant 0$$

puisque $a'_{r+1}, \dots, a'_n \leqslant 0$.

Dès lors,

$$V \cap W = \{0\}$$

car si l'intersection $V \cap W$ contenait un vecteur $u \neq 0$ on devrait avoir d'une part $q(u) > 0$ car $u \in V$, $u \neq 0$, d'autre part $q(u) \leqslant 0$ car $u \in W$.

Par la relation qui lie $\dim(V+W)$ et $\dim(V \cap W)$ (Proposition 5.7, p. 34), on en déduit

$$\dim(V+W) = \dim V + \dim W = r + (n - r').$$

Or, $V + W \subset \mathbb{R}^n$, donc

$$\dim(V+W) \leqslant \dim \mathbb{R}^n = n.$$

En combinant les deux dernières relations, on obtient $r + n - r' \leqslant n$, c'est-à-dire

$$r \leqslant r'.$$

En échangeant les rôles de D et D', on obtiendrait de même

$$r' \leqslant r,$$

donc

$$r = r'.$$

□

Le nombre de coefficients > 0 dans une diagonalisation quelconque d'une forme quadratique q est appelé *indice de positivité* de q et noté $\operatorname{ind}_+ q$. De même, le nombre de coefficients < 0 dans une diagonalisation quelconque de q est appelé *indice de négativité* de q et noté $\operatorname{ind}_- q$. Comme on l'a observé dans le théorème précédent,

$$\operatorname{ind}_+ q + \operatorname{ind}_- q = \operatorname{rang} q.$$

Les indices permettent de détecter certaines propriétés des formes quadratiques : une forme quadratique réelle q en n indéterminées est dite *définie positive* si

$$q(x_1, \dots, x_n) > 0 \qquad \text{pour tout } (x_1, \dots, x_n) \in \mathbb{R}^n \smallsetminus \{(0, \dots, 0)\};$$

elle est dite *semi-définie positive* si

$$q(x_1, \dots, x_n) \geqslant 0 \qquad \text{pour tout } (x_1, \dots, x_n) \in \mathbb{R}^n.$$

On dit aussi que q est *définie négative* si

$$q(x_1, \dots, x_n) < 0 \qquad \text{pour tout } (x_1, \dots, x_n) \in \mathbb{R}^n \smallsetminus \{(0, \dots, 0)\}$$

et qu'elle est *semi-définie négative* si

$$q(x_1, \dots, x_n) \leqslant 0 \qquad \text{pour tout } (x_1, \dots, x_n) \in \mathbb{R}^n.$$

14.5. Proposition. *Soit q une forme quadratique réelle en n indéterminées. La forme q est définie positive (resp. définie négative) si et seulement si $\operatorname{ind}_+ q = n$ (resp. $\operatorname{ind}_- q = n$). Elle est semi-définie positive (resp. semi-définie négative) si et seulement si $\operatorname{ind}_- q = 0$ (resp. $\operatorname{ind}_+ q = 0$).*

DÉMONSTRATION. Soit
$$d(x_1, \ldots, x_n) = a_1 x_1^2 + \cdots + a_n x_n^2$$
une forme quadratique diagonale équivalente à q. Si A est la matrice symétrique associée à q, on a donc
$$P^t \cdot A \cdot P = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$
pour une certaine matrice régulière P. En particulier, si p_{*i} est la i-ème colonne de P, alors
$$q(p_{*i}) = p_{*i}^t \cdot A \cdot p_{*i} = a_i.$$
Dès lors, si q est définie positive on a $a_i > 0$ pour tout i, donc $\mathrm{ind}_+ q = n$. De même, si q est semi-définie positive on a $a_i \geqslant 0$ pour tout i, donc $\mathrm{ind}_- q = 0$.

Réciproquement, si $\mathrm{ind}_+ q = n$, alors $a_i > 0$ pour tout i, donc
$$d(x_1, \ldots, x_n) > 0 \qquad \text{pour tout } (x_1, \ldots, x_n) \in \mathbb{R}^n \smallsetminus \{(0, \ldots, 0)\}$$
et si $\mathrm{ind}_- q = 0$, alors $a_i \geqslant 0$ pour tout i, donc
$$d(x_1, \ldots, x_n) \geqslant 0 \qquad \text{pour tout } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$
Or, pour $X \in \mathbb{R}^n$ on a $q(X) = d(P^{-1} \cdot X)$, donc q est définie positive si et seulement si d l'est. La démonstration est semblable pour les conditions de définie négativité et de semi-définie négativité. □

14.6. COROLLAIRE. *Soit $q(X) = X^t \cdot A \cdot X$ une forme quadratique réelle en n indéterminées, de matrice symétrique associée A. Si q est définie positive, alors*[1] $\mathrm{dét}\, A > 0$.

DÉMONSTRATION. Comme dans la proposition précédente, on peut trouver une matrice inversible P telle que la matrice $P^t \cdot A \cdot P$ est diagonale, toutes ses entrées diagonales étant strictement positives. On a alors $\mathrm{dét}\,(P^t \cdot A \cdot P) > 0$, donc
$$\mathrm{dét}\, P^t \,\mathrm{dét}\, A \,\mathrm{dét}\, P = \mathrm{dét}\, A\, (\mathrm{dét}\, P)^2 > 0,$$
ce qui entraîne $\mathrm{dét}\, A > 0$. □

Les techniques de diagonalisation indiquées au numéro précédent permettent de déterminer les invariants des formes quadratiques :

14.7. PROPOSITION. *Soit $q(X) = X^t \cdot A \cdot X$ une forme quadratique réelle en n indéterminées, de matrice symétrique associée $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$.*

1. *L'indice de positivité (resp. de négativité) de q est le nombre de valeurs propres strictement positives (resp. strictement négatives) de A.*

Pour $k = 1, \ldots, n$, soit
$$A_k = (a_{ij})_{1 \leqslant i,j \leqslant k} \qquad \text{et} \qquad \delta_k = \mathrm{dét}\, A_k.$$
Soit aussi $\delta_0 = 1$.

2. *Si $\delta_k \neq 0$ pour $k = 1, \ldots, n$, l'indice de positivité (resp. de négativité) de q est le nombre de réels strictement positifs (resp. strictement négatifs) parmi les quotients δ_k/δ_{k-1} pour $k = 1, \ldots, n$.*
3. *La forme q est définie positive si et seulement si $\delta_k > 0$ pour tout $k = 1, \ldots, n$.*

DÉMONSTRATION. 1. La diagonalisation de q obtenue par une base orthonormée de vecteurs propres de A est
$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$
où $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A. Cela prouve la première partie.
2. Si $\delta_k \neq 0$ pour tout $k = 1, \ldots, n$, alors la Proposition 14.2 produit la diagonalisation suivante de q :
$$\frac{\delta_1}{\delta_0} x_1^2 + \frac{\delta_2}{\delta_1} x_2^2 + \cdots + \frac{\delta_n}{\delta_{n-1}} x_n^2.$$
La deuxième partie en découle.

[1] La réciproque n'est pas vraie ; voir la Proposition 14.7 ci-dessous.

3. Si $\delta_k > 0$ pour tout $k = 1, \ldots, n$, alors la deuxième partie montre que $\mathrm{ind}_+ q = n$, donc q est définie positive. Réciproquement, il est clair que si q est définie positive, alors pour tout $k = 1, \ldots, n$ la forme quadratique en n variables

$$q_k(x_1, \ldots, x_k) = q(x_1, \ldots, x_k, 0, \ldots, 0)$$

est définie positive. Comme la matrice associée à la forme quadratique q_k est A_k, le Corollaire 14.6 montre que $\delta_k > 0$. □

La proposition suivante indique que les indices de positivité et de négativité d'une forme quadratique suffisent à la déterminer de manière unique à équivalence près :

14.8. PROPOSITION. *Toute forme quadratique réelle en n indéterminées d'indice de positivité r et d'indice de négativité s est équivalente à la forme*

$$x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2.$$

En particulier, toute forme quadratique réelle définie positive en n indéterminées est équivalente à la forme

$$x_1^2 + \cdots + x_n^2.$$

DÉMONSTRATION. Par hypothèse, la forme quadratique q est équivalente à une forme diagonale

$$d(y_1, \ldots, y_n) = a_1 y_1^2 + \cdots + a_n y_n^2$$

dont le nombre de coefficients $a_i > 0$ est r et dont le nombre de coefficients $a_i < 0$ est s (et donc le nombre de coefficients $a_i = 0$ est $n - (r + s)$). Quitte à changer la numérotation des indéterminées, on peut supposer

$$\begin{array}{rcll} a_i & > & 0 & \text{pour } i = 1, \ldots, r \\ a_i & < & 0 & \text{pour } i = r+1, \ldots, r+s \\ a_i & = & 0 & \text{pour } i = r+s+1, \ldots, n. \end{array}$$

On effectue alors le changement de variables suivant :

$$\begin{array}{rcll} y_i & = & (a_i)^{1/2} x_i & \text{pour } i = 1, \ldots, r \\ y_i & = & (-a_i)^{1/2} x_i & \text{pour } i = r+1, \ldots, r+s \end{array}$$

et la forme quadratique $d(y_1, \ldots, y_n)$ devient

$$x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2.$$

□

Exercices

(14.1) (prérequis) Donnez un exemple de forme quadratique réelle en trois variables dont le rang est 3 et dont l'indice de positivité est 1.

(14.2) L'indice de positivité d'une forme quadratique peut-il être strictement plus grand que son rang ?

(14.3) Pour chacune des formes quadratiques q ci-dessous,

(a) donnez la matrice symétrique A associée ;

(b) cherchez les valeurs propres et vecteurs propres de A ;

(c) trouvez une matrice orthogonale P telle que $P^t \cdot A \cdot P$ soit diagonale ;

(d) réduisez q à une forme diagonale par la méthode de complétion des carrés, et déterminez la matrice de changement de variables correspondante ;

(e) déterminez le rang, l'indice de positivité et l'indice de négativité de q.

1. $3x_1^2 + 2x_1x_2 + x_2^2$.
2. $2x_1^2 + x_1x_2 - 3x_2^2$.
3. $5x_1^2 + 3(x_2^2 - x_3^2) - 4x_2x_3$.
4. $-7x_1^2 + 25x_2^2 + 7x_3^2 - 48x_1x_3$.
5. $x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3$.
6. $x_1^2 - x_3^2 + 6x_2^2 + 4x_1x_2$.
7. $2x_1^2 - 4x_1x_2 + 2x_1x_3 + x_2^2 - 2x_2x_3 + \frac{1}{2}x_3^2$.
8. $4x_1x_2 + 2x_1x_3 + x_3^2 + 2x_2x_3$.
9. $x_1^2 + 2x_1x_3 + 2x_3^2 + x_2x_3$.
10. $x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 + \alpha x_2x_4$, où α est un paramètre réel.

(14.4) Déterminez le rang, l'indice de positivité et l'indice de négativité de la forme quadratique suivante d'après la valeur du paramètre réel α :

$$x_1^2 - 2x_1x_2 + 2\alpha x_1x_3 + 2x_2^2 + 8x_3^2.$$

(14.5) Soit A une matrice symétrique réelle d'ordre n. Montrez que la forme quadratique $q(X) = X^t \cdot A \cdot X$ est définie positive si et seulement si $A = P^t \cdot P$ pour une certaine matrice régulière P. Déterminez une matrice P telle que

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = P^t \cdot P.$$

(14.6) Deux matrices symétriques congruentes ont-elles le même déterminant ? Ont-elles les mêmes valeurs propres ? Ont-elles le même rang ? Si oui, justifiez ; si non, donnez un contre-exemple.

(14.7) Montrez que le signe (> 0, < 0 ou $= 0$) du déterminant de la matrice symétrique associée à une forme quadratique réelle est un invariant de la forme.

(14.8) (pour aller plus loin) Soit q une forme quadratique réelle en n variables. Prouvez que si l'équation $q(x_1, \ldots, x_n) = 0$ n'a pas d'autre solution que la solution triviale $x_1 = \cdots = x_n = 0$, alors la forme q est définie positive ou définie négative.

15. Espaces quadratiques

15.1. Formes bilinéaires symétriques

Soit E un espace vectoriel réel. Une *forme bilinéaire symétrique* sur E est une application

$$b\colon\ E \times E \to \mathbb{R}$$

satisfaisant les conditions suivantes :

1. Bilinéarité : pour $x, y, z \in E$ et $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} b(x\alpha + y\beta, z) &= \alpha\, b(x,z) + \beta\, b(y,z). \\ b(z, x\alpha + y\beta) &= b(z,x)\,\alpha + b(z,y)\,\beta. \end{aligned}$$

2. Symétrie : pour $x, y \in E$,

$$b(y,x) = b(x,y).$$

Les produits scalaires sur E sont donc des formes bilinéaires symétriques ; ce sont les formes bilinéaires symétriques qui satisfont de plus la condition

$$b(x,x) > 0 \qquad \text{pour tout } x \neq 0.$$

Il est important de remarquer que cette condition n'est *pas* imposée ici. L'espace E peut donc contenir des vecteurs non nuls x tels que $b(x,x) = 0$ ou même $b(x,x) < 0$.

Un espace muni d'une forme bilinéaire symétrique est appelé *espace quadratique*. Les espaces euclidiens sont donc des espaces quadratiques particuliers.

Exemple : Formes bilinéaires symétriques sur $\mathbb{R}^n$. À toute matrice symétrique $A \in \mathbb{R}^{n\times n}$ on associe une forme bilinéaire b_A sur $\mathbb{R}^n$ définie par

$$b_A(x,y) = x^t \cdot A \cdot y \qquad \text{pour } x, y \in \mathbb{R}^n \ (= \mathbb{R}^{n\times 1}).$$

Comme $x^t \cdot A \cdot y \in \mathbb{R}^{1\times 1}$, on a

$$x^t \cdot A \cdot y = (x^t \cdot A \cdot y)^t = y^t \cdot A \cdot x.$$

Or, $A^t = A$, donc

$$x^t \cdot A \cdot y = y^t \cdot A \cdot x,$$

et la forme bilinéaire b_A est donc symétrique.

Inversement, à toute forme bilinéaire symétrique b sur $\mathbb{R}^n$ on associe la matrice $A_b \in \mathbb{R}^{n\times n}$ définie par

$$A_b = (b(c_i, c_j))_{1\leqslant i,j\leqslant n}$$

où $(c_1, \ldots, c_n)$ est la base canonique de $\mathbb{R}^n$. Comme $b(c_j, c_i) = b(c_i, c_j)$, la matrice A_b est symétrique.

15.1. Proposition. *Les applications $A \mapsto b_A$ et $b \mapsto A_b$ définissent des bijections réciproques entre l'ensemble des matrices symétriques réelles d'ordre n et l'ensemble des formes bilinéaires symétriques sur $\mathbb{R}^n$.*

Démonstration. Il suffit de prouver que pour toute matrice symétrique A et pour toute forme bilinéaire symétrique b,

$$A_{b_A} = A \qquad \text{et} \qquad b_{A_b} = b.$$

Or, si $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, alors

$$b_A(c_i, c_j) = c_i^t \cdot A \cdot c_j = a_{ij},$$

donc

$$A_{b_A} = (b_A(c_i, c_j))_{1\leqslant i,j\leqslant n} = (a_{ij})_{1\leqslant i,j\leqslant n} = A.$$

D'autre part, si b est une forme bilinéaire symétrique sur $\mathbb{R}^n$, alors pour $x, y \in \mathbb{R}^n$,

$$\begin{aligned} b_{A_b}(x,y) &= x^t \cdot A_b \cdot y \\ &= x^t (b(c_i, c_j))_{1\leqslant i,j\leqslant n} \cdot y \\ &= \sum_{i,j=1}^n x_i\, b(c_i, c_j)\, y_j \\ &= b(x,y) \end{aligned}$$

donc $b_{A_b} = b$. □

Cette proposition montre que les formes bilinéaires symétriques sur $\mathbb{R}^n$ sont en correspondance biunivoque avec les matrices symétriques réelles d'ordre n. Comme celles-ci correspondent aux formes quadratiques réelles en n variables, comme on l'a vu au début de la section précédente, on peut conclure que les formes bilinéaires symétriques sur $\mathbb{R}^n$ et les formes quadratiques réelles en n variables se correspondent bijectivement : la forme quadratique $q(X) = X^t \cdot A \cdot X$ correspond à la forme bilinéaire $b(X,Y) = X^t \cdot A \cdot Y$ (par l'intermédiaire de la matrice symétrique $A \in \mathbb{R}^{n\times n}$).

Le point de vue des espaces quadratiques donne cependant un éclairage plus géométrique sur la théorie des formes quadratiques, comme on se propose de le faire voir dans les sections suivantes.

15.2. Représentation matricielle

Soit $e = (e_1, \ldots, e_n)$ une base d'un espace E de dimension finie. À toute forme bilinéaire symétrique b sur E on fait correspondre une matrice $(b)_e \in \mathbb{R}^{n\times n}$ définie par

$$(b)_e = (b(e_i, e_j))_{1\leqslant i,j\leqslant n}.$$

Cette matrice est appelée *matrice de Gram* de la forme b par rapport à la base e. Comme

$$b(e_j, e_i) = b(e_i, e_j) \qquad \text{pour } i, j = 1, \ldots, n,$$

cette matrice est symétrique.

Par exemple, si A est une matrice symétrique d'ordre n et si b_A est la forme bilinéaire symétrique correspondante sur $\mathbb{R}^n$ (voir l'exemple de la section précédente), alors la Proposition 15.1 montre que, par rapport à la base canonique c,

$$(b_A)_c = A.$$

Propriétés.

1. Pour $x, y \in E$,

$$b(x,y) = {}_e(x)^t \cdot (b)_e \cdot {}_e(y).$$

2. Si f est une autre base de E, alors

$$(b)_f = {}_e(I_E)_f^t \cdot (b)_e \cdot {}_e(I_F)_f.$$

Démonstration. 1. Soient $x = \sum_{i=1}^n e_i x_i$ et $y = \sum_{j=1}^n e_j y_j$, de sorte que

$${}_e(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad {}_e(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

En utilisant la bilinéarité de la forme b, on obtient

$$b(x,y) = \sum_{i,j=1}^n x_i\, b(e_i, e_j)\, y_j.$$

Le membre de droite est aussi le résultat de la multiplication matricielle

$$\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot \begin{pmatrix} b(e_1,e_1) & \cdots & b(e_1,e_n) \\ \vdots & & \vdots \\ b(e_n,e_1) & \cdots & b(e_n,e_n) \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

ce qui prouve la première partie.
2. Soit $f = (f_1, \ldots, f_n)$ et

$$f_j = \sum_{i=1}^n e_i a_{ij} \qquad \text{pour } j = 1, \ldots, n,$$

de sorte que ${}_e(I_E)_f = (a_{ij})_{1 \leqslant i,j \leqslant n}$. En utilisant la bilinéarité de b, on obtient

$$\begin{aligned} b(f_i, f_j) &= b\left(\sum_{k=1}^n e_k a_{ki}, \sum_{\ell=1}^n e_\ell a_{\ell j}\right) \\ &= \sum_{k,\ell=1}^n a_{ki}\, b(e_k, e_\ell)\, a_{\ell j}. \end{aligned}$$

Le membre de droite est l'entrée d'indices i, j du produit matriciel

$$\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} b(e_1,e_1) & \cdots & b(e_1,e_n) \\ \vdots & & \vdots \\ b(e_n,e_1) & \cdots & b(e_n,e_n) \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$
$$= {}_e(I_E)_f^t \cdot (b)_e \cdot {}_e(I_E)_f.$$

□

La deuxième propriété ci-dessus montre que les matrices de Gram associées à une même forme bilinéaire symétrique par rapport à des bases différentes sont congruentes.

On peut aussi décrire cette situation de la manière suivante : à toute base e d'un espace quadratique E, on peut associer une forme quadratique

$$q_e(X) = X^t \cdot (b)_e \cdot X$$

qui donne l'expression du produit d'un vecteur par lui-même par rapport à la base e :

$$q_e(x_1, \ldots, x_n) = b(e_1x_1 + \cdots + e_nx_n, e_1x_1 + \cdots + e_nx_n).$$

La deuxième propriété ci-dessus montre que les formes quadratiques q_e et $q_{e'}$ associées à deux bases différentes sont équivalentes. Inversement, si une forme quadratique q est équivalente à la forme q_e, soit

$$q(X) = X^t \cdot A \cdot X$$

où la matrice A est congruente à la matrice $(b)_e$ de la forme q_e :

$$A = P^t \cdot (b)_e \cdot P$$

pour une certaine matrice régulière P, alors d'après la Proposition 9.7 il existe une base f de E telle que $P = {}_e(I_E)_f$. D'après la deuxième propriété ci-dessus, on a

$$(b)_f = {}_e(I_E)_f^t \cdot (b)_e \cdot {}_e(I_E)_f = A,$$

donc la forme quadratique q donnée est la même que la forme q_f associée à la base f :

$$q = q_f.$$

En conclusion, un espace quadratique correspond à une *classe d'équivalence* de formes quadratiques : la donnée d'une base de l'espace détermine une forme quadratique, et un changement de base fait passer d'une forme quadratique à une forme équivalente. Dès lors, les invariants des formes quadratiques à équivalence près doivent pouvoir être définis de manière géométrique en termes d'espaces quadratiques. C'est ce que l'on se propose d'illustrer dans la section suivante.

15.3. Invariants

Rang. Soit E un espace quadratique muni d'une forme bilinéaire symétrique b. On définit le *radical* de b par

$$\operatorname{rad} b = \{x \in E \mid b(x, y) = 0 \text{ pour tout } y \in E\}.$$

On dit que la forme bilinéaire b (ou l'espace quadratique E) est *régulière* (ou *non dégénérée*) si $\operatorname{rad} b = \{0\}$.

15.2. Proposition. *Soit* $\dim E = n$ *et soit* $e = (e_1, \dots, e_n)$ *une base de* E*. Le radical de* b *est l'ensemble des vecteurs* $x \in E$ *dont les coordonnées* ${}_e(x)$ *par rapport à la base* e *sont solutions du système*

$$(b)_e \cdot {}_e(x) = 0.$$

En particulier, $\dim \operatorname{rad} b = n - \operatorname{rang}(b)_e$.

Démonstration. Pour $x, y \in E$, on a

$$b(y, x) = {}_e(y)^t \cdot (b)_e \cdot {}_e(x);$$

donc si $(b)_e \cdot {}_e(x) = 0$, alors $b(y, x) = 0$ pour tout $y \in E$ et par conséquent $x \in \operatorname{rad} b$. Inversement, supposons $x \in \operatorname{rad} b$ et

$$(b)_e \cdot {}_e(x) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Pour $i = 1, \dots, n$, on a $b(e_i, x) = 0$ car $x \in \operatorname{rad} b$ et

$$\begin{array}{c} \\ {}_e(e_i)^t \cdot (b)_e \cdot {}_e(x) = \end{array} \begin{array}{c} \begin{array}{ccccc} & & i & & \\ & & \downarrow & & \end{array} \\ \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} \end{array} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_i,$$

donc $\alpha_i = 0$. Dès lors, $(b)_e \cdot {}_e(x) = 0$ si $x \in \operatorname{rad} b$.

L'isomorphisme ${}_e\gamma$ qui envoie tout vecteur $x \in E$ sur ses coordonnées ${}_e(x) \in \mathbb{R}^n$ met donc en bijection $\operatorname{rad} b$ et l'ensemble des solutions du système homogène

$$(b)_e \cdot X = 0.$$

D'après la Proposition 6.9 (page 41), cet ensemble est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n - \operatorname{rang}(b)_e$. □

Cette proposition montre que $\operatorname{rang}(b)_e$ est un invariant de l'espace quadratique E : il ne dépend que de la forme b et non de la base e. C'est une manière géométrique de prouver la Proposition 14.3 (p. 101).

Indice. Un sous-espace $V \subset E$ est dit *défini positif* si

$$b(v,v) > 0 \qquad \text{pour tout } v \in V,\ v \neq 0.$$

On définit l'*indice de positivité* de l'espace quadratique E par

$$\operatorname{ind}_+ E = \max\{\dim V \mid V \subset E \text{ sous-espace défini positif}\}.$$

15.3. Proposition. *Pour toute base e de E, la forme quadratique q_e associée à e par la formule*

$$q_e(x_1,\ldots,x_n) = b(e_1x_1 + \cdots + e_nx_n, e_1x_1 + \cdots + e_nx_n)$$

a le même indice de positivité que l'espace quadratique E :

$$\operatorname{ind}_+ E = \operatorname{ind}_+ q_e.$$

Démonstration. Soit q une forme quadratique diagonale équivalente à q_e. Quitte à changer l'ordre des variables, on peut supposer

$$q(x_1,\ldots,x_n) = a_1x_1^2 + \cdots + a_nx_n^2$$

avec $a_1,\ldots,a_r > 0$ et $a_{r+1},\ldots,a_n \leqslant 0$, de sorte que

$$\operatorname{ind}_+ q_e = r.$$

Comme q est équivalente à q_e, on peut trouver une base $f = (f_1,\ldots,f_n)$ de E telle que $q = q_f$, comme on l'a vu à la fin de la section précédente.

Soit $V = \operatorname{sev}\langle f_1,\ldots,f_r\rangle$. Pour $v \in V$, soit $v = f_1\alpha_1 + \cdots + f_r\alpha_r$, on a

$$\begin{aligned} b(v,v) &= b(f_1\alpha_1 + \cdots + f_r\alpha_r, f_1\alpha_1 + \cdots + f_r\alpha_r) \\ &= q_f(\alpha_1,\ldots,\alpha_r,0,\ldots,0) \\ &= a_1\alpha_1^2 + \cdots + a_r\alpha_r^2. \end{aligned}$$

Comme $a_1,\ldots,a_r > 0$, on a donc $b(v,v) > 0$ si $v \neq 0$, ce qui montre que V est défini positif. Comme $\operatorname{ind}_+ E$ est la plus grande dimension d'un sous-espace défini positif, on doit avoir

$$\operatorname{ind}_+ E \geqslant \dim V = \operatorname{ind}_+ q_e.$$

Considérons par ailleurs $W = \operatorname{sev}\langle f_{r+1},\ldots,f_n\rangle$. Le même calcul que précédemment montre que pour $w \in W$, soit $w = f_{r+1}\beta_{r+1} + \cdots + f_n\beta_n$, on a

$$b(w,w) = a_{r+1}\beta_{r+1}^2 + \cdots + a_n\beta_n^2 \leqslant 0.$$

Dès lors, l'intersection de W avec tout sous-espace défini positif doit être réduite à $\{0\}$. Si U est un sous-espace défini positif de dimension maximale : $\dim U = \operatorname{ind}_+ E$, alors de la relation $U \cap W = \{0\}$ on déduit

$$\dim U + W = \dim U + \dim W$$

par la Proposition 5.7, donc

$$\dim E \geqslant \dim U + \dim W.$$

Comme $\dim W = n - r = \dim E - \operatorname{ind}_+ q_e$ et $\dim U = \operatorname{ind}_+ E$, on en déduit

$$\operatorname{ind}_+ q_e \geqslant \operatorname{ind}_+ E.$$

$\square$

Cette proposition est une version géométrique de la loi d'inertie de Sylvester.

Exercices

(15.1) (prérequis) On considère l'application $b\colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ définie par

$$b\Big((x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4)\Big) = x_1y_1 + x_2y_2.$$

(a) Vérifiez que b est une forme bilinéaire symétrique.

(b) Donnez la matrice représentant b par rapport à la base canonique de $\mathbb{R}^4$.

(c) Déterminez le radical $\operatorname{rad} b$.

(d) Trouvez un sous-espace de $\mathbb{R}^4$ défini positif de dimension maximale.

(15.2) (prérequis) Soit $A = \begin{pmatrix} 2 & 5 \\ 5 & -1 \end{pmatrix}$. Écrivez la forme bilinéaire b_A sur $\mathbb{R}^2$ associée à la matrice A. Écrivez la matrice de Gram de la forme b_A par rapport à la base $e = ((1,1),(1,-1))$ de $\mathbb{R}^2$. Écrivez la forme quadratique associée à la forme bilinéaire b_A et à la base e.

(15.3) Soit b une forme bilinéaire sur un espace E et soit e une base de E. Démontrez que la forme b est non dégénérée si et seulement si $\operatorname{dét}(b)_e \neq 0$.

(15.4) Sur l'espace $\mathbb{R}[X]_{\leqslant 2}$ on définit une forme bilinéaire symétrique par

$$b(P,Q) = \int_0^1 P(x)Q(x)\,dx.$$

(a) Donnez la matrice représentant b par rapport à la base usuelle $u = (1, X, X^2)$.

(b) Montrez que b est un produit scalaire.

(c) Trouvez une base orthonormée de $\mathbb{R}[X]_{\leqslant 2}$ pour ce produit scalaire.

(15.5) Sur l'espace $\mathbb{R}[X]_{\leqslant 2}$ on définit une forme bilinéaire symétrique par

$$b(P,Q) = \sum_{i=1}^{5} P(a_i)Q(a_i),$$

où $a_1 = -2$, $a_2 = -1$, $a_3 = 0$, $a_4 = 1$, $a_5 = 2$.

(a) Donnez la matrice représentant b par rapport à la base usuelle $u = (1, X, X^2)$.

(b) Montrez que b est un produit scalaire.

(c) Trouvez une base orthonormée de $\mathbb{R}[X]_{\leqslant 2}$ pour ce produit scalaire.

(15.6) Montrez que tout espace quadratique admet une base orthogonale, c'est-à-dire une base $(e_1, \ldots, e_n)$ telle que $b(e_i, e_j) = 0$ pour $i \neq j$.

(15.7) Soit b la forme bilinéaire sur $\mathbb{R}^3$ définie par

$$b\Big((x_1, x_2, x_3), (y_1, y_2, y_3)\Big) = \frac{1}{2}\Big(x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2\Big).$$

Trouvez une base orthogonale de $\mathbb{R}^3$ pour cette forme bilinéaire. Trouvez la dimension du radical de b et son indice de positivité.

(15.8) Soit b le produit scalaire sur $\mathbb{R}^3$ dont une base orthonormée est

$$e = ((1,1,1),(0,1,1),(0,0,1)).$$

Trouvez la matrice de b par rapport à la base canonique de $\mathbb{R}^3$.

Appendice : Rappels et compléments

A.1 Nombres complexes

Un nombre complexe est une expression de la forme

$$a + bi$$

où $a, b \in \mathbb{R}$ et i est une indéterminée. Tout nombre réel $a \in \mathbb{R}$ est identifié au nombre complexe $a + 0i$, et on note simplement bi pour $0 + bi$ et i pour $1i$.

Sur l'ensemble $\mathbb{C}$ des nombres complexes, on définit une addition et une multiplication qui prolongent les opérations correspondantes de $\mathbb{R}$ en posant, pour $a, a', b, b' \in \mathbb{R}$:

$$(a + bi) + (a' + b'i) = (a + a') + (b + b')i$$

et

$$(a + bi) \cdot (a' + b'i) = (aa' - bb') + (ab' + ba')i.$$

On a donc $i^2 = -1$ et le nombre complexe bi est bien, comme la notation le suggère, le produit de $b \in \mathbb{R}$ et de i.

L'addition et la multiplication des nombres complexes sont donc définies comme l'addition et la multiplication des polynômes en i, en tenant compte de la relation supplémentaire $i^2 = -1$.

On définit le *conjugué* du nombre complexe $a + bi$ (où $a, b \in \mathbb{R}$) par

$$\overline{a + bi} = a - bi$$

et sa *norme* par

$$N(a + bi) = (a + bi)\overline{(a + bi)} = a^2 + b^2.$$

La norme de tout nombre complexe est donc un nombre réel positif, qui n'est nul que si le nombre complexe est nul. Pour $z \in \mathbb{C}$, $z \neq 0$, on a donc

$$z \cdot (N(z)^{-1}\overline{z}) = 1,$$

ce qui montre que tout nombre complexe non nul est inversible. Une vérification directe montre que pour $z, z' \in \mathbb{C}$,

$$\overline{z + z'} = \overline{z} + \overline{z'}, \qquad \overline{z \cdot z'} = \overline{z} \cdot \overline{z'}$$

et par conséquent

$$N(z \cdot z') = N(z) \cdot N(z').$$

On définit aussi le *module* d'un nombre complexe $z \in \mathbb{C}$ comme la racine carrée de sa norme ; on le note $|z|$:

$$|z| = \sqrt{N(z)}.$$

Si $z \in \mathbb{R}$, alors $\overline{z} = z$, donc $N(z) = z^2$ et le module de z est sa valeur absolue, ce qui justifie l'utilisation de la même notation pour le module d'un nombre complexe que pour la valeur absolue d'un nombre réel.

A.2 Structures algébriques

Un *groupe* est un ensemble G muni d'une loi de composition

$$\begin{array}{rccc} *\colon & G \times G & \to & G \\ & (x,y) & \mapsto & x * y \end{array}$$

soumise aux conditions suivantes :

– associativité : pour $x, y, z \in G$,

$$(x * y) * z = x * (y * z)$$

– existence d'un élément neutre : il existe $e \in G$ tel que pour tout $x \in G$,

$$x * e = x = e * x$$

– inversibilité : pour tout $x \in G$ il existe un élément $x' \in G$ tel que

$$x * x' = e = x' * x.$$

Si de plus $x * y = y * x$ pour $x, y \in G$, le groupe est dit *commutatif.*

Tout groupe contient un seul élément neutre : en effet, si e et $e' \in G$ sont deux neutres, alors $e * e' = e$ car e' est un neutre, et $e * e' = e'$ car e est un neutre, donc $e = e'$.

Les notations usuelles sont la notation additive, où la loi de composition $*$ est notée $+$, le neutre 0 et l'inverse de x est noté $-x$; et la notation multiplicative où $*$ est noté $\cdot$, le neutre 1 et l'inverse de x est noté x^{-1}.

Un *corps* (commutatif) est un ensemble K muni de deux lois de composition, l'une notée additivement, l'autre multiplicativement, qui satisfont les conditions suivantes :

1. K est un groupe commutatif pour l'addition ;
2. $K \smallsetminus \{0\}$ est un groupe commutatif pour la multiplication ;
3. pour $x, y, z \in K$,

$$(x + y)z = xz + yz.$$

Par exemple, $\mathbb{C}$, $\mathbb{R}$ et l'ensemble $\mathbb{Q}$ des nombres rationnels sont des corps. D'autres exemples peuvent être construits de la manière suivante : si K est un corps, l'ensemble des fractions rationnelles en une indéterminée X à coefficients dans K est l'ensemble des quotients de polynômes :

$$K(X) = \left\{ \frac{P(X)}{Q(X)} \;\middle|\; P(X), Q(X) \in K[X],\ Q(X) \neq 0 \right\}.$$

Cet ensemble est un corps pour les opérations usuelles.

Il y a aussi des corps qui n'ont qu'un nombre fini d'éléments, comme par exemple

$$\mathbb{F}_2 = \{0, 1\},$$

les tables d'addition et de multiplication étant données par

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad\qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

On a donc $1 + 1 = 0$ dans $\mathbb{F}_2$. Ce corps est souvent utilisé en mathématiques discrètes.

A.3 Espaces vectoriels à gauche et à droite

Les espaces vectoriels définis en (3.1) sont plus précisément appelés *espaces vectoriels à droite* car la notation $x\alpha$ a été utilisée pour le produit d'un vecteur x et d'un scalaire α. On définit de même la notion d'*espace vectoriel à gauche* : un groupe additif E est appelé espace vectoriel à gauche sur un corps K s'il est muni d'une opération

$$. \quad : \quad \begin{aligned} K \times E &\to E \\ (\alpha, x) &\mapsto \alpha x \end{aligned}$$

satisfaisant les conditions suivantes : pour $x, y \in E$ and $\alpha, \beta \in K$,

$$\begin{aligned} (\alpha+\beta)x &= \alpha x + \beta x \\ \alpha(x+y) &= \alpha x + \alpha y \\ (\alpha\beta)x &= \alpha(\beta x) \\ 1.x &= x. \end{aligned}$$

Le corps K étant commutatif, tout espace vectoriel à gauche sur K admet aussi une structure d'espace vectoriel à droite sur K (et réciproquement) : il suffit de définir une multiplication scalaire à droite en posant

$$x\alpha = \alpha x \qquad \text{pour } x \in E \text{ et } \alpha \in K.$$

La commutativité de K intervient pour établir la condition $x(\alpha\beta) = (x\alpha)\beta$ pour $x \in E$ et $\alpha, \beta \in K$. En effet, d'après la définition ci-dessus de la multiplication scalaire à droite, on a $x(\alpha\beta) = (\alpha\beta)x$ tandis que $(x\alpha)\beta = \beta(\alpha x)$.

A.4 Espaces vectoriels non finiment engendrés

La notion de combinaison linéaire d'une suite de vecteurs $v_1, \ldots, v_n$ d'un espace vectoriel E sur un corps K peut s'étendre aux suites infinies : si $(v_i)_{i\in I}$ est une famille quelconque de vecteurs de E, où I est un ensemble d'indices arbitraire, une combinaison linéaire de $(v_i)_{i\in I}$ est, par définition, un vecteur $x \in E$ qui peut s'écrire sous la forme

$$x = \sum_{i\in I_0} v_i\alpha_i$$

pour certains scalaires $\alpha_i \in K$ et un certain sous-ensemble *fini* $I_0 \subset I$. Il faut en effet garder à l'esprit que la somme d'un nombre infini de termes n'est *pas définie !*

L'ensemble des combinaisons linéaires de la famille $(v_i)_{i\in I}$ est un sous-espace de E que l'on note $\text{sev}\langle (v_i)_{i\in I}\rangle$ et que l'on appelle *sous-espace vectoriel engendré par* $(v_i)_{i\in I}$.

On peut de même étendre les notions de suite génératrice, de suite libre et de base, et montrer que tout espace vectoriel admet une base. La notion de dimension peut elle aussi être généralisée, car on peut prouver que si $(v_i)_{i\in I}$ et $(w_j)_{j\in J}$ sont deux bases d'un même espace vectoriel E, alors les ensembles d'indices I et J sont en bijection. Les démonstrations de ces énoncés utilisent des notions non élémentaires de théorie des ensembles qui sortent du cadre de ce cours.