ALGÈBRE LINÉAIRE AVANCÉE I

Polycopié élaboré par

Daniel Kressner

Automne 2017

EPF Lausanne

4 décembre 2017

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

– Pour la préparation de ce polycopié, les sources suivantes ont été utilisées :
  1. Abdulle, Assyr. Algèbre linéaire avancée I, Polycopié EPF Lausanne, 2015.
  2. Grifone, Joseph. Algèbre linéaire, 5e Edition. Éditions Cépaduè, 2015.
  3. Gutknecht, Martin. Lineare Algebra für Informatiker. Polycopié EPF Zurich, 2009.
  4. Liesen, Jörg and Mehrmann, Volker. Linear Algebra. Springer, 2015.
  5. Meyberg, Kurt. Algebra. Carl Hanser Verlag, 1975.

  S.V.P. ne pas prendre cette liste comme une suggestion de littérature pour se préparer à l'examen.
– Quelques-une des illustrations ont été créées par Michael Steinlechner.
– Des corrections ? Des suggestions ? N'hésitez pas à m'écrire un mail :

daniel.kressner@epfl.ch.

Chapitre 0

Systèmes d'équations linéaires

Afin de commencer en douceur, ce chapitre rappelle des concepts élémentaires des systèmes linéaires.

Une équation à une inconnue. Une équation linéaire à une inconnue x,

$$ax = b\,,$$

a exactement une solution

$$x = \frac{b}{a}\,,$$

sauf si $a = 0$. Dans cette situation exceptionnelle on a deux cas possibles:

$a = 0,\ b = 0$: une infinité de solutions,
$a = 0,\ b \neq 0$: pas de solution.

Deux équations à deux inconnues. Alors, on considère un système de deux équations linéaires. Par exemple:

$$\begin{array}{rcl} x_1 + x_2 & = & 4 \\ 6x_1 - 2x_2 & = & 8\,. \end{array} \tag{0.1}$$

Est-ce qu'il existe une solution, et si oui, est-elle unique ?

Dans le cas de deux variables, on peut répondre à ces questions facilement au moyen de la géométrie. À cette fin, les équations (0.1) sont transformées comme suit :

$$x_2 = -x_1 + 4, \quad x_2 = 3x_1 - 4\,. \tag{0.2}$$

Chaque équation décrit une droite dans le plan (x_1,x_2). Alors, la solution commune de ces deux équations est le point d'intersection $(x_1,x_2) = (2,2)$ des deux droites, voir figure 0.1. On peut bien sûr résoudre (0.2) sans géométrie : De

$$-x_1 + 4 = x_2 = 3x_1 - 4$$

on obtient $4x_1 = 8$ et ainsi $x_1 = 2$, $x_2 = 2$.

Évidemment, le système suivant n'a pas de solution :

$$\begin{array}{rcl} 4x_1 - 2x_2 & = & -2 \\ 2x_1 - x_2 & = & 4\,. \end{array} \tag{0.3}$$

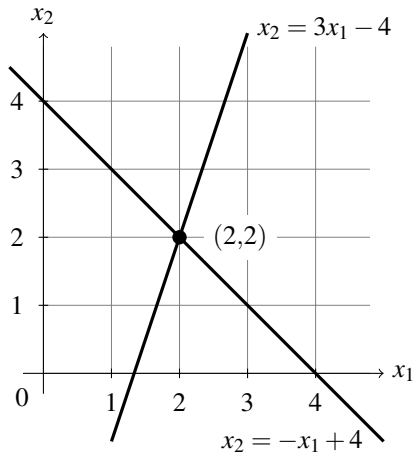


FIG. 0.1 – Interprétation géométrique de (0.2) : La solution est donnée par le point d'intersection.

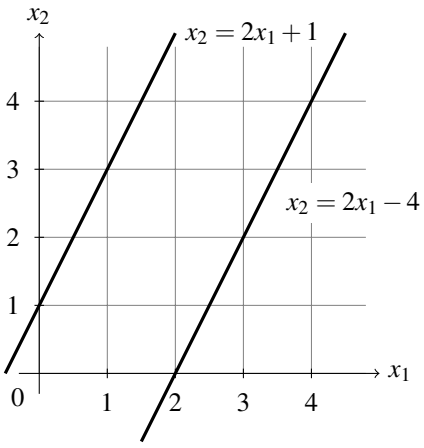


FIG. 0.2 – Interprétation géométrique de (0.3) : Les deux droites sont parallèles et ainsi, il n'existe pas de solution.

Les droites correspondantes sont parallèles, voir figure 0.2.

Au contraire, le système suivant admet une infinité de solutions :

$$\begin{aligned} 4x_1 - 2x_2 &= 8 \\ 2x_1 - x_2 &= 4 \end{aligned} \tag{0.4}$$

parce que la première équation est égale à la deuxième multipliée par 2. L'ensemble des solutions est la droite $x_2 = 2x_1 - 4$ dans la figure 0.2.

Les exemples (0.3) et (0.4) sont exceptionnels. Un système de deux équations linéaires à deux inconnues a presque toujours une seule solution.

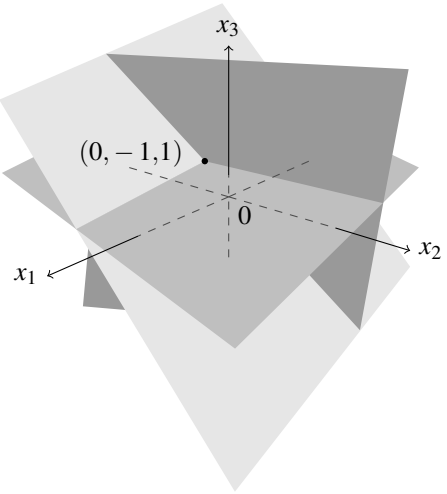


FIG. 0.3 – Interprétation géométrique de (0.5) : La solution est le point d'intersection des trois plans.

Trois équations à trois inconnues. On considére le systéme suivant :

$$\begin{array}{rrrcr} x_1+ & 4x_2+ & 4x_3 & = & 0 \\ 3x_1+ & 4x_2+ & 16x_3 & = & 12 \\ 4x_1+ & 2x_2+ & x_3 & = & -1 \quad . \end{array} \tag{0.5}$$

Chaque équation décrit un plan dans $\mathbb{R}^3$, voir figure 0.3. Avec de la chance, on peut voir que le point d'intersection de ces trois plans est $(x_1,x_2,x_3) = (0,-1,1)$.

Pour calculer la solution de (0.5), on pourrait procéder comme pour les équations (0.1) et éliminer des variables. Mais cela devient rapidement compliqué et brouillon pour trois variables ou plus. Il faut une procédure systématique ! À cette fin, on soustrait la première équation multipliée par 3 (resp. 4) de la deuxième (resp. la troisième), ce qui donne

$$\begin{array}{l} \\ 3 \\ 4 \end{array}\!\!\left\{\begin{array}{rrrcr} 1x_1+ & 4x_2+ & 4x_3 & = & 0 \\ 3x_1+ & 4x_2+ & 16x_3 & = & 12 \\ 4x_1+ & 2x_2+ & x_3 & = & -1 \end{array}\right. \quad \Longrightarrow \quad \begin{array}{rrrcr} 1x_1+4x_2+ & & 4x_3 & = & 0 \\ -8x_2+ & & 4x_3 & = & 12 \\ -14x_2- & & 15x_3 & = & -1 \quad . \end{array}$$

Les multiplicateurs ont été choisis dans le but d'éliminer x_1 de la deuxième et la troisième équations. Alors, on élimine la variable x_2 dans la troisième équation en lui soustrayant la deuxième équation multipliée par $7/4$:

$$\begin{array}{l} \\ \frac{7}{4} \\ \end{array}\!\!\left\{\begin{array}{rrrcr} 1x_1+4x_2+ & & 4x_3 & = & 0 \\ -8x_2+ & & 4x_3 & = & 12 \\ -14x_2- & & 15x_3 & = & -1 \end{array}\right. \quad \Longrightarrow \quad \begin{array}{rrrcr} 1x_1+4x_2+4x_3 & = & 0 \\ -8x_2+4x_3 & = & 12 \\ -22x_3 & = & -22 \quad . \end{array}$$

Cette forme réduite permet de resoudre les équations « de bas en haut ». La dernière équation $-22x_3 = -22$ donne $x_3 = 1$. En substituant dans la deuxième équation on obtient $-8x_2 + 4 \cdot 1 = 12$ et ainsi $x_2 = -1$. En substituant les valeurs connues de x_1,x_2 dans la première équation on obtient $x_1 + 4 \cdot (-1) + 4 \cdot 1 = 0$, c'est-à-dire $x_1 = 0$. Ainsi, $(x_1,x_2,x_3) = (0,-1,1)$ est la solution de (0.5).

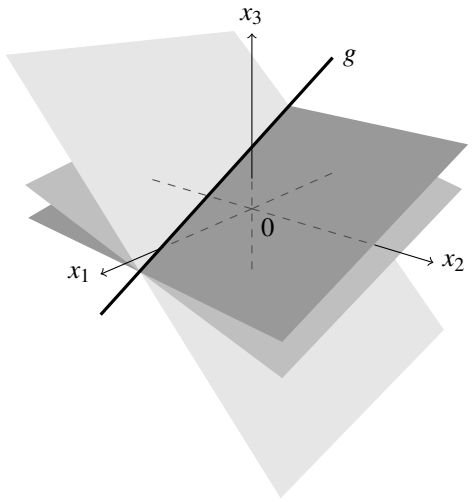


FIG. 0.4 – Interprétation géométrique de (0.6) : L'intersection de trois plans est une droite, qui répresente toutes les solutions de (0.6).

Il n'est pas difficile de construire des systèmes à trois inconnues qui n'ont pas de solution ou qui admettent une infinité de solutions. Par exemple :

$$\begin{array}{rrrcl} 3x_1+ & 4x_2+ & 16x_3 & = & 12 \\ 6x_1+ & x_2+ & 25x_3 & = & 24 \\ & 4x_2+ & 4x_3 & = & 0\,. \end{array} \tag{0.6}$$

L'intersection des trois plans determinés par les trois équations est une droite au lieu d'un point. Tous les points de cette droite sont des solutions de (0.6). Pour les calculer on procède par analogie avec (0.5). En soustrayant la première équation multipliée par 2 à la deuxieme on obtient

$$\begin{array}{rrrcl} 3x_1+4x_2+ & 16x_3 & = & 12 \\ -7x_2- & 7x_3 & = & 0 \\ 4x_2+ & 4x_3 & = & 0\,. \end{array}$$

La troisème équation est redondante en étant équivalente à la deuxième. On choisit comme paramètre libre x_3 et obtient $x_2 = -x_3$. En substituant dans la première équation on obtient $3x_1 - 4x_3 + 16x_3 = 12$, ainsi $x_1 = 4 - 4x_3$. Alors, l'ensemble des solutions est donné par $(4-4x_3, -x_3, x_3)$, qui est la droite g dans la figure 0.4.

Autant d'équations, autant d'inconnues. Comme nous le verrons plus loin, on peut généraliser les observations ci-dessus aux systèmes de m équations linéaires à n inconnues, avec m et n entiers naturels. Cependant avant cela, nous allons introduire des matrices, qui nous permettent de traiter les systèmes linéaires de manière plus élégante.

Chapitre 1
Calcul matriciel

Les matrices jouent un rôle fondamental en algèbre linéaire et plus généralement en mathématiques. À première vue, une matrice n'est rien d'autre qu'un tableau, comme une feuille de calcul Excel ou Google Docs. Mais quand on définit des opérations comme le produit matriciel, les matrices deviennent beaucoup plus intéressant !

Si vous le souhaitez, vous pouvez passer toutes les parties qui sont concernées par MATLAB**. Ce logiciel est très utile pour calculer avec des matrices, mais ce n'est pas nécessaire pour suivre le cours.**

1.1 Matrices, vecteurs colonnes, vecteurs lignes

Définition 1.1 *On appelle* ***matrice*** $m \times n$ *[matrix] (ou matrice de* ***taille*** *[size]* $m \times n$*) un tableau rectangulaire de nombres, arrangés en* m ***lignes*** *[rows] et* n ***colonnes*** *[columns] :*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Pour l'instant, les **éléments** ou **coefficients** a_{ij} $(i = 1,\ldots,m,\ j = 1,\ldots,n)$ d'une matrice A sont des nombres réels. Mais une matrice peut contenir des éléments d'un autre type, comme des nombres complexes ou des polynômes. On va définir plus tard un anneau commutatif K (voir section 2.4) de façon axiomatique et considerer des matrices à coefficients dans A. Toutes les propriétés et définitions du calcul matriciel discutées dans ce chapitre seront valables pour des corps quelconques. On note $M_{m\times n}(K)$ l'ensemble des matrices $m \times n$ à coefficients dans K. Comme dit au-dessus, pour l'instant $K = \mathbb{R}$.

Exemple 1.2

$$A = \begin{pmatrix} 5 & 3 & 1 \\ 4 & -1 & 4 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$$

est une matrice 2×3 à coefficients dans $\mathbb{R}$. Le coefficient $(1,2)$ est $a_{12} = 3$.

$$A = \begin{pmatrix} 1+\mathrm{i} & 2 \\ -2 & 1-\mathrm{i} \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$$

est une matrice 2×2 à coefficients dans $\mathbb{C}$. ♦

MATLAB

Le logiciel MATLAB permet de manipuler facilement des matrices. On entre les commandes suivantes pour l'exemple 1.2 :

```
>> A = [  5   3   1; 4 -1   4 ]
A =  5     3     1
     4    -1     4
>> A(1,2)
ans =
     3
>> A = [ 1+1i, 2; -2, 1-1i ]
A =
   1.0000 + 1.0000i   2.0000 + 0.0000i
  -2.0000 + 0.0000i   1.0000 - 1.0000i
```

Les éléments d'une matrice doivent être mis entre crochets []. Les éléments d'une ligne sont séparés par des espaces ou des virgules et les lignes sont séparées par des point-virgules. Afin d'éviter des erreurs il est recommandé d'utiliser toujours une virgule pour séparer les expressions composées dans une ligne. La commande `size` renvoie la taille (m,n) d'une matrice $m \times n$.

Définition 1.3 *Une matrice $m \times 1$ s'appelle une* ***vecteur colonne de taille*** *m [column vector of length m], une matrice $1 \times n$ s'appelle une* ***vecteur ligne de taille*** *n [row vector of length n].*

On préfère souvent utiliser des vecteurs colonnes. Si le contexte est clair, on dire simplement vecteur. Pour accèder aux éléments d'une matrice générale, on utilise deux indices, un indice ligne et un indice colonne. Seul un indice suffit pour des vecteurs. Par exemple :

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad w = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix}.$$

MATLAB

Le j-ième vecteur colonne d'une matrice A est désigné par `A(:,j)`. Bien entendu, la i-ième ligne est désignée par `A(i,:)`. La commande `length` renvoie la taille d'un vecteur.

```
>> A =  [  8  1  6; 3  5  7; 4  9  2 ];
>> A(:,1),
ans = 8
      3
      4
>> A(2,:),
ans = 3   5   7
```

1.2 Quelques matrices particulières

Matrices carrées. Une matrice $n \times n$ s'appelle **matrice carrée** [*square matrix*] de taille n. On dit parfois **matrice rectangulaire** pour une matrice qui n'est pas nécessairement carrée.

Matrices nulles. Une **matrice nulle** [*zero matrix*] est une matrice $m \times n$ dont tous les coefficients sont nuls. On écrit $0_{m\times n}$ ou, si le contexte est clair, simplement 0.

Exemple 1.4

La matrice nulle de taille 3×3 et le vecteur nul de taille 3 prennent la forme suivante :

$$0_{3\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad 0_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

♦

MATLAB

```
>> zeros(3)
ans = 0     0     0
      0     0     0
      0     0     0
>> zeros(3,1)
ans = 0
      0
      0
```

Diagonale d'une matrice et des matrices diagonales. Soit A une matrice $m \times n$. Les éléments a_{ii} $(i = 1,\ldots,\min\{m,n\})$ s'appellent les **éléments diagonaux** [*diagonal elements*] de la matrice.

Une matrice carrée A de taille n est dite **matrice diagonale** [*diagonal matrix*] si les coefficients en dehors de la diagonale sont nuls, c'est-à-dire $a_{ij} = 0$ si $i \neq j$. On note

$$D = \mathrm{diag}\,(d_{11},d_{22},\ldots,d_{nn}) = \begin{pmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

la matrice diagonale avec les éléments $d_{11},\ldots,d_{nn}$ sur la diagonale (et avec des zéros ailleurs).

Exemple 1.5

$$D = \mathrm{diag}\,(1,2,3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

est une matrice diagonale de taille 3. Les éléments diagonaux de

$$A = \begin{pmatrix} 5 & 3 & 1 \\ 4 & -1 & 4 \end{pmatrix}$$

sont $5, -1$.

♦

MATLAB

```
>> diag([ 5  3  1; 4 -1  4 ]),
ans =  5
      -1
>> diag([ 1  2  3 ]),
ans = 1     0     0
      0     2     0
      0     0     3
```

Matrice identité. Une matrice diagonale de taille n s'appelle une **matrice identité** [*identity matrix*] ou **matrice unité** [*unit matrix*] si $a_{ii} = 1$ pour $i = 1,\ldots,n$. Elle peut s'écrire

$$I_n = \mathrm{diag}\,(1,1,\ldots,1).$$

Exemple 1.6

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

♦

MATLAB

```
>> eye(3),
ans = 1     0     0
      0     1     0
      0     0     1
```

Matrices triangulaires. Une **matrice triangulaire supérieure** [*upper triangular matrix*] est une matrice $U \in M_{m\times n}(K)$ dont toutes les coefficients sous la diagonale sont nuls, c-à-d $u_{ij} = 0$ si $i > j$.

Une **matrice triangulaire inférieure** [*lower triangular matrix*] est une matrice $L \in M_{m\times n}(K)$ dont toutes les coefficients au-dessus de la diagonale sont nuls, c-à-d $\ell_{ij} = 0$ si $i < j$.

On utilise les symboles suivants pour désigner des matrices *carrées* triangulaires:

$$U = \begin{array}{|ll} \diagdown & \\ & \end{array}, \qquad L = \begin{array}{|ll} & \\ \diagup & \end{array}.$$

Exemple 1.7 Exemple d'une matrice triangulaire supérieure (respectivement inférieure) :

$$U = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

$$L = \begin{pmatrix} 7 & 0 & 0 \\ 1 & 8 & 0 \\ 1 & 5 & 6 \end{pmatrix}.$$

Exemples de matrices triangulaires rectangulaires :

$$U = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 2 & 4 & 6 \end{pmatrix},$$

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

MATLAB

```
>> A = [  8  1  6; 3  5  7;
          4  9  2 ];
% triu extrait la partie
% triangulaire superieure
>> triu(A),
ans = 8     1     6
      0     5     7
      0     0     2
% tril extrait la partie
% triangulaire inferieure
>> tril(A),
ans = 8     0     0
      3     5     0
      4     9     2
```

♦

1.3 Notation

Dans ce cours, les matrices sont désignées par des lettres latines majuscules (A,B,...) et vecteurs (colonnes ou lignes) par des lettres latines minuscules (v,w,x,y,...). Les scalaires (qui ne sont pas des éléments d'une matrice ou d'un vecteur) sont souvent désignés par des minuscules grecques (α,β,γ,...).

1.4 Applications des matrices

Cette section présente deux applications simples des matrices.

1.4.1 Images

Sous MATLAB, la commande `imread` permet de lire une image, par exemple `X = imread('ngc6543a.jpg')`. Dans le cas d'un image en niveaux de gris avec $n \times m$ pixels, cette commande renvoie une matrice X de taille $m \times n$ dont l'élément (i,j) est le niveau de gris du pixel (j,i). Dans le cas d'un image couleur, on obtient un tableau à 3 dimensions de taille $m \times n \times 3$. Pour chaque pixel, les niveaux de rouge, vert et bleu (RVB, RGB en anglais) sont codés entre un minimum de 0 et un maximum de 255, voir tableau 1.1. La première matrice $m \times n$ (accès par `X(:,:,1)`) contient les niveaux de rouge, la deuxième les niveaux de vert, et la troisième les niveaux de bleu.

Exemple 1.8

Le script MATLAB à droite produit l'échiquier figurant dans la figure 1.1. La commande `X(1:2:8,1:2:8)` permet de supprimer une ligne et une colonne sur deux. ♦

MATLAB

```
X = zeros(8,8,'uint8');
X(1:2:8,1:2:8) = 255;
X(2:2:8,2:2:8) = 255;
X(:,:,2) = X;
X(:,:,3) = X(:,:,1);
image(X); axis off, axis square
```

TAB. 1.1 – Exemples du code RVB

valeur de rouge	valeur de vert	valeur de bleu	couleur
0	0	0	noir
255	255	255	blanc
255	0	0	rouge
0	128	0	vert
0	0	255	bleu
255	255	0	jaune
0	255	255	cyan
255	0	255	magenta

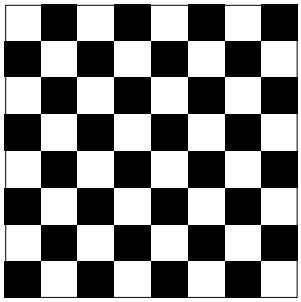


FIG. 1.1 – L'échiquier de l'exemple 1.8

1.4.2 Graphes

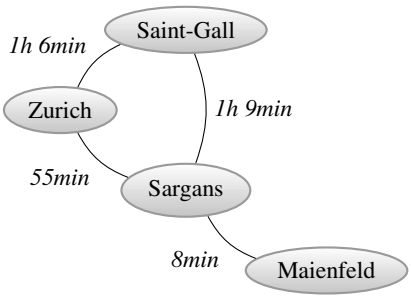
Un **graphe** est un couple $G = (V,E)$ où V est l'ensemble des **nœuds** (ou des **sommets**, [*nodes*] ou [*vertices*]) et E l'ensemble des **arêtes** [*edges*]. Les arêtes peuvent être orientées (arcs ou flèches) ou non orientées (traits).

Graphes non orientés. Dans ce cas, une arête est une paire non ordonnée de nœuds, ainsi $E \subset \mathscr{P}_2(V)$.

Exemple 1.9 (Graphe Heidi)
La figure à droite montre une (petite) partie du réseau CFF, avec les temps du parcours. On peut considérer le reseau comme un graph non orienté avec

$$V = \{\text{Maienfeld}, \text{Saint-Gall}, \text{Sargans}, \text{Zurich}\},$$

$$E = \{ \{\text{Maienfeld}, \text{Sargans}\}, \{\text{Sargans}, \text{Saint-Gall}\}, \{\text{Sargans}, \text{Zurich}\}, \{\text{Saint-Gall}, \text{Zurich}\} \}.$$



♦

On peut représenter un graphe $G = (V,E)$ avec n nœuds par une matrice $n \times n$, la **matrice d'adjacence** [*adjacency matrix*]. Cette matrice A est définie en posant $a_{ij} = 1$, s'il existe une arête entre les sommets i et j, et $a_{ij} = 0$ sinon. Par exemple, pour le graphe Heidi, en

numérotant les sommets dans l'ordre alphabétique croissant, on obtient :

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \tag{1.1}$$

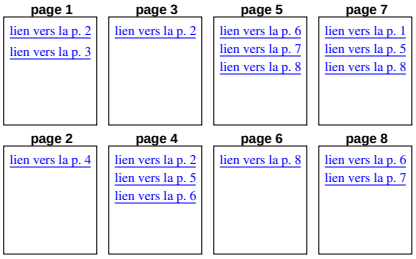
On appelle un **graphe pondéré** [*weighted graph*] un graphe dont les arêtes portent des poids. Par exemple, les arêtes du graphe Heidi portent les temps du parcours. En posant formellement $a_{ij} = \infty$ s'il n'y a pas de connexion directe entre les lieux i et j, on obtient la matrice d'adjacence pondérée

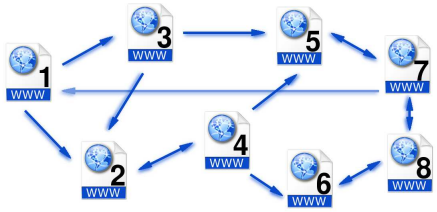
$$A = \begin{pmatrix} 0 & 8 & \infty & \infty \\ 8 & 0 & 69 & 55 \\ \infty & 69 & 0 & 66 \\ \infty & 55 & 66 & 0 \end{pmatrix}. \tag{1.2}$$

Ces matrices sont utilisées pour déterminer le temps de parcours les plus courts.

Graphes orientés. Dans ce cas, un arc est un couple (ordoné) de nœuds, ainsi $E \subset V \times V$.

Exemple 1.10 (Graphe du Web) Le World Wide Web est un graphe gigantesque, dont les sommets sont les pages Web. Un arc du sommet i vers le sommet j indique qu'il existe un lien sur la page i vers la page j. Les illustrations suivantes montrent un réseau Intranet avec 8 pages et le graphe du Web correspondant:





En choisissant comme ensemble des sommets $V = \{1,2,\ldots,8\}$, l'ensemble des arcs est donné par

$$E = \big\{ \begin{array}{l} (1,2),(1,3),(2,4),(3,2),(3,5),(4,2),(4,5),(4,6), \\ (5,7),(6,8),(7,1),(7,5),(7,8),(8,6),(8,7) \big\}. \end{array}$$

♦

La matrice d'adjacence d'un graphe orienté est définie de façon similaire à celle d'un graphe non orienté: $a_{ij} = 1$ s'il existe un arc du sommet i vers le sommet j, $a_{ij} = 0$ si non. Pour le graphe du web ci-dessu on obtient

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

1.5 Opérations sur les matrices et les vecteurs

1.5.1 Opérations élémentaires

Multiplication d'une matrice par un scalaire. Soit $A \in M_{m\times n}(K)$ et $\alpha \in K$. Pour effectuer le produit $\alpha A \in M_{m\times n}(K)$ on multiplie tous les éléments de A par α:

$$(\alpha A)_{ij} = \alpha a_{ij} \qquad (i = 1,\ldots,m;\ j = 1,\ldots,n).$$

Exemple 1.11

$$5\begin{pmatrix} 1 & -3 & 5 \\ -2 & 4 & -6 \end{pmatrix} = \begin{pmatrix} 5 & -15 & 25 \\ -10 & 20 & -30 \end{pmatrix},$$
$$\tfrac{1}{4}\begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

MATLAB
```
>> 5*[ 1 -3  5; -2  4 -6 ],
ans =      5   -15    25
         -10    20   -30
>> [ 4; 8 ] / 4,
ans =      1
           2
```

Somme de matrices. On définit la somme de deux matrices $A,B \in M_{m\times n}(K)$ comme la matrice $C \in M_{m\times n}(K)$ telle que

$$c_{ij} = a_{ij} + b_{ij} \qquad (i = 1,\ldots,m;\ j = 1,\ldots,n).$$

Lemme 1.12 *Soient $A,B,C \in M_{m\times n}(K)$ et $\alpha,\beta \in K$. Alors*

(i) $(\alpha\beta)A = \alpha(\beta A)$,
(ii) $(\alpha+\beta)A = (\alpha A)+(\beta A)$,
(iii) $\alpha(A+B) = (\alpha A)+(\alpha B)$,
(iv) $A+B = B+A$,
(v) $A+0_{m\times n} = A$,
(vi) $A+(-1\cdot A) = 0_{m\times n}$.

DÉMONSTRATION. Pour $K = \mathbb{R}$ les propriétés découlent des propriétés analogues dans $\mathbb{R}$. Par exemple la démonstration de (iv):

$$\begin{aligned} (A+B)_{ij} &= a_{ij}+b_{ij} && \text{(déf. d'addition matricielle)} \\ &= b_{ij}+a_{ij} && \text{(commutativité dans } \mathbb{R}) \\ &= (B+A)_{ij} && \text{(déf. d'addition matricielle)} \end{aligned}$$

Pour un corps les propriétés découlent des axiomes des corps, voir section 2.4. ■

Grâce au point (vi) du lemme 1.12, on définit $-A = -1\cdot A$ et la **soustraction matricielle** par $B-A = B+(-A)$.

Exemple 1.13

$$\begin{pmatrix} 5 & 2 \\ -1 & -5 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 5 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 2 \\ 5 \\ 9 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 12 \end{pmatrix}.$$

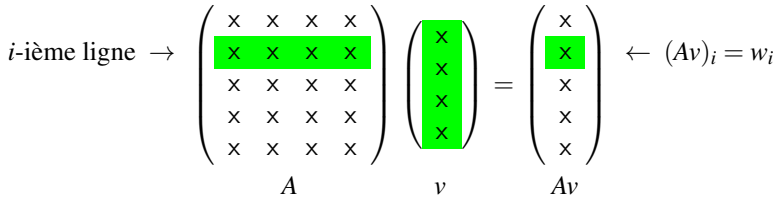
MATLAB
```
>> [  5  2; -1 -5 ] + ...
   [ -1  1;  6  5 ],
ans =     4     3
          5     0
>> [ 2; 5; 9 ] - [ 1; -1; -3 ],
ans =     1
          6
         12
```

1.5.2 Multiplication matrice vecteur

Définition 1.14 *Soit $A \in M_{m\times n}(K)$ et v un vecteur colonne de taille n à coefficients dans K. Le produit matrice-vecteur [matrix-vector product] $w = Av$ est un vecteur colonne de taille m à coefficients dans K dont les coefficients sont donnés par :*

$$w_i = \sum_{k=1}^{n} a_{ik} v_k = a_{i1} v_1 + a_{i2} v_2 + \cdots + a_{in} v_n, \quad (i = 1, \ldots, m). \tag{1.3}$$

La définition (1.3) peut s'exprimer da la manière suivante : le i-ième coefficient de w est donné par la somme des produits des éléments de la i-ième ligne de A par les éléments de v. Voici une illustration :



Exemple 1.15

$$A = \begin{pmatrix} 1 & -3 & 5 \\ -2 & 4 & -6 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},$$

$$w = Av = \begin{matrix} 1\cdot 2 & + & (-3)\cdot 1 & + & 5\cdot(-1) \\ (-2)\cdot 2 & + & 4\cdot 1 & + & (-6)\cdot(-1) \end{matrix} = \begin{pmatrix} -6 \\ 6 \end{pmatrix}.$$

MATLAB

Sous MATLAB on utilise l'operateur $\star$ pour calculer un produit matrice-vecteur:

```
>> A = [ 1 -3  5; -2  4 -6 ];
>> b = [ 2; 1; -1 ];
>> A*b,
ans =   -6
         6
```

Produit matrice-vecteur et système linéaire. La definition de la produit matrice-vecteur permet d'abréger la description d'un système linéaire. On considère par exemple le système linéaire (0.5):

$$\begin{matrix} x_1 + & 4x_2 + & 4x_3 & = & 0 \\ 3x_1 + & 4x_2 + & 16x_3 & = & 12 \\ 4x_1 + & 2x_2 + & x_3 & = & -1 \end{matrix} \quad . \tag{1.4}$$

Si on définit

$$A = \begin{pmatrix} 1 & 4 & 4 \\ 3 & 4 & 16 \\ 4 & 2 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 12 \\ -1 \end{pmatrix},$$

l'équation $Ax = b$ et le système (1.4) sont équivalentes.

Plus généralement, un système linéaire de m èquations à n inconnues prend la forme suivante:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array}$$

Si on définit une matrice A $(m \times n)$ ayant pour éléments les a_{ij}, un vecteur x ayant pour éléments les x_j et un vecteur b ayant comme éléments les b_i, ce système linéaire est encore équivalent à

$$\boxed{Ax = b.} \tag{1.5}$$

On appelle parfois A la **matrice des coefficients du système** [*coefficient matrix, system matrix*], x le **vecteur solution** [*solution vector*] et b le **côté droit** [*right-hand side*].

Suite de Fibonacci. La **suite de Fibonacci** [*Fibonacci sequence*]

$$1,1,2,3,5,8,13,21,34,55,89,144,233,377,\ldots$$

est définie par la relation de récurrence:

$$f_0 = 1, \quad f_1 = 1, \qquad f_{k+2} = f_{k+1} + f_k, \quad k \geq 0. \tag{1.6}$$

Leonardo Fibonacci, en utilisant cette récurrence, a décrit la croissance explosive d'une population de lapins.

Si on définit $b_k = \begin{pmatrix} f_k \\ f_{k+1} \end{pmatrix}$ pour $k = 0,1,\ldots$, la récurrence (1.6) devient un produit matrice-vecteur:

$$b_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad b_{k+1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} b_k, \quad k \geq 0. \tag{1.7}$$

Plus tard, lorsque nous parlerons des valeurs propres, cette représentation donnera une expression explicite de suite de Fibonacci.

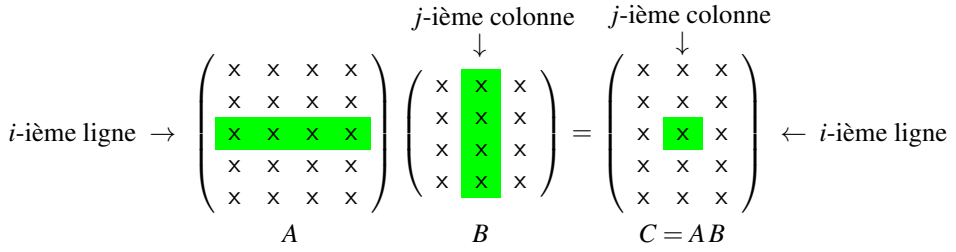
1.5.3 Produit matriciel

Définition 1.16 *Soient $A \in M_{m\times n}(K)$ et $B \in M_{n\times p}(K)$. Le **produit matriciel** [matrix product] $C = AB$ est une matrice $m \times p$ à coefficients dans K dont les coefficients sont donnés par :*

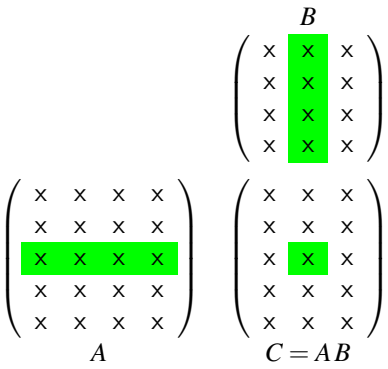
$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}, \quad (i = 1,\ldots,m;\ j = 1,\ldots,p). \tag{1.8}$$

La définition (1.8) peut s'exprimer de la manière suivante : l'élément (i,j) de C est donné par la somme des produits de éléments de la i-ième ligne de A par les éléments de la j-ième

colonne de B. Voici une illustration de cette « règle ligne-colonne »pour $m=5, n=4, p=3$:



Pour calculer le produit matriciel à la main il est recommandé de placer la matrice A à gauche et la matrice B en haut de la matrice C:

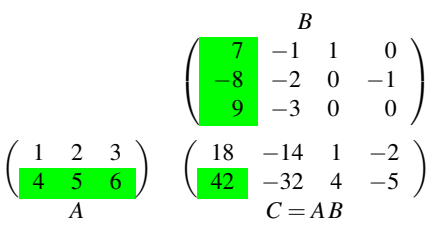


On remarque que le produit matrice-vecteur est un cas particulier du produit matrice-matrice.

Exemple 1.17 Soient

$$A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix},\qquad B=\begin{pmatrix}7&-1&1&0\\-8&-2&0&-1\\9&-3&0&0\end{pmatrix}.$$

On calcule le produit matricielle à l'aide du schema ci-dessus:



Par exemple, on calcule pour l'élément $(2,1)$:

$$\begin{aligned}a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31}&=4\cdot 7+5\cdot(-8)+6\cdot 9\\&=28-40+54=42.\end{aligned}$$

MATLAB

Comme dans le cas de produit matrice-vecteur, on calcule un produit matriciel à l'aide de l'operateur $\star$ sous MATLAB:

```
>> A = [ 1  2  3; ...
         4  5  6 ];
>> B = [ 7 -1  1  0;
        -8 -2  0 -1;
         9 -3  0  0 ];
>> A*B,
ans =
    18   -14     1    -2
    42   -32     4    -5
```

Remarque 1.18 *La définition du produit matriciel correspond à la composition des produits matrice-vecteur. Soient $A\in M_{m\times n}(K)$, $B\in M_{n\times p}(K)$ et $v\in M_{p\times 1}(K)$. On desire calculer le vecteur $x=A(Bv)$, qu'on separe dans les deux produits $w=Bv$ et $x=Aw$. Par* (1.3),

nous avons

$$w_j = \sum_{k=1}^{p} b_{jk} v_k \qquad (j = 1, \ldots, p)$$

et

$$x_i = \sum_{j=1}^{n} a_{ij} w_j = \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{p} b_{jk} v_k = \sum_{k=1}^{p} \underbrace{\sum_{j=1}^{n} a_{ij} b_{jk}}_{=:c_{ik}} v_k. \qquad (i = 1, \ldots, m)$$

Alors, $x = Cv$ òu $C = AB$.

Produit des matrices particulières. Le produit matriciel se simplifie, parfois fortement, dans des cas particuliers.

Multiplication par une matrice nulle. Si A ou B est une matrice nulle (de la taille appropriée), tous les termes de la somme (1.8) sont nuls. Alors, le produit matriciel est nul:

$$A 0_{n\times p} = 0_{m\times p}, \qquad 0_{m\times n} B = 0_{m\times p},$$

pour toutes matrices $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$.

Multiplication par la matrice identité. Si $B = I_n$, les éléments de $C = AB$ satisfont

$$c_{ij} = a_{i1} \underbrace{b_{1j}}_{=0} + \cdots + a_{i,j-1} \underbrace{b_{j-1,j}}_{=0} + a_{ij} \underbrace{b_{jj}}_{=1} + a_{i,j+1} \underbrace{b_{j+1,j}}_{=0} + \cdots + a_{in} \underbrace{b_{nj}}_{=0} = a_{ij}. \tag{1.9}$$

Alors, $A I_n = A$ et, de manière analogue, $I_m B = B$ pour toutes matrices $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$.

Multiplication par une matrice diagonale. Par analogie avec (1.9), le produit matriciel $C = AD$ d'une matrice $A \in M_{m\times n}(K)$ par une matrice diagonale $D = \mathrm{diag}\,(d_{11}, \ldots, d_{nn})$ se calcule comme suit:

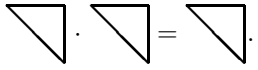
$$c_{ij} = a_{ij} d_{jj}.$$

En notant $a_1, \ldots, a_n$ les colonnes de la matrice A et $c_1, \ldots, c_n$ celles de la matrice C, on a

$$C = \left(\begin{array}{c|c|c|c} c_1 & c_2 & \cdots & c_n \end{array} \right) = \left(\begin{array}{c|c|c|c} d_{11} a_1 & d_{22} a_2 & \cdots & d_{nn} a_n \end{array} \right).$$

Alors, multiplier une matrice A à droite par une matrice diagonale revient à multiplier chaque colonne de A par l'élément diagonal correspondant. Cette opération s'appelle « diagonal scaling »en anglais. De manière analogue, on peut voir que la multiplication à gauche par une matrice diagonale revient à multiplier chaque ligne de A par lélément diagonal correspondant.

Produit de matrices triangulaires. Si A et B sont matrices $n \times n$ triangulaires supérieures (inférieures), le produit matriciel $C = AB$ est encore une matrice triangulaire supérieure (inférieure); voir ci-dessous, lemme 2.8. Voici une illustration de cette assertion par des symboles:



Plus tard, en chapitre 2, on va discuter un autre cas particulier important : la multiplication par une matrice de permutation.

Propriétés du produit matriciel.

Lemme 1.19

$$(AB)C = A(BC) \qquad \forall A \in M_{m\times n}(K),\, B \in M_{n\times p}(K),\, C \in M_{p\times q}(K), \tag{1.10}$$
$$(A+B)C = (AC)+(BC) \qquad \forall A,B \in M_{m\times n}(K),\, C \in M_{n\times p}(K), \tag{1.11}$$
$$A(B+C) = (AB)+(AC) \qquad \forall A \in M_{m\times n}(K),\, B,C \in M_{n\times p}(K). \tag{1.12}$$

DÉMONSTRATION. Voir les exercises. ■

Quelques remarques:

1. La propriété associative (1.10) permet de supprimer les parenthèses: On peut écrire $A(BC) = ABC$. En plus, les multiplications sont prioritaires sur les additions et soustractions (comme dans $\mathbb{R}$ ou $\mathbb{C}$). Par exemple, on a $(AB)+(CD) = AB+CD$.
2. Lemme 1.19 est un résultat théorique. En pratique, il peut y avoir une grande différence entre le coût numerique des deux opérations équivalentes. Par exemple, si $A \in M_{n\times 1}(\mathbb{R})$, $B \in M_{1\times n}(\mathbb{R})$, $C \in M_{n\times n}(\mathbb{R})$, l'ordre des opérations $(AB)C$ est beaucoup plus cher que $A(BC)$.
3. Remarque 1.18 est un cas particulier de (1.10).

Le produit matriciel n'est, en général, pas commutatif:

$$AB \neq BA\,.$$

Par exemple, soient

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

On a

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 3 & 3 \end{pmatrix},$$

mais le produit BA n'est pas même défini. Même quand AB et BA sont définis, le produit ne commute pas en général.

Exemple 1.20

$$A = \begin{pmatrix} 2 & 6 \\ 1 & 7 \end{pmatrix}, \qquad B = \begin{pmatrix} -3 & -1 \\ 2 & 1 \end{pmatrix}$$
$$AB = \begin{pmatrix} 6 & 4 \\ 11 & 6 \end{pmatrix} \neq BA = \begin{pmatrix} -7 & -25 \\ 5 & 19 \end{pmatrix}.$$

Lorsque $AB = BA$, on dit que les matrices A et B **commutent**. Des matrices qui commutent sont l'exception plutôt que la règle. La matrice identité I_n commute avec toutes les matrices $n \times n$: $I_n A = A = A I_n$. On a l'assertion plus forte : Une matrice $A \in M_{n\times n}(K)$ commute avec toutes les matrices $n \times n$ si et seulement si A est scalaire. [1] (Voir exercices)

Remarque 1.21 *Pour deux matrices $A,B \in M_{m\times n}(K)$ le produit de Hadamard (ou le produit de Schur) de A et B est une matrice $m \times n$ dont les coefficients sont $a_{ij}b_{ij}$ ($i = 1,\ldots,m$, $j = 1,\ldots,n$). Contrairement au produit matriciel classique, ce produit terme à terme est commutatif mais, malheureusement, il est beaucoup moins utile! Pour réaliser ces opérations terme à terme sous* MATLAB*, on met un point avant l'operateurs:* `A.*B`, `A./B`, `A.^B`.

1. Une matrice scalaire est une matrice diagonale dont tous les coefficients diagonaux sont égaux.

Une autre formulation du produit matrice-vecteur/matrice. Soit $A \in M_{m\times n}(K)$. On considère les colonnes de A:

$$A = \left(\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_n \end{array} \right),$$

où $a_1, a_2, \ldots, a_n \in M_{m\times 1}(K)$. Le lemme suivant permet de formuler un produit matrice-vecteur comme une combinaison linéaire (voir chapitre 4) des colonnes de A.

Lemme 1.22 *Soit $A \in M_{m\times n}(K)$ et $a_1, a_2, \ldots, a_n \in M_{m\times 1}(K)$ les colonnes de A comme ci-dessus, et $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}(K)$. Alors,*

$$\boxed{Ax = a_1x_1 + a_2x_2 + \cdots + a_nx_n = x_1a_1 + x_2a_2 + \cdots + x_na_n\,.} \tag{1.13}$$

DÉMONSTRATION. Pour vérifier (1.13), on considère le i-ième élément :

$$(Ax)_i = \sum_{k=1}^{n} a_{ik}x_k = \sum_{k=1}^{n} (a_k)_i x_k = \sum_{k=1}^{n} (a_k x_k)_i = \sum_{k=1}^{n} (x_k a_k)_i.$$

■

Un cas particulier interessant de (1.13) est celui où $x = e_j$, la j-ième colonne de la matrice identité I_n :

$$\boxed{Ae_j = a_j\,.} \tag{1.14}$$

Lemme 1.23 *Soit $A \in M_{m\times n}(K)$ et $B = \begin{pmatrix} b_1 & b_2 & \cdots & b_p \end{pmatrix} \in M_{n\times p}(K)$. Alors,*

$$AB = \left(\begin{array}{c|c|c|c} Ab_1 & Ab_2 & \cdots & Ab_p \end{array} \right).$$

DÉMONSTRATION. Par (1.14), on a $b_j = Be_j$, $j = 1, \ldots, p$. Pour la même raison, la j-ième colonne du produit AB est $(AB)e_j$. Grâce à l'associativité de la multiplication matricielle, le lemme 1.22 implique

$$(AB)e_j = A(Be_j) = Ab_j$$

pour $j = 1, \ldots, p$. ■

1.6 La transposée d'une matrice

Définition 1.24 *Soit $A \in M_{m\times n}(K)$. On appelle* ***transposée*** *[transpose] de A la matrice $A^\mathsf{T} \in M_{n\times m}(K)$ dont les éléments sont*

$$(A^\mathsf{T})_{ij} = a_{ji}.$$

Exemple 1.25

$$A = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}, A^\mathsf{T} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}.$$

MATLAB

Pour transposer une matrice (reelle), on utilise le symbole apostrophe sous MATLAB:

```
>> A = [ 1 5; 2 6; 3 7; 4 8 ]; A'
ans =
     1     2     3     4
     5     6     7     8
```

On remarque que A^T est la matrice dont les lignes sont les colonnes de A. En particulier, la tranposée d'un vecteur colonne est un vecteur ligne et vice-versa.

Lemme 1.26 *(i)*

$$\boxed{(A^\mathsf{T})^\mathsf{T} = A} \qquad \forall A \in M_{m\times n}(K). \tag{1.15}$$

(ii)

$$\boxed{(\alpha A)^\mathsf{T} = \alpha A^\mathsf{T}} \qquad \forall A \in M_{m\times n}(K), \alpha \in K. \tag{1.16}$$

(iii)

$$\boxed{(A+B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T}} \qquad \forall A,B \in M_{m\times n}(K). \tag{1.17}$$

(iv)

$$\boxed{(AB)^\mathsf{T} = B^\mathsf{T}A^\mathsf{T}} \qquad \forall A \in M_{m\times n}(K), B \in M_{n\times p}(K). \tag{1.18}$$

DÉMONSTRATION. Les propriétés (i)–(iii) sont des exercises faciles. Pour montrer (iv), on remarque tout d'abord que la matrice AB est de taille $m \times p$ et ainsi $(AB)^\mathsf{T}$ est une matrice $p \times m$. Le produit $B^\mathsf{T}A^\mathsf{T}$ est aussi une matrice $p \times m$, avec les éléments suivants:

$$\begin{aligned}((AB)^\mathsf{T})_{ij} = (AB)_{ji} &= \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ki}a_{jk} \\ &= \sum_{k=1}^{n} (B^\mathsf{T})_{ik}(A^\mathsf{T})_{kj} = (B^\mathsf{T}A^\mathsf{T})_{ij}.\end{aligned}$$

■

1.7 Matrices symétriques

Définition 1.27 *On dit qu'une matrice carrée A est* ***symétrique*** *[symmetric] si*

$$A^\mathsf{T} = A, \quad c\text{-}\grave{a}\text{-}d \quad a_{ij} = a_{ji} \quad \forall i,j = 1,\ldots,n.$$

Nous avons déjà fait la connaissance des matrices symétriques: La matrice d'adjacence (1.1) d'un graphe non orienté. La définition suivante est quasiment le contraire de la symétrie.

Définition 1.28 *On dit qu'une matrice carrée A est* ***antisymétrique*** *[skew-symmetric] si $A^\mathsf{T} = -A$.*

Exemple 1.29 Soient

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

La matrice A est symétrique et la matrice B est antisymétrique.

Remarque 1.30 *Les éléments diagonaux d'une matrice antisymétrique sont toujours nuls: $a_{ii} = 0$ $(i = 1,\ldots,n)$. Toute matrice carrée s'écrit, de façon unique, comme la somme d'une matrice symétrique et d'une matrice antisymétrique.*

Le produit de deux matrices symétriques n'est, en général, pas symmetrique! Par exemple,

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

Mais une autre relation est vraie.

Lemme 1.31 *Soient $A \in M_{m\times n}(K)$ et $B \in M_{m\times m}(K)$. Si B est symétrique, alors $A^\mathsf{T}BA$ est aussi symétrique.*

DÉMONSTRATION. Voir les exercices. ■

Un cas particulier ($B = I_m$) de lemme 1.31 : $A^\mathsf{T}A$ est symétrique pour toute matrice $A \in M_{m\times n}$. De façon analogue, AA^T est symétrique.

1.8 L'inverse d'une matrice

Définition 1.32 *Une matrice $A \in M_{n\times n}(K)$ est dite **inversible** s'il existe une matrice $X \in M_{n\times n}(K)$ telle que*

$$AX = XA = I_n. \tag{1.19}$$

*Une matrice non-inversible est **singulière**.*

Tout nombre réel non nul possède un inverse. Au contraire, si $n \geq 2$, une matrice carrée non nulle n'est pas toujours inversible. Par exemple, pour la matrice

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

il est impossible de trouver une matrice $X \in M_{2\times 2}(\mathbb{R})$ telle que $AX = I_2$:

$$AX = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix} \neq I_2.$$

En général, ce n'est pas facile de déterminer si une matrice est inversible, voir chapitre 3.

Lemme 1.33 *L'inverse d'une matrice s'il existe (c-à-d il existe X avec (1.19)) est unique.*

DÉMONSTRATION. Si $AX = XA = I_n = A\tilde{X} = \tilde{X} = I_n$, alors $\tilde{X} = \tilde{X}\cdot I_n = \tilde{X}AX = I_n \cdot X = X$. ■

On note A^{-1} l'inverse d'une matrice $A \in M_{n\times n}(K)$ (s'il existe) :

$$\boxed{AA^{-1} = I_n = A^{-1}A\,.}$$

Exemple 1.34 On a

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

Cela signifie que ces matrices sont inverses l'une de l'autre.
L'inverse d'une matrice diagonale est encore une matrice diagonale :

$$\mathrm{diag}\,(d_{11},\ldots,d_{nn})^{-1} = \mathrm{diag}\,(1/d_{11},\ldots,1/d_{nn}).$$

En particulier la matrice d'identité I_n est bien sûr inversible : I_n est sa propre inverse !

MATLAB

Sous MATLAB on calcule l'inverse par `inv`.

```
>> A = [ 2 2; 1 2 ]; inv(A)
ans =
    1.0000   -1.0000
   -0.5000    1.0000
```

Mais attention ! Dans la pratique on a rarement besoin de utiliser `inv` ! En lieu de calculer explicitement l'inverse on a assez souvent besoin de multiplier l'inverse d'une matrice par une autre matrice ou un vecteur. Dans ce cas les commandes `A \ B` resp. `B / A` sont moins coûteuses en calcul et plus précises que `inv(A)*B` resp. `B*inv(A)`.

```
>> b = [ 1; 0 ]; A \ b
ans =
    1.0000
   -0.5000
```

Il se trouve qu'une de deux conditions (1.19) suffit à determiner l'inverse.

Lemme 1.35 *Soit $A \in M_{n\times n}(K)$. Les énoncés suivants sont équivalents :*

i) *A est inversible.*

ii) *Il existe une matrice* $X \in M_{n\times n}(K)$ *telle que* $AX = I_n$.
iii) *Il existe une matrice* $X \in M_{n\times n}(K)$ *telle que* $XA = I_n$.

DÉMONSTRATION. Voir chapitre 3. ■

Le lemme suivant contient quelques propriétés de l'inverse.

Lemme 1.36 *Soient* $A,B \in M_{n\times n}(K)$ *des matrices inversibles. Alors :*

i) A^{-1} *est inversible et*

$$\boxed{(A^{-1})^{-1} = A\,.} \tag{1.20}$$

ii) AB *est inversible et*

$$\boxed{(AB)^{-1} = B^{-1}A^{-1}\,.} \tag{1.21}$$

iii) A^{T} *est inversible et*

$$\boxed{(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}\,.} \tag{1.22}$$

DÉMONSTRATION. i) découle directement de la définition (1.19).

ii) Par $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n$ et par le lemme 1.35, la matrice $B^{-1}A^{-1}$ est l'inverse de AB.

iii) La règle (1.18) de transposition signifie que $A^{\mathsf{T}}(A^{-1})^{\mathsf{T}} = (A^{-1}A)^{\mathsf{T}} = I_n^{\mathsf{T}} = I_n$. Ainsi, $(A^{-1})^{\mathsf{T}}$ est l'inverse de A^{T}. ■

1.9 Sous-matrices

En ne gardant que certaines lignes ou colonnes d'une matrice on obtient une nouvelle matrice que l'on appelle sous-matrice.

Définition 1.37 *Soit* A *une matrice* $m\times n$ *et soient* $\mathscr{I} = \{i_1,\ldots,i_k\}$, $\mathscr{J} = \{j_1,\ldots,j_\ell\}$ *sous-ensembles de* $\mathbb{N}$ *avec*

$$1 \le i_1 < \cdots < i_k \le m, \quad 1 \le j_1 < \cdots < j_\ell \le n\,.$$

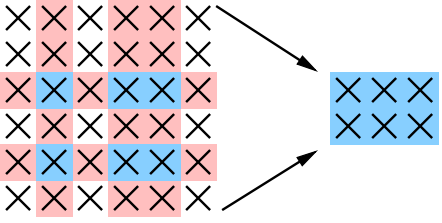
La matrice $k\times\ell$

$$A(\mathscr{I},\mathscr{J}) = \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_\ell} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_\ell} \\ \vdots & \vdots & & \vdots \\ a_{i_m,j_1} & a_{i_m,j_2} & \cdots & a_{i_m,j_\ell} \end{pmatrix}$$

est la ***sous-matrice*** *[submatrix] correspondante de* A. *On dit qu'une sous-matrice est* ***principale*** *si* $\mathscr{I} = \mathscr{J}$.

Exemple 1.38

Une illustration des sous-matrices pour $m = n = 6$, $\mathscr{I} = \{3,5\}$, $\mathscr{J} = \{2,4,5\}$:



MATLAB

```
>> A = magic(6)
A =
   35    1    6   26   19   24
    3   32    7   21   23   25
   31    9    2   22   27   20
    8   28   33   17   10   15
   30    5   34   12   14   16
    4   36   29   13   18   11
>> A([3,5],[2,4,5])
ans =
    9   22   27
    5   12   14
```

Chapitre 2
Structures algébriques

Nous avons vu que beaucoup de règles des opérations sur les nombres réels sont aussi valables pour les matrices, voir les lemmes 1.12 et 1.19. Mais en même temps nous avons vu que le passage des nombres aux matrices créée des différences importantes, notamment la perte de la commutativité de la multiplication. Dans ce chapitre nous allons discuter des structures algébriques qui non seulement capturent les diffrences entre les nombres et les matrices mais aussi couvrent beaucoup d'autres objets (fonctions, polynômes, division euclidienne, ...). Vous avez déjà vu une partie de ce chapitre dans le cours de Géométrie I.

2.1 Groupes

Définition 2.1 *Un* ***groupe*** *[group] est un ensemble G muni d'une loi de composition*

$$\begin{array}{rccc} \star : & G\times G & \to & G \\ & (a,b) & \mapsto & a\star b \end{array}$$

satisfaisant les axiomes suivants :

1. *La loi $\star$ est associative, c-à-d*
$$a\star(b\star c)=(a\star b)\star c \quad \forall a,b,c\in G.$$
2. *Il existe un élément $e\in G$ (appelé* ***élément neutre*** *ou* ***identité*** *[identity]) tel que*
$$a=e\star a=a\star e \quad \forall a\in G.$$
3. *Pour tout $a\in G$ il existe un élément $a^{-1}\in G$ (appelé* ***l'inverse*** *de a) tel que*
$$a^{-1}\star a=a\star a^{-1}=e.$$

Définition 2.2 *Un groupe $(G,\star)$ est dit* ***abélien*** *ou* ***commutatif*** *si*

$$a\star b=b\star a \quad \forall a,b\in G.$$

Lemme 2.3 *Soit $(G,\star)$ un groupe. Alors :*

1. *l'élément neutre e est unique,*
2. *l'inverse de $a\in G$ est unique,*
3. *$(a^{-1})^{-1}=a$ pour tout $a\in G$,*
4. *$(a\star b)^{-1}=b^{-1}\star a^{-1}$ pour tous $a,b\in G$.*

DÉMONSTRATION. Voir les exercices de Géométrie I. ■

2.1.1 Exemples des groupes

Les nombres. $(\mathbb{Z},+),(\mathbb{Q},+),(\mathbb{R},+),(\{+1,-1\},\cdot),(\mathbb{Q}\setminus\{0\},\cdot),(\mathbb{R}\setminus\{0\},\cdot)$ sont des groupes abéliens.

Le plus petit groupe. $G=\{e\}$ avec $e\star e=e$.

Les matrices. Soit K un corps, alors $(M_{m\times n}(K),+)$ est un groupe abélien. En effet, la matrice nulle $0_{m\times n}$ est l'élément neutre. Les axiomes des définitions 2.1 et 2.2 découlent du lemme 1.12.

$(M_{n\times n}(K),\cdot)$ n'est pas un groupe si $n\geq 2$, car il existe des matrices $n\times n$ non nulle, non inversibles. Mais, l'ensemble des matrices inversibles $n\times n$ muni du produit matriciel forme un groupe noté $GL_n(K)$ et appelé **groupe général linéaire**. C'est une conséquence du résultat suivant plus général.

Lemme 2.4 *Soit $(H,\star)$ un **monoïde**, c-à-d la stabilité ainsi que les axiomes 1 et 2 de la définition 2.1 sont satisfaits. Alors, l'énsemble*

$$H^*=\{a\in H\,|\,\text{il existe } a^{-1}\in H \text{ avec } a^{-1}\star a=a\star a^{-1}=e\}$$

muni de la loi de composition $\star$ est un groupe.

DÉMONSTRATION. Il faut vérifier que H^* est stable. Tout d'abord $H^*\neq\emptyset$, car $e\in H^*$. Or, soient $a,b\in H^*$ avec les inverses a^{-1},b^{-1}. On a

$$(b^{-1}\star a^{-1})\star(a\star b)=b^{-1}\star(a^{-1}\star a)\star b=b^{-1}\star b=e$$

et, de façon similaire, $(a\star b)\star(b^{-1}\star a^{-1})=e$. Alors, $a\star b\in H^*$. La validité des axiomes 1 et 2 dans H^* vient de leur validité dans H. ■

$(M_{n\times n}(K),\cdot)$ est un monoïde : l'associativité découle de (1.10) et l'élément neutre est la matrice identité I_n.

Applications. Soit E un ensemble non vide, on considère

$$\mathsf{App}(E):\{f:E\to E\,|\,f \text{ est une application de } E \text{ vers } E\}.$$

On définit pour $f,g\in\mathsf{App}(E)$ une loi de composition $f\circ g$ comme suit :

$$(f\circ g)(x)=f(g(x))\quad\forall x\in E.$$

On a alors que $(\mathsf{App}(E),\circ)$ est un monoïde. L'application identité $\mathsf{id}(x)=x$ pour tout $x\in E$ est l'élément neutre :

$$(\mathsf{id}\circ f)(x)=\mathsf{id}(f(x))=f(x)=f(\mathsf{id}(x))=(f\circ\mathsf{id})(x),$$

donc $\mathsf{id}\circ f=f\circ\mathsf{id}=f$.

D'après le lemme 2.4 $(\mathsf{App}(E)^*,\circ)$ est un groupe. Ce groupe est donné par

$$\mathsf{App}(E)^*=\{f:E\to E\,|\,f \text{ bijective}\}.$$

En effet : si $f\in\mathsf{App}(E)$ est inversible alors il existe $g\in\mathsf{App}(E)$ telle que $g\circ f=f\circ g=\mathsf{id}$, ceci implique que f est bijective. Réciproquement : si $f\in\mathsf{App}(E)$ est bijective, alors la réciproque de l'application f est

$$f^{-1}:E\to E,\quad y\mapsto f^{-1}(y)=x \text{ tel que } f(x)=y.$$

On a $f^{-1}\circ f = f\circ f^{-1} = \mathrm{id}$, donc f est inversible.

Si E est un ensemble fini, le groupe $\mathrm{App}(E)^*$ est appelé **groupe symétrique**, dénoté $S(E)$. Pour $E=\{1,2,\ldots,n\}$ on dénote $S(E)$ par S_n et $(S_n,\circ)$ est appelé le **groupe des permutations**. Les éléments de S_n s'appellent des permutations. On remarque que S_n n'est pas un groupe abélien.

La table suivante simplifie le traitement des permutations :

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}, \quad \pi \in S_n. \tag{2.1}$$

La composition $\pi\circ\sigma$ de $\pi,\sigma\in S_n$ prend la forme

$$\begin{pmatrix} 1 & \cdots & n \\ \pi(1) & \cdots & \pi(n) \end{pmatrix} \circ \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix} = \begin{pmatrix} 1 & \cdots & n \\ \pi(\sigma(1)) & \cdots & \pi(\sigma(n)) \end{pmatrix}.$$

L'inverse ou la réciproque π^{-1} est l'application $\pi(i)\mapsto i$ pour $i=1,\ldots,n$. On obtient la table correspondante en échangeant les deux lignes,

$$\begin{pmatrix} \pi(1) & \cdots & \pi(n) \\ 1 & \cdots & n \end{pmatrix},$$

et en réordonnant la première (et son image dans la deuxième) par ordre croissant.

Exemple 2.5 Soient

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix},$$
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}.$$

L'inverse de σ :

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}.$$

La composition $\pi\circ\sigma$:

$$\pi\circ\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \pi(1) & \pi(4) & \pi(2) & \pi(3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

Sous MATLAB on traite les permutations comme des vecteurs :

```
>> pi    = [ 4 2 3 1 ];
>> sigma = [ 1 4 2 3 ];
```

Par `pi(sigma)` on obtient le vecteur `pi(sigma(1))`, `pi(sigma(2))`, ..., qui constitue la composition de π et de σ.

```
>> pi(sigma)
ans =
     4   1   2   3
```

L'inverse σ^{-1} satisfait $\sigma^{-1}(\sigma(1)) = 1$, $\sigma^{-1}(\sigma(2)) = 2$, ... Cette relation permet d'obtenir l'inverse sous MATLAB :

```
>> r = [];
>> r(sigma) = 1:4,
r =
     1   3   4   2
```

Groupes finis. Voir Géométrie I.

2.1.2 Sous-groupes

Définition 2.6 *Soit $(G,\star)$ un groupe et $H\subseteq G$. Alors $(H,\star)$ est un sous-groupe de G si*

1. *H est non vide,*
2. *si $a,b\in H$ alors $a\star b\in H$ (H est stable par $\star$),*
3. *$a^{-1}\in H$ pour tout $a\in H$.*

Lemme 2.7 *Si $(H,\star)$ est un sous-groupe de $(G,*)$ alors $(H,*)$ est un groupe.*

DÉMONSTRATION. Exercice. ■

Soit $(G,\star)$ un groupe et soit e son élément neutre. Alors, $(\{e\},\star)$ et $(G,\star)$ sont des sous-groupes de $(G,\star)$. Les sous-groupes entre les deux extrêmes sont plus intéressants.

Groupe orthogonal. On considère les **matrices de rotation**

$$SO(2) := \left\{ G(\phi) = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \;\middle|\; \phi \in [0,2\pi[\right\}.$$

Alors, $SO(2)$ est un sous-groupe de $GL_2(\mathbb{R})$ (exercices).

Matrices 2×2 particulières. On considère

$$M_2 := \left\{ A(a,b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \;\middle|\; a,b \in \mathbb{R} \right\}.$$

Alors, $(M_2,+)$ est un sous-groupe de $(M_{2,2}(\mathbb{R}),+)$ et $(M_2 \setminus \{A(0,0)\},\cdot)$ un sous-groupe de $GL_2(\mathbb{R})$ (exercices).

Matrices triangulaires. On considère l'ensemble des matrices triangulaires supérieures à éléments diagonaux non nuls :

$$\widetilde{\mathrm{triu}}(n) = \left\{ U = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ & \ddots & \vdots \\ 0 & & u_{nn} \end{pmatrix} \;\middle|\; u_{11} \neq 0, \ldots, u_{nn} \neq 0 \right\}.$$

Lemme 2.8 *Soient $R,S \in \widetilde{\mathrm{triu}}(n)$. Alors, le produit $T = RS$ est encore dans $\widetilde{\mathrm{triu}}(n)$.*

DÉMONSTRATION. Étant donné que les matrices R,S sont triangulaires supérieures, on a $r_{ik} = 0$ si $i > k$ et $s_{kj} = 0$ si $k > j$. Si $i > j$,

$$t_{ij} = \sum_{k=1}^{n} r_{ik}s_{kj} = t_{ij} = \sum_{k=1}^{i-1} \underbrace{r_{ik}}_{=0} s_{kj} + \sum_{k=i}^{n} r_{ik} \underbrace{s_{kj}}_{=0} = 0.$$

Alors, T est triangulaire supérieure. En outre,

$$t_{ii} = \sum_{k=1}^{i-1} \underbrace{r_{ik}}_{=0} s_{ki} + r_{ii}s_{ii} + \sum_{k=i+1}^{n} r_{ik} \underbrace{s_{ki}}_{=0} = r_{ii}s_{ii} \neq 0$$

pour $i = 1,\ldots,n$. ■

Lemme 2.9 *Soit $U \in \widetilde{\mathrm{triu}}(n)$. Alors, U est inversible et $U^{-1} \in \widetilde{\mathrm{triu}}(n)$.*

DÉMONSTRATION. Par récurrence sur n. Si $n = 1$ l'assertion est triviale. On suppose que l'assertion est vraie pour $n-1$. En partitionnant

$$U = \begin{pmatrix} U_{11} & u_n \\ 0 & u_{nn} \end{pmatrix},$$

on a $U_{11} \in \widetilde{\mathrm{triu}}(n-1)$ et $u_{nn} \neq 0$. Par hypothèse de récurrence, U_{11} est inversible. En définissant

$$U^{-1} = \begin{pmatrix} U_{11}^{-1} & -U_{11}^{-1}u_n u_{nn}^{-1} \\ 0 & u_{nn}^{-1} \end{pmatrix},$$

on vérifie sans peine que $U^{-1}U = UU^{-1} = I_n$. Alors, U est inversible et $U^{-1} \in \widetilde{\text{triu}}(n)$. ■

Les lemmes 2.8 et 2.9 montrant que $\widetilde{\text{triu}}(n)$ est un sous-groupe de $GL_n(K)$. Idem pour des matrices triangulaires inférieures à éléments diagonaux non nuls (exercices).

2.1.3 Morphismes, isomorphismes de groupes

Définition 2.10 *Soient $(G,\star)$ et $(H,\circ)$ deux groupes. Un* ***morphisme de groupes*** *[group homomorphism] est une application $f : G \to H$ telle que*

$$f(a \star b) = f(a) \circ f(b) \quad \forall a,b \in G.$$

Si de plus f est bijective, on dit que f est un ***isomorphisme de groupes*** *[group isomorphism].*

Lemme 2.11 *Soit $f : G \to H$ un morphisme du groupe $(G,\star)$ dans le groupe $(H,\circ)$. Alors,*

(i) $f(e_G) = e_H$, où e_G et e_H sont les éléments neutres de G et H respectivement

(ii) $f(a^{-1}) = \big(f(a)\big)^{-1}$ pour tout $a \in G$.

DÉMONSTRATION. (i) En appliquant $f(e_G)^{-1}$ aux deux côtés de l'équation $f(e_G) = f(e_G \star e_G) = f(e_G) \circ f(e_G)$ on obtient

$$e_H = f(e_G)^{-1} \circ f(e_G) = f(e_G)^{-1} \circ f(e_G) \circ f(e_G) = e_H \circ f(e_G) = f(e_G).$$

(ii) $f(a^{-1}) \circ f(a) = f(a^{-1} \star a) = f(e_G) = f(a \star a^{-1}) = f(a) \circ f(a^{-1})$. ■

On dit que les groupes G et H sont isomorphes s'il existe un isomorphisme de G dans H.

2.2 Anneaux

Contrairement aux groupes, les anneaux comportent deux lois de composition. Ceci permet de décrire des interactions entre les deux lois de composition. La distributivité du produit matriciel par rapport à l'addition matricielle (voir le lemme 1.19) est un bon exemple.

Définition 2.12 *Un* ***anneau*** *[ring] $(A,+,\cdot)$ est un ensemble muni de deux lois de composition $+$ et $\cdot$:*

$$\begin{array}{rccc+\quad\cdot\ :} A\times A & \to & A & A\times A & \to & A \\ (a,b) & \mapsto & a+b & (a,b) & \mapsto & a\cdot b \end{array} \tag{2.2}$$

satisfaisant les axiomes suivants

(1) $a+b=b+a, \quad \forall a,b \in A.$ *(commutativité+)*

(2) $a+(b+c)=(a+b)+c, \quad \forall a,b,c \in A.$ *(associativité+)*

(3) il existe $0 \in A$ tel que $0+a=a$ pour tout $a \in A$. *(élément neutre+)*

(4) pour tout $a \in A$ il existe $-a \in K$ tel que $a+(-a)=0$. *(inverse+)*

(5) $a\cdot(b\cdot c)=(a\cdot b)\cdot c, \quad \forall a,b,c \in A.$ *(associativité·)*

(6) il existe $1 \in A$ tel que $1\cdot a = a\cdot 1 = a$ pour tout $a \in A$. *(élément neutre·)*

(7) $(a+b)\cdot c=(a\cdot c)+(b\cdot c), \quad \forall a,b,c \in A.$ *(distributivité I)*

(8) $a\cdot(b+c)=(a\cdot b)+(a\cdot c), \quad \forall a,b,c \in A.$ *(distributivité II)*

En d'autres termes, $(A,+)$ est un groupe commutatif par (1)–(4) et $(A,\cdot)$ est un monoïde par (5)–(6). Cachée dans (2.2), la stabilité des lois de composition est une propriété essentielle d'un anneau.

Remarque 2.13 *On dit parfois de l'anneau ci-dessous que c'est un* ***anneau unitaire****. On peut aussi définir des anneaux sans l'identité* 1*, l'élément neutre de* $\cdot$.

Définition 2.14 *Un anneau* $(A,+,\cdot)$ *dans lequel la loi* $\cdot$ *est commutative est appelé* ***anneau commutatif*** *[commutative ring].*

Exemples:

- $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, où $+$ et $\cdot$ sont les opérations usuelles, sont des anneaux commutatifs.
- Soit $n\mathbb{Z}=\{nz : z\in\mathbb{Z}\}$ pour $n\in\mathbb{N}$, $n\geq 1$. Alors, $(n\mathbb{Z},+,\cdot)$ n'est pas un anneau pour $n>2$, car il n'existe pas de 1.
- Soit K un anneau commutatif. Alors $(M_{n\times n}(K),+,\cdot)$ avec l'addition matricielle $+$ et le produit matriciel $\cdot$ est un anneau (non-commutatif). (Exercices)
  D'autre part $(GL_n(K),+,\cdot)$ n'est pas un anneau. Par exemple, il n'existe pas de 0 dans $GL_n(K)$.
- On définit sur $\mathbb{R}\cup\{\infty\}$ les opérations

$$a\oplus b=\min\{a,b\},\quad a\odot b=a+b.$$

  $\big(\mathbb{R}\cup\{\infty\},\oplus,\odot\big)$ n'est pas un anneau. (Exercises: Quels axiomes ne sont pas satisfaits?)
- Soit E un ensemble non vide et $(A,+,\cdot)$ un anneau. On définit

$$\mathsf{App}(E,A):=\big\{f\,|\,f \text{ est une application de } E \text{ vers } A\big\}$$

  et les opérations

$$(f+g)(x):=f(x)+g(x),\qquad (f\cdot g)(x):=f(x)g(x).$$

  Alors, $(\mathsf{App}(E,A),+,\cdot)$ est un anneau (non commutatif).
- Soit $(A,+,\cdot)$ un anneau. Un **polynôme** à coefficients dans A est une expression formelle

$$p=a_0\cdot t^0+a_1\cdot t^1+a_2\cdot t^2+\cdots+a_n\cdot t^n,\quad a_0,a_1,\ldots,a_n\in A. \tag{2.3}$$

  On écrit souvent $a_0\cdot t^0=a_0$ et $a_1\cdot t^1=a_1\cdot t$. Le **degré d'un polynôme** $\deg(p)$ est le plus grand entier j tel que $a_j\neq 0$. Si tous les coefficients a_i sont nuls on définit le degré par $-\infty$. On note l'ensemble des polynômes à coefficients dans A par $A[t]$.
  Attention ! Il es important de noter qu'à ce stade, l'expression (2.3) est formelle, t occupe la place d'un objet indéfini. De même, la puissance t^j ainsi que le signe $+$ sont formels.
  Soient $p,q\in A[t]$,

$$p=a_0+a_1\cdot t+\cdots+a_m\cdot t^m,\quad q=b_0+b_1\cdot t+\cdots+b_n\cdot t^n.$$

  On définit les opérations

$$\begin{aligned} p+q&:=(a_0+b_0)+(a_1+b_1)\cdot t+\cdots+(a_{\max\{m,n\}}+b_{\max\{m,n\}})\cdot t^{\max\{m,n\}}\\ p\cdot q&:=c_0+c_1\cdot t+\cdots+c_{m+n}\cdot t^{m+n},\qquad c_k:=\sum_{i+j=k}a_ib_j. \end{aligned}$$

Lemme 2.15 *L'ensemble* $A[t]$ *avec les opérations* $+$ *et* $\cdot$ *comme définies ci-dessus est un anneau. Si* A *est un anneau commutatif, alors* $A[t]$ *est aussi un anneau commutatif.*

DÉMONSTRATION. Exercices. ■

Lemme 2.16 *Soit $(A,+,\cdot)$ un anneau. Alors,*

(i) $0\cdot a = a\cdot 0 = 0$ *pour tout* $a \in A$,

(ii) $(-a)b = a(-b) = -(ab)$ *pour tous* $a,b \in A$,

(iii) $(-a)(-b) = ab$ *pour tous* $a,b \in A$.

DÉMONSTRATION.

(i) Par l'axiome de distributivité I,

$$0\cdot a = 0\cdot a + 0 = 0\cdot a + 0\cdot a + (-(0\cdot a)) = (0+0)\cdot a + (-(0\cdot a)) = 0\cdot a + (-(0\cdot a)) = 0.$$

De même on montre que $a\cdot 0 = 0$.

(ii) Par les axiomes de distributivité I+II et la partie (i),

$$ab + (-a)b = (a+(-a))b = 0\cdot b = 0$$

et

$$ab + a(-b) = a(b+(-b)) = a\cdot 0 = 0.$$

(iii) $(-a)(-b) = -(a(-b) = -(-(ab)) = ab$.

■

Le lemme 2.16 permet d'écrire $-ab$ sans ambiguité.

Définition 2.17 *Soient $(A,+,\cdot)$ et $(B,\oplus,\odot)$ deux anneaux. Un* ***morphisme d'anneaux*** *[ring homomorphism] est une application $f : A \to B$ telle que*

$$f(a+b) = f(a)\oplus f(b), \qquad f(a\cdot b) = f(a)\odot f(b), \qquad \forall a,b \in A,$$

et

$$f(1_A) = 1_B.$$

Si de plus f est bijective, on dit que f est un ***isomorphisme d'anneaux*** *et que les anneaux $(A,+,\cdot)$ et $(B,\oplus,\odot)$ sont isomorphes. On note $(A,+,\cdot) \cong (B,\oplus,\odot)$.*

Définition 2.18 *Soit $(A,+,\cdot)$ un anneau et $U \subseteq A$. On dit que $(U,+,\cdot)$ est un* ***sous-anneau*** *[subring] de A si*

(i) $(U,+)$ *est un sous-groupe de* $(A,+)$,

(ii) si $a,b \in U$ *alors* $a\cdot b \in U$,

(iii) L'élément neutre multiplicatif (1) *de A appartient à U.*

Par exemple, $(M_2,+,\cdot)$ (la définition de M_2 trouvée sur la page 24) est un sous-anneau de $(M_{2\times 2}(\mathbb{R}),+,\cdot)$.

Lemme 2.19 *$(A,+,\cdot)$ un anneau et $U \subseteq A$. Alors, les assertions suivantes sont équivalentes:*

(i) $(U,+,\cdot)$ *est un sous-anneau de* $(A,+,\cdot)$

(ii) $1 \in U$ *et pour tout* $a,b \in U$, *on a* $a-b \in U$ *et* $a\cdot b \in U$.

DÉMONSTRATION. (i) $\Rightarrow$ (ii) découle de la définition d'un sous-anneau.

(ii) $\Rightarrow$ (i) Comme $1 \in U$, U est non vide et comme $1-1=0$, on a que $0 \in U$. Si $b \in U$ alors $-b = 0-b \in U$. Si $a,b \in U$ alors $a+b = a-(-b) \in U$. Donc $(U,+)$ est un sous-groupe de $(A,+)$. Si $a,b \in U$ alors $a\cdot b \in U$ d'après (ii) et donc $(U,+,\cdot)$ est bien un sous-anneau de $(A,+,\cdot)$. ■

2.3 Matrices à coefficients dans un anneau

On peut généraliser la définition d'une matrice en permettant des coefficients dans un anneau.

Théorème 2.20 *Soit $(A,+,\cdot)$ un anneau. Alors, $M_{n\times n}(A)$ est un anneau.*

DÉMONSTRATION. Les axiomes (1), (2), (5), (7), (8) d'anneau découlent des généralisations du lemme 1.12 et du lemme 1.19. Soient 0_A et 1_A les éléments neutres de A par rapport à $+$ et $\cdot$, respectivement. Alors la matrice nulle $0_{n\times n}$ et la matrice identité I_n, définies par

$$0_{n\times n} = \begin{pmatrix} 0_A & \cdots & 0_A \\ \vdots & & \vdots \\ 0_A & \cdots & 0_A \end{pmatrix}, \quad I_n = \mathrm{diag}\,(1_A, 1_A, \ldots, 1_A),$$

sont les éléments neutres de $M_{n\times n}(A)$. ∎

Exemples:

– Les coefficients de $P \in M_{n\times n}(A[t])$ sont des polynômes. Par exemple,

$$P(t) = \begin{pmatrix} t^2+2 & t+1 \\ 3t+1 & 4t^2+t+2 \end{pmatrix} \in M_{2\times 2}(A[t])$$

On peut également écrire

$$P(t) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} t^2 + \begin{pmatrix} 0 & 1 \\ 3 & 1 \end{pmatrix} t + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

et regarder $P(t)$ comme un élément de $\big(M_{2\times 2}(A)\big)[t]$.
En général, soit $(A,+,\cdot)$ un anneau et $n \geq 1$. Alors l'anneau $M_{n\times n}(A[t])$, muni de l'addition/multiplication matricielle, et l'anneau $\big(M_{n\times n}(A)\big)[t]$, muni de l'addition/multiplication polynomiale, sont isomorphes. (Démonstration: Voir les exercises.)

– Les coefficients de $B \in M_{m\times n}(\mathbb{Z})$ sont des nombres entiers. Par exemple,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in M_{2\times 2}(\mathbb{Z}).$$

On remarque que A est inversible vu comme un élément de $M_{2\times 2}(\mathbb{R})$:

$$A^{-1} = \frac{1}{2}\begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R}).$$

Au contraire, A est singulaire comme un élément de $M_{2\times 2}(\mathbb{Z})$. L'existence d'une matrice $X \in M_{2\times 2}(\mathbb{Z})$ avec $AX = XA = I_n$ contrarie l'unicité d'inverse dans $M_{2\times 2}(\mathbb{R})$.

– $M_{m\times m}(M_{n\times n}(A))$ est un exemple des matrices à coefficients dans un anneau non-commutatif. $M_{n\times n}(M_{n\times n}(A))$ et $M_{mn\times mn}(A)$ sont isomorphes (multiplication de matrices par blocs).

Après ce chapitre, nous ne traiterons pas des matrices à coefficients dans un anneau non-commutatif (sauf dans la démonstration du théorème de Cayley-Hamilton au chapitre 7). Quelques concepts avancés (rang, déterminant, ...) n'ont pas des définitions raisonables si A est non-commutatif.

2.4 Corps

Un corps (commutatif) est un anneau commutatif dans lequel $0 \neq 1$ et tout élément non nul est inversible (par rapport à $\cdot$).

Définition 2.21 *Un* ***corps*** *[field]* $(K,+,\cdot)$ *est un anneau commutatif tel que:*

(i) $K \neq \{0\}$

(ii) pour tuot $a \in K \setminus \{0\}$ *il existe* $a^{-1} \in K$ *tel que* $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

On remarque que $(K,+,\cdot)$ est un corps si et seulement si $(K,+)$ et $(K \setminus \{0\},\cdot)$ sont des groupes abéliens et $(a+b)\cdot c = a\cdot c + b\cdot c$ pour tous $a,b,c \in K$.

Une liste de tous les axiomes d'un corps $(K,+,\cdot)$:

$a+b \in K, \quad a\cdot b \in K \quad \forall a,b \in K.$	(stabilité)
$a+b = b+a, \quad \forall a,b \in K.$	(commutativité+)
$a+(b+c) = (a+b)+c, \quad \forall a,b,c \in K.$	(associativité+)
il existe $0 \in K$ tel que $0+a = a$ pour tout $a \in K$.	(élément neutre+)
pour tout $a \in K$ il existe $-a \in K$ tel que $a+(-a) = 0$.	(inverse+)
$a\cdot b = b\cdot a, \quad \forall a,b \in K.$	(commutativité$\cdot$)
$a\cdot(b\cdot c) = (a\cdot b)\cdot c, \quad \forall a,b,c \in K.$	(associativité$\cdot$)
il existe $1 \in K$ tel que $1\cdot a = a\cdot 1 = a$ pour tout $a \in K$.	(élément neutre$\cdot$)
pour tout $a \in K \setminus \{0\}$ il existe $a^{-1} \in K$ tel que $a\cdot a^{-1} = 1$.	(inverse$\cdot$)
$(a+b)\cdot c = (a\cdot c)+(b\cdot c), \quad \forall a,b,c \in K.$	(distributivité I)
$a\cdot(b+c) = (a\cdot b)+(a\cdot c), \quad \forall a,b,c \in K.$	(distributivité II)

En fait, le commutativité de $\cdot$ implique que les deux lois de distributivité I et II sont équivalentes.

Exemples :

- $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, où $+$ et $\cdot$ sont les opérations usuelles, sont des corps.
- $(\mathbb{Z},+,\cdot)$ n'est pas un corps parce qu'il n'y a pas d'inverse multiplicatif en general.
- L'ensemble M_2 de la page 24 muni de la somme et le produit matriciel est un corps (exercices).

Dans les deux sections suivantes on va voir deux autres exemples de corps.

Un morphisme (isomorphisme) de corps est simplement un morphisme (isomorphisme) d'anneaux sous-jacents.

2.5 Le corps des nombres complexes

Un **nombre complexe** est une paire ordonnée (couple) (x,y) où $x,y \in \mathbb{R}$. En définissant l'**unité imaginaire** [*imaginary unit*] i on écrit

$$x + \mathrm{i}y$$

au lieu de (x,y). L'ensemble des nombres complexes se note

$$\mathbb{C} = \{x+\mathrm{i}y : \ x,y \in \mathbb{R}\}.$$

Quelques conventions : Soit $z = x+\mathrm{i}y \in \mathbb{C}$.

- On note $x = \mathsf{Re}(z)$ et on dit que x est la **partie réelle** [*real part*] de z.
- On note $y = \mathsf{Im}(z)$ et on dit que y est la **partie imaginaire** [*imaginary part*] de z.
- Si $y = 0$ il est usuel d'identifier le nombre complexe z avec le nombre réel x. Cela justifie l'inclusion $\mathbb{R} \subset \mathbb{C}$.

– Si $x = 0$ on dit que z est **imaginaire pur** ou **totalement imaginaire** [*purely imaginary*].

On définit une loi $+$ (addition des nombres complexes) et une loi $\cdot$ (multiplication des nombres complexes) sur $\mathbb{C}$:

$$\begin{aligned} +: &\quad \mathbb{C}\times\mathbb{C}\to\mathbb{C}, \quad (x_1+y_1\mathrm{i})+(x_2+y_2\mathrm{i}) := (x_1+x_2)+(y_1+y_2)\mathrm{i}, \\ \cdot: &\quad \mathbb{C}\times\mathbb{C}\to\mathbb{C}, \quad (x_1+y_1\mathrm{i})\cdot(x_2+y_2\mathrm{i}) := (x_1x_2-y_1y_2)+(x_1y_2+y_1x_2)\mathrm{i}. \end{aligned}$$

On constate que $\mathrm{i}^2 = \mathrm{i}\cdot\mathrm{i} = -1$. En effet, cela suffit pour retrouver la loi de multiplication :

$$\begin{aligned} (x_1+\mathrm{i}y_1)\cdot(x_2+\mathrm{i}y_2) &= x_1x_2+\mathrm{i}y_1x_2+\mathrm{i}x_1y_2+\mathrm{i}^2y_1y_2 \\ &= x_1x_2+\mathrm{i}y_1x_2+\mathrm{i}x_1y_2-y_1y_2 \\ &= (x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+y_1x_2). \end{aligned}$$

Exemple 2.22

On a

$$(1+\mathrm{i})+(-2+\mathrm{i}) = -1+2\mathrm{i}$$

et

$$(1+\mathrm{i})(-2+\mathrm{i}) = -3-\mathrm{i}.$$

MATLAB

Sous MATLAB l'unité imaginaire est `i` (ou `j`). Mais il est recommandé d'utiliser `1i` au lieu de `i`, qui peut être une autre variable.

```
>> (1+1i) + (-2+1i),
ans =
  -1.0000 + 2.0000i
>> (1+1i) * (-2+1i),
ans =
  -3.0000 - 1.0000i
```

Héritant des propriétés de $(\mathbb{R},+)$, $(\mathbb{C},+)$ est un groupe abélien. L'élément neutre est $0 = 0+0\mathrm{i}$ et l'inverse additif de $z = x+\mathrm{i}y \in \mathbb{C}$ est $-z := -x-\mathrm{i}y$. On définit la soustraction des nombres complexes par

$$z_1 - z_2 := z_1 + (-z_2) = (x_1-x_2)+\mathrm{i}(y_1-y_2).$$

C'est laborieux de vérifier directement les propriétés de la multiplication. Au lieu de cela, on utilise un morphisme de corps. À cette fin, soit

$$\varphi : \mathbb{C}\to M_2, \quad \varphi : x+\mathrm{i}y \mapsto \begin{pmatrix} x & y \\ -y & x \end{pmatrix}. \tag{2.4}$$

Voir la page 24 pour la définition de M_2. Évidemment, φ est bijective et, ainsi, φ est un isomorphisme du groupe $(\mathbb{C},+)$ dans $(M_2,+)$. En outre,

$$\begin{aligned} \varphi(x_1+\mathrm{i}y_1)\varphi(x_2+\mathrm{i}y_2) &= \begin{pmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{pmatrix}\begin{pmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1x_2-y_1y_2 & x_1y_2+y_1x_2 \\ -y_1x_2-x_1y_2 & -y_1y_2+x_1x_2 \end{pmatrix} = \varphi\big((x_1+y_1\mathrm{i})(x_2+y_2\mathrm{i})\big). \end{aligned}$$

Comme $(M_2\setminus\{0_{2\times2}\},\cdot)$ est un groupe, $(\mathbb{C}\setminus\{0\},\cdot)$ est un groupe. L'élément neutre est $1 = \varphi^{-1}(I_2) = 1+0\mathrm{i}$. On obtient l'inverse multiplicatif dans $(\mathbb{C}\setminus\{0\},\cdot)$ par l'inverse matriciel :

$$\begin{aligned} z^{-1} = \varphi^{-1}(\varphi(z)^{-1}) &= \varphi^{-1}\left(\begin{pmatrix} x & y \\ -y & x \end{pmatrix}^{-1}\right) \\ &= \varphi^{-1}\left(\frac{1}{x^2+y^2}\begin{pmatrix} x & -y \\ y & x \end{pmatrix}\right) = \frac{x}{x^2+y^2} - \mathrm{i}\frac{y}{x^2+y^2}. \end{aligned}$$

En utilisant l'isomorphisme φ, $(\mathbb{C},+,\cdot)$ hérite des lois de distributivité de $(M_2,+,\cdot)$. En outre, la multiplication matricielle est commutative dans M_2 et, ainsi, la multiplication complexe est commutative. En résumé, on a montré le résultat suivant.

Théorème 2.23 *L'ensemble $\mathbb{C}$ muni des opérations $+$ et $\cdot$ définies ci-dessus est un corps.*

Définition 2.24 *Le conjugué d'un nombre complexe $z = x+\mathrm{i}y$ est le nombre complexe $\overline{z}$ défini par $\overline{z} := x-\mathrm{i}y$.*

Le conjugué d'un nombre complexe correspond à la transposée d'une matrice :

$$\varphi(\overline{z}) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}^{\mathsf{T}} = \varphi(z)^{\mathsf{T}}. \tag{2.5}$$

Lemme 2.25 *Soient $z_1,z_2,z \in \mathbb{C}$. Alors,*

(i) $\overline{z_1+z_2} = \overline{z_1}+\overline{z_2}$
(ii) $\overline{z_1\cdot z_2} = \overline{z_1}\cdot\overline{z_2}$
(iii) $\overline{\overline{z}} = z$
(iv) $\mathsf{Re}(z) = \frac{1}{2}(z+\overline{z})$
(v) $\mathsf{Im}(z) = \frac{1}{2\mathrm{i}}(z-\overline{z})$.

DÉMONSTRATION. Ces propriétés sont des exercices faciles. En effet, (i)–(iii) découlent des propriétés de la transposée (voir le lemme 1.26). Par exemple

$$\varphi(\overline{z_1\cdot z_2}) = \varphi(z_1z_2)^{\mathsf{T}} = \varphi(z_2)^{\mathsf{T}}\varphi(z_1)^{\mathsf{T}} = \varphi(z_1)^{\mathsf{T}}\varphi(z_2)^{\mathsf{T}} = \varphi(\overline{z_1})\varphi(\overline{z_2}) = \varphi(\overline{z_1}\cdot\overline{z_2}),$$

ce qui montre (ii). ∎

Les parties (i)–(iii) du lemme 2.25 impliquent que la conjugaison est un isomorphisme du corps $(\mathbb{C},+,\cdot)$ dans lui-même.[2]

Définition 2.26 *Le **module** [absolute value, modulus, magnitude] d'un nombre complexe $z = x+\mathrm{i}y$ est le nombre réel positif $|z|$ défini par $|z| := \sqrt{x^2+y^2}$.*

Lemme 2.27 *Soient $z_1,z_2,z \in \mathbb{C}$. Alors,*

(i) $z\overline{z} = |z|^2$
(ii) $z^{-1} = \frac{\overline{z}}{|z|^2}$ $\quad (z\neq 0)$
(iii) $\overline{z^{-1}} = \overline{z}^{-1}$ $\quad (z\neq 0)$
(iv) $|z_1\cdot z_2| = |z_1|\cdot|z_2|$
(v) $|z_1+z_2| \leq |z_1|+|z_2|$ *avec égalité si et seulement si il existe $\alpha \geq 0$ tel que $z_1 = \alpha z_2$ ou $z_2 = \alpha z_1$.*

DÉMONSTRATION. (i). $z\overline{z} = (x+\mathrm{i}y)(x-\mathrm{i}y) = x^2+y^2+\mathrm{i}(xy-yx) = x^2+y^2 = |z|^2$.
(ii) découle de (i) et (iii) découle de (ii).
(iv). En utilisant le lemme 2.25 et la commutativité de la multiplication complexe on obtient

$$|z_1\cdot z_2|^2 = z_1\cdot z_2\cdot\overline{z_1\cdot z_2} = z_1\cdot z_2\cdot\overline{z_1}\cdot\overline{z_2} = z_1\cdot\overline{z_1}\cdot z_2\cdot\overline{z_2} = |z_1|^2\cdot|z_2|^2.$$

(v). L'inégalité découle de

$$\begin{aligned} |z_1+z_2|^2 &= (z_1+z_2)\overline{(z_1+z_2)} = (z_1+z_2)(\overline{z_1}+\overline{z_2}) \\ &= |z_1|^2+z_2\overline{z_1}+z_1\overline{z_2}+|z_2|^2 = |z_1|^2+2\,\mathsf{Re}(z_1\overline{z_2})+|z_2|^2 \\ &\leq |z_1|^2+2\,|z_1||\overline{z_2}|+|z_2|^2 = |z_1|^2+2\,|z_1||z_2|+|z_2|^2 = (|z_1|+|z_2|)^2. \end{aligned}$$

2. Un isomorphisme d'une structure algébrique dans elle-même est dit **automorphisme**.

On a utilisé le fait que le module est toujours supérieur à la partie reélle. L'inégalité ci-dessus devient une égalité si $\mathrm{Re}(z_1\overline{z_2}) = |z_1\overline{z_2}|$ c-à-d si $\beta = z_1\overline{z_2}$ est reél positif. Si $z_2 = 0$ on a bien $z_2 = \alpha z_1$ avec $\alpha = 0$. Si $z_2 \neq 0$, alors

$$z_1\overline{z_2}z_2 = \beta z_2 \quad \Rightarrow \quad z_1 = \frac{\beta z_2}{|z_2|^2} = \alpha z_2 \text{ avec } \alpha = \frac{\beta}{|z_2|^2} \geq 0.$$

La réciproque est évidente. ■

La division d'un nombre complexe z_1 par un nombre complexe $z_2 \neq 0$ est définie par $z_1/z_2 := z_1 z_2^{-1}$. D'aprés le lemme 2.25 (ii), on a

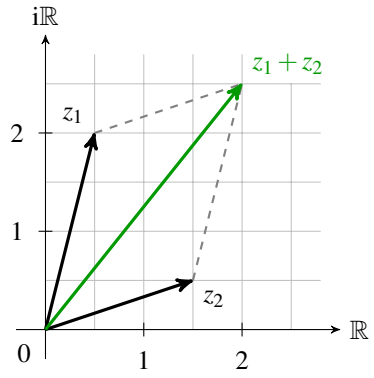
$$\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{|z_2|^2}.$$

Par exemple,

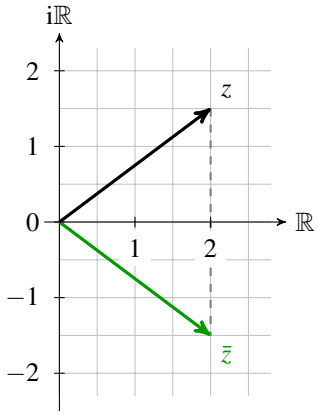
$$\frac{2+3\mathrm{i}}{1+\mathrm{i}} = \frac{(2+3\mathrm{i})(1-\mathrm{i})}{1+1} = \frac{5+\mathrm{i}}{2} = \frac{5}{2} + \frac{1}{2}\mathrm{i}.$$

2.5.1 Plan complexe et forme polaire

Par définition, les nombres complexes $\mathbb{C}$ sont des couples de nombres réels. Pour cette raison, tout nombre complexe correspond uniquement à un vecteur dans la plan (qui s'appele **plan complexe** [*complex plane*]). La somme des nombres complexes correspond à la somme des vecteurs et la conjugaison correspond à la réflexion par rapport à l'axe réel, voir figure 2.1.



l'addition des deux nombres complexes



le conjugué

FIG. 2.1 – L'addition et le conjugué dans la plan complexe.

Soit $z = x + \mathrm{i}y \in \mathbb{C} \setminus \{0\}$. En notant $r = \sqrt{x^2+y^2} > 0$ la longueur et $\theta = \arctan\frac{y}{x} \in]-\pi,\pi]$ l'angle du vecteur (x,y) dans la plan complexe, on peut écrire

$$(x,y) = (r\cos\theta, r\sin\theta).$$

Ainsi on a

$$z = x + \mathrm{i}y = r\cos\theta + \mathrm{i}r\sin\theta = r(\cos\theta + \mathrm{i}\sin\theta),$$

où θ est défini à $2k\pi$ près avec $k \in \mathbb{Z}$. On l'appelle la **forme polaire** [*polar form*] de z. $\theta = \arg$ est l'argument de z. La **fonction exponentielle complexe** [*complex exponential*] permet de faire une représentation plus compacte.

Définition 2.28 *Pour $z = x + \mathrm{i}y \in \mathbb{C}$ on définie*

$$e^z = \exp(z) := e^x(\cos y + \mathrm{i}\sin y) = e^{\mathsf{Re}z}(\cos(\mathsf{Im}z) + \mathrm{i}\sin(\mathsf{Im}z))$$

où e^x est la fonction exponentielle réelle usuelle.

Propriétés de l'exponentielle:

1. $|e^z| = e^x = e^{\mathsf{Re}z}$
2. $\arg(e^z) = \mathsf{Im}z$ (à $2k\pi$ près avec $k \in \mathbb{Z}$)
3. si $\mathsf{Im}z = 0$ on a $e^z = e^{\mathsf{Re}z}$
4. $e^{z+2k\pi\mathrm{i}} = e^x\big(\cos(y+2k\pi) + \mathrm{i}\sin(y+2k\pi)\big)$ pour tout $k \in \mathbb{Z}$
5. $e^{w+z} = e^w \cdot e^z$ pour tous $w,z \in \mathbb{C}$

La **formule d'Euler** s'écrit, pour $\theta \in \mathbb{R}$,

$$e^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta.$$

En particulier pour $\theta = \pi$ on obtient **l'identité d'Euler** (« l'étalon-or de la beauté mathématique »)

$$e^{\mathrm{i}\pi} + 1 = 0.$$

Par les identités trigonométriques, la forme polaire permet de multiplier facilement deux nombres complexes :

$$\begin{aligned} z_1 z_2 &= \rho_1(\cos\varphi_1 + \mathrm{i}\sin\varphi_1)\cdot\rho_2(\cos\varphi_2 + \mathrm{i}\sin\varphi_2) \\ &= \rho_1\rho_2\big(\cos(\varphi_1+\varphi_2) + \mathrm{i}\sin(\varphi_1+\varphi_2)\big). \end{aligned} \tag{2.6}$$

Alors, le produit $z_1 z_2$ représente géométriquement une multiplication de la longueur de z_1 par ρ_2 et une rotation anti-horaire de z_1 d'angle φ_2.

Lemme 2.29 (Formule de Moivre) *Pour tous $r > 0$, $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$ on a*

$$\big(r(\cos\theta + \mathrm{i}\sin\theta)\big)^n = r^n\big(\cos(n\theta) + \mathrm{i}\sin(n\theta)\big).$$

DÉMONSTRATION. Par récurrence utilisant (2.6). Voir exercices. ■

2.5.2 Matrices à coefficients complexes

Dans cette section, on considère $M_{m\times n}(\mathbb{C})$, l'ensemble des matrices $m \times n$ complexes. Le conjuguée d'une matrice $A \in M_{m\times n}(\mathbb{C})$ est la matrice (notée $\overline{A}$) formée des éléments de A conjugués :

$$\overline{A} \in M_{m\times n}(\mathbb{C}) \quad \text{avec} \quad (\overline{A})_{ij} := \overline{a_{ij}}, \quad i = 1,\dots,m,\ j = 1,\dots,n.$$

Exemple 2.30 Soient

$$A = \begin{pmatrix} 1+2\mathrm{i} & 1-\mathrm{i} \\ 3 & -\mathrm{i} \\ 2-\mathrm{i} & 2+\mathrm{i} \end{pmatrix},$$

$$B = \begin{pmatrix} 1+\mathrm{i} & 1-\mathrm{i} \\ -1-\mathrm{i} & 1+\mathrm{i} \end{pmatrix}.$$

Alors,

$$\overline{A} = \begin{pmatrix} 1-2\mathrm{i} & 1+\mathrm{i} \\ 3 & \mathrm{i} \\ 2+\mathrm{i} & 2-\mathrm{i} \end{pmatrix},$$

$$AB = \begin{pmatrix} -3+3\mathrm{i} & 5+\mathrm{i} \\ 2+4\mathrm{i} & 4-4\mathrm{i} \\ 2-2\mathrm{i} & 2 \end{pmatrix}.$$

MATLAB

```
>> A = [1+2i 1-1i;
        3     -1i;
        2-1i 2+1i];
>> B = [1+1i 1-1i;
       -1-1i 1+1i];
>> conj(A),
ans =
   1 - 2i   1 + 1i
   3 - 0i  -0 + 1i
   2 + 1i   2 - 1i
>> A*B,
ans =
  -3 + 3i   5 + 1i
   2 + 4i   4 - 4i
   2 - 2i   2 + 0i
```

Définition 2.31 *La* ***matrice adjointe*** A^* *(aussi appelée* ***matrice transposée conjuguée*** *[conjugate transpose, Hermitian transpose]) d'une matrice* $A \in M_{m\times n}(\mathbb{C})$ *est la matrice transposée de la matrice conjuguée de* A *:*

$$A^* \in M_{n\times m}(\mathbb{C}) \quad avec \quad (\overline{A})_{ij} := \overline{a_{ji}}, \quad i = 1,\ldots,n,\ j = 1,\ldots,m.$$

Dans le cas particulier où les coefficients de A sont réels, on a $A^* = A^T$.

Exemple 2.32

Soit A comme dans l'exemple 2.30. Alors,

$$A^* = \begin{pmatrix} 1-2\mathrm{i} & 3 & 2+\mathrm{i} \\ 1+\mathrm{i} & \mathrm{i} & 2-\mathrm{i} \end{pmatrix},$$

$$A^{\mathsf{T}} = \begin{pmatrix} 1+2\mathrm{i} & 3 & 2-\mathrm{i} \\ 1-\mathrm{i} & -\mathrm{i} & 2+\mathrm{i} \end{pmatrix}.$$

MATLAB

Étant donné une matrice à coefficients complexes, `A'` retourne la matrice adjointe et `A.'` retourne la matrice transposée.

```
>> A'
ans =
   1 - 2i   3 - 0i   2 + 1i
   1 + 1i  -0 + 1i   2 - 1i
>> A.'
ans =
   1 + 2i   3 + 0i   2 - 1i
   1 - 1i  -0 - 1i   2 + 1i
```

D'après la définition 2.31, on a

$$A^* = (\overline{A})^{\mathsf{T}} = \overline{A^{\mathsf{T}}}. \tag{2.7}$$

Lemme 2.33 *(i)*

$$\boxed{(A^*)^* = A} \qquad \forall A \in M_{m\times n}(\mathbb{C}).$$

(ii)

$$\boxed{(\alpha A)^* = \overline{\alpha} A^*} \qquad \forall A \in M_{m\times n}(\mathbb{C}), \alpha \in \mathbb{C}.$$

(iii)

$$\boxed{(A+B)^* = A^* + B^*} \qquad \forall A,B \in M_{m\times n}(\mathbb{C})$$

(iv)

$$\boxed{(AB)^* = B^* A^*} \qquad \forall A \in M_{m\times n}(\mathbb{C}), B \in M_{n\times p}(\mathbb{C}).$$

DÉMONSTRATION. Exercices. Indication: utiliser le lemme 1.26 et (2.7). ■

Définition 2.34 *On dit qu'une matrice* $A \in M_{n\times n}(\mathbb{C})$ *est* ***hermitienne*** *[Hermitian matrix] si*

$$A^* = A \quad c\text{-}\grave{a}\text{-}d \quad a_{ij} = \overline{a_{ji}} \quad \forall i,j = 1,\ldots,n.$$

Par exemple,

$$A = \begin{pmatrix} 1 & 2+3\mathrm{i} & 4+5\mathrm{i} \\ 2-3\mathrm{i} & 6 & 7+8\mathrm{i} \\ 4-5\mathrm{i} & 7-8\mathrm{i} & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 1+\mathrm{i} & 2+3\mathrm{i} & 4+5\mathrm{i} \\ 2+3\mathrm{i} & 6+2\mathrm{i} & 7+8\mathrm{i} \\ 4+5\mathrm{i} & 7+8\mathrm{i} & 9-3\mathrm{i} \end{pmatrix}.$$

La matrice A est hermitienne. La matrice B au contraire est une matrice (complexe) symétrique mais elle n'est pas hermitienne. Les éléments diagonaux d'une matrice hermitienne sont réels.

2.6 Corps finis

En Géométrie vous avez déjà vu $\mathbb{Z}/p\mathbb{Z}$ pour un nombre entier p. Dans cette section nous rappelons cette structure algébrique.

Soit $p \geq 2$ un nombre entier et $\mathbb{N}_{<p} = \{0,1,2,\ldots,p-1\}$. On définit deux lois de compositions $\oplus$ et $\odot$ sur $\mathbb{N}_{<p}$ par

$$a \oplus b = \text{reste dans la division euclidienne de } a+b \text{ par } p,$$
$$a \odot b = \text{reste dans la division euclidienne de } a \cdot b \text{ par } p.$$

Héritant de la structure de $\mathbb{Z}$, $(\mathbb{N}_{<p},\oplus)$ et $(\mathbb{N}_{<p},\odot)$ sont des monoïdes.

Exemple 2.35 Les tables de Cayley pour $p = 2,3,4$:

$\mathbb{N}_{<2}$:

$\oplus$	0	1
0	0	1
1	1	0

$\odot$	0	1
0	0	0
1	0	1

$\mathbb{N}_{<3}$:

$\oplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\odot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

$\mathbb{N}_{<4}$:

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\odot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

La commutativité des $\oplus$ et $\odot$ signifie que les tables ci-dessus sont symétriques. L'inversibilité d'un élément signifie que la ligne (ou la colonne) correspondante contient 1. En effet, par la « règle Sudoku », un monoïde fini est un groupe si et seulement si toute ligne et toute colonne contient une fois et une fois seulement chaque élément du monoïde. Alors, $(\mathbb{N}_{<2},\oplus)$, $(\mathbb{N}_{<2}\setminus\{0\},\odot)$, $(\mathbb{N}_{<3},\oplus)$, $(\mathbb{N}_{<3}\setminus\{0\},\odot)$, $(\mathbb{N}_{<4},\oplus)$ sont des groupes (abéliens). Au contraire, $(\mathbb{N}_{<4}\setminus\{0\},\odot)$ n'est pas un groupe, par exemple l'élément 2 n'est pas inversible. ♦

Lemme 2.36 *$a \in \mathbb{N}_{<p}$ tel que $a \neq 0$ est inversible (par rapport à $\cdot$) si et seulement si a et p sont premiers entre eux.*

DÉMONSTRATION. L'inversibilité de a dit qu'il existe $1 \leq b \leq p-1$, $q \in \mathbb{Z}$ tels que

$$ab + pq = 1. \tag{2.8}$$

Le plus grand diviseur commun (PGCD) de a,p divise $ab+pq$. Par (2.8), le PGCD doit être 1, c-à-d a et p sont premiers entre eux.

Réciproquement, on suppose que le PGCD de a,p soit 1. Alors, la division euclidienne permet de trouver $1 \leq b \leq p-1$, $q \in \mathbb{Z}$ satisfaisant (2.8). ■

Théorème 2.37 *Soit $p \geq 2$ un nombre entier. Alors :*

(i) $(\mathbb{N}_{<p},\oplus)$ *est un groupe abélien.*

(ii) $(\mathbb{N}_{<p},\odot)$ *est un monoïde commutatif.*

(iii) $(\mathbb{N}_{<p}\setminus\{0\},\odot)$ *est un groupe abélien si et seulement si p est un nombre premier.*

(iv) $(\mathbb{N}_{<p},\oplus,\odot)$ *est un anneau commutatif.*

(v) $(\mathbb{N}_{<p},\oplus,\odot)$ *est un corps si et seulement si p est un nombre premier.*

DÉMONSTRATION. Les parties (i) et (ii) sont des exercices.

Pour la partie (iii) il reste à montrer que tout élément de $\mathbb{N}_{<p} \setminus \{0\}$ est inversible si et seulement si p est un nombre premier. On suppose que p est un nombre premier. Alors, tous les nombres entre 1 et $p-1$ sont premiers avec p. Par le lemme 2.36, ils sont inversibles et, ainsi, le monoïde $\mathbb{N}_{<p} \setminus \{0\}$ devient un groupe. Réciproquement, on suppose que p est divisible par un nombre q avec $2 \leq q \leq p-1$. Par le lemme 2.36, q n'est pas inversible et, ainsi, $\mathbb{N}_{<p} \setminus \{0\}$ n'est pas un groupe.

La partie (iv) est un excercice. La partie (v) découle des parties (iii) et (iv). ■

Si p est premier on écrit souvent $\mathbb{F}_p$ en lieu de $(\mathbb{N}_{<p}, \oplus, \odot)$.

Le produit matriciel $C = AB$ de deux matrices $A \in M_{m\times n}(\mathbb{F}_p)$, $B \in M_{n\times p}(\mathbb{F}_p)$ est défini comme d'habitude. Pour calculer à la main ou sur un ordinateur, il peut être plus pratique de faire d'abord les additions et les multiplications usuelles; puis de faire la division euclidienne par p.

Exemple 2.38

Soient $A \in \mathbb{F}_5^{3\times 2}$, $B \in \mathbb{F}_5^{2\times 2}$ définies par

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 1 \\ 2 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}.$$

Alors,

$$C = AB = \begin{pmatrix} 0 & 1 \\ 0 & 2 \\ 1 & 4 \end{pmatrix}.$$

MATLAB

MATLAB n'a pas de fonctions pour matrices à coefficients dans un corps finis. Malgré cela, le produit matriciel est facile à realiser :

```
>> A = [ 2 3; 4 1; 2 4];
>> B = [ 1 1; 1 3 ];
>> mod( A*B, 5 ),
ans =
   0   1
   0   2
   1   4
```

2.7 Polynômes à coefficients dans un corps

Si K est un corps en particulier $(K, +, \cdot)$ est un anneau commutatif et on sait d'après le lemme 2.15 que l'ensemble des polynômes à coefficients dans $K[t]$ muni de l'addition et de la multiplication des polynômes est un anneau commutatif.

Définition 2.39 *Soit K un corps, sous-anneau d'un anneau A. Soit $p \in K[t]$ avec $p(t) = a_0 + a_1 t + \cdots + a_n t^n$. **L'évaluation de p** en $s \in A$ notée $p(s)$ est*

$$a_0 + a_1 \cdot s + \cdots + a_n \cdot s^n, \quad \text{où} \quad s^j := \underbrace{s \cdot s \cdots s}_{j \text{ fois}}.$$

Exemple 2.40

Soit $K = \mathbb{R}$, $A = \mathbb{C}$, $p(t) = t^2 + 1 \in \mathbb{R}[t]$.

$$\begin{aligned} p(\mathrm{i}) &= \mathrm{i}^2 + 1 = -1 + 1 = 0 \\ p(\mathrm{i}+1) &= (\mathrm{i}+1)^2 + 1 \\ &= \mathrm{i}^2 + 2\mathrm{i} + 1 + 1 = 2\mathrm{i} + 1 \end{aligned}$$

MATLAB

Sous MATLAB un polynôme $a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + a_n t^n$ est representé par le vecteur $(a_n, a_{n-1}, \ldots, a_1, a_0)$. La commande polyval permet d'évaluer un polynôme :

```
>> polyval([1 0 1],1i)
ans =
     0
>> polyval([1 0 1],1i+1)
ans =
   1.0000 + 2.0000i
```

Théorème 2.41 (division euclidienne des polynômes) *Soient $p, q \in K[t]$ avec $q \neq 0$. Alors, il existe un unique couple de polynômes $g, r \in K[t]$ tels que*

$$p = gq + r \qquad \text{avec} \qquad \deg r < \deg q.$$

DÉMONSTRATION. **Existence.** Par récurrence sur $n = \deg p$. Si $n < \deg q$ alors $g = 0$ et $r = p$ conviennent. Supposons le résultat montré pour tout polynôme de degré strictement inférieur à n et $n \geq m := \deg q$. On pose

$$p(t) = a_0 + a_1 t + \cdots + a_n t^n, \qquad q(t) = b_0 + b_1 t + \cdots b_m t^m.$$

Posons $f(t) = p(t) - a_n/b_m \cdot t^{n-m} q(t)$, alors $\deg f < n$. Par hypothèse de récurrence, on a

$$f = g_1 q + r \qquad \text{avec} \qquad \deg r_1 < \deg q.$$

Alors

$$p(t) = g_1(t)q(t) + r + \frac{a_n}{b_m} \cdot t^{n-m} q(t) = \underbrace{\big(g_1(t) + \frac{a_n}{b_m} \cdot t^{n-m}\big)}_{=:g(t)} q(t) + r,$$

où g,r possèdent les propriétés demandées.

Unicité. Si $p = g_1 q + r_1 = g_2 q + r_2$ alors

$$(g_1 - g_2)q = r_2 - r_1 \qquad \text{avec} \qquad \deg(r_2 - r_1) < \deg q.$$

Si $g_1 - g_2 \neq 0$

$$\deg((g_1 - g_2)q) = \deg(g_1 - g_2) + \deg q \geq \deg q,$$

ce qui est absurde. Donc $g_1 = g_2$ et par suite $r_1 = r_2$. ■

Définition 2.42 *Un élément $c \in K$ s'appelle une* ***racine*** *[root, zero] de $p \in K[t]$ si $p(c) = 0$.*

Corollaire 2.43 *Soit $p \in K[t]$ et $c \in K$. Alors c est une racine de p si et seulement si $t - c$ divise p (sans reste), c-à-d $p(t) = g(t)(t - c)$ pour un certain $g \in K[t]$.*

DÉMONSTRATION. L'assertion découle du théorème 2.41 en posant $q(t) = t - c$. ■

Vocabulaire: Soient p,$q \in K[t]$ avec $q \neq 0$. On dit que

- q **divise** p, q est un **diviseur** de p
- p est **divisible** par q
- p est un **multiple** de q

si le reste de la division de p par q est nul.

Définition 2.44 *Un polynôme $p \in K[t]$ est dit* ***irréductible*** *(sur K) si*

(i) $\deg p \geq 1$

(ii) les seuls diviseurs de p sont les polynômes de degré 0 (les polynômes constants) et $c \cdot p(t)$ avec $c \in K \setminus \{0\}$.

Exemples :

1. tout polynôme de degré 1 est irréductible
2. $t^2 + 1 \in \mathbb{R}[t]$ est irréductible (sur $\mathbb{R}$)
3. $t^2 + 1 \in \mathbb{C}[t]$ est réductible : $t^2 + 1 = (t + \mathrm{i})(t - \mathrm{i})$
4. $at^2 + bt + c \in \mathbb{R}[t]$ est irréductible si et seulement si $b^2 - 4ac < 0$.

Théorème 2.45 *Tout polynôme $p \in K[t]$ de degré ≥ 1 peut s'écrire de manière unique (à permutation des facteurs près)*

$$p = \alpha\, g_1 g_2 \cdots g_r \qquad \text{où} \qquad \alpha \in K \tag{2.9}$$

et g_i, $i = 1, \ldots, r$, sont des polynômes irréductibles unitaires (c-à-d que le coefficient dominant vaut 1).

DÉMONSTRATION. Sans perte de généralité, on peut supposer que le coefficient dominant de p vaille 1.

Existence. Si p est irréductible on obtient directement (2.9). Sinon on peut écrire $p = p_1p_2$, où p_1,p_2 sont des polynômes de degré strictement inférieur à $\deg p$. Ainsi, on obtient (2.9) par la récurrence.

Unicité. Soit $p = g_1g_2\cdots g_r = h_1h_2\cdots h_s$, où h_i, $i = 1,\ldots,m$, sont des polynômes irréductibles unitaires. Comme h_1 est irréductible, h_1 divise un de g_i. Mais, comme g_i est aussi irréductible, $h_1 = g_i$. Soit σ une permutation avec $\sigma(1) = i$. Alors,

$$\prod_{\substack{i=1\\ i\neq\sigma(1)}}^{r} g_i = h_2h_3\cdots h_s.$$

En continuant de cette manière, on obtient $r = s$ et l'existence d'une permutation σ telle que $h_i = g_{\sigma(i)}$, $i = 1,\ldots,r$. ■

On dit qu'un polynôme p de degré ≥ 1 est **scindé** si tous les facteurs irréductibles (dans la décomposition du théorème 2.45) sont de degré 1 :

$$p(t) = \alpha(t-c_1)(t-c_2)\cdots(t-c_n), \qquad \alpha,c_1,\ldots,c_n \in K.$$

Théorème 2.46 (Théorème fondamental de l'algèbre) *Tout polynôme à coefficients dans $\mathbb{C}$ est scindé.*

DÉMONSTRATION. $2^{\text{ème}}$ année. ■

Chapitre 3

Réduction de matrices : forme échelonnée

Pour ce chapitre, $(K,+,\cdot)$ sera au moins un anneau commutatif. En fait, pour trouver la forme échelonnée et donner une définition univoque du rang il sera nécessaire de supposer que $(K,+,\cdot)$ soit un corps (par exemple, $K=\mathbb{R}$, $K=\mathbb{C}$ ou $K=\mathbb{F}_2$).

C'est une idée constante en algèbre linéaire de réduire une matrice $A\in M_{m\times n}(K)$ en une forme plus simple (par exemple, diagonale ou triangulaire). Une telle réduction peut simplifier considérablement l'analyse d'un problème comme la résolution d'un système d'équations. Dans ce chapitre, nous allons voir comment on transforme une matrice en une matrice échelonnée (par la méthode de Gauss).

3.1 Matrices élémentaires

Soit $(K,+,\cdot)$ un anneau commutatif. Trois types de transformations sont utilisés pour la réduction d'une matrice sous forme échelonnée.

Type I : les matrices de permutation P_{ij}. On rappelle (voir section 2.1.1) que $(S_n,\circ)$ désigne le groupe des permutations de $\{1,2,\ldots,n\}$. Soit $e_i\in K^n$ le i-ième **vecteur unité** [*unit vector*] défini par

$$e_i=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}\leftarrow i\text{-ième ligne.}$$

Définition 3.1 *Pour une permutation $\sigma\in S_n$ on appelle la matrice $P_\sigma\in M_{n\times n}(K)$ définie par*

$$P_\sigma=\begin{pmatrix}e_{\sigma(1)}^{\mathsf{T}}\\ e_{\sigma(2)}^{\mathsf{T}}\\ \vdots\\ e_{\sigma(n)}^{\mathsf{T}}\end{pmatrix}$$

*une **matrice de permutation**.*

Exemple 3.2 La permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$$

donne la matrice

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

♦

MATLAB

Des matrices de permutation sous MATLAB:

```
>> sigma = [ 1 4 2 3 ];
>> P = eye(4); P = P(sigma,:)
P =
    1    0    0    0
    0    0    0    1
    0    1    0    0
    0    0    1    0
```

Lemme 3.3 *Une matrice $P \in M_{n\times n}(K)$ est une matrice de permutation si et seulement si elle possède dans chaque ligne et chaque colonne exactement un élément égal à 1 et les autres égaux à 0.*

DÉMONSTRATION. On remarque d'abord qu'une matrice de permutation $P = P_\sigma$ ne possède que des 1 et des zéros comme éléments. En outre, par définition 3.1, chaque ligne de P_σ a la propriété desirée. S'il existe une colonne de P_σ avec plusieurs 1, alors il existe $i \neq j$ avec $\sigma(i) = \sigma(j)$ et par suite σ ne peut pas être une permutation. Par le principe des tiroirs, il n'existe pas une colonne avec que des zéros. Ceci montre la nécéssité de la condition du lemme.

Pour montrer la suffisance, soit P une matrice possédant exactement un élément égal à 1 dans chaque ligne et chaque colonne et tous les autres éléments égaux à zéro. On définit l'application

$$\sigma(\text{indice de la ligne contenant un } 1) = \text{indice de colonne de l'élément } 1.$$

C'est bien une permutation et donc $P = P_\sigma$ est une matrice de permutation. ■

Le produit d'un vecteur $v \in K^n$ par P_σ permute les éléments de v selon σ :

$$P_\sigma v = \begin{pmatrix} e_{\sigma(1)}^\mathsf{T} \\ \vdots \\ e_{\sigma(n)}^\mathsf{T} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_{\sigma(1)} \\ v_{\sigma(2)} \\ \vdots \\ v_{\sigma(n)} \end{pmatrix}.$$

Donc, en posant $\tilde{v} := P_\sigma v$ on obtient $\tilde{v}_i = v_{\sigma(i)}$, $i = 1, \ldots, n$.

Soit P_π une (autre) matrice de permutation. Alors,

$$P_\pi P_\sigma v = P_\pi \tilde{v} = \begin{pmatrix} \tilde{v}_{\pi(1)} \\ \tilde{v}_{\pi(2)} \\ \vdots \\ \tilde{v}_{\pi(n)} \end{pmatrix} = \begin{pmatrix} v_{\sigma(\pi(1))} \\ v_{\sigma(\pi(2))} \\ \vdots \\ v_{\sigma(\pi(n))} \end{pmatrix} = P_{\sigma\circ\pi} v.$$

Comme cette égalité est vraie pour tout $v \in K^n$ on obtient

$$P_\pi P_\sigma = P_{\sigma\circ\pi} \tag{3.1}$$

Attention au renversement de l'ordre de σ et π !

En posant $\pi = \sigma^{-1}$, l'inverse de σ, dans (3.1) on obtient

$$\sigma^{-1}\circ\sigma = \mathsf{id} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} \quad \Rightarrow \quad P_\sigma P_{\sigma^{-1}} = P_{\sigma^{-1}\circ\sigma} = P_{\mathsf{id}} = I_n$$

On montre que $P_{\sigma^{-1}}P_\sigma = I_n$ en échangeant les rôles de σ et σ^{-1}.

On a d'autre part

$$P_\sigma P_\sigma^{\mathsf{T}} = \begin{pmatrix} e_{\sigma(1)}^{\mathsf{T}} \\ \vdots \\ e_{\sigma(n)}^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} e_{\sigma(1)} & \cdots & e_{\sigma(n)} \end{pmatrix} = \begin{pmatrix} e_{\sigma(1)}^{\mathsf{T}} e_{\sigma(1)} & \cdots & e_{\sigma(1)}^{\mathsf{T}} e_{\sigma(n)} \\ \vdots & & \vdots \\ e_{\sigma(n)}^{\mathsf{T}} e_{\sigma(1)} & \cdots & e_{\sigma(n)}^{\mathsf{T}} e_{\sigma(n)} \end{pmatrix} = I_n.$$

Alors,

$$P_\sigma^{-1} = P_{\sigma^{-1}} = P_\sigma^T. \tag{3.2}$$

Lemme 3.4 *L'ensemble des matrices de permutations muni du produit matriciel est un sous-groupe de $GL_n(K)$.*

DÉMONSTRATION. 1) L'ensemble est non-vide car I_n est une matrice de permutation. 2) La stabilité découle de (3.1). 3) Par (3.2) l'inverse d'une matrice de permutation est encore une matrice de permutation. ■

Définition 3.5 *Une **transposition** $\sigma \in S_n$ est une permutation qui échange exactement deux éléments:*

$$\pi = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{pmatrix}, \quad 1 \le i < j \le n.$$

La matrice de permutation correspondante est notée P_{ij}.

Plus tard, on démontrera que toute permutation $\sigma \in S_n$ peut s'écrire comme composition d'au plus $n-1$ transpositions.

La multiplication à droite par P_{ij} échange les colonnes i et j d'une matrice. La multiplication à gauche par P_{ij} échange les lignes i et j.

Exemple 3.6

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \Rightarrow \quad P_{13}A = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}, \quad AP_{13} = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \\ 9 & 8 & 7 \end{pmatrix}.$$

♦

Comme $P_{ij} \cdot P_{ij} = I_n$ on a que $P_{ij} = P_{ij}^{-1} = P_{ij}^{\mathsf{T}}$, donc la matric P_{ij} est symmétrique.

Type II : les matrices diagonales $M_i(\lambda)$. Pour $\lambda \in K$ on définit la matrice $M_i(\lambda)$ par

$$M_i(\lambda) = \mathrm{diag}\,(\underbrace{1,\ldots,1}_{i-1 \text{ fois}},\lambda,\underbrace{1,\ldots,1}_{n-i \text{ fois}})$$

La multiplication à gauche d'une matrice $A \in M_{n\times p}(K)$ par une matrice $M_i(\lambda)$ multiplie la ligne i de A par λ et laisse les autres lignes inchangées. (La multiplication à droite multiplie la colonne i par λ.)

Exemple 3.7

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \Rightarrow \quad M_2(3)A = \begin{pmatrix} 1 & 2 & 3 \\ 12 & 15 & 18 \\ 7 & 8 & 9 \end{pmatrix}$$

♦

On voit facilement que $M_i(\lambda)$ est inversible si λ est inversible et que l'inverse est donné par

$$M_i(\lambda)^{-1} = M_i(\lambda^{-1}).$$

Type III : $G_{ij}(\lambda)$**.** Soit $n \geq 2$, $\lambda \in K$ et $1 \leq i < j \leq n$. On définit alors la matrice

$$G_{ij}(\lambda) = I_n + \lambda e_j e_i^{\mathsf{T}} = \begin{array}{c} \\ \\ j \to \end{array} \begin{array}{c} \begin{array}{ccccc} & & \overset{i}{\downarrow} & & \end{array} \\ \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & \ddots & \\ & & & & 1 \end{pmatrix} \end{array} \in M_{n\times n}(K).$$

La multiplication de $G_{ij}(\lambda)$ et $A \in M_{n\times p}(K)$:

$$G_{ij}(\lambda)A = (I_n + \lambda e_j e_i^{\mathsf{T}})A = A + \lambda e_j e_i^{\mathsf{T}} A = A + \begin{array}{c} \\ j \to \end{array} \begin{pmatrix} 0 & \cdots & 0 \\ \lambda a_{i1} & \cdots & \lambda a_{ip} \\ 0 & \cdots & 0 \end{pmatrix}.$$

Donc, $G_{ij}(\lambda)A$ additionne λ fois la ligne i de la matrice A à la ligne j de cette même matrice et laisse les autres lignes inchangées. De façon analogue on voit que $G_{ij}(\lambda)^{\mathsf{T}}A$ additionne λ fois la ligne j à la ligne i :

$$G_{ij}(\lambda)^{\mathsf{T}}A = (I_n + \lambda e_i e_j^{\mathsf{T}})A = A + \lambda e_i e_j^{\mathsf{T}} A.$$

Exemple 3.8

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \Rightarrow G_{13}(-2)A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 4 & 3 \end{pmatrix}, G_{13}(-2)^{\mathsf{T}}A = \begin{pmatrix} -13 & -14 & -15 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

♦

Lemma 3.9 $G_{ij}(\lambda)^{-1} = G_{ij}(-\lambda)$.

DÉMONSTRATION.

$$\begin{aligned} G_{ij}(\lambda)G_{ij}(-\lambda) &= (I_n + \lambda e_j e_i^{\mathsf{T}})(I_n - \lambda e_j e_i^{\mathsf{T}}) \\ &= I_n + \lambda e_j e_i^{\mathsf{T}} - \lambda e_j e_i^{\mathsf{T}} - \lambda^2 e_j \underbrace{e_i^{\mathsf{T}} e_j}_{=0} e_i^{\mathsf{T}} = I_n. \end{aligned}$$

De même $G_{ij}(-\lambda)G_{ij}(\lambda) = I_n$. ■

3.2 Reduction à la forme échelonnée

Soit $(K, +, \cdot)$ un corps et soit $A \in M_{m\times n}(K)$. On veut construire une matrice inversible $B \in M_{m\times m}(K)$ telle que BA est la plus « simple » possible. La construction est basée sur l'élimination de Gauss, que nous avons déjà utilisée dans le chapitre 0 afin de résoudre un système à trois inconnus. On va exprimer les transformations effectuées par l'élimination de Gauss comme des multiplications à gauche par des matrices élémentaires :

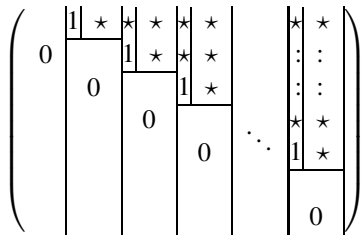
Type I. P_{ij} – Échange des lignes i et j.

Type II. $M_i(\lambda)$ – Multiplication de la ligne i par λ.

Type III. $G_{ij}(\lambda)$ – Ajout de la ligne i multipliée par λ à la ligne j.

On construira la matrice B comme produit des matrices élémentaires.

Définition 3.10 *Une matrice $C \in M_{m\times n}(K)$ est dite* ***échelonnée*** *[(row) echelon form] si elle est de la forme*



(3.3)

où les étoiles $\star$ désignent des éléments quelconques.

Une définition plus formelle de (3.3) : Il existe des entiers $j_1,\dots,j_r \in \mathbb{N}$ tels que $1 \le j_1 < \cdots < j_r \le n$, $1 \le r \le \min\{m,n\}$ et

- $c_{ij} = 0$ si $0 < i \le m$ et $0 < j < j_1$;
- $c_{ij} = 0$ si $k < i \le m$ et $j_k \le j < j_{k+1}$, $k = 1,\dots,r$;
- $c_{kj_k} = 1$, $k = 1,\dots,r$.

Le premier coefficient non-nul sur une ligne non-nulle d'une matrice échelonnée est appelé un **pivot**. Ainsi, les éléments $c_{kj_k} = 1$, $k = 1,\dots,r$, sont les pivots de C. On remarque que les pivots peuvent parfois être des nombres quelconques non nuls.

Définition 3.11 *Une matrice $C \in M_{m\times n}(K)$ est dite* ***échelonnée réduite*** *[reduced (row) echelon form] si elle est échelonnée et si tous ses coefficients au-dessus des pivots sont nuls ($c_{ij_k} = 0$ si $1 \le i < k$, $k = 1,\dots,r$):*

$$C = \left(\begin{array}{c|c|c|c|c|c} 0 & \begin{array}{c|c} 1 & \star \\ \hline 0 & \end{array} & \begin{array}{c|c} 0 & \star \\ 1 & \star \\ \hline 0 & \end{array} & \begin{array}{c|c} 0 & \star \\ 0 & \star \\ 1 & \star \\ \hline 0 & \end{array} & \ddots & \begin{array}{c|c} 0 & \star \\ \vdots & \vdots \\ 0 & \star \\ 1 & \star \\ \hline 0 & \end{array} \end{array}\right).$$

Exemple 3.12 La matrice

$$C = \begin{pmatrix} 0 & 1 & 0 & 3 & 0 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

est échelonnée (réduite) avec $r = 3$ et $j_1 = 2, j_2 = 3, j_3 = 5$. ♦

Théorème 3.13 *Soit $(K,+,\cdot)$ un corps et soit $A \in M_{m\times n}(K)$. Alors il existe une matrice $B \in M_{m\times m}(K)$, produit des matrices élémentaires, telle que $C = BA$ soit échelonnée réduite. Pour $m = n$ on a que A est inversible si et seulement si $C = I_n$. Si A est inversible on a alors $A^{-1} = B$.*

DÉMONSTRATION. Si $A = 0$ alors en prenant $B = I_m$ le théorème est vrai. On suppose donc $A \neq 0$.

Étape 1. On note $A^{(0)} := A$. Soit j_1 l'indice de la première colonne de $A^{(1)}$ avec au moins un élément non nul. On note i_1 la ligne de premier élément non nul de la colonne j,

ainsi $a^{(0)}_{i_1,j_1} \neq 0$ et

$$A^{(0)} = \begin{array}{c} \\ i_1 \to \end{array} \begin{array}{c} \begin{array}{ccc} & \overset{j_1}{\downarrow} & \end{array} \\ \begin{pmatrix} 0 & 0 & \star \\ 0 & a^{(0)}_{i_1,j_1} & \star \\ 0 & \star & \star \end{pmatrix} \end{array}.$$

On échange alors la ligne i_1 avec la ligne 1, puis on divise la ligne par $a^{(0)}_{i_1,j_1}$:

$$\tilde{A}^{(1)} := M_1\big(1/a^{(0)}_{i_1,j_1}\big)P_{1,i_1}A^{(0)} = \left(\begin{array}{c|cc} 0 & 1 & \star \\ 0 & \tilde{a}^{(1)}_{2,j_1} & \star \\ \vdots & \vdots & \vdots \\ 0 & \tilde{a}^{(1)}_{m,j_1} & \star \end{array}\right).$$

Finalement on élimine les termes dans la colonne j_1 en dessous de l'élément 1 en effectant

$$A^{(1)} := G_{1,m}\big(-\tilde{a}^{(1)}_{m,j_1}\big)\cdots G_{1,2}\big(-\tilde{a}^{(1)}_{2,j_1}\big)\tilde{A}^{(1)} = \left(\begin{array}{c|cc} 0 & 1 & \star \\ \hline 0 & 0 & \\ \vdots & \vdots & \hat{A}^{(2)} \\ 0 & 0 & \end{array}\right). \tag{3.4}$$

En accumulant les matrices élémentaires utilisées,

$$B_1 = G_{1,m}\big(-\tilde{a}^{(1)}_{m,j_1}\big)\cdots G_{1,2}\big(-\tilde{a}^{(1)}_{2,j_1}\big)M_1\big(1/a^{(0)}_{i_1,j_1}\big)P_{1,i_1},$$

on obtient

$$A^{(1)} = B_1A.$$

Si $\hat{A}^{(2)} = 0$, la sous-matrice de (3.4), le processus s'arrête, sinon on continue.

Étape 2. On applique le procédé de l'étape 1 à $\hat{A}^{(2)} \neq 0$. Soit $j_2 > j_1$ l'indice de la colonne de $A^{(1)}$ qui correspond à la première colonne non nulle de $\hat{A}^{(2)}$ et soit $i_2 \geq 2$ l'indice de ligne du premier élément non nul de cette colonne. Alors, $a^{(1)}_{i_2,j_2} \neq 0$ et on obtient

$$M_2\big(1/a^{(1)}_{i_2,j_2}\big)P_{2,i_2}A^{(1)} = \left(\begin{array}{c|ccc|c} 0 & 1 & \star & \star & \star \\ \hline 0 & 0 & 0 & 1 & \star \\ & & & \tilde{a}^{(2)}_{3,j_2} & \\ 0 & 0 & 0 & \vdots & \star \\ & & & \tilde{a}^{(2)}_{m,j_2} & \end{array}\right).$$

On note que ces transformations ne modifient pas la première ligne de $A^{(1)}$. Comme auparavant les $m-2$ éléments potentiellement non nuls en dessous de 1 sont éliminés en définissant

$$B_2 = G_{1,m}\big(-\tilde{a}^{(2)}_{m,j_2}\big)\cdots G_{1,3}\big(-\tilde{a}^{(2)}_{3,j_2}\big)M_2\big(1/a^{(1)}_{i_2,j_2}\big)P_{2,i_2}.$$

On obtient

$$A^{(2)} := B_2B_1A = \left(\begin{array}{c|cc|cc} 0 & 1 & \star & \star & \star \\ \hline 0 & 0 & 0 & 1 & \star \\ \hline 0 & 0 & 0 & 0 & \hat{A}^{(3)} \end{array}\right).$$

Si $\hat{A}^{(3)} \neq 0$ on applique le même procédé à la matrice $\hat{A}^{(3)}$. Après au plus $r \leq \min\{m,n\}$ étapes le procédé s'arrête et l'on obtient la forme échelonnée

$$A^{(r)} := B_r \cdots B_2 B_1 A = \begin{pmatrix} 0 & \boxed{1} & \star & \star & \star & \star & \star & \star \\ & & 0 & \boxed{1} & \star & \star & \star & \star \\ & & & & 0 & \boxed{1} & \star & \star \\ & & & & & & \ddots & \vdots \\ & & & & & & & \boxed{1} \\ & & & & & & & 0 \end{pmatrix}. \tag{3.5}$$

Les positions des pivots se trouvent aux lignes et colonnes données par

$$(1,j_1),\ (2,j_2),\ \ldots,\ (r,j_r). \tag{3.6}$$

Pour avoir une matrice échelonnée réduite, on doit encore éliminer les éléments au dessus des pivots. En définissant $C^{(r)} := A^{(r)}$ et de façon recursive

$$C^{(k-1)} := \tilde{B}_k C^{(k)} \quad \text{avec} \quad k = r, r-1, \ldots, 2,$$
$$\tilde{B}_k := \big(G_{1,k}(-c^{(k)}_{1,j_k})\big)^{\mathsf{T}} \big(G_{2,k}(-c^{(k)}_{2,j_k})\big)^{\mathsf{T}} \cdots \big(G_{k-1,k}(-c^{(k)}_{k-1,j_k})\big)^{\mathsf{T}},$$

on obtient alors que $C := C^{(1)}$ est échelonée réduite. Cela montre la première assertion du théorème avec la matrice inversible

$$B := \tilde{B}_2 \cdots \tilde{B}_r B_r \cdots B_1.$$

Pour montrer la deuxième assertion supposons que $m = n$. Si nous supposons que A est inversible alors $C = BA$ est inversible. Comme une matrice inversible ne peut pas avoir de ligne ou de colonne nulle on a forcément $C = I_n$. Réciproquement si $C = I_n$, comme $I_n = BA$ et que B est inversible on a $A = B^{-1} I_n = B^{-1}$ et donc A est inversible. ■

Exemple 3.14 Soit

$$A = \begin{pmatrix} 0 & 2 & 1 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{pmatrix} \in M_{3\times 4}(\mathbb{Q}).$$

Selon la preuve ci-dessus on obtient

$$B_1 : \overset{M_1(1/2)}{\Rightarrow} \begin{pmatrix} 0 & 1 & 1/2 & 3/2 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{pmatrix} \overset{\substack{G_{12}(-2)\\ G_{13}(-2)}}{\Rightarrow} \begin{pmatrix} 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -1 & -1 \end{pmatrix}$$
$$B_2 : \overset{M_2(-1)}{\Rightarrow} \begin{pmatrix} 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & -1 \end{pmatrix} \overset{G_{23}(1)}{\Rightarrow} \begin{pmatrix} 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$\check{B}_3 : \overset{\substack{G_{23}(-2)^{\mathsf{T}}\\ G_{13}(-3/2)^{\mathsf{T}}}}{\Rightarrow} \begin{pmatrix} 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$\check{B}_2 : \overset{G_{12}(-1/2)^{\mathsf{T}}}{\Rightarrow} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = C.$$

La matrice B est le produit de matrices élémentaires :

$$\begin{aligned} B &= \check{B}_2 \check{B}_3 B_2 B_1 \\ &= G_{12}(-1/2)^{\mathsf{T}} G_{13}(-3/2)^{\mathsf{T}} G_{23}(-2)^{\mathsf{T}} G_{23}(1) M_2(-1) G_{13}(-2) G_{12}(-2) M_1(1/2) \\ &= \begin{pmatrix} 0 & 1 & -1/2 \\ 1 & 1 & -2 \\ 0 & -1 & 1 \end{pmatrix}. \end{aligned}$$

On vérifie que $BA = C$ est vrai. ♦

3.3 Matrices équivalentes

Tout d'abord on rappelle la notion de relation d'équivalence.

Définition 3.15 *Une* ***relation binaire*** *sur un ensemble E est un sous-ensemble R de $E \times E$. On note aussi xRy à la place de $(x,y) \in R$. Une relation binaire R est appelée une* ***relation d'équivalence*** *si elle est à la fois:*

- ***réflexive*** *: xRx pour tout $x \in E$*
- ***symétrique*** *: xRy implique yRx*
- ***transitive*** *: xRy et yRz impliquent xRz.*

Dans ce cas l'ensemble

$$[x] = \{y \in E \,|\, yRx\} \tag{3.7}$$

est appelé la ***classe d'équivalence*** *de x pour la relation R.*

La forme échelonée réduite est obtenue en multipliant une matrice $A \in M_{m\times n}(K)$ par des matrices élémentaires à gauche. Si l'on effectue aussi des opérations sur les colonnes (c-à-d on multiplie par des matrices élémentaires à droite) on est amené à la définition suivante.

Définition 3.16 *Deux matrices $A,B \in M_{m\times n}(K)$ sont dites* ***équivalentes*** *s'il existe des matrices* inversible *$Q \in M_{m\times m}(K)$, $Z \in M_{n\times n}(K)$ telles que $A = QBZ$.*

On voit facilement que la « l'équivalence des matrices » est une relation d'équivalence sur $M_{m\times n}(K)$:

- réflexive : A est équivalente à elle-meme avec $Q = I_m$, $Z = I_n$.
- symétrique : $A = QBZ$ implique $B = Q^{-1}AZ^{-1}$.
- transitive : $A = Q_1BZ_1$ et $B = Q_2CZ_2$ impliquent $A = (Q_1Q_2)C(Z_2Z_1)$.

La classe d'équivalence de A est :

$$[A] = \big\{QAZ : Q \in M_{m\times m}(K), Z \in M_{n\times n}(K) \text{ inversibles}\big\}.$$

Théorème 3.17 *Soit $(K,+,\cdot)$ un corps.*

(i) Soit $A \in M_{m\times n}(K)$. Alors A est équivalente à la matrice

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

où r est le nombre de pivots de la forme échelonnée réduite de A.

(ii) Deux matrices $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m\times n}(K)$ et $\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \in M_{m\times n}(K)$ sont équivalentes si et seulement si $r = s$.

DÉMONSTRATION. (i) D'après le théorème 3.13 il existe une matrice Q inversible telle que $C = QA$ soit échelonnée réduite. Soient $(1,j_1),(2,j_2),\ldots,(r,j_r)$ les positions des pivots de C. On considère la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & r & r+1 & \cdots & n \\ j_1 & j_2 & \cdots & j_r & \star & \cdots & \star \end{pmatrix}. \tag{3.8}$$

En multipliant C par la matrice de permutation P_σ^T à droite, on met devant les colonnes contenant les pivots [3] :

$$CP_\sigma^\mathsf{T} = \left(\begin{array}{c|c} I_r & \star \\ \hline 0 & 0 \end{array}\right) =: \left(\begin{array}{c|c} I_r & X \\ \hline 0 & 0 \end{array}\right), \qquad X \in M_{r\times(n-r)}(K).$$

En posant $Z_0 = \left(\begin{array}{c|c} I_r & -X \\ \hline 0 & I_{n-r} \end{array}\right)$ on remarque que Z_0 est inversible, d'inverse $Z_0^{-1} = \left(\begin{array}{c|c} I_r & X \\ \hline 0 & I_{n-r} \end{array}\right)$.
En effet,

$$Z_0 Z_0^{-1} = \left(\begin{array}{c|c} I_r & I_r X - X I_{n-r} \\ \hline 0 & I_{n-r} \end{array}\right) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & I_{n-r} \end{array}\right).$$

Donc,

$$QAP_\sigma^\mathsf{T} Z_0 = CP_\sigma^\mathsf{T} Z_0 = \left(\begin{array}{c|c} I_r & X \\ \hline 0 & 0 \end{array}\right)\left(\begin{array}{c|c} I_r & -X \\ \hline 0 & I_{n-r} \end{array}\right) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Posant $Z = P_\sigma^\mathsf{T} Z_0$, on conclut la preuve de la partie (i).

(ii). Si $r = s$ les deux matrices sont identiques donc équivalentes. Il reste à montrer que la condition $r = s$ est nécessaire pour l'équivalence de deux matrices. Montrons cela par l'absurde : Supposons que $r \neq s$, donc que $r < s$ sans perte de généralité, et supposons qu'il existe Q,Z inversibles telles que

$$\begin{pmatrix} I_r & 0 & 0 \\ 0 & I_{s-r} & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix}\begin{pmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{pmatrix}$$
$$= \begin{pmatrix} Q_{11}Z_{11} & Q_{11}Z_{12} & Q_{11}Z_{13} \\ Q_{21}Z_{11} & Q_{21}Z_{12} & Q_{21}Z_{13} \\ Q_{31}Z_{11} & Q_{31}Z_{12} & Q_{31}Z_{13} \end{pmatrix},$$

où Q et Z sont partitionnées de façon compatible (pour que les produits aient un sens). Comme $Q_{11}Z_{11} = I_r$ la matrice Q_{11} est inversible. Comme $Q_{11}Z_{12} = 0$ avec Q_{11} inversible on obtient $Z_{12} = 0$, mais c'est une contradiction avec $Q_{21}Z_{12} = I_{s-r}$. ■

Remarque 3.18 *Dans la langage des « classes d'équivalence », voir* (3.7)*, on peut exprimer le théorème 3.17 ainsi :*

$$M_{m\times n}(K) = \bigcup_{r=0}^{\min\{m,n\}} \left[\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\right]$$

où

$$\left[\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\right] \cap \left[\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}\right] = \emptyset \quad si \quad r \neq s.$$

Corollaire 3.19 *Soit $A \in M_{m\times n}(K)$ et C_1, C_2 deux formes échelonnées réduites de A, alors le nombre de pivots de C_1 et C_2 est identique.*

DÉMONSTRATION. Il existe Q_1,$Q_2 \in M_{m\times m}(K)$ inversibles telles que $C_1 = Q_1 A$ et $C_2 = Q_2 A$. D'après le théorème 3.17, point (i), C_1 et C_2 sont équivalentes à $E_1 = \begin{pmatrix} I_{r_1} & 0 \\ 0 & 0 \end{pmatrix}$ et $E_2 = \begin{pmatrix} I_{r_2} & 0 \\ 0 & 0 \end{pmatrix}$, où r_1 et r_2 sont les nombres de pivots de C_1 et C_2, respectivement. Comme l'équivalence est transitive, on a que E_1 et E_2 sont équivalentes. D'après le théorème 3.17, point (ii), $r_1 = r_2$. ■

3. Pour montrer ceci il convient de considérer la transposée $(CP_\sigma^\mathsf{T})^\mathsf{T} = P_\sigma C^\mathsf{T}$.

Avec un peu plus de travail on peut en fait montrer que la forme échelonnée réduite d'une matrice est unique.

Définition 3.20 *Soit $(K,+,\cdot)$ un corps et soit $A \in M_{m\times n}(K)$. Le* ***rang*** *[rank] de A, noté $\mathrm{rang}(A)$, est le nombre de pivots r de A.*

Théorème 3.21 *Soit $(K,+,\cdot)$ un corps et soit $A \in M_{m\times n}(K)$. Alors :*

(i) Pour $Q \in M_{m\times m}(K)$, $Z \in M_{n\times n}(K)$ inversibles, on a

$$\mathrm{rang}(QAZ) = \mathrm{rang}(A).$$

(ii) Pour $A = BC$ avec $B \in M_{m\times p}(K)$, $C \in M_{p\times n}(K)$ on a

$$\mathrm{rang}(A) \leq \mathrm{rang}(B), \qquad \mathrm{rang}(A) \leq \mathrm{rang}(C).$$

(iii) $\mathrm{rang}(A^{\mathsf{T}}) = \mathrm{rang}(A)$.

DÉMONSTRATION. (i) découle directement du théorème 3.17 et du corollaire 3.19.

(iii). D'après le théorème 3.17 il existe Q, Z inversibles telles que $QAZ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ et donc

$$Z^{\mathsf{T}}A^{\mathsf{T}}Q^{\mathsf{T}} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Alors, A^{T} est équivalente à une matrice de rang r et, ainsi, $\mathrm{rang}(A^{\mathsf{T}}) = r = \mathrm{rang}(A)$.

(iia) Montrons d'abord que $\mathrm{rang}(A) \leq \mathrm{rang}(B)$. Soit Q une matrice inversible telle que QB est en forme échelonnée réduite. Alors les $m - \mathrm{rang}(B)$ dernières lignes de QB sont nulles. Mais alors les $m - \mathrm{rang}(B)$ dernières lignes de $QA = (QB)C$ sont nulles. Le rang d'une matrice $m \times n$ contenant $m - \mathrm{rang}(B)$ lignes nulles ne peut être strictement supérieur à $\mathrm{rang}(B)$. Alors, d'après (i), $\mathrm{rang}(A) = \mathrm{rang}(QA) \leq \mathrm{rang}(B)$.

(iib) Pour montrer $\mathrm{rang}(A) \leq \mathrm{rang}(C)$, on utilise (iii) et (iib) : $\mathrm{rang}(A) = \mathrm{rang}(A^{\mathsf{T}}) = \mathrm{rang}(C^{\mathsf{T}}B^{\mathsf{T}}) \leq \mathrm{rang}(C^{\mathsf{T}}) = \mathrm{rang}(C)$. ■

3.4 Solutions de systèmes linéaires

On va voir qu'à l'aide de la forme échelonnée (réduite) d'une matrice, l'on peut résoudre aisément des systèmes linéaires et décrire l'ensemble des solutions de ces systèmes.

On considère le système linéaire suivant sur un corps $(K,+,\cdot)$:

$$\begin{array}{ccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & +\cdots+ & a_{2n}x_n & = & b_2, \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1 & +\cdots+ & a_{mn}x_n & = & b_m. \end{array}$$

On a vu que ce système s'écrit

$$Ax = b, \tag{3.9}$$

avec $A \in M_{m\times n}(K)$, $x \in K^n$ et $b \in K^m$. Si $b = 0$ on dit que le système linéaire est **homogène** [*homogeneous*], sinon on dit qu'il est **inhomogène** [*inhomogeneous*].

Définition 3.22 *On note par $S(A,b) = \{x \in K^n : Ax = b\}$ l'ensemble des solutions d'un système linéaire.*

Lemme 3.23 *Soit $A \in M_{m\times n}(K)$, $b \in K^n$ et soit $x_p \in S(A,b)$. Alors*

$$S(A,b) = \{x_p + x_h \,|\, x_h \in S(A,0)\}.$$

DÉMONSTRATION. (a) $\{x_p + x_h \,|\, x_h \in S(A,0)\} \subseteq S(A,b)$: Soit $x_h \in S(A,0)$. Alors

$$A(x_p + x_h) = Ax_p + Ax_h = b + 0 \quad \Rightarrow \quad x_p + x_h \in S(A,b).$$

(b) $S(A,b) \subseteq \{x_p + x_h \,|\, x_h \in S(A,0)\}$: Soit $x \in S(A,b)$. De

$$A(x - x_p) = Ax - Ax_p = b - b = 0,$$

on obtient $x = x_p + x_h$ avec $x_h = x - x_p \in S(A,0)$. ∎

Le lemme ci-dessus montre une propriété fondamentale des systèmes linéaires : pour trouver l'ensemble des solutions de $Ax = b$ il suffit de trouver l'ensemble des solutions de système homogène associé $Ax = 0$ et d'additionner une solution particulière de $Ax = b$.

Soit $Q \in M_{m\times m}(K)$ inversible. Alors

$$Ax = b \quad \Leftrightarrow \quad QAx = Qb,$$

c-à-d

$$S(A,b) = S(QA,Qb). \tag{3.10}$$

En particulier d'après le théorème 3.13 il existe une matrice Q inversible telle que QA est en forme échelonnée :

$$\left(\begin{array}{c|c|c|c|c|c} 0 & \boxed{1}\ \star & \boxed{0}\ \star & \boxed{0}\ \star & & \boxed{0}\ \star \\ & \boxed{1}\ \star & \boxed{0}\ \star & \boxed{0}\ \star & & \vdots\ \ \vdots \\ & 0 & \boxed{1}\ \star & \boxed{0}\ \star & & \vdots\ \ \vdots \\ & & 0 & \boxed{1}\ \star & \ddots & \boxed{0}\ \star \\ & & & 0 & & \boxed{1}\ \star \\ & & & & & 0 \end{array}\right) x = \tilde{b},$$

On a vu dans la preuve du théorème 3.17, voir (3.8), qu'à l'aide d'une multiplication à droite par une matrice de permutation on obtient

$$\tilde{A} := QAP_\sigma^\mathsf{T} = \left(\begin{array}{c|c} I_r & \tilde{A}_{12} \\ \hline 0 & 0 \end{array}\right).$$

Au contraire de la multiplication à gauche, celle à droite modifie l'ensemble des solutions :

$$Ax = b \quad \Leftrightarrow \quad QAP_\sigma^\mathsf{T} P_\sigma x = Qb \quad \Leftrightarrow \quad \tilde{A}\tilde{x} = \tilde{b},$$

où $\tilde{x} = P_\sigma x$. Une solution de $Ax = b$ devient une solution de $\tilde{A}\tilde{x} = \tilde{b}$ en permutant ses élements, et vice versa. Ecrivons le système $\tilde{A}\tilde{x} = \tilde{b}$ sous la forme partitionée

$$\left(\begin{array}{c|c} I_r & \tilde{A}_{12} \\ \hline 0 & 0 \end{array}\right)\left(\begin{array}{c} \tilde{x}_1 \\ \hline \tilde{x}_2 \end{array}\right) = \left(\begin{array}{c} \tilde{b}_1 \\ \hline \tilde{b}_2 \end{array}\right), \quad \tilde{x}_1 \in K^r, \tilde{x}_2 \in K^{n-r}, \quad \tilde{b}_1 \in K^r, \tilde{b}_2 \in K^{m-r}. \tag{3.11}$$

Ce système peut s'écrire comme

$$\tilde{x}_1 + \tilde{A}_{12}\tilde{x}_2 = \tilde{b}_1, \qquad 0 = \tilde{b}_2.$$

Donc si $\tilde{b}_2 \neq 0$ le système linéaire $\tilde{A}x = \tilde{b}$ n'a pas de solution, on dit qu'il est **incompatible**. Si en revanche $\tilde{b}_2 = 0$ alors en posant

$$\tilde{x}_p = \begin{pmatrix} \tilde{b}_1 \\ 0 \end{pmatrix}, \tag{3.12}$$

on obtient une solution particulière, c-à-d $\tilde{x}_p \in S(\tilde{A},\tilde{b})$. On a trouvé

$$S(\tilde{A},\tilde{b}) \neq \emptyset \qquad \Leftrightarrow \qquad \tilde{b}_2 = 0.$$

En considérant la **matrice augmentée** [*augmented matrix*] $\left(\begin{array}{c|c} A & b \end{array}\right) \in M_{m\times(n+1)}(K)$ on obtient un critère élégant.

Lemme 3.24 *Soit $(K,+,\cdot)$ un corps, soit $A \in M_{m\times n}(K)$ et $b \in K^m$. Alors $S(A,b) \neq \emptyset$ si et seulement si*

$$\mathrm{rang}\left(\begin{array}{c|c} A & b \end{array}\right) = \mathrm{rang}(A).$$

DÉMONSTRATION. D'abord, on considère la matrice augmentée du système réduit (3.11) :

$$\left(\begin{array}{c|c} \tilde{A} & \tilde{b} \end{array}\right) = \left(\begin{array}{ccc|c} I_r & \tilde{A}_{12} & 0 & \tilde{b}_1 \\ 0 & 0 & 0 & \tilde{b}_2 \end{array}\right).$$

Si $\tilde{b}_2 = 0$ alors $\mathrm{rang}(\tilde{A}|\tilde{b}) = \mathrm{rang}(\tilde{A}) =: r$. Si en revanche $\tilde{b}_2 \neq 0$ alors la matrice a un pivot de plus que $\tilde{A}$ et donc $\mathrm{rang}(\tilde{A}|\tilde{b}) = r+1 \neq \mathrm{rang}(\tilde{A})$. Selon la discussion ci-dessus, $\tilde{b}_2 = 0$ si et seulement si $S(A,b) \neq \emptyset$. Alors la conclusion du lemme est vraie pour (3.11).

Le théorème 3.21, point (i), montre la conclusion du lemme pour le système d'origine $Ax = b$:

$$\mathrm{rang}(\tilde{A}) = \mathrm{rang}(QAP_\sigma^T) = \mathrm{rang}(A),$$
$$\mathrm{rang}\left(\left(\begin{array}{c|c} \tilde{A} & \tilde{b} \end{array}\right)\right) = \mathrm{rang}\left(Q\left(\begin{array}{c|c} AP_\sigma^T & b \end{array}\right)\right) = \mathrm{rang}\left(\left(\begin{array}{c|c} A & b \end{array}\right)\right).$$

■

Pour décrire l'ensemble des solutions de $\tilde{A}\tilde{x} = \tilde{b}$ on utilise le lemme 3.23. On sait déjà, voir (3.12), que $\tilde{x}_p = \begin{pmatrix} \tilde{b}_1 \\ 0 \end{pmatrix}$ est une solution particulière si $\tilde{b}_2 = 0$. Dans le cas homogène ($\tilde{b}_1 = 0$) on a

$$S(\tilde{A},0) = \left\{ \begin{pmatrix} \tilde{x}_{h1} \\ \tilde{x}_{h2} \end{pmatrix} : \tilde{x}_{h2} \in K^{n-r}, \tilde{x}_{h1} = -\tilde{A}_{12}\tilde{x}_{h2} \right\},$$

et donc

$$S(\tilde{A},\tilde{b}) = \left\{ \begin{pmatrix} \tilde{b}_1 - \tilde{A}_{12}\tilde{x}_{h2} \\ \tilde{x}_{h2} \end{pmatrix} : \tilde{x}_{h2} \in K^{n-r} \right\}.$$

Comme $\tilde{x}_{h2}$ peut être choisi librement, on a

- une solution unique si $n = r$ et $\tilde{b}_2 = 0$,
- plus d'une solution si $n > r$ et $\tilde{b}_2 = 0$,
- pas de solution si $\tilde{b}_2 \neq 0$.

Comme $\tilde{x} \in S(\tilde{A},\tilde{b})$ si et seulement si $P_\sigma^T\tilde{x} \in S(A,b)$, on obtient la caractérisation suivante.

Solutions de $Ax = b$ avec $A \in M_{m\times n}(K)$, $b \in K^m$

1. Si $\mathrm{rang}\left(\begin{array}{c|c} A & b \end{array}\right) > \mathrm{rang}(A)$ alors $S(A,b) = \emptyset$.
2. Si $\mathrm{rang}\left(\begin{array}{c|c} A & b \end{array}\right) = \mathrm{rang}(A) = n$ il existe une solution unique à $Ax = b$.
3. Si $\mathrm{rang}\left(\begin{array}{c|c} A & b \end{array}\right) = \mathrm{rang}(A) < n$ il existe plus qu'une solution à $Ax = b$.

Attention ! Dans un corps $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ on a une infinité de solutions dans le cas 3. Dans un corps fini comme (par exemple $\mathbb{F}_p$) on a seulement un nombre fini de solutions dans le cas 3.

Exemple 3.25 Soit $K = \mathbb{Q}$ et

$$A = \begin{pmatrix} 0 & 2 & 1 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 3 \end{pmatrix}.$$

Afin de déterminer l'ensemble des solutions de $Ax = b$, on va d'abord réduire A sous sa forme échelonnée QA. C'est une bonne idée qu'on appliquer les transformations correspondantes à b pendant la réduction au lieu de calculer explicitement Qb après la réduction :

$$\begin{aligned} \left(\begin{array}{c|c} A & b \end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 1/2 & 3/2 & 1 \\ 0 & 0 & -1 & -2 & 1 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 1/2 & 3/2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \\ &\rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 1/2 & 0 & 5/2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \\ &= \left(\begin{array}{c|c} QA & Qb \end{array}\right). \end{aligned}$$

On rearrange les colonnes par

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \left(\begin{array}{c|c} \tilde{A} & \tilde{b} \end{array}\right) = \left(\begin{array}{c|c} QAP^{\mathsf{T}} & Qb \end{array}\right) = \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Finalement

$$S(\tilde{A},\tilde{b}) = \left\{ \begin{pmatrix} 2 \\ 1 \\ -1 \\ \tilde{x}_4 \end{pmatrix} : \tilde{x}_4 \in \mathbb{Q} \right\} \quad \Rightarrow \quad S(A,b) = \left\{ \begin{pmatrix} x_1 \\ 2 \\ 1 \\ -1 \end{pmatrix} : x_1 \in \mathbb{Q} \right\}.$$

♦

Chapitre 4
Espaces vectoriels

La notion d'espace vectoriel est la structure de base de l'algèbre linéaire.

4.1 Définitions

$(K,+,\cdot)$ désigne un corps dans ce chapitre.

Définition 4.1 *Un* ***K-espace vectoriel*** *[vector space, linear space] est un ensemble muni de deux lois*

$$\begin{array}{lll} +: V\times V\to V, & (v,w)\mapsto v+w, & \textit{(addition de vecteurs)} \\ \cdot: K\times V\to V, & (\alpha,v)\mapsto \alpha\cdot v, & \textit{(multiplication par un scalaire)} \end{array} \tag{4.1}$$

vérifiant :

(i) $(V,+)$ *est un groupe abélien.*
(ii) $\alpha\cdot(\beta\cdot v)=(\alpha\beta)\cdot v, \quad \forall\alpha,\beta\in K, v\in V.$ *(compatibilité·)*
(iii) $1\cdot v=v, \quad \forall v\in V.$ *(neutralité 1)*
(iv) $(\alpha+\beta)\cdot v=(\alpha\cdot v)+(\beta\cdot v), \quad \forall\alpha,\beta\in K, v\in V.$ *(distributivité I)*
(v) $\alpha\cdot(v+w)=(\alpha\cdot v)+(\alpha\cdot w), \quad \forall\alpha\in K, v,w\in V.$ *(distributivité II)*

Quelques remarques:

- La définition 4.1 utilise les mêmes symboles $+$ / $\cdot$ aussi bien pour l'addition / la multiplication dans K que pour l'addition / la multiplication par un scalaire dans V. On comprend normalement la signification de ces symboles à partir du contexte.
- On omet souvent le $\cdot$, par exemple on écrit $\alpha\cdot v=\alpha v$. On peut écrire $\alpha\beta v=\alpha(\beta v)=(\alpha\beta)v$ grace à la compatibilité de $\cdot$. En plus, la multiplication par un scalaire est prioritaire sur l'addition, par exemple $\alpha v+\beta w=(\alpha v)+(\beta w)$.
- La stabilité des deux lois de composition est une hypothèse importante cachée dans (4.1).
- Comme d'habitude on écrit $v-w:=v+(-w)$.
- Les éléments de V s'appellent des **vecteurs**[4] et les éléments de K des **scalaires**.

Exemples d'espaces vectoriels

Matrices. $M_{m\times n}(K)$ est un K-espace vectoriel avec l'addition des matrices et avec la multiplication par un scalaire comme définies au chapitre 1.

4. On ne peut pas confondre les vecteurs colonnes ou lignes du chapitre 1 avec la notion plus générale de vecteur comme un élément d'un espace vectoriel.

Vecteurs colonnes. En particulier, $K^n = M_{n\times 1}(K)$ est un K-espace vectoriel. C'est le prototype d'un espace vectoriel. En effet, on va voir dans la section 5.2 que tout espace vectoriel de dimension finie est isomorphe à K^n pour un certain n.

Polynômes. L'ensemble de polynômes $K[t]$ est un K-espace vectoriel avec l'addition des polynômes comme définie à la page 26 et avec la multiplication par un scalaire comme suit :

$$\cdot : K \times K[t] \to K[t] \qquad \cdot : (\lambda, p) \mapsto \lambda \cdot p = \lambda p,$$

où $\lambda p(t) = \lambda\alpha_0 + \lambda\alpha_1 t + \lambda\alpha_2 t^2 + \cdots + \lambda\alpha_n t^n$ pour un pôlynome $p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n$.

Suites. Soient $v = \{v_n\}_{n=1}^{\infty}$, où $v_n \in K\ \forall n \geq 1$, et $w = \{w_n\}_{n=1}^{\infty}$, où $w_n \in K\ \forall n \geq 1$, deux suites. En définissant

$$v + w := \{v_n + w_n\}_{n=1}^{\infty}, \quad \alpha \cdot v := \{\alpha v_n\}_{n=1}^{\infty}, \quad \alpha \in K,$$

cet ensemble des suites d'éléments de K devient un K-espace vectoriel.

Applications. Soit E un ensemble non vide et soit $\mathsf{App}(E,K)$ l'ensemble des applications $f : E \to K$ défini dans la section 2.2. Muni de deux lois

$$(f+g)(x) := f(x) + g(x), \quad (\alpha f)(x) := \alpha f(x), \quad \forall f,g \in \mathsf{App}(E,K),\ \alpha \in K,$$

$\mathsf{App}(X,K)$ est un K-espace vectoriel.

Le lemme suivant contient quelques propriétés qui semblent triviales.

Lemme 4.2 *Soit V un K-espace vectoriel. Alors,*

(i) $\underbrace{0}_{\in K} \cdot v = \underbrace{0}_{\in V}$ *pour tout* $v \in V$,

(ii) $\alpha \cdot \underbrace{0}_{\in V} = \underbrace{0}_{\in V}$ *pour tout* $\alpha \in K$,

(iii) $(-1)\cdot v = -v$ *pour tout* $v \in V$,

(iv) $-(\alpha \cdot v) = (-\alpha)\cdot v = \alpha \cdot (-v)$ *pour tout* $\alpha \in K$, $v \in V$,

(v) $\alpha \cdot v = 0$ *si et seulement si* $\alpha = 0$ *ou* $v = 0$.

DÉMONSTRATION. (i) $0 \cdot v = (0+0)\cdot v = 0\cdot v + 0\cdot v$. En ajoutant l'inverse de $0\cdot v$ de chaque côté, il vient $0\cdot v = 0$.
(ii) $\alpha\cdot 0 = \alpha\cdot(0+0) = \alpha\cdot 0 + \alpha\cdot 0$. En ajoutant l'inverse de $\alpha\cdot 0$ de chaque côté, il vient $\alpha\cdot 0 = 0$.
(iii) $v + (-1)\cdot v = 1\cdot v + (-1)\cdot v = (1-1)\cdot v = 0\cdot v = 0$.
(iv) $(-\alpha)\cdot v = (-1)\cdot(\alpha\cdot v) = -(\alpha\cdot v)$. $\quad (-\alpha)\cdot v = \alpha\cdot((-1)\cdot v) = \alpha\cdot(-v)$.
(v) On suppose que $\alpha\cdot v = 0$. Si $\alpha = 0$ alors on a fini. Sinon $v = \alpha^{-1}\cdot 0 = 0$. Réciproquement si $\alpha = 0$ ou $v = 0$ alors (ii) et (i) montrent que $\alpha\cdot v = 0$. ■

4.2 Sous-espaces vectoriels

Définition 4.3 *Soit V un K-espace vectoriel. Une partie W de V s'appelle un* ***sous-espace vectoriel*** *[subspace] de V si W muni des deux lois de composition de V (restreintes à W) fait de W un K-espace vectoriel.*

Lemme 4.4 *Soit V un K-espace vectoriel et $W \subseteq V$, $W \neq \emptyset$. Alors W est un sous-espace vectoriel de V si et seulement si*

(i) $v + w \in W$ *pour tous* $v,w \in W$, *et*

(ii) $\alpha v \in W$ *pour tous* $\alpha \in K$, $v \in W$.

DÉMONSTRATION. Que les conditions (i) et (ii) soient nécessaires découle directement de la définition d'un K-sous-espace vectoriel.

La suffisance des conditions : Comme $W \neq \emptyset$ on prend un $v \in W$. Par (ii) $(-1)\cdot v = -v \in W$ et donc $-v+v = 0 \in W$, ceci montre que $(W,+)$ est un groupe abélien (la commutativité et l'associativité sont heritées de celle de V). Les propriétés (ii) à (v) de la définition 4.1 sont vraies dans W car elles sont vraies dans V. ■

Il est recommandé de vérifier d'abord $0 \in W$, où 0 est le vecteur nul de V. En même temps ceci vérifie la première condition du lemme 4.4, que W soit non-vide. En fait, $\{0\}$ lui-même est un sous-espace vectoriel de V, ainsi que V. Bien sûr, les cas interessants se situent entre ces deux extrêmes.

Exemples des sous-espaces vectoriels

Solutions d'un système linéaire homogène. Soit $A \in M_{m\times n}(K)$. L'ensemble des solutions $S(A,0) = \{x \in K^n \,|\, Ax = 0\}$ est un sous-espace vectoriel de K^n. En effet $S(A,0)$ est non-vide car $A\cdot 0 = 0$ et ainsi $0 \in S(A,0)$. Les conditions (i) et (ii) du lemme 4.4 :

$$x,y \in S(A,0) \;\Rightarrow\; Ax = 0, Ay = 0 \;\Rightarrow\; A(x+y) = 0 \;\Rightarrow\; x+y \in S(A,0),$$
$$x \in S(A,0) \;\Rightarrow\; Ax = 0 \;\Rightarrow\; \alpha Ax = A(\alpha x)0 \;\Rightarrow\; \alpha x \in S(A,0),$$

sont verifiées.

Matrices symétriques. L'ensemble des matrices symétriques est un sous-espace vectoriel de $M_{n\times n}(K)$.

Polynômes. Étant donné un entier n, on définit

$$K_n[t] := \{p = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n \,|\, \alpha_0,\ldots,\alpha_n \in K\},$$

l'ensemble des pôlynomes de degré inférieur ou égal à n. Comme l'addition des deux pôlynomes $p,q \in K_n$ et la multiplication de $p \in K_n$ par un scalaire sont encore des pôlynomes de degré inférieur ou égal à n, on a que $K_n[t]$ est un sous-espace vectoriel de $K[t]$. En outre, $K_m[t]$ est un sous-espace vectoriel de $K_n[t]$ si $m \leq n$.

Suites convergentes. On reprend V le K-espace vectoriel des suites sur K. Soient deux suites convergentes $\{v_n\}_{n=1}^{\infty} \in V$ avec $v_n \overset{n\to\infty}{\to} \bar{v} \in K$ et $\{w_n\}_{n=1}^{\infty} \in V$ avec $w_n \overset{n\to\infty}{\to} \bar{w} \in K$. Alors

$$v_n + w_n \overset{n\to\infty}{\to} \bar{v}+\bar{w} \quad \text{et} \quad \alpha v_n \overset{n\to\infty}{\to} \alpha\bar{v},$$

c-à-d les suites convergentes forment un sous-espace vectoriels de V. (On remarque que la suite $\{0\}_{n=1}^{\infty}$ est convergente.)

Attention ! $\mathbb{R}^2$ n'est pas un sous-espace vectoriel de $\mathbb{R}^3$ car $\mathbb{R}^2$ n'est pas inclus dans $\mathbb{R}^3$. Mais

$$W = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} \,|\, v_1,v_2 \in \mathbb{R} \right\}$$

est un sous-espace vectoriel de $\mathbb{R}^3$.

En rajoutant les vecteurs qui manquent, selon le lemme 4.4, on peut transformer n'importe quelle famille de vecteurs en un sous-espace vectoriel.

Définition 4.5 *Soit V un K-espace vectoriel et $v_1,\dots,v_n \in V$ et $\alpha_1,\dots,\alpha_n \in K$. Un vecteur*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V,$$

s'appelle une ***combinaison linéaire*** *[linear combination] de $v_1,\dots,v_n$. On dénote par*

$$\mathsf{span}(v_1,\dots,v_n) := \Big\{ \sum_{i=1}^{n} \alpha_i v_i \,|\, \alpha_1,\dots,\alpha_n \in K \Big\}$$

l'ensemble des combinaisons linéaires de $v_1,\dots,v_n$.

On rappelle que tout produit matrice vecteur est une combinaison linéaire des colonnes de la matrice, voir (1.13).

On peut élargir la définition 4.5 en prennant une infinité de vecteurs. À cette fin soit une famille de vecteurs[5] $(v_i)_{i\in I}$, où I est un ensemble d'indices fini ou infini. Alors on définit $\mathsf{span}(v_i)_{i\in I}$ par l'ensemble de toutes les combinaisons linéaires possibles d'un nombre fini de vecteurs :

$$\mathsf{span}(v_i)_{i\in I} := \big\{ \alpha_1 v_{i_1} + \cdots + \alpha_n v_{i_n} :\ n \in \mathbb{N}, \{i_1,\dots,i_n\} \subseteq I, \alpha_1,\dots,\alpha_n \in K \big\}.$$

Lemme 4.6 *Soit V un K-espace vectoriel et $(v_i)_{i\in I} \subset V$. Alors $\mathsf{span}(v_i)_{i\in I}$ est un sous-espace vectoriel de V.*

DÉMONSTRATION. $\mathsf{span}(v_i)_{i\in I}$ est non vide car il contient 0. Soient $x,y \in \mathsf{span}(v_i)_{i\in I}$, ainsi il existe deux ensembles d'indices finis $I_x,I_y \subset I$ et des coefficients $\alpha_i,\beta_i \in K$ tels que

$$x = \sum_{i\in I_x} \alpha_i v_i, \qquad y = \sum_{i\in I_y} \beta_i v_i.$$

Alors,

$$x+y = \sum_{i\in I_x\cup I_y} (\alpha_i+\beta_i) v_i \in \mathsf{span}(v_i)_{i\in I}, \quad \lambda x = \sum_{i\in I_x} (\lambda\alpha_i) v_i \in \mathsf{span}(v_i)_{i\in I},$$

où on prend $\alpha_i = 0$ si $i \notin I_x$ et $\beta_i = 0$ si $i \notin I_y$. ■

Le sous-espace vectoriel $\mathsf{span}(v_1,\dots,v_n)$ est appelé le **sous-espace vectoriel engendré** *[linear hull, span]* par $v_1,\dots,v_n$. De façon analogue pour une famille générale $(v_i)_{i\in I}$. On pose $\{0\} \subset V$ le sous-espace vectoriel engendré par une famille vide (I est vide).

Exemple 4.7 (i) Soit $V = K^n$. Tout vecteur colonne $x \in K^n$ est un combinaison linéaire des vecteurs colonnes $e_1,\dots,e_n$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{i=1}^{n} x_i e_i.$$

En particulier $K^n = \mathsf{span}\{e_1,\dots,e_n\}$.

(iii) On considère l'espace vectoriel des suites sur K et les suites

$$\begin{aligned} z_1 &= (1,0,0,0,0,\dots) \\ z_2 &= (0,1,0,0,0,\dots) \\ z_3 &= (0,0,1,0,0,\dots) \\ &\vdots \quad \vdots \end{aligned}$$

5. Au contraire d'un ensemble, une famille peut contenir des éléments repétés, c-à-d deux indices distincts dans I peuvent correspondre au même vecteur.

Une suite peut s'écrire comme une combinaison linéaire de $z_1, z_2, \ldots$ si et seulement si elle contient un nombre *fini* de coefficients non nuls. Par exemple, la suite $(1,1,1,\ldots)$ n'est pas une combinaison linéaire de $z_1, z_2, \ldots$.

Lemme 4.8 *Soit V un K-espace vectoriel est soit U, W deux sous-espaces vectoriels de V. Alors $U \cap W$ est un sous-espace vectoriel de V.*

DÉMONSTRATION. Exercices. ■

4.3 Indépendance linéaire, bases, dimensions

On a vu précédemment qu'une famille de vecteurs engendre un sous-espace vectoriel. Dans cette section on va dans la direction opposée : Étant donné un (sous-)espace vectoriel on cherche une famille de vecteurs, aussi petite que possible, qui l'engendre.

Définition 4.9 *Soit V un K-espace vectoriel et $v_1, \ldots, v_r \in V$. On dit que $v_1, \ldots, v_r$* ***engendrent*** *[generate] V ou que la famille $(v_1, \ldots, v_r)$ est une* ***famille génératrice*** *[generator] de V si $V = \mathsf{span}(v_1, \ldots, v_r)$. Plus généralement, on dit qu'une famille (eventuellement infinie) $(v_i)_{i \in I} \subset V$ est une* ***famille génératrice*** *de V si $V = \mathsf{span}(v_i)_{i \in I}$.*

Une famille génératrice n'est pas unique. Par exemple, les deux familles

$$(e_1, e_2, e_3) = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

et

$$(v_1, v_2, v_3, v_4) = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

engendrent $\mathbb{R}^3$. On trouve que tout vecteur s'écrit de façon unique comme une combinaison linéaire de la première famille. En particulier, on a

$$0 = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \alpha_3 = 0.$$

C'est different pour la deuxième famille. Par exemple,

$$0 = 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + 0 \cdot v_4 = 1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3 + (-1) \cdot v_4.$$

Définition 4.10 *Soit V un K-espace vectoriel. Une famille $(v_1, \ldots, v_r)$ de V est dite* ***linéairement indépendante*** *[linearly independent] ou* ***libre*** *si $\forall \alpha_1, \ldots, \alpha_r \in K$ l'équation (vectorielle)*

$$0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r,$$

n'admet que la solution triviale

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Une famille (eventuellement infinie) $(v_i)_{i \in I} \subset V$ est dite linéairement indépendante ou libre si toutes ses parties finies sont libres.

Si une famille $(v_1, \ldots, v_r)$ ne satisfait pas les exigences de la définition 4.10, il existe $\alpha_1, \ldots, \alpha_n \in K$, dont au moins un coefficient α_j est non nul, tels que $0 = \sum_{i=1}^{n} \alpha_i v_i$. Une famille qui n'est pas linéairement indépendante est dite **linéairement dépendante** ou **liée**.

Exemple 4.11 Soient $a_1, \ldots, a_r \in K^n$. En définissant $A = (a_1, \ldots, a_r) \in K^{n \times r}$, ces vecteurs sont linéairement indépendants si et seulement si

$$Ax = 0 \quad \Rightarrow \quad x = 0,$$

ainsi $S(A, 0) = \{0\}$. Selon la section 3.4, cette condition est équivalente à $\mathsf{rang}(A) = r$. ♦

Trivialement, une famille contenant le vecteur nul ne peut pas être linéairement indépendante. Tout comme une famille contenant deux fois le même vecteur n'est pas linéairement indépendante. Alors, on peut considérer une famille linéairement indépendante comme un ensemble non-ordonné.

Lemme 4.12 *Soit V un K-espace vectoriel et $(v_i)_{i\in I}$ une famille de vecteurs de V. Alors, les deux énoncés suivants sont équivalents :*

(i) *$(v_i)_{i\in I}$ est linéairement indépendante.*

(ii) *Tout vecteur $v \in \mathsf{span}(v_i)_{i\in I}$ s'écrit de façon unique comme une combinaison linéaire de $(v_i)_{i\in I}$.*

DÉMONSTRATION. (i)⇒(ii) : On considère deux combinaisons linéaires (finies)

$$v = \sum_{i\in I_1} \alpha_i v_i = \sum_{i\in I_2} \beta_i v_i.$$

Alors

$$0 = v - v = \sum_{i\in I_1} \alpha_i v_i - \sum_{i\in I_2} \beta_i v_i = \sum_{i\in I_1\cup I_2} (\alpha_i - \beta_i) v_i,$$

où $\alpha_i := 0$ si $i \in I_2 \setminus I_1$ et $\beta_i := 0$ si $i \in I_1 \setminus I_2$. Comme $(v_i)_{i\in I}$ est libre, on obtient $\alpha_i - \beta_i = 0$ pour tout i et, ainsi, les deux combinaisons sont les mêmes.
(ii)⇒(i) se montre en choisissant $v = 0$. ■

Le lemme suivant donne une variation de la définition d'indépendance linéaire. Cette variation est peut être plus intiutive mais, d'autre coté, elle est moins pratique.

Lemme 4.13 *Soit V un K-espace vectoriel. Alors une famille de vecteurs de V est linéairement dépendante si et seulement si au moins un vecteur de la famille est une combinaison linéaire des autres vecteurs de la famille.*

DÉMONSTRATION. Soit $\{v_i\}_{i\in I}$ une famille linéairement dépendante. Alors il existe une partie finie $I_0 \subset I$ et des coefficients $\alpha_i \in K$, $i \in I_0$, tels que $\sum_{i\in I_0} \alpha_i v_i = 0$, où $\alpha_j \neq 0$ pour au moins un indice $j \in I_0$. Ceci permet d'écrire

$$v_j = -\sum_{i\in I_0\setminus\{j\}} \frac{\alpha_i}{\alpha_j} v_i.$$

Réciproquement : Si un vecteur est une combinaison linéaire des autres, il existe $j \in I$, une partie finie $I_1 \subset I \setminus \{j\}$, et des coefficients $\beta_i \in K$ ($i \in I_1$) tels que $v_j = \sum_{i\in I_1} \beta_i v_i$. Alors $1 \cdot v_j - \sum_{i\in I_1} \beta_i v_i = 0$, c-à-d $\{v_i\}_{i\in I}$ est linéairement dépendante. ■

La définition suivante introduit le concept le plus fondamental des espaces vectoriels.

Définition 4.14 *Une famille $\mathscr{B} = (v_i)_{i\in I}$ d'un K-espace vectoriel V s'appelle une* ***base*** *[basis] de V si*

(i) *$\mathscr{B}$ est une famille génératrice de V, et*

(ii) *$\mathscr{B}$ est linéairement indépendante.*

Exemples des bases

Vecteurs colonnes. Soit e_i la i-ième colonne de I_n. Alors, $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ est une base de K^n. On dit que c'est la **base canonique** [*canonical basis*] de K^n.
Plus généralement les colonnes d'une matrice inversible quelconque forment une base de K^n.

Polynômes. Les monômes $1, t, t^2, \ldots, t^n$ forment une base de $K_n[t]$, le K-espace vectoriel de polynômes de degré $\leq n$.

La base la plus petite. Si V ne contient que le vecteur nul, $\mathscr{B} = \emptyset$ est la base (unique).

Construction des bases

Dans ce qui suit, on ne considère que des familles *finies*. Plus tard, dans la section 4.3.1, on traitera le cas infini.

Lemme 4.15 *Étant donné un K-espace vectoriel V, soient $v_1,\dots,v_r,w_1,\dots,w_s \in V$ tels que $v_1,\dots,v_r$ sont linéairement indépendants et $\mathsf{span}(v_1,\dots,v_r,w_1,\dots,w_s) = V$. Alors, on peut former une base de V en ajoutant certains vecteurs parmi $w_1,\dots,w_s$ à $v_1,\dots,v_r$.*

DÉMONSTRATION. Par récurrence sur s. Si $s = 0$, $(v_1,\dots,v_r)$ est déjà une base de V par hypothèse. On suppose que l'assertion est vraie pour $s-1 \geq 0$ et r quelconque. Alors on doit démontrer l'assertion pour s. Si $(v_1,\dots,v_r)$ est une base, la preuve est finie. Sinon on a $\mathsf{span}(v_1,\dots,v_r) \neq V$. Alors il existe $w_j \neq 0$, $1 \leq j \leq s$ tel que $w_j \notin \mathsf{span}(v_1,\dots,v_r)$. En particulier, l'équation

$$\sum_{i=1}^{r} \alpha_i v_i + \beta w_j = 0,$$

implique $\beta = 0$ et, par l'indépendance linéaire de $v_1,\dots,v_r$, $\alpha_1 = \cdots = \alpha_r = 0$. Alors, $(v_1,\dots,v_r,w_j)$ est linéairement indépendante. Par l'hypothèse de récurrence, on obtient une base en ajoutant certains vecteurs parmi $w_1,\dots,w_{j-1},w_{j+1},\dots,w_s$ (une famille de $s-1$ vecteurs) à $v_1,\dots,v_r,w_j$. Cela démontre l'assertion pour s. ■

Corollaire 4.16 *Soit $A_1 \in K^{m\times r}$ de rang r. Alors, il existe $A_2 \in K^{m\times(m-r)}$ telle que $A = \left(\begin{array}{c|c} A_1 & A_2 \end{array}\right)$ est inversible.*

DÉMONSTRATION. Exercice. ■

On va démontrer que toute base d'un espace vectoriel comporte le même nombre de vecteurs. Le lemme et le théoreme suivant joueront un rôle important dans la preuve.

Lemme 4.17 *Soit V un K-espace vectoriel et $w_1,\dots,w_n \in V$. Soit $v \in \mathsf{span}(w_1,\dots,w_n)$, on écrit $v = \sum_{i=1}^{n} \alpha_i w_i$. S'il existe un $k \in \{1,\dots,n\}$ tel que $\alpha_k \neq 0$, alors*

$$\mathsf{span}(w_1,\dots,w_n) = \mathsf{span}(w_1,\dots,w_{k-1},v,w_{k+1},\dots,w_n). \tag{4.2}$$

DÉMONSTRATION. Quitte à renuméroter les vecteurs w_i on peut supposer $k = 1$. On a alors

$$w_1 = \frac{1}{\alpha_1} v - \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} w_i$$

Soit $w \in \mathsf{span}(w_1,\dots,w_n)$, ainsi il existe $\beta_1,\dots,\beta_n \in K$ tels que

$$w = \sum_{i=1}^{n} \beta_i w_i = \frac{\beta_1}{\alpha_1} v + \sum_{i=2}^{n} \left(\beta_i - \frac{\beta_1 \alpha_i}{\alpha_1}\right) w_i.$$

Alors, $w \in \mathsf{span}(v,w_2,\dots,w_n)$. Comme w est arbitraire, on obtient

$$\mathsf{span}(w_1,\dots,w_n) \subseteq \mathsf{span}(v,w_2,\dots,w_n).$$

L'inclusion $\mathsf{span}(v,w_2,\dots,w_n) \subseteq \mathsf{span}(w_1,\dots,w_n)$ est trivialement vraie et le lemme est donc démontré. ■

Théorème 4.18 (Lemme de Steinitz) *Soit V un K-espace vectoriel et soit $v_1,\dots,v_m \in V$ une famille linéairement indépendante et $w_1,\dots,w_n \in V$. On suppose que*

$$\mathsf{span}(v_1,\dots,v_m) \subset \mathsf{span}(w_1,\dots,w_n). \tag{4.3}$$

Alors :

(i) $m \leq n$

(ii) on peut remplacer m vecteurs parmi $w_1,\dots,w_n$ par $v_1,\dots,v_m$ sans changer l'espace engendré.

Avant de démontrer le théorème 4.18, une explication de l'assertion du deuxième point : Il existe des indices $i_1,\dots,i_m \in \{1,\dots,n\}$ tels que l'on peut remplacer w_{i_1} par v_1, w_{i_2} par v_2, …, w_{i_m} par v_m, sans changer l'espace engendré par $w_1,\dots,w_n$. Quitte à renuméroter on peut supposer que $i_1 = 1, i_2 = 2, \dots, i_m = m$, donc

$$\mathsf{span}(v_1,\dots,v_m,w_{m+1},\dots,w_n) = \mathsf{span}(w_1,\dots,w_n). \tag{4.4}$$

DÉMONSTRATION. Grâce à (4.3) on peut écrire $v_1 = \alpha_1 w_1 + \cdots + \alpha_n w_n$. Comme $v_1 \neq 0$ (par l'indépendance de $v_1,\dots,v_n$), il existe $i_1 \in \{1,\dots,n\}$ tel que $\alpha_{i_1} \neq 0$. En utilisant le lemme 4.17 on obtient

$$\mathsf{span}(w_1,\dots,w_n) = \mathsf{span}(w_1,\dots,w_{i_1-1},v_1,w_{i_1+1},\dots,w_n).$$

Ou, quitte à renuméroter,

$$\mathsf{span}(w_1,\dots,w_n) = \mathsf{span}(v_1,w_2,\dots,w_n).$$

On répète ce procédé comme suit. Par récurrence on suppose que les vecteurs $w_1,\dots,w_r$, où $1 \leq r \leq m-1$, sont déjà remplacés par $v_1,\dots,v_r$:

$$\mathsf{span}(w_1,\dots,w_n) = \mathsf{span}(v_1,\dots,v_r,w_{r+1},\dots,w_n).$$

Évidemment $r \leq n$. Grâce à (4.3) il existe $\beta_1,\dots,\beta_n \in K$ tels que

$$v_{r+1} = \sum_{i=1}^{r} \beta_i v_i + \sum_{i=r+1}^{n} \beta_i w_i.$$

Comme $v_{r+1} \notin \mathsf{span}(v_1,\dots,v_r)$ (par l'indépendance de $v_1,\dots,v_n$, voir le lemme 4.13), il existe $i_{r+1} \in \{r+1,\dots,n\}$ tel que $\beta_{i_{r+1}} \neq 0$. En particulier, $r+1 \leq n$. Encore quitte à renuméroter on suppose que $i_{r+1} = r+1$. En utilisant le lemme 4.17 on obtient

$$\mathsf{span}(w_1,\dots,w_n) = \mathsf{span}(v_1,\dots,v_{r+1},w_{r+2},\dots,w_n).$$

En répétant ce procédé jusqu'à $r = m-1$, ça donne $m \leq n$ et (4.4). ∎

Corollaire 4.19 *Soit V un K-espace vectoriel. Alors :*

(i) si V a une base finie, tout autre base est aussi finie,

(ii) deux bases finies de V ont le même nombre d'éléments.

DÉMONSTRATION. (i). Soit $v_1,\dots,v_n$ une base (finie) de V. On suppose qu'il existe une base infinie de V. Alors il existe $n+1$ vecteurs $w_1,\dots,w_{n+1} \in V$ qui sont linéairement indépendants. Mais alors

$$\mathsf{span}(w_1,\dots,w_{n+1}) \subseteq \mathsf{span}(v_1,\dots,v_n) = V,$$

ce qui contredit le théorème 4.18.

(ii). Soient $\{v_1,\dots,v_m\}$ et $\{w_1,\dots,w_n\}$ deux bases de V. Par le théorème 4.18, on a $m \leq n$ et (en echangeant les rôles de v et w) $n \leq m$. Donc $m = n$. ∎

Du théorème 4.18 découle également l'assertion suivante : si V a une base infinie, V ne peut pas avoir de bas finie. Par exemple, l'espace vectoriel des suites n'admet pas de base finie.

Le résultat du corollaire 4.19 permet la définition suivante.

Définition 4.20 *Soit V un K-espace vectoriel. On définit la* ***dimension*** *de V par*

$$\dim(V) := \begin{cases} n, & \text{si } V \text{ a une base de } n < \infty \text{ vecteurs}, \\ \infty, & \text{sinon.} \end{cases}$$

Si $\dim(V) < \infty$ on dit que V est de dimension finie, sinon on dit que V est de dimension infinie.

Exemples :

- K^n est un K-espace vectoriel de dimension n.
- $K_n[t]$ est un K-espace vectoriel de dimension $n+1$.
- $\{0\}$ est un K-espace vectoriel de dimension 0.
- Dans l'espace vectoriel des suites on trouve une infinité de vecteurs linéairement indépendants : $(1,0,0,\ldots)$, $(0,1,0,\ldots)$, …. Alors, il est de dimension infinie.
- L'espace vectoriel $M_{m\times n}(K)$ (muni de l'addition matricielle et de la multiplication par un scalaire) est de dimension mn. Si $m = n$, l'ensemble des matrices symétriques est un sous-espace vectoriel de dimension $n(n+1)/2$. Démonstration: Voir exercices.

Finalement, les deux résultats suivants apportent une compréhension additionnelle de la notion de dimension.

Lemme 4.21 *Soit V un K-espace vectoriel de dimension $n < \infty$ et soit $(v_1,\ldots,v_p)$ une famille de V. Si cette famille est libre alors $p \leq n$.*

DÉMONSTRATION. Si $p > n$, par le lemme 4.15, on peut ajouter $(v_1,\ldots,v_p)$ à une base de plus de n vecteurs. Mais, cela contredit le résultat du corollaire 4.19. ■

Lemme 4.22 *Soit V un K-espace vectoriel de dimension finie et $W \subseteq V$ un sous-espace vectoriel de V. Alors :*

(i) $\dim(W) \leq \dim(V)$,

(ii) si $\dim(W) = \dim(V)$ *alors* $W = V$.

DÉMONSTRATION. (i). Par le lemme 4.21, W est de dimension finie car toute famille libre de W est aussi une famille libre de V. Soit $\{w_1,\ldots,w_r\}$ une base de W. Par le lemme de Steinitz (Théorème 4.18) on peut construire une base de V en ajoutant des vecteurs à $\{w_1,\ldots,w_r\}$. Alors $\dim(W) \leq \dim(V)$.

(ii). Soit $n = \dim(W) = \dim(V)$ et $w_1,\ldots,w_n$ une base de W. On suppose que $V \neq W$. Alors il existe $v \in V$ tel que $v \notin \mathsf{span}(w_1,\ldots,w_n)$. En particulier, $(w_1,\ldots,w_n,v)$ est une famille libre. Mais cela contredit le résultat du lemme 4.21. ■

4.3.1 Le cas de dimension infinie

La procédé du lemme de Steinitz est restreint à la dimension finie. Dans le cas de dimension infinie la construction d'une base n'est pas facile du tout. Par exemple, on considère l'espace vectoriel des suites : les suites $(1,0,0,\ldots)$, $(0,1,0,\ldots)$, … sont linéairement indépendantes mais elles ne forment pas une base ! Dans le cas infini la démonstration de l'existence d'une base n'est pas constructive et elle recourt aux concepts de la théorie des ensembles.

Définition 4.23 *Soit E un ensemble. Une* ***relation d'ordre*** *[partial order] sur E est une relation binaire R satisfaisant :*

(i) xRx *pour tout* $x \in E$,

(ii) $(xRy$ *et* $yRx) \Rightarrow x = y$,

(iii) $(xRy \text{ et } yRz) \Rightarrow xRz$.

Exemples typiques :

- la relation $\leq$ sur $\mathbb{R}$,
- la relation $\subseteq$ sur $P(M)$, l'ensemble des parties d'un ensemble M.

Définition 4.24 *Soit $\leq$ une relation d'ordre sur X. Un ensemble non vide $A \subseteq X$ est* ***totalement ordonné*** *[totally ordered] si deux éléments quelconques $x,y \in A$ sont toujours comparables, c-à-d on a toujours $x \leq y$ ou $y \leq x$.*

L'ensemble $\mathbb{R}$ muni de la relation d'ordre habituelle est totalement ordonné. Au contraire $P(M)$ muni de la relation d'ordre $\subset$ n'est pas totalement ordonné. Mais un sous-ensemble $A = \{E_1, E_2, E_3, \ldots\} \subset P(E)$, où $E_1 \subset E_2 \subset E_3 \subset \cdots$, est totalement ordonné.

Définition 4.25 *Soit $\leq$ une relation d'ordre sur X et $A \subseteq X$ avec $A \neq \emptyset$. Alors*

- *$s(A) \in X$ est un* ***majorant*** *de A si $a \leq s(A)$ pour tout $a \in A$.*
- *$m(A) \in A$ est un* ***élement maximal*** *de A, si l'implication $m(A) \leq a \Rightarrow m(A) = a$ et vraie pour tout $a \in A$.*
- *X est appelé* ***ensemble inductif*** *si tout sous-ensemble totalement ordonné possède un majorant.*

Lemme 4.26 (Lemme de Zorn ou lemme de Kuratowski-Zorn) *Tout ensemble inductif admet (au moins) un élément maximal.*

La preuve du lemme de Zorn utilise l'axiome du choix. En fait, l'axiome du choix et le lemme de Zorn sont équivalents.

Enfin, soit V un K-espace vectoriel ! On considère la collection de tous les ensembles linéairement indépendants :

$$X := \{E \subseteq V : \text{ les éléments de } E \text{ sont linéairement indépendants}\}. \tag{4.5}$$

Comme $\{\emptyset\} \in X$, l'ensemble X est non vide. Le lemme suivant montre que l'on peut appliquer le lemme de Zorn pour X.

Lemma 4.27 *L'ensemble X défini par* (4.5) *est inductif par rapport à la relation d'ordre $\subseteq$.*

DÉMONSTRATION. Soit $A \subset X$ un ensemble totalement ordonné dans lequel tout élément est un sous-ensemble (de V) qui est linéairement indépendant. En définissant

$$\overline{A} = \bigcup_{E \in A} E,$$

il est évident que $\overline{A}$ est un majorant de $P(A)$. Il reste à montrer que $\overline{A} \in X$, c-à-d que les éléments de $\overline{A}$ sont linéairement indépendants.

À cette fin on choisit une famille quelconque de vecteurs

$$v_1, \ldots, v_n \in \overline{A}.$$

Par récurrence sur n on montre qu'il existe un ensemble $E_n \in A$ tel que $v_1, \ldots, v_n \in E_n$. Pour $n = 1$ ceci découle de la définition de $\overline{A}$. Supposons l'assertion vraie pour $n-1$, alors il existe $E_{n-1} \in A$ tel que $v_1, \ldots, v_{n-1} \in E_{n-1}$. Par la définition de $\overline{A}$, il existe $E' \in A$ tel que $v_n \in E'$. Comme A est totalement ordonné on a $E_{n-1} \subseteq E'$ ou $E' \subseteq E_{n-1}$. Dans le premiér cas l'assertion devient vraie pour n en posant $E_n = E'$, dans le deuxième en

posant $E_n = E_{n-1}$. Comme E_n est linéairement indépendant, la famille $v_1,\dots,v_n$ est aussi linéairement indépendante. Alors, $\overline{A} \in X$ est un majorant de A. ■

L'application du lemme de Zorn sur X établit l'existence d'un *sous-ensemble linéairement indépendant maximal* de V. Le théoreme suivant montre que c'est en fait une base. On dit qu'un ensemble générateur $\mathscr{B} \subset V$ est **minimal** s'il n'existe pas un ensemble générateur A tel que $A \subsetneq \mathscr{B}$.

Théorème 4.28 *Soit V un K-espace vectoriel et $\mathscr{B} \subset V$. Alors, les énoncés suivants sont équivalents :*

(i) *$\mathscr{B}$ est une base.*

(ii) *$\mathscr{B}$ est un ensemble générateur minimal.*

(iii) *$\mathscr{B}$ est linéairement indépendant maximal.*

DÉMONSTRATION. $(i) \Rightarrow (ii)$. Soit $\mathscr{B}$ une base et $A \subsetneq \mathscr{B}$. Par l'indépendance linéaire de $\mathscr{B}$ un vecteur $v \in \mathscr{B} \setminus A$ ne peut pas s'écrire comme une combinaison linéaire d'éléments de A (voir le lemme 4.13). En particulier, A n'engendre pas V et ainsi $\mathscr{B}$ est un ensemble générateur minimal.

$(ii) \Rightarrow (iii)$. Soit $\mathscr{B}$ un ensemble générateur minimal. Supposons que $\mathscr{B}$ soit linéairement dépendant. Par le lemme 4.13 il existe $v \in \mathscr{B}$ qui peut s'écrire comme une combinaison linéaire d'éléments de $\mathscr{B} \setminus \{v\}$. Alors $\mathscr{B} \setminus \{v\}$ engendre V, ce qui contredit la minimalité de $\mathscr{B}$. Ainsi $\mathscr{B}$ est linéairement indépendant. Il est aussi maximal : Soit $\mathscr{B}' \supseteq \mathscr{B}$ linéairement indépendant. Comme $\mathscr{B}$ est un ensemble générateur, on peut écrire tout élément de V comme une combinaison linéaire de $\mathscr{B}$. Alors $\mathscr{B}' = \mathscr{B}$.

$(iii) \Rightarrow (i)$. Soit $\mathscr{B}$ linéairement indépendant maximal. Il reste à montrer que V est engendré par $\mathscr{B}$. Trivialement, on a $v \in \mathsf{span}(\mathscr{B})$ pour tout $v \in \mathscr{B}$. Ainsi, soit $v \notin \mathscr{B}$. Par la maximalité de $\mathscr{B}$, l'ensemble $\mathscr{B} \cup \{v\}$ ne peut pas être linéairement indépendant. Alors il existe une combinaison linéaire

$$\alpha v + \alpha_1 v_1 + \cdots + \alpha_n v_n = 0, \quad v_1,\dots,v_n \in \mathscr{B},$$

où au moins un des coefficients $\alpha,\alpha_1,\dots,\alpha_n \in K$ est non nul. Si $\alpha = 0$, alors $v_1,\dots,v_n$ sont linéairement depéndants ce qui contredit l'indépendance de $\mathscr{B}$. Alors on a $\alpha \neq 0$ et ainsi le vecteur v est une combinaison linéaire des éléments de $\mathscr{B}$:

$$v = -\sum_{i=1}^{n} \frac{\alpha_i}{\alpha} v_i.$$

Cela montre que $\mathscr{B}$ engendre V. ■

Corollaire 4.29 *Tout espace vectoriel possède une base.*

4.4 Sommes d'espaces vectoriels

Définition 4.30 *Soient $U_1,\dots,U_s$ des sous-espaces vectoriels d'un K-espace vectoriel V. Leur **somme** est définie par*

$$U_1 + \cdots + U_s := \{u_1 + \cdots + u_s : u_1 \in U_1,\dots,u_s \in U_s\}.$$

Exemple 4.31 Soit $V = \mathbb{R}^3$ et $U = \mathsf{span}(u)$, $W = \mathsf{span}(w)$, où $u \neq 0$, $w \neq 0$. Les sous-espaces U,W correspondent aux droites dans $\mathbb{R}^3$. Si u et w sont linéairement indépendants (c-á-d il n'existe pas un scalaire λ tel que $u = \lambda w$), la somme $U + W$ correspond à un plan. ♦

Lemme 4.32 *Soient $U_1,\dots,U_s$ des sous-espaces vectoriels d'un K-espace vectoriel V. Alors*

(i) $U_1+\cdots+U_s$ *est encore un sous-espace vectoriel de* V,

(ii) $U_1+\cdots+U_s=\mathsf{span}(U_1\cup\cdots\cup U_s)$,

(iii) $\dim(U_1+\cdots+U_s)\le\dim(U_1)+\cdots+\dim(U_s)$.

DÉMONSTRATION. Exercice. ∎

Voici un exemple simple montrant que l'inégalité du lemme 4.32 (iii) peut être stricte : La somme d'un espace vectoriel $U\neq\{0\}$ avec lui-même donne $U+U=U$, alors $\dim(U)=\dim(U+U)<\dim(U)+\dim(U)$. Le théorème suivant explique la différence entre les deux côtés de l'inégalité pour $s=2$. On rappelle que l'intersection de deux sous-espace vectoriels est encore un sous-espace vectoriel, voir le lemme 4.8.

Théorème 4.33 (Formule de Grassmann) *Soient U,W deux sous-espaces vectoriels (de dimension finie) d'un K-espace vectoriel V. Alors*

$$\dim(U+W)=\dim(U)+\dim(W)-\dim(U\cap W). \tag{4.6}$$

DÉMONSTRATION. Si $U\cap W=\{0\}$ on prend une base $(u_1,\dots,u_m)$ de U et une base $(w_1,\dots,w_n)$ de W. Alors, $(u_1,\dots,u_m,w_1,\dots,w_n)$ est une base de $U+W$ et la formule (4.6) est vraie.

Supposons $U\cap W\neq\{0\}$, soit $v_1,\dots,v_r$ une base de $U\cap W$. Par le lemme 4.15 on peut la compléter en une base $v_1,\dots,v_r,u_1,\dots,u_{\tilde m}$ de U et en une base $v_1,\dots,v_r,w_1,\dots,w_{\tilde n}$ de W. Pour montrer (4.6), on va montrer que

$$v_1,\dots,v_r,u_1,\dots,u_{\tilde m},w_1,\dots,w_{\tilde n} \tag{4.7}$$

est une base de $U+W$. Par la construction,

$$\mathsf{span}(v_1,\dots,v_r,u_1,\dots,u_{\tilde m},w_1,\dots,w_{\tilde n})=\mathsf{span}(U\cup W)=U+W,$$

et ainsi il reste à montrer leur indépendance linéaire. Soient $\alpha_1,\dots,\alpha_r\in K$, $\beta_1,\dots,\beta_{\tilde m}\in K$, $\gamma_1,\dots,\gamma_{\tilde n}\in K$, tels que

$$0=\sum_{i=1}^{r}\alpha_i v_i+\sum_{i=1}^{\tilde m}\beta_i u_i+\sum_{i=1}^{\tilde n}\gamma_i w_i, \tag{4.8}$$

ainsi

$$v:=\sum_{i=1}^{r}\alpha_i v_i+\sum_{i=1}^{\tilde m}\beta_i u_i=-\sum_{i=1}^{\tilde n}\gamma_i w_i. \tag{4.9}$$

La première équation donne que $v\in U$ et la seconde que $v\in W$, alors $v\in U\cap W$. En particulier on peut écrire v comme une combinaison linéaire de $v_1,\dots,v_r$. Mais (4.9) est aussi une combinaison linéaire et, par l'indépendance de $v_1,\dots,v_r,u_1,\dots,u_{\tilde m}$, on obtient que $\beta_1=\cdots=\beta_{\tilde m}=0$. En les substituant dans (4.8) on obtient, par l'indépendance de $v_1,\dots,v_r,w_1,\dots,w_{\tilde n}$, que $\alpha_1=\cdots=\alpha_r=\gamma_1=\cdots=\gamma_{\tilde n}=0$. Alors, (4.7) est une base de $U+W$. ∎

Le terme de correction $\dim(U\cap W)$ dans (4.6) disparait si et seulement si $U\cap W=\{0\}$. Le lemme suivant donne une autre condition équivalente.

Lemme 4.34 *Soient U,W deux sous-espace vectoriels (de dimension finie) d'un K-espace vectoriel V tels que $U+W=V$. Alors $U\cap W=\{0\}$ si et seulement si tout $v\in V$ s'écrit de façon unique $v=u+w$ avec $u\in U$, $w\in W$.*

DÉMONSTRATION. Soit $U \cap W = \{0\}$. Supposons qu'un vecteur $v \in V$ admette deux décompositions $v = u + w = u' + w'$, où $u,u' \in U$ et $w,w' \in W$. Alors

$$\underbrace{u - u'}_{\in U} = \underbrace{w' - w}_{\in W} \in U \cap W = \{0\},$$

ce qui donne $u = u'$ et $w = w'$.

Réciproquement, soit $u \in U \cap W$. Alors $0 = 0+0$ et $0 = u + (-u)$ sont deux décompositions de 0. Comme la décomposition est unique on obtient $u = 0$. ■

Définition 4.35 *Soit V un K-espace vectoriel et U_1, U_2 deux sous-espaces vectoriels de V. On dit que V est la* ***somme directe*** *[direct sum] de U_1 et U_2, noté $V = U_1 \oplus U_2$, si $V = U_1 + U_2$ et $U_1 \cap U_2 = \{0\}$.*

Tout sous-espace peut se compléter en tout l'espace par une somme directe (en dimension finie).

Théorème 4.36 *Soit V un K-espace vectoriel de dimension finie et U un sous-espace vectoriel de V. Alors il existe un sous-espace vectoriel U' tel que $V = U \oplus U'$.*

DÉMONSTRATION. Soit $v_1,\dots,v_r$ une base U. Par le lemme 4.15 il existe $v_{r+1},\dots,v_n \in V$ tels que $v_1,\dots,v_r,v_{r+1},\dots,v_n$ est une base de V. En définissant $U' = \mathsf{span}(v_{r+1},\dots,v_n)$ on obtient $V = U + U'$ et $U \cap U' = \{0\}$. ■

On conclut ce chapitre par l'extension à plus de deux sous-espaces.

Définition 4.37 *Soient $U_1,\dots,U_r$ des sous-espaces vectoriels d'un K-espace vectoriel V. On dit que V est la* ***somme directe*** *de $U_1,\dots,U_r$ si $V = U_1 + \cdots + U_r$ et*

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r) = \{0\} \tag{4.10}$$

pour $i = 1,\dots,r$.

On ne peut pas remplacer la condition (4.37) par la condition $U_i \cap U_j = \{0\}$ pour $i,j = 1,\dots,r$. Par exemple, $V = \mathbb{R}^2$ n'est pas une somme directe des sous-espaces

$$U_1 = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in K \right\}, \quad U_2 = \left\{ \begin{pmatrix} 0 \\ \alpha \end{pmatrix} : \alpha \in K \right\}, \quad U_3 = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} : \alpha \in K \right\}.$$

Théorème 4.38 *Soient $U_1,\dots,U_r$ des sous-espaces vectoriels d'un K-espace vectoriel V. Alors les énoncés suivants sont équivalents :*

(i) $V = U_1 \oplus \cdots \oplus U_r$.

(ii) *Tout $v \in V$ s'écrit de façon unique $v = u_1 + \cdots + u_r$ avec $u_i \in U_i$, $i = 1,\dots,r$.*

(iii) *Si $v_{i,1},\dots,v_{i,n_i}$ sont des bases de U_i, $i = 1,\dots,r$, alors*

$$\mathscr{B} = \big(v_{1,1},\dots,v_{1,n_1},\, v_{2,1},\dots,v_{2,n_2},\quad \dots,\quad v_{r,1},\dots,v_{r,n_r}\big)$$

est une base de V.

(iv) $V = U_1 + \cdots + U_r$ *et* $\dim(V) = \dim(U_1) + \cdots + \dim(U_r)$.

DÉMONSTRATION. (i) $\Rightarrow$ (ii). Supposons que $V = U_1 \oplus \cdots \oplus U_r$ et que $v \in V$ s'écrive

$$v = u_1 + \cdots + u_r = w_1 + \cdots + w_r, \qquad u_i,w_i \in U_i,\ i = 1,\dots,r.$$

Alors

$$\underbrace{-(u_i - w_i)}_{\in U_i} = \underbrace{u_1 - w_1 + \cdots + u_{i-1} - w_{i-1} + u_{i+1} - w_{i+1} + \cdots + u_r - w_r}_{\in (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r)}.$$

Par la définition (4.10) on obtient $u_i - w_i = 0$ et ainsi $u_i = w_i$ pour $i = 1,\dots,r$.

(ii) $\Rightarrow$ (iii). Tout $v \in V$ s'écrit comme une somme d'éléments de U_i, alors la famille $\mathscr{B}$ engendre V. Pour montrer l'indépendance linéaire, soit

$$0 = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \alpha_{ij} v_{ij}, \qquad \alpha_{ij} \in K.$$

En définissant $w_i := \sum_{j=1}^{n_i} \alpha_{ij} v_{ij} \in U_i$, l'unicité de la décomposition $0 = 0 + \cdots + 0$ donne $w_i = 0$. Par l'indépendance linéaire de $v_{i,1},\dots,v_{i,n_i}$, on obtient que tous les coefficients α_{ij} sont nuls.

(iii) $\Rightarrow$ (i). Soit $v \in U_k \cap (U_1 + \cdots + U_{k-1} + U_{k+1} + \cdots + U_r)$. Alors il existe $\alpha_{ij} \in K$ tels que

$$v = \sum_{j=1}^{n_k} \alpha_{kj} v_{kj} = \sum_{\substack{i=1 \\ i \neq k}}^{r} \sum_{j=1}^{n_i} \alpha_{ij} v_{ij}.$$

Par le lemme 4.12, v s'écrit de façon unique comme une combinaison linéaire des éléments de $\mathscr{B}$. Ceci montre que $\alpha_{ij} = 0$ et ainsi $v = 0$.

(iii) $\Leftrightarrow$ (iv) suit directement des définitions. ∎

Chapitre 5
Applications linéaires

Dans ce chapitre on va considérer des applications entre deux espaces vectoriels V, W, en particulier celles qui sont compatibles avec la structure de l'espace vectoriel.

5.1 Définitions et premières propriétés

Définition 5.1 *Soient V,W deux K-espaces vectoriels. Une* ***application linéaire*** *[linear map] $F : V \to W$ est une application satisfaisant les deux conditions suivantes :*

(i) $F(v_1 + v_2) = F(v_1) + F(v_2)$ *pour tous* $v_1, v_2 \in V$,

(ii) $F(\alpha v) = \alpha F(v)$ *pour tous* $\alpha \in K, v \in V$.

Remarques:

- En posant $v_2 = 0$, le point (ii) de la définition 5.1 implique que $F(0) = 0$.
- On a une définition équivalente en demandant que

$$F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 F(v_1) + \alpha_2 F(v_2) \tag{5.1}$$

  pour tous $\alpha \in K$, $v_1, v_2 \in V$. Quelle définition, la 5.1 (i)+(ii) ou la (5.1), doit-on favoriser ? C'est une question de goût.
- Plus généralement, si $f : V \to W$ est une application linéaire, alors

$$F(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 F(v_1) + \cdots + \alpha_n F(v_n) \tag{5.2}$$

  pour tous $n \geq 1$, tous $\alpha_1, \ldots, \alpha_n \in K$ et tous $v_1, \ldots, v_n \in V$. La preuve de (5.2) procède par récurrence : le cas $n = 1$ est la définition 5.1 (ii). Supposons que la relation (5.2) est vraie pour $n \geq 1$. Alors,

$$\begin{aligned} & F(\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha_{n+1} v_{n+1}) \\ =& F(\alpha_1 v_1 + \cdots + \alpha_n v_n) + \alpha_{n+1} F(v_{n+1}) && \text{(par définition 5.1)} \\ =& \alpha_1 F(v_1) + \cdots + \alpha_n F(v_n) + \alpha_{n+1} F(v_{n+1}). && \text{(par hypothèse de récurrence)} \end{aligned}$$

  Ceci démontre (5.2) pour $n + 1$.

Exemples des applications linéaires

Deux exemples banals : L'application nulle $F : V \to W$, $F : v \mapsto 0$, et l'application identité $F : V \to V$, $F : v \mapsto v$, sont toujours linéaires.

Dans tous les exemples suivants on suppose que $K = \mathbb{R}$.

Fonctions linéaires. Soit $V = W = \mathbb{R}$. Alors, la fonction $g(x) = \beta x$, où $\beta \in \mathbb{R}$, est une application linéaire, car

$$g(\alpha_1 x_1 + \alpha_2 x_2) = \beta\alpha_1 x_1 + \beta\alpha_2 x_2 = \alpha_1 g(x_1) + \alpha_2 g(x_2).$$

On observe qu'une fonction $\tilde{g}(x) = \beta x + \gamma$ n'est pas une application linéaire si $\gamma \neq 0$.

Produits matrice-vecteur. Soient $V = K^n$, $W = K^m$ est soit $A \in M_{m\times n}(K)$. Le produit matrice vecteur

$$F_A : K^n \to K^m, \quad F_A(x) = Ax,$$

est une application linéaire. En effect, on a vu dans le chapitre 1 que

$$F_A(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha F_A(x) + \beta F_A(y).$$

On va voir plus tard que la réciproque est vraie aussi : Si V, W sont de dimension finie, alors on peut regarder toute application linéaire comme un produit matrice-vecteur.

Intégration. Soient $a,b \in \mathbb{R}$, $a < b$, et soit

$$C([a,b]) = \{g : [a,b] \to \mathbb{R} \,|\, g \text{ est continue sur } [a,b]\}.$$

Alors, les deux applications

$$\ell : C([a,b]) \to \mathbb{R}, \qquad \ell(g) := \int_a^b g(t)\,\mathrm{d}t.$$

et

$$\Psi : C([a,b]) \to C([a,b]), \qquad [\Psi(g)](x) := \int_a^x g(t)\,\mathrm{d}t.$$

sont linéaires.

Dérivation. Soit

$$C^1([a,b]) = \{g : [a,b] \to \mathbb{R} \,|\, g \text{ est continûment dérivable sur } [a,b]\}.$$

Alors, la dérivation

$$D : C^1([a,b]) \to C([a,b]), \qquad [D(f)](x) := f'(x)$$

est une application linéaire.

Opérateur de décalage. Soit V l'espace vectoriel des suites réelles. Alors, **l'opérateur de décalage** [*shift operator*]

$$\Sigma : V \to V, \qquad \Sigma(v_0,v_1,v_2,\ldots) := (v_1,v_2,v_3,\ldots)$$

est une application linéaire.

On note $L(V,W)$ l'ensemble des applications linéaires de V vers W.

Définition 5.2 *Soient V,W deux K-espaces vectoriels.*

*(i) Une application linéaire $F \in L(V,W)$ qui est bijective s'appelle un **isomorphisme (d'espace vectoriel)** [(vector space) isomporphism]. S'il existe un isomorphisme entre deux K-espace vectoriel V,W on dit que V et W sont **isomorphes** et on écrit*

$$V \cong W.$$

(ii) Si $V = W$, une application linéaire $F \in L(V,V)$ est appelée un ***endomorphisme****. Si de plus F est bijective on dit que F est un* ***automorphisme****.*

Soit $F_A : K^n \to K^m$, $F_A : x \mapsto Ax$ pour une matrice $A \in M_{m\times n}(K)$. Alors, les résultats de la section 3.4 donnent

$$F_A \text{ isomorphisme} \quad \Leftrightarrow \quad m = n \text{ et } A \text{ inversible} \quad \Leftrightarrow \quad F_A \text{ automorphisme.}$$

Définition 5.3 *Soient V, W deux K-espaces vectoriels et $F \in L(V,W)$. Le* ***noyau*** *[null space, kernel] et* ***l'image*** *de f sont définis comme suit :*

$$\mathrm{Ker}(F) := \{v \in V : F(v) = 0\}, \qquad \mathrm{Im}(F) := \{F(v) : v \in V\}.$$

Plus généralement, pour un sous-ensemble $\tilde{V} \subset V$, on note

$$F(\tilde{V}) := \{F(v) : v \in \tilde{V}\},$$

l'image de V par F. En particulier, $F(V) = \mathrm{Im}(F)$. Pour un sous-ensemble $\tilde{W} \subset W$, on note

$$F^{-1}(\tilde{W}) := \{v \in V : F(v) \in \tilde{W}\},$$

la pré-image [*pre-image*] par F de $\tilde{W}$. Cette notation est utilisée indépendamment du fait qu'il existe une réciproque à F ou non. Si $\tilde{W}$ ne contient qu'un élément w, on peut omettre les accolades : $F^{-1}(w) = F^{-1}(\{w\})$. En particulier, $F^{-1}(0) = \mathrm{Ker}(F)$.

Le lemme suivant rassemble quelques propriétés des sous-espaces vectoriels et des applications linéaires.

Lemme 5.4 *Soient V,W deux K-espaces vectoriels et $F \in L(V,W)$. Alors :*

(i) Si $\tilde{V}$ est un sous-espace vectoriel de V, alors $F(\tilde{V})$ est un sous-espace vectoriel de W.

(ii) Si $\tilde{W}$ est un sous-espace vectoriel de W, alors $F^{-1}(\tilde{W})$ est un sous-espace vectoriel de V.

(iii) Si $v_1,\ldots,v_n \in V$ sont linéairement dépendants, alors $F(v_1),\ldots,F(v_n) \in W$ sont aussi linéairement dépendants.

(iv) Si $v_1,\ldots,v_n \in V$ sont linéairement indépendants et F est injective, alors $F(v_1),\ldots,F(v_n) \in W$ sont aussi linéairement indépendants.

(v) $\mathrm{Ker}(F) = \{0\}$ *si et seulement si F est injective.*

(vi) $\mathrm{Im}(F) = W$ *si et seulement si F est surjective.*

(vii) Si F est un isomporphisme, alors $F^{-1} \in L(W,V)$.

DÉMONSTRATION. (i). Comme $0 \in \tilde{V} \;\Rightarrow\; F(0) \in F(\tilde{V})$, on a que l'ensemble $F(\tilde{V})$ est non vide. Soient $F(v_1),F(v_2) \in F(\tilde{V})$. Comme $\tilde{V}$ est un sous-espace vectoriel, on obtient

$$\alpha_1 F(v_1) + \alpha_2 F(v_2) = F(\underbrace{\alpha_1 v_1 + \alpha_2 v_2}_{\in \tilde{V}}) \in F(\tilde{V})$$

et, ainsi, $F(\tilde{V})$ est un sous-espace vectoriel.

(ii) Comme $F(0) = 0$ on a toujours $0 \in F^{-1}(\tilde{W})$ et, ainsi, l'ensemble $F^{-1}(\tilde{W})$ est non vide. Soient $v_1,v_2 \in F^{-1}(\tilde{W})$, c-à-d $F(v_1),F(v_2) \in \tilde{W}$. Comme $\tilde{W}$ est un sous-espace vectoriel, on obtient

$$F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 F(v_1) + \alpha_2 F(v_2) \in \tilde{W}$$

et, ainsi, $F^{-1}(\tilde{W})$ est un sous-espace vectoriel.

(iii) Si $v_1,\dots,v_n$ sont linéairement dépendants alors il existe $\alpha_1,\dots,\alpha_n \in K$ tels que $\alpha_j \neq 0$ pour certain j et $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Par (5.2) :

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \quad \Rightarrow \quad \alpha_1 F(v_1) + \cdots + \alpha_n F(v_n) = 0.$$

Alors $F(v_1),\dots,F(v_n)$ sont linéairement dépendants.

(iv) Soient $\alpha_1,\dots,\alpha_n \in K$ tels que $\alpha_1 F(v_1) + \cdots + \alpha_n F(v_n) = 0$. Par (5.2), on a $F(\alpha_1 v_1 + \cdots + \alpha_n v_n) = 0$. Mais, $F(0) = 0$ et F est injective, alors $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Comme $v_1,\dots,v_n$ sont linéairement indépendants, on obtient $\alpha_i = 0$ pour $i = 1,\dots,n$.

(v). Par définition, l'injectivité de F implique que $\mathrm{Ker}(F) = \{0\}$. Réciproquement, on suppose que $\mathrm{Ker}(F) = \{0\}$. Soient $v_1, v_2 \in V$ tels que $F(v_1) = F(v_2)$. Alors $0 = F(v_1) - F(v_2) = F(v_1 - v_2)$. Alors, on a $v_1 - v_2 = 0$, donc $v_1 = v_2$ et, ainsi, F est injective.

(vi) découle directement de la définition de surjectivité.

(vii) Comme F est bijective, on peut définir l'application $F^{-1} : W \to V$. Il reste à montrer que F^{-1} est linéaire. À cette fin, soient $w_1, w_2 \in W$. Alors, il existe $v_1, v_2 \in V$ tels que $w_1 = F(v_1)$, $w_2 = F(v_2)$. En utilisant que F est linéaire on obtient

$$\begin{aligned} F^{-1}(\alpha_1 w_1 + \alpha_2 w_2) &= F^{-1}(\alpha_1 F(v_1) + \alpha_2 F(v_2)) = F^{-1}(F(\alpha_1 v_1 + \alpha_2 v_2) \\ &= \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 F^{-1}(w_1) + \alpha_2 F^{-1}(v_2) \end{aligned}$$

et, ainsi, F^{-1} est linéaire. ■

Corollaire 5.5 *Soient V,W deux K-espaces vectoriels et $F \in L(V,W)$. Alors,* $\mathrm{Ker}(F)$ *et* $\mathrm{Im}(F)$ *sont des sous-espaces vectoriels de V et W respectivement.*

Corollaire 5.6 *Soient V,W deux K-espaces vectoriels de dimension finie. Supposons qu'il existe un isomorphisme $F : V \to W$. Alors,* $\dim V = \dim W$.

DÉMONSTRATION. Soit $v_1,\dots,v_n$ une base de V, alors $F(v_1),\dots,F(v_n)$ sont libres par le lemme 5.4 (iv). Soit $w \in W$. En posant $v := F^{-1}(w)$ il existe $\alpha_1,\dots,\alpha_n \in K$ tels que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ et donc $w = \alpha_1 F(v_1) + \cdots + \alpha_n F(v_n)$. Ainsi $F(v_1),\dots,F(v_n)$ est une base de W. ■

5.1.1 Le théorème du rang

Dans cette section on va établir une relation importante entre la dimension du noyau et celle d'image d'une application linéaire $F : V \to W$, lorsque V est de dimension finie.

Définition 5.7 *Soient V, W deux K-espaces vectoriels et $F \in L(V,W)$. On dénote par* ***le rang de F****, noté* $\mathsf{rang}(F)$*, la dimension de l'espace vectoriel* $\mathrm{Im}(F)$.

Lemme 5.8 *Soit $A \in M_{m\times n}(K)$. On considère l'application linéaire $F_A : K^n \to K^m$, $F_A : x \mapsto Ax$. Alors,*

$$\mathsf{rang}(F_A) = \mathsf{rang}(A),$$

où $\mathsf{rang}(A)$ *est le rang de la matrice A comme défini dans le chapitre 4.*

DÉMONSTRATION. Soit $Q \in M_{m\times m}(K)$ une matrice inversible telle que $C = QA$ est sous la forme échelonnée reduite (voir la définition 3.11). Comme $\mathrm{Im}(F_C)$ est constitué de toutes les combinaisons linéaires des colonnes de C, on voit directement que

$$\mathrm{Im}(F_C) = \mathsf{span}(e_1,\dots,e_r).$$

L'application $x \to Qx$ est un isomorphisme de $\mathrm{Im}(F_A)$ vers $\mathrm{Im}(F_C)$ et ainsi, d'après le corollaire 5.6,

$$\mathsf{rang}(F_A) = \dim \mathrm{Im}(F_A) = \dim \mathrm{Im}(F_C) = r = \mathsf{rang}(A).$$

Par le lemme 5.4 on a l'inégalité

$$\dim \mathrm{Im}(F) \leq \dim V \tag{5.3}$$

pour toute $F \in L(V,W)$. Le théorème suivant quantifie la différence entre les deux cotés de (5.3).

Théorème 5.9 (Théorème du rang) *Soient V,W deux K-espaces vectoriels, où $\dim V < \infty$, et soit $F \in L(V,W)$. Alors,*

$$\dim V = \mathsf{rang}(F) + \dim \mathrm{Ker}(F).$$

DÉMONSTRATION. Soit $(v_1,\ldots,v_k)$ une base de $\mathrm{Ker}(F)$. Par le lemme 4.15, on peut la compléter en une base

$$(v_1,\ldots,v_k,v_{k+1},\ldots,v_n)$$

de V, où $n = \dim V$. L'assertion

$$(F(v_{k+1}),\ldots,F(v_n)) \text{ est une base de } \mathrm{Im}(F) \tag{5.4}$$

montre le théorème car elle implique que $\mathsf{rang}(F) = n-k = \dim V - \dim \mathrm{Ker}(V)$.

Afin de montrer l'assertion (5.4), soit $v \in V$. Alors il existe $\alpha_1,\ldots,\alpha_n \in K$ tels que

$$v = \alpha_1 v_1 + \cdots \alpha_k v_k + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n.$$

Comme F est linéaire, on obtient

$$\begin{aligned} F(v) &= \alpha_1 F(v_1) + \cdots \alpha_k F(v_k) + \alpha_{k+1} F(v_{k+1}) + \cdots + \alpha_n F(v_n) \\ &= \alpha_{k+1} F(v_{k+1}) + \cdots + \alpha_n F(v_n) \end{aligned}$$

et donc $\mathrm{Im}(F) = \mathsf{span}(F(v_{k+1}),\ldots,F(v_n))$. Pour montrer l'indépendance linéaire, supposons que

$$0 = \alpha_{k+1} F(v_{k+1}) + \cdots + \alpha_n F(v_n) = F(\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n).$$

Cela signifie que $\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n \in \mathrm{Ker}(F)$. Alors, il existe $\beta_1,\ldots,\beta_k \in K$ tels que

$$\begin{aligned} \beta_1 v_1 + \cdots + \beta_k v_k &= \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n. \\ \text{Donc } \beta_1 v_1 + \cdots + \beta_k v_k - \alpha_{k+1} v_{k+1} - \cdots - \alpha_n v_n &= 0. \\ \text{Donc } \beta_1 = \cdots = \beta_k = \alpha_{k+1} = \cdots = \alpha_n &= 0, \end{aligned}$$

car $\{v_1,\ldots,v_n\}$ est une base de V. Donc la famille $(F(v_{k+1}),\ldots,F(v_n))$ est linéairement indépendante et donc (comme elle engendre $\mathrm{Im}(F)$) c'est une base de $\mathrm{Im}(F)$. ■

Corollaire 5.10 *Soient V, W deux K-espaces vectoriels, où $\dim V = \dim W < \infty$, $f \in L(V,W)$. Alors, les énoncés suivants sont équivalents :*

(i) F est injective
(ii) F est surjective
(iii) F est bijective.

DÉMONSTRATION. Il suffit de montrer l'équivalence entre l'injectivité et la surjectivité :

$$\begin{aligned} f \text{ injective } &\Leftrightarrow \dim \mathrm{Ker}(F) = 0 \\ &\Leftrightarrow \mathsf{rang}(F) = \dim V = \dim W \Leftrightarrow f \text{ surjective.} \end{aligned}$$

■

5.1.2 Composition des applications linéaires

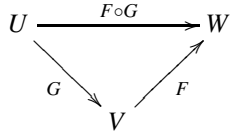
Théorème 5.11 *Soient U, V, W trois K-espaces vectoriels et $F \in L(V,W)$, $G \in L(U,V)$. Alors $F \circ G \in L(U,W)$.*

DÉMONSTRATION.

$$\begin{aligned}(F\circ G)(\alpha_1 v_1+\alpha_2 v_2) &= F(G(\alpha_1 v_1+\alpha_2 v_2))\\ &\overset{G \text{ linéaire}}{=} F(\alpha_1 G(v_1)+\alpha_2 G(v_2))\\ &\overset{F \text{ linéaire}}{=} \alpha_1 F(G(v_1))+\alpha_2 F(G(v_2))\\ &= \alpha_1 (F\circ G)(v_1)+\alpha_2 (F\circ G)(v_2).\end{aligned}$$

■

Le diagramme suivant illustre l'ordre des applications du théorème 5.11 :



$L(V,V)$, l'ensemble des endomorphismes de V, forme un anneau avec les deux lois de composition :

$$(F_1+F_2)(v) := F_1(v)+F_2(v), \qquad (F_1\circ F_2)(v) := F_1(F_2(v)).$$

5.2 Coordonées d'un vecteur, matrice d'une application linéaire

On a vu qu'un produit matrice-vecteur est une application linéaire $x \mapsto Ax$. Dans ce qui suit on va dans la direction opposée : Si on a une application linéaire arbitraire entre deux espaces vectoriels (de dimension finie) existe-t-il une matrice qui permet de représenter cette application ?

Théorème 5.12 *Soient V, W deux K-espaces vectoriels avec $\dim V = \dim W < \infty$. Soient $\mathscr{B}_V = (v_1,\dots,v_n)$ et $\mathscr{B}_W = (w_1,\dots,w_n)$ des bases de V et W respectivement. Alors il existe une unique application linéaire $F : V \to W$ telle que*

$$F(v_i) = w_i, \qquad i = 1,\dots,n. \tag{5.5}$$

Cette application linéaire est un isomorphisme.

DÉMONSTRATION. Soit $v \in V$, alors v s'écrit de façon unique dans la base $\mathscr{B}_V$: $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ avec $\alpha_1,\dots,\alpha_n \in K$. On définit

$$F(v) := \alpha_1 w_1 + \cdots + \alpha_n w_n.$$

Cette application vérifie (5.5) par construction.

Afin de montrer que F est linéaire, soit $\tilde{v} \in V$ tel que $\tilde{v} = \tilde{\alpha}_1 v_1 + \cdots + \tilde{\alpha}_n v_n$ (où $\tilde{\alpha}_1,\dots,\tilde{\alpha}_n \in K$) et $\beta \in K$. Alors

$$F(v+\tilde{v}) = \sum_{i=1}^n (\alpha_i+\tilde{\alpha}_i) w_i = \sum_{i=1}^n \alpha_i w_i + \sum_{i=1}^n \tilde{\alpha}_i w_i = F(v)+F(\tilde{v})$$

$$F(\beta v) = \sum_{i=1}^n \alpha_i \beta w_i = \beta \sum_{i=1}^n \alpha_i w_i = \beta F(v),$$

et ainsi F est linéaire.

Si $0 = F(v) = \alpha_1 w_1 + \cdots + \alpha_n w_n$ on obtient $\alpha_1 = \cdots = \alpha_n = 0$ par l'indépendance de $\mathscr{B}_W$. Alors F est injective par le lemme 5.4 (v) et ainsi F est bijective par le corollaire 5.10.

Il reste à montrer l'unicité d'une telle application linéaire. Supposons qu'il existe une autre application linéaire $\tilde{F} : V \to W$ telle que $F(v_i) = w_i$, $i = 1,\ldots,n$. Pour $v \in V$ arbitraire avec $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, on obtient

$$F(v) - \tilde{F}(v) = F\Big(\sum_{i=1}^{n} \alpha_i v_i\Big) - \tilde{F}\Big(\sum_{i=1}^{n} \alpha_i v_i\Big) = \sum_{i=1}^{n} \alpha_i \big(F(v_i) - \tilde{F}(v_i)\big) = 0.$$

Donc $F = \tilde{F}$. ■

Corollaire 5.13 *Soient V, W deux K-espaces vectoriels et $\dim V < \infty$. Alors, V et W sont isomorphes si et seulement si W est de dimension finie et $\dim V = \dim W$.*

DÉMONSTRATION. $V \cong W$ dit qu'il existe un isomorphisme $F \in L(V,W)$ tel que $\mathrm{Im}(F) = W$ et $\mathrm{Ker}(F) = \{0\}$. Par le théorème 5.9, on a $\dim V = \mathsf{rang}(F) + \dim \mathrm{Ker}(F) = \dim(W)$. L'autre direction découle directement du théorème 5.12. ■

En posant $W = K^n$ et en choisissant la base canonique, on obtient le cas le plus intéressant du théorème 5.12.

Corollaire 5.14 *Soit V un K-espace vectoriel et $\mathscr{B} = (v_1,\ldots,v_n)$ une base de V. Alors il existe un unique isomporphisme*

$$[\cdot]_{\mathscr{B}} : V \to K^n \quad \textit{tel que} \quad [v_i]_{\mathscr{B}} = e_i, \quad i = 1,\ldots,n,$$

où $(e_1,\ldots,e_n)$ est la base canonique de K^n.

Définition 5.15 *L'isomorphisme $[\cdot]_{\mathscr{B}}$ s'appelle un* ***système de coordonnées*** *[coordinate system]. Pour $v \in V$ on note*

$$[v]_{\mathscr{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

les ***coordonnées*** *de v dans la base $\mathscr{B}$.*

Afin de déterminer les coordonnées on écrit v dans la base $\mathscr{B}$:

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n. \tag{5.6}$$

Cela donne $[v]_{\mathscr{B}} = (\alpha_1,\ldots,\alpha_n)^{\mathsf{T}}$. Le calcul de (5.6) n'est simple que pour des bases triviales.

Quelques exemples :

Polynômes. On considère le K-espace vectoriel

$$K_n[t] := \{p = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n \,|\, \alpha_0,\ldots,\alpha_n \in K\},$$

qui possède la base $\mathscr{B} = (1,t,\ldots,t^n)$. D'habitude on représente un polynôme dans cette base :

$$p = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n \quad \Rightarrow \quad [p]_{\mathscr{B}} = (\alpha_0,\alpha_1,\ldots,\alpha_n)^{\mathsf{T}}.$$

Soit

$$\mathscr{C} = (v_0,v_1,\ldots,v_n) = (1,1+t,1+t+t^2,\ldots,1+t+\cdots+t^n)$$

une autre base. Pour calculer les coordonnées dans cette base du polynôme ci-dessus $p = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n$ on pose $p = \beta_0 v_0 + \cdots \beta_n v_n$ et on utilise la relation $v_i = 1 + t + \cdots + t^i$:

$$p = \sum_{j=0}^{n} \beta_j v_j = \sum_{j=0}^{n} \sum_{i=0}^{j} \beta_j t^i = \sum_{i=0}^{n} \Big(\sum_{j=i}^{n} \beta_j \Big) t^i .$$

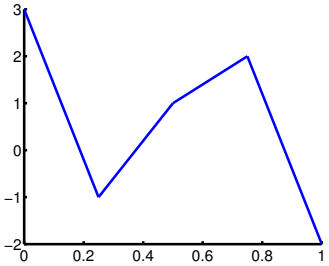
La comparaison des coefficients avec $p = \alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n$ donne le système linéaire

$$\sum_{j=i}^{n} \beta_j = \alpha_i, \quad i = 1, \ldots, n, \quad \text{c-à-d} \quad \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Par simple substitution, on obtient

$$[p]_{\mathscr{C}} = \begin{pmatrix} \alpha_0 - \alpha_1 \\ \alpha_1 - \alpha_2 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix}.$$

Fonctions continues affines par morceaux. On considère une partition de l'intervalle $[0,1]$ en n sous-intervalles $[0,h],[h,2h],\ldots,[(n-1)h,1]$, où $h = 1/n$. Une fonction $f : I \to \mathbb{R}$ est **affine par morceaux** [*piecewise linear*] si sa restriction à chacun de ces sous-intervalles est donnée par une expression affine $(\alpha t + \beta)$. Soit V l'ensemble des fonctions qui sont continues sur I et affines par morceaux. Alors V est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $C(I)$.
La figure à droite montre une de ses fonctions pour $n = 4$. Une base de V est donnée par

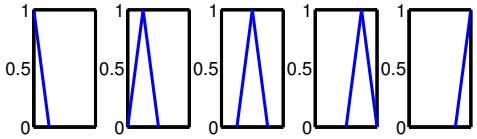


$$b_0(t) := \max\{1 - t/h, 0\},$$

$$b_i(t) := \begin{cases} (t - (i-1)h)/h, & t \in [(i-1)h, ih)], \\ -(t - (i+1)h)/h, & t \in [ih, (i+1)h)], \\ 0, & \text{sinon}, \end{cases}$$

$$b_n(t) := \max\{(t - (n-1)h)/h, 0\},$$

où $i = 1, \ldots, n-1$.
Une illustration des 5 fonctions de base pour $n = 4$:



Une fonction f continue affine par morceaux s'écrit dans cette base

$$f(t) = f(0)b_0(t) + f(h)b_1(t) + \cdots + f\big((n-1)h\big)b_{n-1}(t) + f(1)b_n(t).$$

Alors les coordonnées de f sont $\big(f(0), f(h), \ldots, f(1)\big)^{\mathsf{T}}$. Par exemple, les coordonnées de la fonction montrée ci-dessus sont $(3, -1, 1, 2, -2)^{\mathsf{T}}$.

5.2.1 Matrice d'une application linéaire

Soient V, W deux K-espaces vectoriels et $\mathscr{B}_V = (v_1,\dots,v_n)$ et $\mathscr{B}_W = (w_1,\dots,w_m)$ des bases de V et W respectivement. Soit $F \in L(V,W)$. Considérons le diagramme commutatif suivant :

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ {\scriptstyle [\cdot]_{\mathscr{B}_V}}\downarrow & & \downarrow{\scriptstyle [\cdot]_{\mathscr{B}_W}} \\ K^n & \xrightarrow{F_{\mathscr{B}_V,\mathscr{B}_W}} & K^m \end{array}$$

Par leur définition, on sait que $[\cdot]_{\mathscr{B}_V}$ et $[\cdot]_{\mathscr{B}_W}$ sont des applications linéaires. Comme l'inverse et la composition d'application linéaire sont encore des applications linéaires,

$$F_{\mathscr{B}_V,\mathscr{B}_W} := [\cdot]_{\mathscr{B}_W} \circ F \circ [\cdot]_{\mathscr{B}_V}^{-1} : K^n \to K^m. \tag{5.7}$$

est une application linéaire.

Lemme 5.16 *Soit $G \in L(K^n,K^m)$. On munit K^n, K^m de leurs bases canoniques respectives. Alors il existe une unique matrice $A \in M_{m\times n}(K)$ telle que $G(x) = Ax\ \forall x \in K^n$.*

DÉMONSTRATION. Soit $x = \sum_{i=1}^n \alpha_i e_i$, alors

$$G(x) = \sum_{i=1}^n \alpha_i G(e_i) = \underbrace{\big(G(e_1),G(e_2),\dots,G(e_n)\big)}_{=:A} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = Ax,$$

ce qui montre l'existence d'une telle matrice A.

Afin de montrer l'unicité supposons qu'il existe $\tilde{A} \in M_{m\times n}(K)$ telle que $\tilde{A}x = G(x) = Ax$. En choisissant $x = e_j$ on obtient que les colonnes j de A et $\tilde{A}$ sont égales pour $j = 1,\dots,n$. Alors, $\tilde{A} = A$. ■

Définition 5.17 *On note $[F]_{\mathscr{B}_V,\mathscr{B}_W}$ la matrice A du lemme précédent lorsque $G = F_{\mathscr{B}_V,\mathscr{B}_W}$. On l'appelle* ***matrice de l'application linéaire*** *F [matrix representation of F] par rapport aux bases $\mathscr{B}_V$ et $\mathscr{B}_W$.*

D'après la démonstration du lemme 5.16 on a

$$[F]_{\mathscr{B}_V,\mathscr{B}_W} = \Big(F_{\mathscr{B}_V,\mathscr{B}_W}(e_1), F_{\mathscr{B}_V,\mathscr{B}_W}(e_2), \dots, F_{\mathscr{B}_V,\mathscr{B}_W}(e_n)\Big) \in M_{m\times n}(K),$$

d'où

$$F_{\mathscr{B}_V,\mathscr{B}_W}(e_j) = \big([\cdot]_{\mathscr{B}_W} \circ F \circ [\cdot]_{\mathscr{B}_V}^{-1}\big)(e_j) = \big([\cdot]_{\mathscr{B}_W} \circ F\big)(v_j) = \big[F(v_j)\big]_{\mathscr{B}_W}$$

Cela signifie que les coordonnées de $F(v_j) \in W$ dans la base $\mathscr{B}_W$ donne la j-ième colonne de $F_{\mathscr{B}_V,\mathscr{B}_W}$. Alors on exprime

$$F(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m, \qquad a_{1j},\dots,a_{mj} \in K,$$

et on obtient que

$$F_{\mathscr{B}_V,\mathscr{B}_W}(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \qquad \text{donc} \qquad [F]_{\mathscr{B}_V,\mathscr{B}_W} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \tag{5.8}$$

Exemple 5.18 Soit $V = W = \mathbb{R}_3[t]$. La dérivation est une application linéaire :

$$D \in L(V,V), \quad D(p) := p'.$$

Afin de calculer la matrice de D par rapport à la base $\mathscr{B} = (1,t,t^2,t^3)$ de V, on applique D sur tout élément de base :

$$D(1) = 0, \quad D(t) = 1, \quad D(t^2) = 2t, \quad D(t^3) = 3t^2,$$

et puis on exprime les résultats dans la base $\mathscr{B}$. Dans ce cas c'est facile et on obtient

$$[D]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.9}$$

Théorème 5.19 *Soient V,W deux K-espaces vectoriels de dimension finie et $\mathscr{B}_V$, $\mathscr{B}_W$ des bases de V et W respectivement. Alors l'application*

$$\Psi : L(V,W) \to M_{m\times n}(K), \qquad F \mapsto [F]_{\mathscr{B}_V,\mathscr{B}_W},$$

où $n = \dim V$ et $m = \dim W$, est un isomorphisme d'espace vectoriel. En particulier, $L(V,W) \cong M_{m\times n}(K)$.

DÉMONSTRATION. D'après le lemme 5.16 et par la linéarité des applications concernées,

$$\begin{aligned} \Psi(F+G) &= \Big(\big[F(v_1)+G(v_1)\big]_{\mathscr{B}_W}, \ldots, \big[F(v_n)+G(v_n)\big]_{\mathscr{B}_W} \Big) \\ &= \Big(\big[F(v_1)\big]_{\mathscr{B}_W} + \big[G(v_1)\big]_{\mathscr{B}_W}, \ldots, \big[F(v_n)\big]_{\mathscr{B}_W} + \big[G(v_n)\big]_{\mathscr{B}_W} \Big) \\ &= \Big(\big[F(v_1)\big]_{\mathscr{B}_W}, \ldots, \big[F(v_n)\big]_{\mathscr{B}_W} \Big) + \Big(\big[G(v_1)\big]_{\mathscr{B}_W}, \ldots, \big[G(v_n)\big]_{\mathscr{B}_W} \Big) \\ &= \Psi(F) + \Psi(G) \end{aligned}$$

pour tout $F,G \in L(V,W)$. De même $\Psi(\alpha F) = \alpha\Psi(F)$ pour tous $\alpha \in K$ et toutes $F \in L(V,W)$.

Soit $\Psi(F) = [F]_{\mathscr{B}_V,\mathscr{B}_W} = 0 \in M_{m\times n}(K)$, alors

$$0 = [\cdot]_{\mathscr{B}_W} \circ F \circ [\cdot]^{-1}_{\mathscr{B}_V}, \quad \text{donc} \quad 0 = [\cdot]^{-1}_{\mathscr{B}_W} \circ [\cdot]_{\mathscr{B}_W} \circ F \circ [\cdot]^{-1}_{\mathscr{B}_V} \circ [\cdot]_{\mathscr{B}_V} = F.$$

Par le lemme 5.4 (v), cela signifie que Ψ est injective. Afin de montrer la surjectivité, soit

$$A = (a_1, \ldots, a_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

une matrice arbitraire. On cherche $F : V \to W$ linéaire telle que

$$\Psi(F) = \Big(\big[F(v_1)\big]_{\mathscr{B}_W}, \ldots, \big[F(v_n)\big]_{\mathscr{B}_W} \Big) = (a_1, \ldots, a_n).$$

On pose $F(v_j) := \sum_{i=1}^m a_{ij} w_i$ que l'on étend par linéarité pour tout vecteur $v = \sum_{j=1}^n \alpha_j v_j$ quelconque par $F(v) = \sum_{j=1}^n \alpha_j F(v_j)$. Cette application est linéaire par construction et l'on a bien $\Psi(f) = A$. ∎

Corollaire 5.20 *Soient V,W deux K-espaces vectoriels avec $\dim V = n < \infty$ et $\dim W = m < \infty$. Alors $\dim L(V,W) = m \cdot n$.*

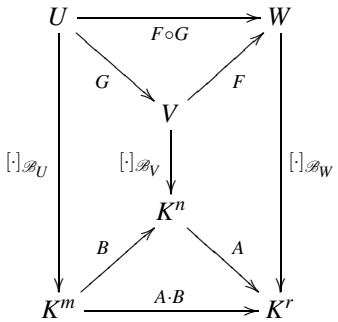
DÉMONSTRATION. La dimension de $M_{m\times n}(K)$ est $m\cdot n$. Par le théorème 5.19, $M_{m\times n}(K)$ et $L(V,W)$ sont isomorphes. Alors l'assertion découle du fait que deux espaces vectoriels qui sont isomorphes ont la même dimension. ■

Le résultat suivant montre que la composition d'applications linéaires correspond à la multiplication matricielle des matrices associées.

Théorème 5.21 *Soient U,V,W des K-espaces vectoriels de dimension finie avec des bases respectives $\mathscr{B}_U,\mathscr{B}_V,\mathscr{B}_W$. Soient $G:U\to V$ et $F:V\to W$ des applications linéaires. Alors*

$$[F\circ G]_{\mathscr{B}_U,\mathscr{B}_W}=[F]_{\mathscr{B}_V,\mathscr{B}_W}\cdot[G]_{\mathscr{B}_U,\mathscr{B}_V}.$$

DÉMONSTRATION. Posons $m=\dim U$, $n=\dim V$, $r=\dim W$ et $A=[F]_{\mathscr{B}_V,\mathscr{B}_W}$, $B=[G]_{\mathscr{B}_U,\mathscr{B}_V}$. Alors le diagramme suivant est commutatif[6] :



Ce diagramme permet de « lire » l'assertion du théorème en choisissant deux chemins différents de K^m vers K^r, ce qui donne $F_{AB}=(F\circ G)_{\mathscr{B}_U,\mathscr{B}_W}$ et donc $AB=[F\circ G]_{\mathscr{B}_U,\mathscr{B}_W}$.

On peut aussi vérifier l'assertion par le calcul suivant :

$$F_{AB}=[\cdot]_{\mathscr{B}_W}\circ F\circ[\cdot]^{-1}_{\mathscr{B}_V}\circ[\cdot]_{\mathscr{B}_V}\circ G\circ[\cdot]^{-1}_{\mathscr{B}_U}=[\cdot]_{\mathscr{B}_W}\circ(F\circ G)\circ[\cdot]^{-1}_{\mathscr{B}_U}=(F\circ G)_{\mathscr{B}_U,\mathscr{B}_W}.$$

■

Corollaire 5.22 *Soient V,W deux K-espaces vectoriels avec des bases $\mathscr{B}_V$ et $\mathscr{B}_W$ respectivement. Supposons que* $\dim(V)=\dim(W)$ *et soit $F\in L(V,W)$ bijective. Alors*

$$\left[F^{-1}\right]_{\mathscr{B}_W,\mathscr{B}_V}=\left([F]_{\mathscr{B}_V,\mathscr{B}_W}\right)^{-1}.$$

DÉMONSTRATION. D'après le théorème 5.21

$$\left[F^{-1}\right]_{\mathscr{B}_W,\mathscr{B}_V}[F]_{\mathscr{B}_V,\mathscr{B}_W}=\left[F^{-1}\circ F\right]_{\mathscr{B}_V,\mathscr{B}_V}=[I]_{\mathscr{B}_V,\mathscr{B}_V}=I_{\dim(V)},$$

où $I:V\to V$ est l'application identité, et donc $\left[F^{-1}\right]_{\mathscr{B}_W,\mathscr{B}_V}$ est l'inverse de $[F]_{\mathscr{B}_V,\mathscr{B}_W}$. ■

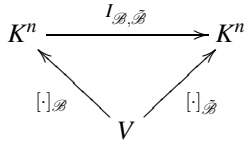
5.2.2 Changement de bases

Soit V un K-espace vectoriel de dimension finie et soient

$$\mathscr{B}=(v_1,\dots,v_n),\qquad \tilde{\mathscr{B}}=(\tilde{v}_1,\dots,\tilde{v}_n)$$

6. En algèbre linéaire un diagramme commutatif est un graphe orienté avec des espaces vectoriels comme nœuds et des applications linéaires comme arcs, tels que, lorsque l'on choisit deux espaces vectoriels, on peut suivre un chemin quelconque à travers le diagramme et obtenir le même résultat par composition des applications linéaires.

deux bases de V. Le diagramme correspondant :



L'application linéaire $I_{\mathscr{B},\tilde{\mathscr{B}}} : K^n \to K^n$ des coordonnées dans $\mathscr{B}$ vers des coordonnées dans $\tilde{\mathscr{B}}$ est donnée par

$$I_{\mathscr{B},\tilde{\mathscr{B}}} = [\cdot]_{\tilde{\mathscr{B}}} \circ [\cdot]_{\mathscr{B}}^{-1}.$$

Alors la multiplication par $[I]_{\mathscr{B},\tilde{\mathscr{B}}} \in M_{n\times n}(K)$ transforme les coordonnées d'un vecteur $v \in V$ dans $\mathscr{B}$ en celles dans $\tilde{\mathscr{B}}$:

$$[v]_{\tilde{\mathscr{B}}} = [I]_{\mathscr{B},\tilde{\mathscr{B}}} [v]_{\mathscr{B}}.$$

Définition 5.23 *La matrice $[I]_{\mathscr{B},\tilde{\mathscr{B}}}$ est appelée* ***matrice de passage*** *(ou* ***matrice de changement de base****) de la base $\mathscr{B}$ à la base $\tilde{\mathscr{B}}$.*

Afin de déterminer $[I]_{\mathscr{B},\tilde{\mathscr{B}}} \in M_{n\times n}(K)$ on procéde comme dans la section 5.2.1 :

$$[I]_{\mathscr{B},\tilde{\mathscr{B}}} = \Big(\big[I(v_1)\big]_{\tilde{\mathscr{B}}}, \ldots, \big[I(v_n)\big]_{\tilde{\mathscr{B}}} \Big) = \Big(\big[v_1\big]_{\tilde{\mathscr{B}}}, \ldots, \big[v_n\big]_{\tilde{\mathscr{B}}} \Big)$$

Cela signifie qu'on exprime l'élément v_j de base $\mathscr{B}$ dans la base $\tilde{\mathscr{B}}$:

$$v_j = a_{1j}\tilde{v}_1 + \tilde{v}_{2j} w_2 + \cdots + \tilde{v}_{nj} w_n.$$

Alors $[I]_{\mathscr{B},\tilde{\mathscr{B}}} = [a_{ij}]_{i,j=1}^n$.

Dans le cas particulier $V = K^n$ on peut écrire les éléments des bases $\mathscr{B}$ et $\tilde{\mathscr{B}}$ comme les colonnes des matrices

$$B = \big(v_1, \ldots, v_n\big) \in K^{n\times n}, \qquad \tilde{B} = \big(\tilde{v}_1, \ldots, \tilde{v}_n\big) \in K^{n\times n}.$$

Comme $Bx = [x]_{\mathscr{B}}^{-1}$ et $\tilde{B}x = [x]_{\tilde{\mathscr{B}}}^{-1}$ on obtient

$$[I]_{\mathscr{B},\tilde{\mathscr{B}}} = \tilde{B}^{-1} B.$$

par le théorème 5.21.

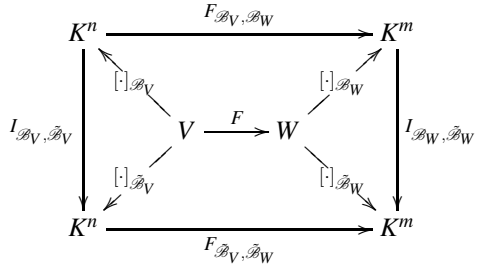
Remarque 5.24 *Par le corollaire 5.22,* $\big([I]_{\mathscr{B},\tilde{\mathscr{B}}}\big)^{-1} = [I]_{\tilde{\mathscr{B}},\mathscr{B}}$.

Le théoreme suivant est le résultat central de cette section. Il décrit le changement de la matrice d'une application linéaire sous un changement de bases.

Théorème 5.25 *Soient V, W deux K-espaces vectoriels de dimension finie. On considère des bases $\mathscr{B}_V, \tilde{\mathscr{B}}_V$ de V et des bases $\mathscr{B}_W, \tilde{\mathscr{B}}_W$ de W. Alors pour $F \in L(V,W)$ on a*

$$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_W} = [I]_{\mathscr{B}_W,\tilde{\mathscr{B}}_W} \cdot [F]_{\mathscr{B}_V,\mathscr{B}_W} \cdot [I]^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V}. \tag{5.10}$$

DÉMONSTRATION. On considère le diagramme suivant :



Ce diagramme est commutatif parce que tout sous-diagramme est commutatif. En particulier, on obtient

$$F_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_W} = I_{\mathscr{B}_W,\tilde{\mathscr{B}}_W} \circ F_{\mathscr{B}_V,\mathscr{B}_W} \circ I^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V}$$

en choisissant les chemins correspondants dans le diagramme. Ainsi, (5.10) découle du théorème 5.21 et de la remarque 5.24. ■

Par le théorème 5.25, deux matrices de la même application linéaire sont toujours équivalentes (voir la définition 3.16).

Corollaire 5.26 *Soient V,W deux K-espaces vectoriels de dimension finie et soit $F \in L(V,W)$. Alors il existe des bases $\tilde{\mathscr{B}}_V$ et $\tilde{\mathscr{B}}_W$ telles que*

$$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_W} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.11}$$

où $r = \mathsf{rang}(F)$.

DÉMONSTRATION. Soient $\mathscr{B}_V = (v_1,\dots,v_n)$, $\mathscr{B}_W = (w_1,\dots,w_m)$ des bases quelconques de V, W. Considérons $[F]_{\mathscr{B}_V,\mathscr{B}_W}$ la matrice de l'application linéaire par rapport aux bases $\mathscr{B}_V$, $\mathscr{B}_W$. Par le théorème 3.17 il existe des matrices inversibles $P \in M_{m\times m}(K)$, $Q \in M_{n\times n}(K)$ telles que

$$P^{-1}[F]_{\mathscr{B}_V,\mathscr{B}_W} Q = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.12}$$

En posant $\tilde{\mathscr{B}}_V = (\tilde{v}_1,\dots,\tilde{v}_n)$ avec $\tilde{v}_j := \big[Q\cdot[v_j]_{\mathscr{B}_V}\big]^{-1}_{\mathscr{B}_V}$ on obtient

$$Qe_j = [\tilde{v}_j]_{\mathscr{B}_V} = \big([\cdot]_{\mathscr{B}_V} \circ [\cdot]^{-1}_{\tilde{\mathscr{B}}_V}\big)(e_j) = [I]_{\tilde{\mathscr{B}}_V,\mathscr{B}_V} e_j \quad \Rightarrow \quad [I]_{\tilde{\mathscr{B}}_V,\mathscr{B}_V} = Q.$$

De façon analogue, on peut choisir $\tilde{\mathscr{B}}_W$ telle que $[I]_{\tilde{\mathscr{B}}_W,\mathscr{B}_W} = P$. Ainsi (5.12) devient

$$[I]_{\mathscr{B}_W,\tilde{\mathscr{B}}_W}[F]_{\mathscr{B}_V,\mathscr{B}_W}[I]^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

et (5.11) découle du théorème 5.25. ■

Pour un endomorphisme $F \in L(V,V)$ on choisit typiquement la même base $\mathscr{B}_V$ pour l'espace de départ et l'espace de arrivée. Par le théorème 5.25, un changement de base de $\mathscr{B}_V$ à $\tilde{\mathscr{B}}_V$ effectue la transformation suivante :

$$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_V} = [I]_{\mathscr{B}_V,\tilde{\mathscr{B}}_V} \cdot [F]_{\mathscr{B}_V,\mathscr{B}_V} \cdot [I]^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V}. \tag{5.13}$$

C'est un cas particulier d'équivalence de matrices.

Définition 5.27 *Deux matrices $A,B \in M_{n\times n}(K)$ sont dites* ***semblables*** *[similar] s'il existe une matrice inversible $P \in M_{n\times n}(K)$ telle que $A = PBP^{-1}$.*

Comme pour les matrices équivalentes, « être semblable » définit une relation d'équivalence sur les matrices $n \times n$.

Exemple 5.28 On continue l'exemple 5.18. On a vu la matrice (5.9) de la dérivation $D \in L(V,V)$, $V = \mathbb{R}_3[t]$, par rapport à $\mathscr{B} = (1,t,t^2,t^3)$. Étant donnée une autre base

$$\tilde{\mathscr{B}} = (1,1+t,1+t+t^2,1+t+t^2+t^3)$$

on a

$$[I]_{\tilde{\mathscr{B}},\mathscr{B}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad [I]_{\mathscr{B},\tilde{\mathscr{B}}} = [I]^{-1}_{\tilde{\mathscr{B}},\mathscr{B}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Alors la matrice de D par rapport à $\tilde{\mathscr{B}}$ est donnée par

$$\begin{aligned} [D]_{\tilde{\mathscr{B}},\tilde{\mathscr{B}}} &= [I]_{\mathscr{B},\tilde{\mathscr{B}}}[D]_{\mathscr{B},\mathscr{B}}[I]^{-1}_{\mathscr{B},\tilde{\mathscr{B}}} \\ &= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Par exemple pour

$$p = 1+t+t^2+t^3, \quad \text{on a} \quad [p]_{\tilde{\mathscr{B}}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad \text{Donc} \quad [D]_{\tilde{\mathscr{B}},\tilde{\mathscr{B}}}[p]_{\tilde{\mathscr{B}}} = \begin{pmatrix} -1 \\ -1 \\ 3 \\ 0 \end{pmatrix}$$

et ainsi

$$\Big[[D]_{\tilde{\mathscr{B}},\tilde{\mathscr{B}}}[p]_{\tilde{\mathscr{B}}}\Big]^{-1}_{\tilde{\mathscr{B}}} = -1-(1+t)+3(1+t+t^2) = 1+2t+3t^2 = p'.$$

♦

Chapitre 6
Déterminants

Les déterminants ont joué un rôle fondamental dans le développement historique de l'algèbre linéaire. Avant le développement moderne de la notation des matrices on a exprimé toutes les assertions et leurs preuves en termes de déterminants (sans matrices). Heureusement, ceci a changé ! Malgré cela, les déterminants gardent un rôle dans la compréhension théorique.

6.1 Définitions

$(K,+,\cdot)$ désigne un anneau commutatif dans ce chapitre. On rappelle que S_n est l'ensemble des permutations de l'ensemble $\{1,2,\ldots,n\}$.

Définition 6.1 *(i) Soit $\sigma \in S_n$, $n \geq 2$. Un couple $(i,j) \in \mathbb{N}\times\mathbb{N}$, $1 \leq i < j \leq n$, s'appelle une **inversion** de σ si $\sigma(i) > \sigma(j)$.*

*(ii) Soit k le nombre d'inversion d'une permutation σ. On appelle **signature** [sign] de la permutation le nombre $\mathsf{sgn}(\sigma) := (-1)^k$. Pour $n = 1$, il existe seulement une permutation $\sigma = \mathsf{id}$ et on définit $\mathsf{sgn}(\mathsf{id}) = 1$.*

Exemple 6.2 La permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$$

a les inversions (2,3), (2,4), alors $\mathsf{sgn}(\sigma) = (-1)^2 = 1$. ♦

Définition 6.3 *Soit K un anneau commutatif et soit $A \in M_{n\times n}(K)$. Le **déterminant** de A est défini par*

$$\det(A) := \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}. \tag{6.1}$$

Dit autrement, étant donnée une permutation σ on choisit le $\sigma(i)$-ième élément de la i-iéme ligne de A pour $i = 1,\ldots,n$. On multiplie par la signature de σ le produit de ces n élements. Enfin, on fait la somme de tous ces produits sur toutes les permutations. On remarque qu'il existe $n!$ permutations, alors le calcul du déterminant devient extrêmement laborieux pour une matrice de grande taille. Dans la section 6.5.1 on developpera un algorithme plus efficace.

Remarque 6.4 *Le déterminant d'une matrice A est un exemple d'une fonction linéaire par rapport à chaque colonne de A. Une telle fonction s'appelle **multilinéaire**. Un autre*

exemple moins important d'une fonction multilinéaire est le ***permanent*** *défini par*

$$\mathrm{perm}(A) := \sum_{\sigma\in S_n}\prod_{i=1}^{n} a_{i,\sigma(i)}.$$

La différence semble faible, on a seulement supprimé la signature, mais elle est en fait énorme ! Par exemple, il n'existe pas un algorithme efficace pour le calcul du permanent d'une matrice de grande taille. Le calcul du permanent est un problème NP-difficile.[7]

Déterminants des matrices particuliers.

Matrices 1×1**.** Par définition, le déterminant d'une matrice $A=(a_{11})\in M_{1\times 1}(K)$ est donné par $\det(A)=a_{11}$.

Matrices 2×2**.** Pour $n=2$ il y a deux permutations :

$$\sigma_1=\begin{pmatrix}1&2\\1&2\end{pmatrix}\Rightarrow \mathrm{sgn}(\sigma_1)=1,\qquad \sigma_1=\begin{pmatrix}1&2\\2&1\end{pmatrix}\Rightarrow \mathrm{sgn}(\sigma_2)=-1.$$

Alors le déterminant d'une matrice A de taille 2×2 est donné par

$$\det(A)=\det\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}=a_{11}a_{22}-a_{12}a_{21}.$$

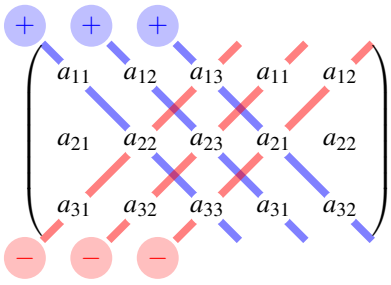
Matrices 3×3**.** Pour $n=3$ il y a six permutations :

$$\sigma_1=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \sigma_2=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\quad \sigma_3=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$
$$\sigma_4=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\quad \sigma_5=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\quad \sigma_6=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},$$

avec $\mathrm{sgn}(\sigma_1)=\mathrm{sgn}(\sigma_2)=\mathrm{sgn}(\sigma_3)=1$ et $\mathrm{sgn}(\sigma_4)=\mathrm{sgn}(\sigma_5)=\mathrm{sgn}(\sigma_6)=-1$. Alors le déterminant d'une matrice A de taille 3×3 est donné par

$$\begin{aligned}\det(A)=\det&\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\\&=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}\\&\quad -a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}\end{aligned}$$

On peut bien visualiser cette formule par la **règle de Sarrus** :



7. En fait, il existe des algorithmes passablement efficaces pour l'approximation du permanent d'une matrice à coefficients positifs, voir, par exemple [M. Jerrum, A. Sinclair, E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. J. ACM 51 (2004), 4, 671–697].

Matrices diagonales. Soit $D = \mathrm{diag}\,(d_{11},d_{22},\ldots,d_{nn}) \in M_{n\times n}(K)$. Alors il existe seulement une permutation σ telle que $d_{1,\sigma(1)}d_{2,\sigma(2)}\cdots d_{n,\sigma(n)} \neq 0$, i.e. $\sigma = \mathrm{id}$. Donc

$$\det\big(\mathrm{diag}\,(d_{11},d_{22},\ldots,d_{nn})\big) = d_{11}d_{22}\cdots d_{nn}.$$

Matrices triangulaires. Lemme 6.5 *Soit $A \in M_{n\times n}(K)$ une matrice triangulaire (supérieure ou inférieure). Alors*

$$\det(A) = a_{11}a_{22}\ldots a_{nn}.$$

DÉMONSTRATION. Soit

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Considérons $\prod_{i=1}^n a_{i,\sigma(i)}$ alors $\sigma(1) = 1$ est nécessaire pour que le produit soit non nul. Par suite $\sigma(2) \in \{1,2\}$ pour que $a_{2,\sigma(2)} \neq 0$. Mais $\sigma(2) = 1$ n'est pas possible car une permutation σ doit être bijective. Alors $\sigma(2) = 2$. En répétent cet argument on obtient que $\sigma = \mathrm{id}$. Donc $\det(A) = a_{11}a_{22}\ldots a_{nn}$.
Le raisonnement est similaire pour une matrice triangulaire supérieure. ■

Matrices de permutation. Soit $\pi \in S_n$, considérons

$$P_\pi = \begin{pmatrix} e_{\pi(1)}^\mathsf{T} \\ \vdots \\ e_{\pi(n)}^\mathsf{T} \end{pmatrix} = \big(p_{ij}\big)_{1\leq i,j\leq n}.$$

Alors $\prod_{i=1}^n a_{i,\sigma(i)} = 0$ si $\sigma \neq \pi$ car toutes les éléments de la ligne i sauf l'élément $\pi(i)$ sont nuls. Alors on a

$$\det(P_\pi) = \mathrm{sgn}(\pi)\prod_{i=1}^n \underbrace{p_{i,\pi(i)}}_{=1} = \mathrm{sgn}(\pi). \tag{6.2}$$

6.2 Propriétés de la signature

Le lemme suivant donne une expression explicite de la signature d'une permutation.

Lemme 6.6 *Soit $\sigma \in S_n$. Alors*

$$\mathrm{sgn}(\sigma) = \prod_{1\leq i<j\leq n} \frac{\sigma(i)-\sigma(j)}{i-j}.$$

DÉMONSTRATION. Pour $n = 1$ la formule est vraie par la définition. Soit $n \geq 2$. Alors on a

$$\prod_{1\leq i<j\leq n} |\sigma(j)-\sigma(i)| = \prod_{1\leq i<j\leq n} (j-i). \tag{6.3}$$

On montre cette relation en calculant

$$\begin{aligned}\prod_{i<j}|\sigma(j)-\sigma(i)| &= \Big(\prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} (\sigma(j)-\sigma(i))\Big)\Big(\prod_{\substack{i<j\\ \sigma(j)<\sigma(i)}} (\sigma(i)-\sigma(j))\Big) \\ &= \Big(\prod_{\substack{\sigma^{-1}(\tilde{i})<\sigma^{-1}(\tilde{j})\\ \tilde{i}<\tilde{j}}} (\tilde{j}-\tilde{i})\Big)\Big(\prod_{\substack{\sigma^{-1}(\tilde{j})<\sigma^{-1}(\tilde{i})\\ \tilde{i}<\tilde{j}}} (\tilde{j}-\tilde{i})\Big) = \prod_{\tilde{i}<\tilde{j}} (\tilde{j}-\tilde{i}),\end{aligned}$$

où la dernière égalité découle du fait qu'on a soit $\sigma^{-1}(\tilde{i}) < \sigma^{-1}(\tilde{j})$, soit $\sigma^{-1}(\tilde{j}) < \sigma^{-1}(\tilde{i})$ si $\tilde{i} \neq \tilde{j}$. Soit k le nombre d'inversions de σ, alors

$$\prod_{1\leq i<j\leq n} \frac{\sigma(j)-\sigma(i)}{|\sigma(j)-\sigma(i)|} = (-1)^k = \mathsf{sgn}(\sigma).$$

Mis en ensemble avec (6.3), cela donne l'assertion. ∎

Théorème 6.7 *Soient $\sigma_1, \sigma_2 \in S_n$. Alors* $\mathsf{sgn}(\sigma_1 \circ \sigma_2) = \mathsf{sgn}(\sigma_1)\mathsf{sgn}(\sigma_2)$.

DÉMONSTRATION. Par le lemme 6.6 on obtient

$$\begin{aligned}
\mathsf{sgn}(\sigma_1\circ\sigma_2) &= \prod_{i<j} \frac{\sigma_1(\sigma_2(i))-\sigma_1(\sigma_2(j))}{i-j} \\
&= \Big(\prod_{i<j} \frac{\sigma_1(\sigma_2(i))-\sigma_1(\sigma_2(j))}{\sigma_2(i)-\sigma_2(j)}\Big)\Big(\prod_{i<j} \frac{\sigma_2(i)-\sigma_2(j)}{i-j}\Big) \\
&\overset{\substack{\tilde{i}=\sigma_2(i)\\ \tilde{j}=\sigma_2(j)}}{=} \mathsf{sgn}(\sigma_2) \prod_{\sigma_2^{-1}(\tilde{i})<\sigma_2^{-1}(\tilde{j})} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}} \\
&= \mathsf{sgn}(\sigma_2)\Big(\prod_{\substack{\tilde{i}<\tilde{j}\\ \sigma_2^{-1}(\tilde{i})<\sigma_2^{-1}(\tilde{j})}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}}\Big)\Big(\prod_{\substack{\tilde{i}>\tilde{j}\\ \sigma_2^{-1}(\tilde{i})<\sigma_2^{-1}(\tilde{j})}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}}\Big).
\end{aligned}$$

En permutant les rôles de $\tilde{i}$ et $\tilde{j}$ dans le deuxième facteur, on obtient

$$\begin{aligned}
\mathsf{sgn}(\sigma_1\circ\sigma_2) &= \mathsf{sgn}(\sigma_2)\Big(\prod_{\substack{\tilde{i}<\tilde{j}\\ \sigma_2^{-1}(\tilde{i})<\sigma_2^{-1}(\tilde{j})}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}}\Big)\Big(\prod_{\substack{\tilde{j}>\tilde{i}\\ \sigma_2^{-1}(\tilde{j})<\sigma_2^{-1}(\tilde{i})}} \frac{\sigma_1(\tilde{j})-\sigma_1(\tilde{i})}{\tilde{j}-\tilde{i}}\Big) \\
&= \mathsf{sgn}(\sigma_2)\Big(\prod_{\substack{\tilde{i}<\tilde{j}\\ \sigma_2^{-1}(\tilde{i})<\sigma_2^{-1}(\tilde{j})}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}}\Big)\Big(\prod_{\substack{\tilde{i}<\tilde{j}\\ \sigma_2^{-1}(\tilde{j})<\sigma_2^{-1}(\tilde{i})}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}}\Big) \\
&= \mathsf{sgn}(\sigma_2)\prod_{\tilde{i}<\tilde{j}} \frac{\sigma_1(\tilde{i})-\sigma_1(\tilde{j})}{\tilde{i}-\tilde{j}} = \mathsf{sgn}(\sigma_1)\mathsf{sgn}(\sigma_2).
\end{aligned}$$

∎

En posant $\sigma_1 = \sigma$ et $\sigma_2 = \sigma^{-1}$ pour $\sigma \in S_n$, le théorème 6.7 donne

$$\mathsf{sgn}\big(\sigma^{-1}\big) = \mathsf{sgn}(\sigma).$$

En d'autres termes, le théorème 6.7 dit que $\mathsf{sgn} : S_n \to \{+1,-1\}$ est un morphisme de groupes de $(S_n,\circ)$ vers $(\{+1,-1\},\cdot)$.

On rappelle le cas particulier d'une transposition (voir la définition 3.5) :

$$\tau = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{pmatrix}, \quad 1 \leq i < j \leq n. \tag{6.4}$$

En comptant simplement on voit que τ a $2(j-i)-1$ inversions, alors $\mathsf{sgn}(\tau) = -1$. Toute

permutation peut s'écrire comme une composition de transpositions, par exemple

$$\begin{aligned}\sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}.\end{aligned}$$

On remarque que le nombre de transpositions n'est pas unique. En revanche, par le théorème 6.7, la parité de ce nombre est unique.

Corollaire 6.8 *Soit $\sigma \in S_n$. Alors*

(i) $\operatorname{sgn}(\sigma) = +1$ *si et seulement si σ peut s'écrire comme la composition d'un nombre pair de transpositions,*

(ii) $\operatorname{sgn}(\sigma) = -1$ *si et seulement si σ peut s'écrire comme la composition d'un nombre impair de transpositions.*

6.3 Propriétés du déterminant

Lemme 6.9 *Soit $A \in M_{n\times n}(K)$. Alors* $\det(A^{\mathsf{T}}) = \det(A)$.

DÉMONSTRATION. Par la définition du déterminant on obtient

$$\begin{aligned}\det(A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \overset{\mu=\sigma^{-1}}{=} \sum_{\mu^{-1}\in S_n} \operatorname{sgn}(\mu^{-1}) \prod_{i=1}^{n} a_{i,\mu^{-1}(i)} \\ &\overset{j=\mu^{-1}(i)}{=} \sum_{\mu\in S_n} \operatorname{sgn}(\mu) \prod_{j=1}^{n} a_{\mu(j),j} = \det(A^{\mathsf{T}}).\end{aligned}$$

■

Lemme 6.10 *Soit $A \in M_{n\times n}(K)$.*

1. *Si au moins une des lignes de A est nulle alors* $\det(A) = 0$.
2. *Si $n \geq 2$ et deux lignes de A sont identiques, alors* $\det(A) = 0$.

DÉMONSTRATION. (i). Exercice.

(ii). Supposons que les lignes i et j, où $i < j$, de A sont identiques. On considére la réunion disjointe

$$S_n = T_n \cup S_n \backslash T_n,$$

où

$$T_n := \{\sigma \in S_n \,|\, \sigma(i) < \sigma(j)\}, \quad S_n\backslash T_n = \{\sigma \in S_n \,|\, \sigma(i) > \sigma(j)\}.$$

En utilisant la transposition τ de (6.4) on peut écrire

$$S_n \backslash T_n = \{\sigma \circ \tau \,|\, \sigma \in T_n\}.$$

En utilisant le théorème 6.7, $\operatorname{sgn}(\tau) = -1$, et $\tau = \tau^{-1}$, on obtient l'identité

$$\begin{aligned}\sum_{\tilde\sigma\in S_n\backslash T_n} \operatorname{sgn}(\tilde\sigma) \prod_{k=1}^{n} a_{k,\tilde\sigma(k)} &= \sum_{\sigma\in T_n} \operatorname{sgn}(\sigma\circ\tau) \prod_{k=1}^{n} a_{k,\sigma(\tau(k))} \\ &\overset{\ell=\tau(k)}{=} -\sum_{\sigma\in T_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\tau(\ell),\sigma(\ell)} \\ &= -\sum_{\sigma\in T_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell,\sigma(\ell)}.\end{aligned}$$

La dernière égalité découle de l'égalité des lignes i et j. Alors on a

$$\begin{aligned}\det(A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)} = \sum_{\sigma\in T_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)} + \sum_{\sigma\in S_n\setminus T_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)}\\ &= \sum_{\sigma\in T_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)} - \sum_{\sigma\in T_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)} = 0.\end{aligned}$$

■

En combinant le lemme 6.9 et le lemme 6.10 on obtient qu'une matrice avec une colonne nulle a un déterminant nul.

Le lemme suivant décrit l'effet de la multiplication par une des matrices élémentaires (voir la section 3.1) sur le déterminant.

Lemme 6.11 *Soit $A \in M_{n\times n}(K)$ et $n \geq 2$. Alors*

(i) $\det(M_i(\lambda)A) = \lambda\cdot\det(A) = \det(M_i(\lambda))\cdot\det(A)$ *pour* $\lambda \in K$, $1 \leq i \leq n$,

(ii) $\det(G_{ij}(\lambda)A) = \det(A) = \det(G_{ij}(\lambda))\cdot\det(A)$ *et*
$\det(G_{ij}^{\mathsf{T}}(\lambda)A) = \det(A) = \det(G_{ij}^{\mathsf{T}}(\lambda))\cdot\det(A)$ *pour* $\lambda \in K$,$1 \leq i < j \leq n$;

(iii) $\det(P_{ij}A) = -\det(A) = \det(P_{ij})\cdot\det(A)$ *pour* $1 \leq i < j \leq n$.

DÉMONSTRATION. (i).

$$\begin{aligned}\det(M_i(\lambda)A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\left(\lambda a_{i,\sigma(i)}\prod_{k\neq i} a_{k,\sigma(k)}\right)\\ &= \lambda\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)} = \lambda\cdot\det(A).\end{aligned}$$

Comme $M_i(\lambda)$ est une matrice diagonale, on obtient $\det(M_i(\lambda)) = \lambda$ et donc $\lambda\cdot\det(A) = \det(M_i(\lambda))\cdot\det(A)$.

(ii).

$$\begin{aligned}\det(G_{ij}(\lambda)A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\left(\left(a_{j,\sigma(j)} + \lambda a_{i,\sigma(j)}\right)\prod_{k\neq j} a_{k,\sigma(k)}\right)\\ &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{k=1}^{n} a_{k,\sigma(k)} + \lambda\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\left(a_{i,\sigma(j)}\prod_{k\neq j} a_{k,\sigma(k)}\right).\end{aligned}$$

On remarque que le second terme correspond au déterminant de la matrice A avec la ligne j remplacée par la ligne i. Par le lemme 6.10, ce déterminant est nul est, ainsi, $\det(G_{ij}(\lambda)A) = \det(A)$. Comme $G_{ij}(\lambda)$ est une matrice triangulaire inférieure et tous ses éléments diagonaux sont 1, on obtient $\det(G_{ij}(\lambda)) = 1$ et donc $\det(A) = \det(G_{ij}(\lambda))\cdot\det(A)$. Le raisonnement est similaire pour $G_{ij}^{\mathsf{T}}(\lambda)$.

(iii). L'équation $\det(P_{ij}A) = -\det(A)$ est montrée comme dans la preuve du lemme 6.10. L'équation $\det(P_{ij})\cdot\det(A) = -\det(A)$ découle du $\det(P_{ij}) = \operatorname{sgn}(\tau) = -1$. ■

Le lemme 6.11 nous dit que le déterminant d'un produit d'une matrice élémentaire et une matrice quelconque est égale au produit des déterminants de ces deux matrices. En fait, cette propriété est vraie en général.

Théorème 6.12 *Soit $(K,+,\cdot)$ un anneau commutatif et soient $A,B \in M_{n\times n}(K)$. Alors*

$$\det(AB) = \det(A)\cdot\det(B).$$

DÉMONSTRATION. La démonstration ici utilise la forme échelonnée (réduite). Dans ce but, on suppose que K soit un corps. Mais le résultat du théorème reste vrai si K est un anneau commutatif. [8]

On sait d'après le théorème 3.13 qu'il existe des matrices élémentaires E_i, $i = 1,\dots,m$, telles que $\tilde{A} := E_m E_{m-1} \cdots E_1 A$ est sous forme échelonnée réduite. Comme l'inverse d'une matrice élémentaire est encore une matrice élémentaire, le lemme 6.11 donne

$$\begin{array}{rcl} \det(A) & = & \det\left(E_m^{-1}\cdots E_1^{-1}\tilde{A}\right) = \det(E_m^{-1})\cdot\det\left(E_{m-1}^{-1}\cdots E_1^{-1}\tilde{A}\right) \\ & = & \cdots = \det(E_m^{-1})\cdots\det(E_1^{-1})\det(\tilde{A}) \end{array} \tag{6.5}$$

ainsi que

$$\det(AB) = \det(E_m^{-1})\cdots\det(E_1^{-1})\det(\tilde{A}B). \tag{6.6}$$

Si A n'est pas inversible, les dernières lignes de $\tilde{A}$ et $\tilde{A}B$ sont nulles, alors $\det(\tilde{A}) = \det(\tilde{A}B) = 0$. Par (6.5)–(6.6), on obtient $\det(AB) = 0 = \det(A) = \det(A)\cdot\det(B)$. Si A est inversible, $\tilde{A} = I$ (voir le théorème 3.13) et donc l'assertion découle directement de (6.5)–(6.6). ■

Corollaire 6.13 *Soit $A \in M_{n\times n}(K)$, $P \in M_{n\times n}(K)$, dont P soit inversible. Alors*

(i) $\det(P)$ *est inversible et* $\det(P^{-1}) = (\det(P))^{-1}$,

(ii) $\det\left(P^{-1}AP\right) = \det(A)$.

DÉMONSTRATION. (i). Par le théorème 6.12, $1 = \det(I) = \det\left(PP^{-1}\right) = \det(P)\cdot\det\left(P^{-1}\right)$ et $1 = \det(I) = \det\left(P^{-1}P\right) = \det\left(P^{-1}\right)\cdot\det(P)$. Alors $\det(P) \in K$ est inversible et son inverse est donné par $\det\left(P^{-1}\right)$.

(ii).

$$\det\left(P^{-1}AP\right) = \det\left(P^{-1}\right)\cdot\det(A)\cdot\det(P) = \frac{\det(P)}{\det(P)}\det(A) = \det(A).$$

■

Remarque 6.14 *Soit V un K-espace vectoriel de dimension finie et soit $F : V \to V$ une application linéaire (i.e., un endomorphisme). On choisit n'importe quelle base $\mathscr{B}_V$ et on définit*

$$\det(F) := \det\left([F]_{\mathscr{V},\mathscr{V}}\right). \tag{6.7}$$

Cette définition est indépendante du choix de la base. En effet soit $\hat{\mathscr{B}}_V$ une autre base. Par la relation (5.13) *et le corollaire 6.13, on a*

$$\det\left([F]_{\hat{\mathscr{B}}_V,\hat{\mathscr{B}}_V}\right) = \det\left([I]_{\mathscr{B}_V,\hat{\mathscr{B}}_V}\right)\cdot\det\left([F]_{\mathscr{B}_V,\mathscr{B}_V}\right)\cdot\det\left([I]^{-1}_{\mathscr{B}_V,\hat{\mathscr{B}}_V}\right) = \det\left([F]_{\mathscr{B}_V,\mathscr{B}_V}\right).$$

Corollaire 6.15 *Soient $A_{11} \in K^{n_1\times n_1}$, $A_{12} \in K^{n_1\times n_2}$, $A_{22} \in K^{n_2\times n_2}$. Alors*

$$\det\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} = \det(A_{11})\cdot\det(A_{22}).$$

DÉMONSTRATION. Exercice. ■

8. Voir, par exemple, la section 5.13 dans le livre Nicholas Loehr, *Advanced Linear Algebra*, CRC press, 2014.

6.4 Comatrice et formules de Laplace

Soit K un anneau commutatif et soit $A \in M_{n\times n}(K)$, $n \geq 2$. On note $A(k,\ell) \in M_{(n-1)\times(n-1)}(K)$ la matrice obtenue de A en suppriment la ligne k et la colonne ℓ de A. Par exemple, pour

$$A = \begin{pmatrix} 16 & 2 & 3 & 13 \\ 5 & 11 & 10 & 8 \\ 9 & 7 & 6 & 12 \\ 4 & 14 & 15 & 1 \end{pmatrix}, \tag{6.8}$$

on obtient

$$A(2,3) = \begin{pmatrix} 16 & 2 & \cancel{3} & 13 \\ \cancel{5} & \cancel{11} & \cancel{10} & \cancel{8} \\ 9 & 7 & \cancel{6} & 12 \\ 4 & 14 & \cancel{15} & 1 \end{pmatrix} = \begin{pmatrix} 16 & 2 & 13 \\ 9 & 7 & 12 \\ 4 & 14 & 1 \end{pmatrix}.$$

Définition 6.16 *Soit $A \in M_{n\times n}(K)$ et $n \geq 2$. La* ***comatrice*** *[cofactor matrix] de A, notée $\mathsf{com}(A)$, est la matrice $B = \mathsf{com}(A) \in M_{n\times n}(K)$ dont les éléments sont donnés par*

$$b_{ij} = (-1)^{i+j} \det\big(A(i,j)\big), \qquad i,j = 1,\ldots,n. \tag{6.9}$$

Si $n = 1$ on pose $\mathsf{com}(A) = 1$. Les coefficients b_{ij} de $\mathsf{com}(A)$ sont appelés ***cofacteurs*** *de A.*

Les facteurs $(-1)^{i+j}$ dans (6.9) se peuvent lire dans la « matrice échiquier »

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

La comatrice de la matrice A de (6.8) est donnée par

$$\mathsf{com}(A) = \begin{pmatrix} -136 & -408 & 408 & 136 \\ -408 & -1224 & 1224 & 408 \\ 408 & 1224 & -1224 & -408 \\ 136 & 408 & -408 & -136 \end{pmatrix}.$$

On observe que $\mathsf{com}(A)^{\mathsf{T}} A = A\,\mathsf{com}(A)^{\mathsf{T}} = 0$ pour cette matrice (non inversible). En général, on a le résultat suivant.

Théorème 6.17 *Soit K un anneau commutatif et soit $A \in M_{n\times n}(K)$. Alors*

$$\mathsf{com}(A)^{\mathsf{T}} A = A\,\mathsf{com}(A)^{\mathsf{T}} = \det(A)\cdot I_n.$$

DÉMONSTRATION. On a évidemment le résultat pour $n = 1$. Alors, soit $n \geq 2$.

Soit $A = \big(a_1, a_2, \ldots, a_n\big)$, où $a_1,\ldots,a_n \in K^{n\times 1}$. Pour $1 \leq i \leq n$ et $1 \leq j \leq n$ on considère la matrice

$$B(i,j,\alpha) = \big(\; a_1 \quad \ldots \quad a_{j-1} \;\big|\; \alpha e_i \;\big|\; a_{j+1} \quad \cdots \quad a_n \;\big), \qquad \alpha \in K.$$

En posant

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & n \\ i & 1 & \cdots & i-2 & i-1 & i+1 & \cdots & n \end{pmatrix},$$
$$\mu = \begin{pmatrix} 1 & 2 & \cdots & j-1 & j & j+1 & \cdots & n \\ j & 1 & \cdots & j-2 & j-1 & j+1 & \cdots & n \end{pmatrix},$$

on obtient

$$P_\pi B(i,j,\alpha)P_\mu^T = \begin{pmatrix} \alpha & \star \\ 0 & A(i,j) \end{pmatrix}, \quad \det(P_\pi) = (-1)^{i-1}, \quad \det(P_\mu) = (-1)^{j-1}.$$

Par le lemme 6.9 et le corollaire 6.15, on a

$$\det\big(B(i,j,\alpha)\big) = \alpha \det\big(A(i,j)\big)\det(P_\pi)\det(P_\mu) = \alpha(-1)^{i+j}\det\big(A(i,j)\big).$$

D'après la définition de comatrice on a pour la matrice $C = \mathrm{com}(A)^{\mathsf{T}}A$ que

$$c_{jk} = \sum_{i=1}^{n}(-1)^{i+j}\det\big(A(i,j)\big)a_{ik} = \sum_{i=1}^{n}\det\big(B(i,j,a_{ik})\big). \tag{6.10}$$

D'après la multlinéarité du déterminant par rapport aux colonnes, on a

$$c_{jk} = \det\big(\ a_1 \quad \dots \quad a_{j-1} \ \big|\ a_k \ \big|\ a_{j+1} \quad \cdots \quad a_n\ \big).$$

Pour $j \neq k$ la matrice à droite a deux colonnes identiques et donc $c_{jk} = 0$. Pour $j = k$ la matrice à droite est égale à A et donc $c_{jj} = \det(A)$. Alors $\mathrm{com}(A)^{\mathsf{T}}A = \det(A)\cdot I_n$.

Pour la relation $A\,\mathrm{com}(A)^{\mathsf{T}} = \det(A)\cdot I_n$ on utilise

$$A\,\mathrm{com}(A)^{\mathsf{T}} = \big(\mathrm{com}(A)A^{\mathsf{T}}\big)^{\mathsf{T}} = \big(\mathrm{com}(A^{\mathsf{T}})^{\mathsf{T}}A^{\mathsf{T}}\big)^{\mathsf{T}} = \big(\det(A^{\mathsf{T}})\cdot I_n\big)^{\mathsf{T}} = \det(A)\cdot I_n,$$

où l'on a utilisé que $\mathrm{com}(A)^{\mathsf{T}} = \mathrm{com}(A^{\mathsf{T}})$, ce qui se vérfie facilement. ■

Comme corollaire du théorème 6.17 on obtient la preuve du lemme 1.35.

Corollaire 6.18 *Soit K un anneau commutatif et $A \in M_{n\times n}(K)$. Les énoncés suivants sont équivalents :*

i) *A est inversible.*

ii) *$\det(A)$ est inversible.*

iii) *Il existe une matrice $X \in M_{n\times n}(K)$ telle que $AX = I_n$.*

iv) *Il existe une matrice $X \in M_{n\times n}(K)$ telle que $XA = I_n$.*

DÉMONSTRATION. i) $\Rightarrow$ ii) Voir le corollaire 6.13.

ii) $\Rightarrow$ i) Par le théorème 6.17, la matrice $(\det(A))^{-1}\mathrm{com}(A)^{\mathsf{T}}$ est l'inverse de A.

iii) $\Rightarrow$ ii) On suppose que $AX = I_n$. D'après le théorème 6.12, $\det(A)\det(X) = \det(AX) = 1$. Alors, $\det(A)$ est inversible. On montre de façon analogue iv) $\Rightarrow$ ii).

Les implications i) $\Rightarrow$ iii) et i) $\Rightarrow$ iv) sont triviales. ■

Une autre conséquence du théorème 6.17 est la formule de Laplace, qui permet de bien calculer le déterminant, surtout si des nombreux coefficients sont nuls.

Corollaire 6.19 (Formules de Laplace) *Soit $A \in M_{n\times n}(K)$ et $n \geq 2$. Alors, on a*

(i) le développement de A par rapport à la i-ième ligne pour $1 \leq i \leq n$:

$$\det(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det\big(A(i,j)\big),$$

(ii) le développement de A par rapport à la j-ième colonne pour $1 \leq j \leq n$:

$$\det(A) = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det\big(A(i,j)\big).$$

DÉMONSTRATION. (i). Le i-ième élément diagonal de la relation $\det(A)\cdot I_n = A\,\mathrm{com}(A)^{\mathsf{T}}$ donne

$$\det(A) = \sum_{j=1}^{n} a_{ij}\big(\mathrm{com}(A)\big)_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det\big(A(i,j)\big).$$

(ii). Le j-ième élément diagonal de la relation $\det(A)\cdot I_n = \mathrm{com}(A)^{\mathsf{T}} A$ donne

$$\det(A) = \sum_{i=1}^{n} \big(\mathrm{com}(A)\big)_{ij} a_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det\big(A(i,j)\big).$$

■

Exemple 6.20 Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 1 & 4 \end{pmatrix}$. Le développement par rapport à la première ligne :

$$\begin{aligned} \det A &= (-1)^2 \det A(1,1) + (-1)^3 2\cdot \det A(1,2) + (-1)^4 3\cdot \det A(1,3) \\ &= \det\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} - 2\det\begin{pmatrix} 0 & 2 \\ 3 & 4 \end{pmatrix} + 3\det\begin{pmatrix} 0 & 1 \\ 3 & 1 \end{pmatrix} \\ &= 2 - 2\cdot(-6) + 3\cdot(-3) = 5. \end{aligned}$$

Le développement par rapport à la première colonne :

$$\begin{aligned} \det A &= (-1)^2 \det A(1,1) + (-1)^3 \cdot 0 \cdot \det A(2,1) + (-1)^4 3 \cdot \det A(3,1) \\ &= \det\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} + 3\det\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = 2 + 3 = 5. \end{aligned}$$

Le développement par rapport à la deuxième ligne :

$$\begin{aligned} \det A &= (-1)^3 \cdot 0 \cdot \det A(2,1) + (-1)^4 \cdot 1 \cdot \det A(2,2) + (-1)^5 \cdot 2 \cdot \det A(2,3) \\ &= \det\begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix} - 2\det\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = -5 + 10 = 5. \end{aligned}$$

♦

Théorème 6.21 (Règle de Cramer) *Soit $A = (a_1, a_2, \ldots, a_n) \in M_{n\times n}(K)$ une matrice inversible. Considerérons le système linéaire $Ax = b$ pour $b \in K^n$. Alors le solution du sysème est donnée par $x = (x_1, \ldots, x_n)^{\mathsf{T}}$ avec*

$$x_i = \frac{\det\begin{pmatrix} a_1 & \ldots & a_{i-1} & b & a_{i+1} & \cdots & a_n \end{pmatrix}}{\det(A)}, \quad 1 \le i \le n. \tag{6.11}$$

DÉMONSTRATION. Exercice. ■

La règle de Cramer n'est pas utilisée numériquement, parce qu'elle est trop coûteuse et amène des erreurs d'arrondi catastrophiques.

6.5 Aspects pratiques

6.5.1 Calculer des déterminants

La commande MATLAB `det` calcule le déterminant d'une matrice $A \in M_{n\times n}(K)$, pour $K = \mathbb{R}$ ou $K = \mathbb{C}$, à l'aide de la décomposition LU

$$PA = LU,$$

où P est une matrice de permutation, L est une matrice triangulaire inférieure, dont tous les éléments diagonaux sont égaux à 1, et U est une matrice triangulaire supérieure. Le calcul de cette décomposition est proche du calcul de la forme échélonnée (vous en parlerez en détail en analyse numérique). En fait, la matrice de permutation P est composée des transpositions utilisées dans la réduction à la forme échélonnée. Alors on a que

$$\det(A) = \pm u_{11}u_{22}\cdots u_{nn},$$

où on choisit le signe $+$ si le nombre de transpositions est pair et -1 sinon.

6.5.2 Le déterminant et l'inversibilité

Le corollaire 6.18 *pourrait* suggérer que le valeur absolue du déterminant soit une bonne mesure de la « bonne » inversibilité d'une matrice réelle. Par exemple, on pourrait conclure d'un petit déterminant que la matrice est proche d'être singulière. On met en garde contre telles conjectures ! En effet, la documentation de la commande MATLAB `det` dit :

```
DET    Determinant.
   DET(X) is the determinant of the square matrix X.

   Use COND instead of DET to test for matrix singularity.
```

(Le commande `COND` calcule le nombre de conditionnement d'une matrice, que l'on va voir en analyse numérique.)

Par exemple, on considère la matrice $T_n = \frac{1}{2}I_n$. Cette matrice est bien inversible, $T^{-1} = 2I_n$. Mais le déterminant devient très petit pour des grandes valeurs de n : $\det(T_{100}) = 2^{-100} \approx 8\cdot 10^{-31}$. D'un autre coté on considère

$$W_n = \begin{pmatrix} 1 & -1 & \cdots & -1 \\ & 1 & \ddots & \vdots \\ & & \ddots & -1 \\ & & & 1 \end{pmatrix} \in M_{n\times n}(K).$$

Malgré le fait que le déterminant se comporte bien ($\det(W_n) = 1$), cette matrice est (numériquement) difficile à inverser. L'élément $(1,n)$ de W_n^{-1} est donné par 2^{n-2}, c'est $\approx 3\cdot 10^{29}$ pour $n = 100$. En faisant une petite modification de la dernière ligne de W_n, on obtient la matrice *singulière* suivante :

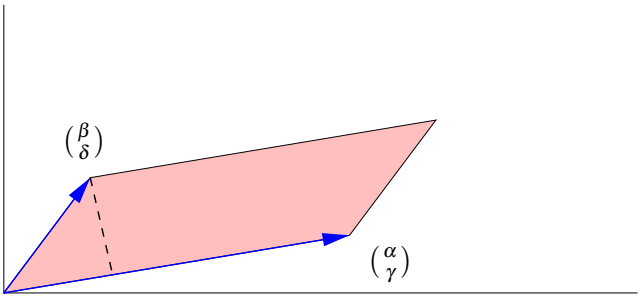
$$\widetilde{W}_n = \begin{pmatrix} 1 & -1 & \cdots & -1 \\ & 1 & \ddots & \vdots \\ & & \ddots & -1 \\ \alpha & \cdots & \alpha & 1 \end{pmatrix}, \quad \alpha = -\frac{1}{2^{n-1}-1}.$$

Pour $n = 100$, la différence entre les coefficients de la matrice inversible W_n et ceux de la matrice singulière $\tilde{W}_n$ est seulement $\approx 2\cdot 10^{-30}$.

6.5.3 Interprétation géometrique du déterminant

Il y a une relation intime entre les déterminants et les volumes de parallélépipèdes dans $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^4$, …. On peut facilement voir cette relation pour $\mathbb{R}^2$. Soit $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. On

considère le parallélogramme déterminé par les colonnes de A :



On note l'aire du parallélogramme (produit de la longueur de la base par la hauteur)

$$\text{Aire}\Big\{\binom{\alpha}{\gamma},\binom{\beta}{\delta}\Big\}.$$

Dans le cas particulier $\tilde{A} = \begin{pmatrix} \tilde{\alpha} & \tilde{\beta} \\ 0 & \tilde{\delta} \end{pmatrix}$ la base se situe sur l'axe des abscisses et on obtient que

$$\text{Aire}\Big\{\binom{\tilde{\alpha}}{0},\binom{\tilde{\beta}}{\tilde{\delta}}\Big\} = |\tilde{\alpha}|\,|\tilde{\delta}| = \left|\det\begin{pmatrix} \tilde{\alpha} & \tilde{\beta} \\ 0 & \tilde{\delta} \end{pmatrix}\right| = |\det(\tilde{A})|. \tag{6.12}$$

Pour tout vecteur $\binom{\alpha}{\gamma}$ il existe une rotation $Q = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$ telle que $Q\binom{\alpha}{\gamma} = \binom{\tilde{\alpha}}{0}$, $\tilde{\alpha} \in \mathbb{R}$. On peut montrer que l'aire du parallélogramme n'est pas modifiée par la rotation :

$$\text{Aire}\Big\{\binom{\alpha}{\gamma},\binom{\beta}{\delta}\Big\} = \text{Aire}\Big\{\binom{\tilde{\alpha}}{0},\binom{\tilde{\beta}}{\tilde{\delta}}\Big\}, \quad \text{où} \quad \binom{\tilde{\beta}}{\tilde{\delta}} := Q\binom{\beta}{\delta}.$$

Par (6.12) on obtient

$$\text{Aire}\Big\{\binom{\alpha}{\gamma},\binom{\beta}{\delta}\Big\} = |\det(\tilde{A})| = |\det(QA)| = |\det(Q)|\,|\det(A)| = |\det(A)|, \tag{6.13}$$

où on a utilisé que $\det(Q) = \cos^2\phi + \sin^2\phi = 1$.

La relation (6.13) se generalise en dimension quelconque :

Soit $A \in M_{d\times d}(\mathbb{R})$.

Alors le volume du parallélépipède engendré par les colonnes de A est égal à $|\det(A)|$.

On peut montrer ce résultat par la décomposition QR (en algèbre linéaire 2).

Il existe une autre interprétation de (6.13) : Un carré de longueur de coté ℓ est engendré par $\binom{\ell}{0}$, $\binom{0}{\ell}$, de plus son aire est égale à ℓ^2. En transformant ce carré par la matrice $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ on obtient le parallélépipède engendré par $\binom{\ell\alpha}{\ell\gamma},\binom{\ell\beta}{\ell\delta}$ avec une aire de $|\det(A)|\ell^2$. Alors la valeur absolue du déterminant, $|\det(A)|$, décrit la modification de l'aire par la multiplication par A. On peut généraliser ce résultat à domaines quelconques $\Omega \subset \mathbb{R}^d$:

$$\widetilde{\Omega} = \{Ax :\ x \in \Omega\} \quad \Rightarrow \quad \text{Volume}(\widetilde{\Omega}) = |\det(A)| \times \text{Volume}(\Omega).$$

Cette formule joue un rôle important dans le contexte de l'intégration.

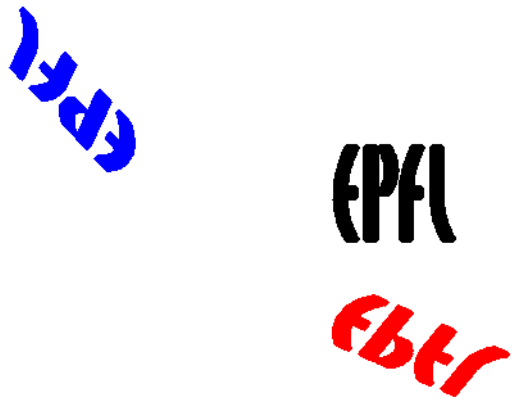


FIG. 6.1 – Le logo EPFL (en noir), transformé par les matrices A_1 (en bleu) et A_2 (en rouge) de (6.14).

Le signe du déterminant $\det(A)$ signifie un change d'orientation. Dans la figure 6.1 on voit les transformations par les matrices

$$A_1 = \begin{pmatrix} -1 & -1/4 \\ 1 & -1/2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1/4 & 1 \\ 1 & 1/2 \end{pmatrix}, \tag{6.14}$$

avec $\det(A_1) = 3/4$ et $\det(A_2) = -9/8$.

Chapitre 7

Valeurs propres et vecteurs propres

7.1 Définitions

Définition 7.1 *Soit $(K,+,\cdot)$ un corps et $A \in M_{n\times n}(K)$. Un scalaire $\lambda \in K$ s'appelle une* ***valeur propre*** *[eigenvalue] de A s'il existe un vecteur $x \in K^n$, $x \neq 0$, tel que*

$$Ax = \lambda x.$$

Le vecteur x s'appelle un ***vecteur propre*** *[eigenvector] de A associé à la valeur propre λ.*

Définition 7.2 *Soit V un K-espace vectoriel et $F \in L(V,V)$. Un scalaire $\lambda \in K$ s'appelle une valeur propre de F s'il existe un vecteur $v \in V$, $v \neq 0$, tel que $f(v) = \lambda v$. Le vecteur v s'appelle un vecteur propre de F associé à la valeur propre λ.*

Définition 7.3 *L'ensemble des valeurs propres de A (resp. de F) s'appelle le* ***spectre*** *[spectrum] de A (resp. de f), noté* $\mathsf{spec}(A)$ *(resp.* $\mathsf{spec}(F)$*).*

Exemple 7.4

1. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$. Alors
   - $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_1 = 1$,
   - $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_2 = 0$,
   - $v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ n'est pas un vecteur propre.
2. Soit $A = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ pour $\phi \in \mathbb{R}$.
   - Si $\phi \neq k\pi$, $k \in \mathbb{N}$, alors A n'a pas de valeur propre (réelle).
   - Si $\phi = (2k+1)\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ a une valeur propre $\lambda = -1$ et tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.
   - Si $\phi = 2k\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ a une valeur propre $\lambda = 1$ et encore tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.

   On va voire que si on considère A comme une matrice complexe, alors on a toujours les valeurs propres $\cos\phi + i\sin\phi$ et $\cos\phi - i\sin\phi$.

3. Soit $V = C^\infty(\mathbb{R})$ et $D^2 : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $f \mapsto D^2 f := f''$. Pour tout $k \in \mathbb{Z}$ les fonctions $g_k(\xi) = \cos(k\xi)$ et $h_k(\xi) = \sin(k\xi)$ (si $k \neq 0$) sont des vecteurs (fonctions) propres de D^2 associés à la valeur propre $\lambda_k = -k^2$.

♦

Si x est un vecteur propre de A et $\beta \in K \setminus \{0\}$, alors

$$Ax = \lambda x \quad \Rightarrow \quad A(\beta x) = \beta Ax = \beta \lambda x = \lambda(\beta x),$$

c-à-d βx est aussi un vecteur propre. De même pour un vecteur propre v d'un endomorphisme F.

7.2 Le polynôme caractéristique

Soit $A \in M_{n\times n}(K)$ et soit $x \in K^n$ un vecteur propre de A associé à la valeur propre $\lambda \in K$. Alors, puisque $x \neq 0$,

$$\begin{aligned} & Ax = \lambda x \\ \Leftrightarrow\ & (\lambda I - A)x = 0 \\ \Leftrightarrow\ & \lambda I - A \text{ n'est pas inversible} \\ \Leftrightarrow\ & \det(\lambda I - A) = 0. \end{aligned}$$

Donc les valeurs propres sont les zéros de la fonction $t \mapsto \det(tI - A)$ dans K.

La matrice $tI - A$ prend la forme

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & t - a_{nn} \end{pmatrix} \in M_{n\times n}(K[t]).$$

Comme $K[t]$ est un anneau, le déterminant de $tI - A$ est aussi dans $K[t]$.

Définition 7.5 *Le **polynôme caractéristique** [characteristic polynomial] d'une matrice $A \in M_{n\times n}(K)$ est le polynôme $p_A(t) := \det(tI - A)$.*

Pour $n = 1$ le polynôme caractéristique de $A = (a_{11})$ est donné par $p_A(t) = t - a_{11}$. Pour $n = 2$

$$p_A(t) = \det \begin{pmatrix} t - a_{11} & a_{12} \\ a_{21} & t - a_{22} \end{pmatrix} = t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}).$$

En général on n'a pas des formules simples pour les coefficients de ce polynôme. On peut néanmoins caractériser quelques termes.

Lemme 7.6 *Soit $A \in M_{n\times n}(K)$. Alors, le polynôme caractéristique p_A de A est de degré n et*

$$p_A(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1}t^{n-1} + t^n,$$

où

$$\alpha_{n-1} = -(a_{11} + a_{22} + \cdots + a_{nn}), \qquad \alpha_0 = (-1)^n \det(A).$$

DÉMONSTRATION. On définit le **symbole de Kronecker**

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

On a alors

$$p_A(t) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)\prod_{i=1}^{n}\left(\delta_{i,\sigma(i)}t - a_{i,\sigma(i)}\right) \tag{7.1}$$

et donc

$$\alpha_0 = p_A(0) = (-1)^n \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)} = (-1)^n \det(A).$$

Ensuite on remarque que l'on peut écrire (7.1) comme

$$p_A(t) = \prod_{i=1}^{n}(t-a_{ii}) + \sum_{\substack{\sigma\in S_n\\ \sigma\neq \mathrm{id}}} \mathrm{sgn}(\sigma)\prod_{i=1}^{n}\left(\delta_{i,\sigma(i)}t - a_{i,\sigma(i)}\right), \tag{7.2}$$

où le premier terme prend la forme

$$\prod_{i=1}^{n}(t-a_{ii}) = t^n - (a_{11}+\cdots+a_{nn})t^{n-1} + \text{polynôme de degré} \leq n-2.$$

Le deuxième terme à droite de (7.2) est un polynôme de degré inférieur à $n-2$, car si $\sigma \neq \mathrm{id}$ il existe au moins un couple i,j tel que $i\neq j$, $\sigma(i)\neq i$ et $\sigma(j)\neq j$. Alors on peut choisir au plus $n-2$ éléments diagonaux de $tI-A$. Donc

$$p_A(t) = t^n - (a_{11}+\cdots+a_{nn})t^{n-1} + \text{polynôme de degré} \leq n-2.$$

■

La somme des termes diagonaux d'une matrice $A\in M_{n\times n}(K)$ s'appelle la **trace** de A, i.e.,

$$\mathrm{trace}(A) := a_{11}+a_{22}+\cdots+a_{nn}.$$

Le lemme suivant va dans la direction opposé. Etant donné un polynôme unitaire p il trouve une matrice A telle que p est le polynôme caractéristique de A.

Lemme 7.7 *Soit $p\in K[t]$ de la forme $p = \alpha_0+\alpha_1 t+\cdots+\alpha_{n-1}t^{n-1}+t^n \in K[t]$. Alors p est le polynôme caractéristique de la matrice*

$$A = \begin{pmatrix} 0 & & & -\alpha_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -\alpha_{n-2} \\ & & 1 & -\alpha_{n-1} \end{pmatrix}.$$

DÉMONSTRATION. Exercice. ■

La matrice A du lemme 7.7 s'appelle la **matrice compagnon** [*companion matrice*] de p.

D'après la caractérisation

$$\lambda \text{ v.p. de } A\in M_{n\times n}(K) \quad \Leftrightarrow \quad p_A(\lambda I - A) = 0,$$

pour déterminer le spectre de A il faut déterminer les zéros de son polynôme caractéristique. En général ($K=\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$), il n'est pas possible de trouver une formule explicite pour les zéros d'un polynôme de degré 5 ou plus (théorie de Galois) et on doit recourir aux méthodes numériques.

Par le théorème fondamental de l'algèbre (voir le théorème 2.46) une matrice complexe de taille $n\times n$ a n valeurs propres (comptées avec leurs multiplicités).

Corollaire 7.8 *Soit $A\in M_{n\times n}(K)$ une matrice triangulaire (inférieure ou supérieure). Alors les valeurs propres de A sont les éléments diagonaux de A.*

DÉMONSTRATION. Découle directement du lemme 6.5 :

$$p_A(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}).$$

■

7.2.1 Théorème de Hamilton-Cayley

Soit K un anneau commutatif et soit $p \in K[t]$ un polynôme à coefficients dans R,

$$p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t_n, \qquad \alpha_0, \ldots, \alpha_n \in R.$$

Soit $A \in M_{n\times n}(K)$. On peut alors évaluer p en A de la manière suivante :

$$p(A) = \alpha_0 I_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_n A^n \in M_{n\times n}(R),$$

où

$$A^j = \underbrace{A \cdot A \cdots A}_{j \text{ fois}}.$$

On pose $A^0 := I_n$. Comme le déterminant est defini pour des matrices sur un anneau, on peut définir le polynôme charactéristique de $A \in M_{n\times n}(R)$ comme $p_A(t) = \det(tI_n - A)$ et donc $p_A \in R[t]$.

Théorème 7.9 (Hamilton-Cayley) *Soit $A \in M_{n\times n}(R)$ et p_A le polynôme charactéristique de A. Alors $p_A(A) = 0$.*

DÉMONSTRATION.[9] Pour $n = 1, A = \alpha \in R$ et $p_A(A) = \det(\alpha \cdot 1 - \alpha) = 0$. Soit alors $n \geq 2$. Pour une matrices fixe $A \in M_{n\times n}(R)$ on considère l'ensemble

$$R[A] := \{p(A) : p \in R[t]\} \subset M_{n\times n}(R),$$

contenant tous les polynômes evalués en A. On vérifie que cet ensemble (muni de l'addition et la multiplication matricielle) est un anneau commutatif. On définit

$$\mathscr{A} := \begin{pmatrix} A - a_{11}I_n & -a_{21}I_n & \cdots & -a_{n1}I_n \\ -a_{12}I_n & A - a_{22}I_n & \cdots & -a_{n2}I_n \\ \vdots & \vdots & & \vdots \\ -a_{1n}I_n & -a_{2n}I_n & \cdots & A - a_{nn}I_n \end{pmatrix} \in M_{n\times n}\big(M_{n\times n}(R)\big).$$

Le déterminant de $\mathscr{A}$ (par rapport à l'anneau $R[A]$!) est donné par

$$\det(\mathscr{A}) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \big(\delta_{i,\sigma(i)} A - a_{\sigma(i),i} I_n\big) = p_{A^\mathsf{T}}(A) = p_A(A) \in R[A].$$

Soit $x = \begin{pmatrix} e_1^\mathsf{T} & e_2^\mathsf{T} & \cdots & e_n^\mathsf{T} \end{pmatrix}^\mathsf{T} \in R^{n^2}$. Alors

$$\begin{aligned} \begin{pmatrix} A - a_{11}I_n & -a_{21}I_n & \cdots & -a_{n1}I_n \end{pmatrix} x &= (A - a_{11}I_n)e_1 - a_{21}I_n e_2 + \cdots + -a_{n1}I_n e_n \\ &= Ae_1 - a_{11}e_1 - a_{21}e_2 + \cdots + a_{n1}e_n \\ &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} - \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = 0 \end{aligned}$$

9. [Higham, N. Functions of matrices. SIAM, 2008] (S. 7): *It is incorrect to prove the Cayley-Hamilton theorem by $p_A(A) = \det(A * I - A) = 0$.*

et, ainsi, la première ligne bloc de $\mathscr{A}$ multipliée par x donne 0. De façon similaire, la i-ième ligne bloc de $\mathscr{A}$ multipliée par x donne 0 pour $i = 2,\ldots,n$. Alors

$$\mathscr{A}x = 0, \tag{7.3}$$

mais on doit faire attention qu'on voit $\mathscr{A}$ comme élément de $M_{n^2\times n^2}(R)$ (et non de $M_{n\times n}\big(M_{n\times n}(R)\big)$) dans le produit $\mathscr{A}x$. Par le Théorème 6.17, on a la formule

$$\mathsf{com}(\mathscr{A})^{\mathsf{T}}\mathscr{A} = \begin{pmatrix} \det(\mathscr{A}) & & \\ & \ddots & \\ & & \det(\mathscr{A}) \end{pmatrix}.$$

En regardant les deux côtés de cette égalité comme des éléments de $M_{n^2\times n^2}(R)$ et utilisant (7.3) on obtient

$$\begin{pmatrix} \det(\mathscr{A})e_1 \\ \vdots \\ \det(\mathscr{A})e_n \end{pmatrix} = \mathsf{com}(\mathscr{A})^{\mathsf{T}}\mathscr{A}x = 0.$$

Alors chaque colonne de la matrice $\det(\mathscr{A})$ est nulle et, ainsi, $p_A(A) = \det(\mathscr{A}) = 0$. ■

Exemple 7.10 Le polynôme caractéristique de $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ est $p_A(t) = (t-1)(t-2)$. On a

$$p_A(A) = (A - I_n)(A - 2I_n) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Pour la matrices $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, on a bien sur que $p_A(A) = 0$ pour $p_A(t) = (t-2)^2$. Mais il existe un polynôme de degré 1, $q(t) = t-2$, tel que $q(A) = 0$. Le **polynôme minimal** de A est le polynôme unitaire de degré minimum parmi ceux qui annulent A, c-à-d $q(A) = 0$. ♦

Les resultats suivants donnent des utilisations typiques du théorème 7.9.

Corollaire 7.11 *Soit $A \in M_{n\times n}(R)$.*

(i) Toute puissance A^k avec $k \in \mathbb{N}$ peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \ldots, A^{n-1}$.

(ii) Si A est inversible, alors l'inverse A^{-1} peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \ldots, A^{n-1}$.

DÉMONSTRATION. (i). Trivialement, l'assertion est vraie pour $k = 0,1,\ldots,n-1$. On montre le cas $k = n$. Par le théorème 7.9 :

$$0 = p_A(A) = \alpha_0 I + \alpha_1 A \cdots + \alpha_{n-1}A^{n-1} + A^n \quad \Rightarrow \quad A^n = -\alpha_0 I - \alpha_1 A \cdots - \alpha_{n-1}A^{n-1}.$$

De façon similaire, on montre le cas $k > n$ par récurrence, utilisant $0 = A^{k-n}p_A(A)$.

(ii). Si A est inversible alors $\alpha_0 = (-1)^n \det(A)$ est inversible. De $0 = p_A(A)$ on obtient que

$$I = -\frac{\alpha_1}{\alpha_0}A \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-1} - \frac{1}{\alpha_0}A^n = A\Big(-\frac{\alpha_1}{\alpha_0}I \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-2} - \frac{1}{\alpha_0}A^{n-1}\Big)$$

et donc $A^{-1} = -\frac{\alpha_1}{\alpha_0}I \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-2} - \frac{1}{\alpha_0}A^{n-1}$. ■

7.2.2 Matrices semblables, endomorphismes

Deux matrices semblables ont le même spectre.

Théorème 7.12 *Soit $A \in M_{n\times n}(K)$ et $P \in M_{n\times n}(K)$ inversible. Alors*

(i) les spectres de A et celui de $P^{-1}AP$ sont identiques,

(ii) $x \in K^n$ est un vecteur propre de A si et seulement si $P^{-1}x$ est un vecteur propre de $P^{-1}AP$.

DÉMONSTRATION. (i). Posons $B = P^{-1}AP$,

$$\begin{aligned} p_B(t) &= \det(tI - B) = \det(tI - P^{-1}AP) = \det\left(P^{-1}(tI - A)P\right) \\ &= \underbrace{\det\left(P^{-1}\right)}_{=1/\det P} \det(tI - A) \det P = \det(tI - A) = p_A(t), \end{aligned}$$

donc les racines de p_B et p_A sont les mêmes et donc les spectres de A et B sont identiques.

(ii). Comme P^{-1} est inversible, on remarque que x est non nul si et seulement si $P^{-1}x$ est non nul.

Si $x \in K^n \setminus \{0\}$ est un vecteur propre de A, il existe $\lambda \in K$ tel que

$$Ax = \lambda x \quad \text{donc} \quad P^{-1}A\underbrace{PP^{-1}}_{=I}x = \lambda P^{-1}x \quad \text{donc} \quad BP^{-1}x = \lambda P^{-1}x.$$

Réciproquement si $P^{-1}x$ est un vecteur propre de $P^{-1}AP$, alors $P^{-1}Ax = \lambda P^{-1}x$ et donc $Ax = \lambda x$. ■

Définition 7.13 *Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. On définit le polynôme caractéristique de f comme le polynôme caractéristique de la matrice de l'application linéaire dans une certaine base.*

D'après le théorème 7.12 la définition 7.13 est bien indépendante du choix de la base : Soient $\mathscr{B}_V$, $\tilde{\mathscr{B}}_V$ deux bases de V. Comme

$$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_V} = [I]_{\mathscr{B}_V,\tilde{\mathscr{B}}_V} \cdot [F]_{\mathscr{B}_V,\mathscr{B}_V} \cdot [I]^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V},$$

$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_V}$ et $[F]_{\mathscr{B}_V,\mathscr{B}_V}$ ont le même polynôme caractéristique.

Théorème 7.14 *Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. Soit $\mathscr{B}_V$ une base de V. Alors si $\lambda \in \mathsf{spec}(F)$ associé au vecteur propre $v \in V$ alors $\lambda \in \mathsf{spec}\left([F]_{\mathscr{B}_V,\mathscr{B}_V}\right)$ associé au vecteur propre $[v]_{\mathscr{B}_V} \in K^n$. Réciproquement si $\lambda \in \mathsf{spec}\left([F]_{\mathscr{B}_V,\mathscr{B}_V}\right)$ associé au vecteur propre $x \in K^n$, alors $\lambda \in \mathsf{spec}(F)$ associé au vecteur propre $[x]^{-1}_{\mathscr{B}_V} \in V$.*

DÉMONSTRATION. Découle des équivalences suivantes avec $[v]_{\mathscr{B}_V} = x$:

$$\begin{aligned} & F(v) = \lambda v \\ \Leftrightarrow\ & \left(F \circ [\cdot]^{-1}_{\mathscr{B}_V}\right)(x) = \lambda [x]^{-1}_{\mathscr{B}_V} \\ \Leftrightarrow\ & \left([\cdot]_{\mathscr{B}_V} \circ F \circ [\cdot]^{-1}_{\mathscr{B}_V}\right)(x) = \lambda x \\ \Leftrightarrow\ & [F]_{\mathscr{B}_V,\mathscr{B}_V} x = \lambda x. \end{aligned}$$

■

Exemple 7.15 Soit $V = \mathbb{R}_3[t]$ et $D \in L(V,V)$, $D : p \mapsto p'$. Par l'exemple 5.18 on connaît déjà la matrice de l'application linéaire D par rapport à $\mathscr{B} = \{1,t,t^2,t^3\}$:

$$[D]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La seule valeur propre de cette matrice est $\lambda = 0$ associé au vecteur propre $x = (1,0,0,0)^\mathsf{T}$ (et à ses multiples scalaires). Par le théorème 7.14, $\lambda = 0$ est la seule valeur propre de D et tout vecteur propre et un polynôme constant non nul. ♦

Définition 7.16 *Soit $A \in M_{n\times n}(K)$ et $\lambda \in \mathsf{spec}(A)$. **L'espace propre** [eigenspace] associé à λ est défini par*

$$E_\lambda(A) = \{x \in K^n \,|\, Ax = \lambda x\}.$$

Soit V un K-espace vectoriel de dimension finie, $F \in L(V,V)$, et $\lambda \in \mathsf{spec}(F)$. L'espace propre associé à λ est défini par

$$E_\lambda(F) = \{v \in V \,|\, F(v) = \lambda v\}.$$

On voit facilement que $E_\lambda(A)$ et $E_\lambda(F)$ sont des sous-espaces vectoriels de K^n et V respectivement. Par définition

$$E_\lambda(A) = \mathrm{Ker}(\lambda I_n - A), \qquad E_\lambda(F) = \mathrm{Ker}(\lambda \cdot \mathsf{id} - F),$$

donc, par le théorème du rang,

$$\dim E_\lambda(A) = n - \mathsf{rang}(\lambda I_n - A), \qquad \dim E_\lambda(F) = \dim V - \mathsf{rang}(\lambda \cdot \mathsf{id} - F).$$

7.3 Diagonalisation

Le but de cette section est de trouver une matrice P inversible telle que $P^{-1}AP$ est diagonale. Ce n'est pas toujours possible, même si $K = \mathbb{C}$. Par exemple, soit

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{C}), \quad P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \in M_{2\times 2}(\mathbb{C}), \tag{7.4}$$

où P est inversible. Supposons qu'il existe une matrice diagonale $D = \mathsf{diag}\,(d_{11},d_{22})$ telle que $P^{-1}AP = D$. Alors, $d_{11} = d_{22} = 0$ car D et A ont les mêmes valeurs propres par le lemme 7.12. Mais c'est une contradiction : $0 \neq A = PDP^{-1} = 0$.

Définition 7.17 *Soit $A \in M_{n\times n}(K)$. On dit que A est **diagonalisable** s'il existe une matrice $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP$ est diagonal.*

Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. On dit que F est diagonalisable si $[F]_{\mathscr{B}_V,\mathscr{B}_V}$ est diagonalisable, où $\mathscr{B}_V$ est une base quelconque de V.

Si A est diagonalisable les éléments diagonaux de $P^{-1}AP$ sont les valeurs propres de A : $P^{-1}AP = \mathsf{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

Dans ce qui suit on va déduire une caractérisation des matrices diagonalisables. Le résultat suivant joue un rôle clé.

Théorème 7.18 *$A \in M_{n\times n}(K)$ est diagonalisable si et seulement s'il existe une base de K^n formée de vecteurs propres de A.*

DÉMONSTRATION. Supposons A est diagonalisable, alors il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP = \text{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n)$. En notant $P = \big(x_1,x_2,\ldots,x_n\big)$, cette rélation est équivalente à

$$\begin{aligned} & AP = P\cdot \text{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n) \\ \Leftrightarrow\ & A\big(x_1,x_2,\ldots,x_n\big) = \big(x_1,x_2,\ldots,x_n\big)\cdot \text{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n) \\ \Leftrightarrow\ & Ax_1 = \lambda_1 x_1,\, Ax_2 = \lambda_2 x_2,\ \ldots,\, Ax_n = \lambda_n x_n. \end{aligned}$$

Comme $x_1,x_2,\ldots,x_n$ est une famille linéairement indépendante, on a bien une base de vecteurs propres de K^n.

Réciproquement, supposons qu'il existe une base de K^n $\{x_1,x_2,\ldots,x_n\}$ constituée de vecteurs propres de A, i.e., $Ax_i = \lambda_i x_i$ pour $i = 1,\ldots,n$. Alors on pose $P = \big(x_1,x_2,\ldots,x_n\big)$. P est inversible et, par les équivalences ci-dessus, on obtient que $P^{-1}AP = \text{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n)$. ■

Lemme 7.19 *Soit $A \in M_{n\times n}(K)$ et $\lambda_1,\ldots,\lambda_m \in K$ des valeurs propres distinctes de A associées aux vecteurs propres $x_1,\ldots,x_m \in K^n$. Alors $(x_1,\ldots,x_m)$ est une famille libre de K^n.*

DÉMONSTRATION. Par récurrence sur m. Pour $m = 1$ le résultat est vrai car un vecteur propre est non nul. Soit alors $m \geq 2$ et supposons l'assertion vraie pour $m-1$. Considérons une combinaison linéaire

$$0 = \alpha_1 x_1 + \cdots + \alpha_{m-1}x_{m-1} + \alpha_m x_m, \quad \alpha_1,\ldots,\alpha_m \in K. \tag{7.5}$$

Alors

$$0 = \alpha_1 Ax_1 + \cdots + \alpha_{m-1}Ax_{m-1} + \alpha_m Ax_m = \alpha_1\lambda_1 x_1 + \cdots + \alpha_{m-1}\lambda_{m-1}x_{m-1} + \alpha_m\lambda_m x_m.$$

En ajoutant la relation (7.5) multipliée par $-\lambda_m$

$$0 = \alpha_1(\lambda_1 - \lambda_m)x_1 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)x_{m-1}.$$

Par hypothèse de récurrence, $(x_1,\ldots,x_{m-1})$ est libre et donc $\alpha_i(\lambda_i - \lambda_m) = 0$ pour $i = 1,\ldots,m-1$. Comme $\lambda_i \neq \lambda_m$ on obtient

$$\alpha_1 = \cdots = \alpha_{m-1} = 0.$$

Par suite comme $x_m \neq 0$, l'équation (7.5) montre que $\alpha_m = 0$. ■

La même démonstration montre aussi le résultat suivant : Soit V un K-espace vectoriel de dimension finie, $f \in L(V,V)$ et $\lambda_1,\ldots,\lambda_r \in K$ des valeurs propres distinctes de f associées aux vecteurs propres $v_1,\ldots,v_m \in V$. Alors $(v_1,\ldots,v_m)$ est une famille libre de V.

Corollaire 7.20 *Soit $A \in M_{n\times n}(K)$ et $\lambda_1,\ldots,\lambda_r$ les valeurs propres distinctes de A. Alors A est diagonalisable si et seulement si*

$$E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus \cdots \oplus E_{\lambda_r}(A) = K^n.$$

DÉMONSTRATION. Par le théorème 4.38, $E_{\lambda_1}(A) \oplus \cdots \oplus E_{\lambda_r}(A) = K^n$ si et seulement si il existe une base de K^n qui est la réunion des bases de $E_{\lambda_i}(A)$, $i = 1,\ldots,r$. Alors l'assertion découle du théorème 7.18. ■

Corollaire 7.21 *Soit $A \in M_{n\times n}(K)$ et supposons que A possède n valeurs propres distinctes. Alors A est diagonalisable.*

DÉMONSTRATION. Découle directement du lemme 7.19 et du théorème 7.18. ■

Avoir n valeurs propres distinctes est une condition suffisante pour être diagonalisable mais *pas* nécéssaire. Par exemple, la matrice identité I_n possède une seule valeur propre et elle est évidemment diagonalisable.

Soit $A \in M_{n\times n}(K)$ et $p_A(t) = \det(tI_n - A)$. Pour $\lambda \in \mathsf{spec}(A)$ on a $p_A(\lambda_i) = 0$, alors on peut écrire

$$p_A(t) = (t-\lambda)^m g(t) \quad \text{avec} \quad g(\lambda) \neq 0.$$

L'entier m s'appelle la multiplicité du zéro λ de p_A.

Définition 7.22 *La **multiplicité algébrique** d'une valeur propre $\lambda \in \mathsf{spec}(A)$ est la multiplicité de λ comme zéro de p_A. On la note $m_{\text{alg}}(\lambda)$.*

*La **multiplicité géometrique** d'une valeur propre $\lambda \in \mathsf{spec}(A)$ est la dimension de $E_\lambda(A)$. On la note $m_{\text{géom}}(\lambda)$.*

On a les mêmes définitions pour un endomorphisme $f \in L(V,V)$.

Exemple 7.23 Soient

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}), \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}),$$

$$C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Pour la valeur propre 2 de A, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 1$. Pour la valeur propre 2 de B, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 2$. Pour la valeur propre 2 de C, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 3$. ♦

Lemme 7.24 *Soit $\lambda \in \mathsf{spec}(A)$. Alors $m_{\text{géom}}(\lambda) \leq m_{\text{alg}}(\lambda)$.*

DÉMONSTRATION. Soit $m = m_{\text{géom}}(\lambda)$ et $(y_1,\dots,y_m) \subset K^n$ une base de $E_\lambda(A)$. On complète cette base en une base de K^n : $(y_1,\dots,y_m,y_{m+1},\dots,y_n)$. On pose la matrice inversible $Y = \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}$. Alors

$$\begin{aligned} AY = A\begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} &= \begin{pmatrix} Ay_1 & \cdots & Ay_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda y_1 & \cdots & \lambda y_m & Ay_{m+1} & \cdots & Ay_n \end{pmatrix} \end{aligned}$$

et par suite

$$P^{-1}AP = \begin{pmatrix} \lambda I_m & B_{12} \\ 0 & B_{22} \end{pmatrix}.$$

Par le théorème 7.12 et le corollaire 6.15, on obtient que

$$\det(tI - A) = \det(tI_m - \lambda I_m)\det(tI - B_{22}) = (t-\lambda)^m \det(tI - B_{22}).$$

Donc $m_{\text{alg}}(\lambda) \geq m$. ■

Théorème 7.25 (théorème de diagonalisation) *Soit $A \in K^{n\times n}$. Alors A est diagonalisable si et seulement si*

(i) p_A est scindé (sur K), et

(ii) $m_{\text{alg}}(\lambda) = m_{\text{géom}}(\lambda)$ pour tout $\lambda \in \mathsf{spec}(A)$.

DÉMONSTRATION. Supposons A diagonalisable, alors il existe P inversible telle que $P^{-1}AP$ est diagonale. En regroupant les valeurs propres identiques on peut supposer que

$$P^{-1}AP = \operatorname{diag}\big(\underbrace{\lambda_1,\ldots,\lambda_1}_{m_{\text{alg}}(\lambda_1)\text{ fois}},\underbrace{\lambda_2,\ldots,\lambda_2}_{m_{\text{alg}}(\lambda_2)\text{ fois}},\ldots,\underbrace{\lambda_r,\ldots,\lambda_r}_{m_{\text{alg}}(\lambda_r)\text{ fois}}\big), \qquad \lambda_i \neq \lambda_j \text{ pour } i \neq j.$$

Alors

$$m_{\text{géom}}(\lambda_i) = \dim\operatorname{Ker}(\lambda_i I - A) = \dim P\cdot\operatorname{Ker}(\lambda_i - P^{-1}AP) = \dim\operatorname{Ker}(\lambda_i - P^{-1}AP) = m_{\text{alg}}(\lambda_i).$$

Réciproquement, supposons que (i) et (ii) sont vraies, alors

$$p_A(t) = \prod_{i=1}^{r}(t-\lambda_i)^{m_{\text{alg}}(\lambda_i)}.$$

Posons $W := E_{\lambda_1}(A)+\cdots+E_{\lambda_r}(A)$, on obtient $W := E_{\lambda_1}(A)\oplus\cdots\oplus E_{\lambda_r}(A)$ par le lemme 7.19. Comme $m_{\text{alg}}(\lambda_i) = m_{\text{géom}}(\lambda_i)$ on a

$$\dim W = \sum_{i=1}^{r}\dim E_{\lambda_i}(A) = \sum_{i=1}^{r} m_{\text{géom}}(\lambda_i) = \sum_{i=1}^{r} m_{\text{alg}}(\lambda_i) = n,$$

et donc $W = K^n$. Par le corollaire 7.20, ceci implique que A est diagonalisable. ■

Le resultat (et la preuve) du théorème 7.25 est anlogue pour $F \in L(V,V)$, V un K-espace vectoriel de dimension finie.

Exemple 7.26 (i). Soit

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

On calcule $p_A(t) = \det(tI - A) = \cdots = (\lambda-1)(\lambda+2)^2$, alors les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = -2$, et p_A est scindé. Par le lemme 7.24, $m_{\text{géom}}(1) = m_{\text{alg}}(1) = 1$. $m_{\text{alg}}(-2) = 2$ et il reste à calculer $m_{\text{géom}}(-2)$. On cherche

$$Ax = -2x \quad\Leftrightarrow\quad (-2I - A)x = 0 \quad\Leftrightarrow\quad \begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Leftrightarrow\quad x_1 + x_2 + x_3 = 0,$$

c-à-d deux variables sont libres et $m_{\text{géom}}(-2) = \dim\operatorname{Ker}(-2I - A) = 2 = m_{\text{alg}}(-2)$. Alors A est diagonalisable.

Si on veut calculer une matrice P telle que $P^{-1}AP$ soit sous forme diagonale, on calcule une base des espaces propres :

$$Ax = \lambda_1 x \quad\Leftrightarrow\quad \begin{pmatrix} 0 & -3 & -3 \\ 3 & 6 & 3 \\ -3 & -3 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

Ceci donne $x_1 = x_3 = -x_2$. Par exemple, on peut choisir $x_2 = -1$, $x_1 = x_3 = 1$, et donc $E_1(A) = \operatorname{span}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Par le calcul ci-dessus, $E_{-2}(A) = \operatorname{span}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$. Finalement, en posant

$$P = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{on a bien} \quad P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

(ii). Soit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \det(t-A) = (t-1)(t-2)^2.$$

Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$, et $m_{\text{géom}}(1) = m_{\text{alg}}(1) = 1$, $m_{\text{alg}}(2) = 2$. Pour la multiplicité géometrique de 2 on calcule

$$Ax = 2x \quad \Leftrightarrow \quad (2I-A)x = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{array}{l} x_1 = 0 \\ x_2 \text{ libre} \\ x_3 = 0. \end{array}$$

Alors $m_{\text{géom}}(2) = \dim \operatorname{Ker}(2I-A) = 1 \neq m_{\text{alg}}(2)$ et donc A n'est pas diagonalisable. ♦

Algorithme pour la diagonalisation *à la main* de $A \in M_{n\times n}(K)$:

1. On calcule $p_A(\lambda)$ et on cherche les racines de ce polynôme.
2. Si on trouve n racines (comptées avec leurs multiplicités) on va à l'étape 3, sinon A n'est pas diagonalisable.
3. Soient $\lambda_1, \ldots, \lambda_r$ les racines distinctes du polynôme scindé $p_A(t) = \prod_{i=1}^r (t-\lambda_i)^{m_i}$m où $m_i = m_{\text{alg}}(\lambda_i)$. On calcule de bases de $E_{\lambda_i}(A) = \operatorname{Ker}(\lambda_i I - A)$ et $m_{\text{géom}}(\lambda_i) = \dim E_{\lambda_i}(A)$. Si $m_{\text{géom}}(\lambda_i) < m_{\text{alg}}(\lambda_i)$ pour un $i \in \{1, \ldots, n\}$, alors A n'est pas diagonalisable.
4. En posant

$$P = \big(\underbrace{p_1, \ldots, p_{m_1}}_{\text{base de } E_{\lambda_1}(A)}, \underbrace{p_{m_1+1}, \ldots, p_{m_1+m_2}}_{\text{base de } E_{\lambda_2}(A)}, \ldots, \underbrace{p_{n-m_r+1}, \ldots, p_n}_{\text{base de } E_{\lambda_r}(A)} \big),$$

on obtient que P est inversible et

$$P^{-1}AP = \operatorname{diag}\big(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_{\text{alg}}(\lambda_1) \text{ fois}}, \underbrace{\lambda_2, \ldots, \lambda_2}_{m_{\text{alg}}(\lambda_2) \text{ fois}}, \ldots, \underbrace{\lambda_r, \ldots, \lambda_r}_{m_{\text{alg}}(\lambda_r) \text{ fois}} \big)$$

Le même algorithme s'applique par la diagonalisation de $F \in L(V,V)$, V de dimension finie. D'abord on choisit une base $\mathscr{B}_V$ de V et on calcule $A = [F]_{\mathscr{B}_V, \mathscr{B}_V}$. Ensuite, on applique l'algorithme ci-dessus à A.

La diagonalisation de matrices (lorsque c'est possible) est un outil puissant. La caractérisation suivante de la commutativité de matrices donne un premièr aperçu de cette puissance.

Lemme 7.27 *Soient $A, B \in M_{n\times n}(K)$ des matrices diagonalisable. Alors $AB = BA$ si et seulement s'il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP$ et $P^{-1}BP$ soient diagonales (on dit dans ce cas que A et B sont* ***simultanément diagonalisables****).*

DÉMONSTRATION. (i). Supposons qu'il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP = \Lambda_A$ et $P^{-1}BP = \Lambda_B$, où Λ_A, Λ_B sont diagonales. Alors

$$AB = P\Lambda_A P^{-1} P\Lambda_B P^{-1} = P\Lambda_A\Lambda_B P^{-1} = P\Lambda_B\Lambda_A P^{-1} = P\Lambda_B P^{-1} P\Lambda_A P^{-1} = BA,$$

où l'on utilisé que deux matrices diagonales commutent.

(ii). Supposons que A et B soient diagonalisables et $AB = BA$. Alors il existe $\tilde{P}$ inversible telle que

$$\tilde{P}^{-1}A\tilde{P} = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} =: \Lambda_A, \quad \text{avec } \lambda_i \neq \lambda_j \text{ pour } i \neq j.$$

Alors comme $\tilde{P}\Lambda_A\tilde{P}^{-1}B = AB = BA = B\tilde{P}\Lambda_A\tilde{P}^{-1}$ on obtient

$$\Lambda_A\tilde{B} = \tilde{B}\Lambda_A, \quad \text{avec} \quad \tilde{B} = \tilde{P}^{-1}B\tilde{P}. \tag{7.6}$$

On écrit $\tilde{B}$ sous forme partitionée $\tilde{B} = \left(B_{ij}\right)_{i,j=1}^{k}$ avec $B_{ij} \in M_{n_i\times n_j}(K)$. Alors d'après (7.6)

$$\begin{pmatrix} \lambda_1 B_{11} & \cdots & \lambda_1 B_{1k} \\ \vdots & & \vdots \\ \lambda_k B_{k1} & \cdots & \lambda_k B_{kk} \end{pmatrix} = \begin{pmatrix} \lambda_1 B_{11} & \cdots & \lambda_k B_{1k} \\ \vdots & & \vdots \\ \lambda_1 B_{k1} & \cdots & \lambda_k B_{kk} \end{pmatrix},$$

donc comme $\lambda_i \neq \lambda_j$ pour $i \neq j$ on obtient $B_{ij} = 0$ pour $i \neq j$. Donc $\tilde{B}$ est une matrice diagonale par bloc. Comme B est diagonalisable, $\tilde{B} = \tilde{P}^{-1}B\tilde{P}$ est diagonalisable et donc les blocs diagonales B_{ii}, $i = 1,\ldots,r$, sont diagonalisables. Soit $P_i \in M_{n_i\times n_i}(K)$ inversible telle que $P_i^{-1}B_{ii}P_i$ soit diagonale pour $i = 1,\ldots,k$. On pose

$$P := \tilde{P}\begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_k \end{pmatrix}.$$

Alors $P^{-1}BP$ est diagonale par construction et

$$P^{-1}AP = \begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_k^{-1} \end{pmatrix}\tilde{P}^{-1}A\tilde{P}\begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_k \end{pmatrix} = \Lambda_A.$$

■

7.4 Dynamique discrète

Dans la section 1.5.2 on a vu la suite de Fibonacci

$$f_0 = 1, \qquad f_1 = 1, \qquad f_{k+2} = f_{k+1} + f_k, \quad k \geq 0, \tag{7.7}$$

qui peut s'écrire comme produit matrice-vecteur. En général, on considère une **suite récurrente linéaire d'ordre** n :

$$f_{k+n} = \alpha_{n-1}f_{k+n-1} + \cdots + \alpha_1 f_{k+1} + \alpha_0 f_k \tag{7.8}$$

où $\alpha_0,\alpha_1,\ldots,\alpha_{n-1}$ sont des scalaries fixés d'un corps K. Pour définir une suite à partir de (7.8) il faut fixer n conditions initiales

$$f_0 = s_0, \quad f_1 = s_1, \quad \ldots, \quad f_{n-1} = s_{n-1}, \tag{7.9}$$

où $s_0,s_1,\ldots,s_{n-1} \in K$ sont des scalaries. En définissant

$$u_k := \begin{pmatrix} f_k \\ f_{k+1} \\ \vdots \\ f_{k+n-1} \end{pmatrix},$$

on voit que (7.8)–(7.9) est équivalente à

$$u_{k+1} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \alpha_0 & \cdots & \alpha_{n-2} & \alpha_{n-1} \end{pmatrix} u_k, \qquad u_0 = \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_{n-1} \end{pmatrix}. \tag{7.10}$$

À part les signes des coefficients α_i, la matrice dans (7.10) est la transposée de la matrice compagnon du lemme 7.7 !

Lemme 7.28 *Soit $A \in M_{n\times n}(K)$ une matrice comme dans* (7.10) *avec $\alpha_0, \alpha_1, \ldots, \alpha_{n-1} \in K$. Alors :*

(i) le polynôme caractéristique de A est $p_A = t^n - \alpha_{n-1}t^{n-1} - \cdots - \alpha_1 t - \alpha_0$.

(ii) Si $\lambda \in K$ est une valeur propre de A alors $\big(1, \lambda, \lambda^2, \ldots, \lambda^{n-1}\big)^\mathsf{T}$ est un vecteur propre associé.

(iii) A est diagonalisable si et seulement si A possède n valeurs propres distinctes.

DÉMONSTRATION. (i) découle du lemme 7.7.

(ii).

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \alpha_0 & \cdots & \alpha_{n-2} & \alpha_{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda^{n-1} \\ \lambda^n - p_A(\lambda) \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \vdots \\ \lambda^{n-2} \\ \lambda^{n-1} \end{pmatrix}.$$

(iii). Par le corollaire 7.21, une matrice avec n valeurs propres distinctes est diagonalisable. Réciproquement, supposons que A soit diagonalisable. Trivialement, les $n-1$ dernières colonnes de la matrice $tI - A$ sont toujours linéairement indépendantes. Si $\lambda \in K$ est une valeur propre on obtient alors que $m_{\text{géom}}(\lambda) = \dim \operatorname{Ker}(\lambda I - A) = 1$. Par le théorème 7.25, $m_{\text{alg}}(\lambda) = m_{\text{géom}}(\lambda) = 1$. Alors les n valeurs propres de A sont distinctes. ■

En notant A la matrice $n \times n$ dans (7.10) supposons que A possède n valeur propres distinctes $\lambda_1, \ldots, \lambda_n \in K$. Alors la matrice des vecteurs propres associés,

$$P = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}, \tag{7.11}$$

est inversible.

En général un **système dynamique linéaire discret** (homogène) prend la forme

$$u_{k+1} = Au_k, \qquad u_0 = s, \tag{7.12}$$

où $A \in M_{n\times n}(K)$, $u_k \in K^n$, et $s \in K^n$ est le vecteur des conditions initiales. Si A est diagonalisable, on trouve une formule explicite des vecteurs $u_k = A^k s$ définies par (7.12). Soit $P = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \in M_{n\times n}(K)$ inversible tel que

$$P^{-1}AP = \operatorname{diag}\big(\lambda_1, \ldots, \lambda_n\big) =: \Lambda.$$

Alors

$$A^k = \big(P\Lambda P^{-1}\big)\big(P\Lambda P^{-1}\big)\cdots\big(P\Lambda P^{-1}\big) = P\Lambda^k P^{-1}.$$

En posant $c := P^{-1}u_0$ la solution de (7.12) peut s'écrire comme suit :

$$u_k = P\Lambda^k c = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n. \quad (7.13)$$

On note que le vecteur c représente les coordonnées des conditions initiales s par rapport à la base des vecteurs propres (les colonnes de P). Si s est un vecteur propre x_j (ou un multiple) on obtient la solution particulière $\lambda_j^k x_j$. D'après (7.13) on a que u_k est en fait une combinaison linéaire de telles solutions.

Exemple 7.29 On considère le système associé à la suite de Fibonacci (7.7) :

$$u_{k+1} = Au_k, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad u_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Les valeurs propres de A sont données par

$$0 = \det(\lambda I - A) = \lambda^2 - \lambda - 1 \quad \Rightarrow \quad \lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2},$$

et d'après (7.11)

$$P = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \quad \Rightarrow \quad c = P^{-1}u_0 = \frac{1}{\lambda_1 - \lambda_2}\begin{pmatrix} 1-\lambda_2 \\ \lambda_1 - 1 \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda_1 \\ -\lambda_2 \end{pmatrix}.$$

Par (7.13) on obtient

$$u_k = \frac{\lambda_1}{\sqrt{5}}\lambda_1^k \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{\lambda_2}{\sqrt{5}}\lambda_2^k \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.$$

Le premier coefficient donne la formule explicite

$$f_k = \frac{1}{\sqrt{5}}(\lambda_1^{k+1} - \lambda_2^{k+1}) = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right).$$

♦

Dans ce qui suit on analyse le comportement asymptotiqe de la solution u_k de (7.12) lorsque $k \to \infty$.

Définition 7.30 *Le* ***rayon spectral*** *[spectral radius] $\rho(A)$ d'une matrice $A \in \mathbb{C}^{n\times n}$ est definie par*

$$\rho(A) := \max\{|\lambda| : \lambda \in \mathbb{C} \text{ est une valeur propre de } A\}.$$

Supposons que A soit diagonalisable. On ordonne les valeurs propres $\lambda_1, \ldots, \lambda_n$ comme suit :

$$\rho(A) = |\lambda_1| = \cdots = |\lambda_m| > |\lambda_{m+1}| \geq \cdots \geq |\lambda_n|. \quad (7.14)$$

Par (7.13),

$$u_k = \rho(A)^k \left(c_1 \frac{\lambda_1^k}{\rho(A)^k} x_1 + \cdots + c_m \frac{\lambda_m^k}{\rho(A)^k} x_m + c_{m+1} \frac{\lambda_{m+1}^k}{\rho(A)^k} x_{m+1} + \cdots + c_n \frac{\lambda_n^k}{\rho(A)^k} x_n \right). \quad (7.15)$$

Les valeur absolues des m premiers termes de cette somme sont constantes (par rapport à k). Les $n-m$ derniers termes convergent vers 0 quand $k \to \infty$. Ainsi la convergence de u_k est determinée par $\rho(A)$.

Lemme 7.31 *Soit $A \in M_{n\times n}(\mathbb{C})$ diagonalisable. Alors les vecteurs u_k définis par* (7.12) *convergent vers* 0 *pour toute condition initiale u_0 si et seulement si $\rho(A) < 1$.*

L'assertion du lemme (7.31) est aussi vraie si A est non diagonalisable; mais il nous manque les outils pour prouver ce fait. Si $\rho(A) > 1$ la formule (7.15) implique que les vecteur u_k divergent (sauf si u_0 est tel que $c_1 = \cdots = c_m = 0$). Dans le cas limite $\rho(A) = 1$ le vecteur u_k peut converger ou diverger. En particulier, si

$$1 = \lambda_1 = \cdots = \lambda_m \tag{7.16}$$

alors

$$u_k \overset{k\to\infty}{\to} u_\infty = c_1 x_1 + \cdots + c_m x_m, \tag{7.17}$$

à condition que A soit diagonalisable. La situation est plus compliquée si A n'est pas diagonalisable. Par exemple,

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

ce qui implique que la suite $u_k = A^k u_0$ peut être non bornée.

7.5 Matrices stochastiques

Soit $P \in M_{n\times n}(\mathbb{R})$ une matrice telle que $p_{ij} \geq 0$ pour $i,j = 1,\ldots,n$ et $\sum_{i=1}^n p_{ij} = 1$ pour $j = 1,\ldots,n$. De telles matrices sont appelées **matrices stochastiques** (colonnes). Ces matrices jouent un rôle important en la théorie de probabilité. Soit $\{X_k\}_{k\geq 0}$ une suite de variables aléatoires discrètes définies sur l'ensemble des éventualités $E = \{1,2,\ldots,n\}$. Alors la **matrice de transition de probabilité** $P^{(k)} \in M_{n\times n}(\mathbb{R})$, donnée par

$$\left(P^{(k)}\right)_{ij} = \mathbb{P}(X_{k+1} = i \,|\, X_k = j),$$

est une matrice stochastique.

Une matrice $A \in M_{n\times n}(\mathbb{R})$ est dite **strictement positive**, notée $A > 0$, si $a_{ij} > 0$ pour tous $i,j = 1,\ldots,n$. Pour deux matrices $A,B \in M_{n\times n}(\mathbb{R})$ on dit que $A > B$ si $A - B > 0$. Pour un vecteur $x \in \mathbb{R}^n$ on dit que x est (strictement) positif, noté $x \geq 0$ ($x > 0$), si $x_i \geq 0$ ($x_i > 0$) pour tous $i = 1,\ldots,n$.

Théorème 7.32 (Perron) *Soit $A \in \mathbb{R}^{n\times n}$ une matrice positive. Alors :*

(i) Le rayon spectral $r = \rho(A)$ est strictement positif.
(ii) r est une valeur propre de A.
(iii) r est la seule valeur propre[10] *de A de module r.*
(iv) $m_{\text{alg}}(r) = 1$.
(v) Il y a un vecteur propre x associé à r tel que x est réel et strictement positif.
(vi) À part les multiples positives de x, il n'y a plus des vecteurs propres de A qui sont réels et positifs.

DÉMONSTRATION. On fait la preuve à condition que A soit diagonalisable. Les résultats du théorème sont vrais si A n'est pas diagonalisable mais il nous manque les moyens d'en prouver.

(i). Raisonnement par l'absurde : Soit $r = 0$. Alors toutes les valeurs propres de A sont nulles et le polynôme caractéristique est $p_A = t^n$. Par le théorème 7.9 on a que $A^n = 0$, mais ceci contredit que A^n est positive.

Dans ce qui suit on peut supposer que $r = 1$, en remplaçant A par $\frac{1}{\rho(A)}A$.

10. On regarde A comme une matrice complexe et prend toutes les valeurs propres complexes ou réelles.

(ii) et (v). Comme $r = \rho(A) = 1$, il y a une valeur propre $\lambda \in \mathbb{C}$ telle que $|\lambda| = 1$. Soit $y \in \mathbb{C}^{n\times n}$ le vecteur propre associé. Pour une matrice $C \in \mathbb{C}^{m\times n}$ on note $|C|$ la matrice dont les coefficients sont $(|C|)_{ij} := |c_{ij}|$. Par l'inégalité triangulaire, on obtient que

$$|y| = |\lambda y| = |Ay| \leq |A| \cdot |y| = A|y|. \tag{7.18}$$

Alors, en posant $z := A|y|$ le vecteur $b := z - |y|$ est positif. Supposons que $b \neq 0$, c-à-d au moins un coefficient de b est strictement positif. Alors $Ab > 0$. Comme $z = A|y| > 0$, il existe $\varepsilon > 0$ tel que $Ab > \varepsilon z$. En substituant $b = z - |y|$ on obtient

$$Bz > z, \quad \text{avec} \quad B := \frac{1}{1+\varepsilon} A.$$

Alors $B^k z > z$ pour tout $k \geq 0$. Mais, c'est une contradiction car $\rho(B) = \rho(A)/(1+\varepsilon) = 1/(1+\varepsilon) < 1$ et ainsi $B^k z \to 0$ par le lemme 7.31. Alors $b \neq 0$ et l'inegalité (7.18) est en fait une egalité. Ceci montre que $x = |y|$ est un vecteur propre réel et positif associé à la valeur propre $r = 1$. En outre, $x = |y| = A|y| > 0$.

(iii). Il reste à montrer que toute valeur propre λ de module 1 est en fait égale à 1. Soit $Ay = \lambda y$ alors $|y| = A|y| > 0$ et

$$\Big|\sum_{j=1}^{n} a_{ij} y_j\Big| = |y_i| = \sum_{j=1}^{n} \big|a_{ij} y_j\big|.$$

L'égalité implique que tous les termes de la somme de gauche ont le même signe. Alors il existe un vecteur $p = (1, p_2, \ldots, p_n)^\mathsf{T}$ tel que $p > 0$ et $y = y_1 p$. De $Ay = \lambda y$, on obtient que

$$\lambda p = Ap = |Ap| = |\lambda p| = p \quad \Rightarrow \quad \lambda = 1.$$

(iv). Comme $\rho(A^\mathsf{T}) = \rho(A) = 1$ il existe $w \in \mathbb{R}^{n\times 1}$ tel que $w > 0$ et $A^\mathsf{T} w = w$. En multipliant w par un scalaire on peut supposer que $w^\mathsf{T} x = 1$. Soit $X_\perp \in \mathbb{R}^{n\times(n-1)}$ telle que les colonnes de $X_\perp$ forment une base de $\mathrm{Ker}(w^\mathsf{T})$. Comme $w^\mathsf{T} x \neq 0$ on a que $x \notin \mathrm{Ker}(w^\perp)$. Alors la matrice $P = (x, X_\perp)$ est inversible est l'inverse est de la forme $P^{-1} = \begin{pmatrix} w^\mathsf{T} \\ W_\perp^\mathsf{T} \end{pmatrix}$ avec une matrice $W_\perp \in \mathbb{R}^{n\times(n-1)}$ telle que $W_\perp^\mathsf{T} x = 0$. Donc

$$P^{-1}AP = \begin{pmatrix} w^\mathsf{T}Ax & w^\mathsf{T}AX_\perp \\ W_\perp^\mathsf{T}Ax & W_\perp^\mathsf{T}AX_\perp \end{pmatrix} = \begin{pmatrix} w^\mathsf{T}x & w^\mathsf{T}X_\perp \\ W_\perp^\mathsf{T}x & W_\perp^\mathsf{T}AX_\perp \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & W_\perp^\mathsf{T}AX_\perp \end{pmatrix}.$$

Supposons que $m_{\mathrm{alg}}(1) > 1$ alors $W_\perp^\mathsf{T} A X_\perp$ possède un valeur propre 1 avec un vecteur propre $\tilde{y}_2$. Alors A possède deux vecteurs propres x et $\tilde{x} = P^{-1}\begin{pmatrix} 1 \\ \tilde{y}_2 \end{pmatrix}$ qui sont linéairement indépendants. Choisissons i tel que $\tilde{x}_i \neq 0$. Le vecteur $y = x - \frac{x_i}{\tilde{x}_i}\tilde{x}$ est non nul (grâce à l'indépendance linéaire de x et $\tilde{x}$) et un vecteur propre associé à 1. Par la démonstration de la partie (ii), ceci implique que $|y|$ est un vecteur propre : $|y| = A|y| > 0$. Mais, c'est une contradiction : le i-ième coefficient de y est nul par construction. Alors $m_{\mathrm{alg}}(1) = 1$.

(vi). Soit (encore) $w > 0$ tel que $w^\mathsf{T} A = w$ et $w^\mathsf{T} x = 1$. Si $\tilde{x} \geq 0$ est un vecteur propre de A associé à une valeur propre λ, $A\tilde{x} = \lambda\tilde{x}$, on obtient que $\lambda = \lambda w^\mathsf{T}\tilde{x} = w^\mathsf{T}A\tilde{x} = w^\mathsf{T}\tilde{x} = 1$. D'après (iv), $m_{\mathrm{géom}}(1) = m_{\mathrm{alg}}(1) = 1$ et ainsi $\tilde{x}$ est un multiple positif de x. ∎

Corollaire 7.33 *Une matrice stochastique $P \in M_{n\times n}(\mathbb{R})$ a une valeur propre égale à 1 et $\rho(P) = 1$. Si, en plus, P est strictement positive alors $m_{\mathrm{alg}}(1) = 1$ il existe un vecteur propre strictement positif associé à 1.*

DÉMONSTRATION. Soit $e = (1, \dots, 1)^\mathsf{T}$ alors $P^\mathsf{T} e = e$ donc $1 \in \mathsf{spec}(P^\mathsf{T}) = \mathsf{spec}(P)$. Soit $\lambda \in \mathbb{C}$ une valeur propre de A^T et soit y un vecteur propre associé. Soit j tel que $y_j \neq 0$. Par $A^\mathsf{T} y = \lambda y$ on obtient que

$$|\lambda||y_j| = \left|\sum_{i=1}^{n} a_{ij} y_i\right| \leq \sum_{i=1}^{n} a_{ij}|y_i| \leq |y_j|$$

et donc $|\lambda| \leq 1$. Alors $\rho(A) = 1$. Le reste du corollaire découle directement du théorème 7.32. ■

Table des matières

0 Systèmes d'équations linéaires **1**

1 Calcul matriciel **5**
1.1 Matrices, vecteurs colonnes, vecteurs lignes 5
1.2 Quelques matrices particulières 6
1.3 Notation 8
1.4 Applications des matrices 8
1.4.1 Images 8
1.4.2 Graphes 9
1.5 Opérations sur les matrices et les vecteurs 11
1.5.1 Opérations élémentaires 11
1.5.2 Multiplication matrice vecteur 12
1.5.3 Produit matriciel 13
1.6 La transposée d'une matrice 17
1.7 Matrices symétriques 18
1.8 L'inverse d'une matrice 19
1.9 Sous-matrices 20

2 Structures algébriques **21**
2.1 Groupes 21
2.1.1 Exemples des groupes 22
2.1.2 Sous-groupes 23
2.1.3 Morphismes, isomorphismes de groupes 25
2.2 Anneaux 25
2.3 Matrices à coefficients dans un anneau 28
2.4 Corps 29
2.5 Le corps des nombres complexes 29
2.5.1 Plan complexe et forme polaire 32
2.5.2 Matrices à coefficients complexes 33
2.6 Corps finis 35
2.7 Polynômes à coefficients dans un corps 36

3 Forme échelonnée **39**
3.1 Matrices élémentaires 39
3.2 Reduction à la forme échelonnée 42
3.3 Matrices équivalentes 46
3.4 Solutions de systèmes linéaires 48

4 Espaces vectoriels **53**
4.1 Définitions . . . 53
4.2 Sous-espaces vectoriels . . . 54
4.3 Indépendance linéaire, bases, dimensions . . . 57
4.3.1 Le cas de dimension infinie . . . 61
4.4 Sommes d'espaces vectoriels . . . 63

5 Applications linéaires **67**
5.1 Définitions et premières propriétés . . . 67
5.1.1 Le théorème du rang . . . 70
5.1.2 Composition des applications linéaires . . . 72
5.2 Coordonées, matrice d'une application . . . 72
5.2.1 Matrice d'une application linéaire . . . 75
5.2.2 Changement de bases . . . 77

6 Déterminants **81**
6.1 Définitions . . . 81
6.2 Propriétés de la signature . . . 83
6.3 Propriétés du déterminant . . . 85
6.4 Comatrice et formules de Laplace . . . 88
6.5 Aspects pratiques . . . 90
6.5.1 Calculer des déterminants . . . 90
6.5.2 Le déterminant et l'inversibilité . . . 91
6.5.3 Interprétation géometrique du déterminant . . . 91

7 Valeurs propres **95**
7.1 Définitions . . . 95
7.2 Le polynôme caractéristique . . . 96
7.2.1 Théorème de Hamilton-Cayley . . . 98
7.2.2 Matrices semblables, endomorphismes . . . 100
7.3 Diagonalisation . . . 101
7.4 Dynamique discrète . . . 106
7.5 Matrices stochastiques . . . 109