Algèbre linéaire dans $\mathbb{R}^n$
Licence de Sciences et Technologie
Deuxième semestre

Rejeb Hadiji, Stéphane Seuret

Année 2010-2011

Table des matières

Chapitre 1

Systèmes linéaires

Au lycée, vous avez appris à résoudre des systèmes de 2, 3 voire 4 équations à 2, 3 ou 4 inconnues. Ce chapitre est consacré à la théorie des systèmes linéaires comportant un nombre arbitraire d'équations et d'inconnues. Il s'agit notamment de présenter une méthode générale de résolution de tels systèmes.

1.1 Quelques exemples élémentaires

Donnons d'abord quelques exemples de résolution de systèmes de deux équations à deux inconnues.

(i) Résolution de $(S_1)\ :\ \begin{cases} 2x+3y=8 & (L_1) \\ x-y=-1 & (L_2) \end{cases}$

Il s'agit de déterminer l'ensemble des couples de réels (x,y) qui satisfont les deux lignes du système (S_1). Pour ce faire, on peut procéder ainsi : on retranche $2(L_2)$ à (L_1), et on obtient le système

$$(S_1')\ :\ \begin{cases} 5y=10 & (L_1') \\ x-y=-1 & (L_2') \end{cases}$$

dont l'ensemble des solutions est le même que celui de (S_1).
La ligne (L_1') est équivalente à $y=2$, et en reportant dans (L_2'), on obtient $x=1$.
En conclusion, (S_1) a pour unique solution le couple $(1,2)$.

(ii) Résolution de $(S_2)\ :\ \begin{cases} 2x-2y=8 & (L_1) \\ x-y=-1 & (L_2) \end{cases}$

Cette fois-ci, si l'on retranche $2(L_2)$ à (L_1), on obtient le système

$$(S_2')\ :\ \begin{cases} 0=10 & (L_1') \\ x-y=-1 & (L_2') \end{cases}$$

La première ligne ne peut jamais être réalisée, et l'on conclut que (S_2) n'a pas de solution.

(iii) Résolution de $(S_3)\ :\ \begin{cases} 2x-2y=-2 & (L_1) \\ x-y=-1 & (L_2) \end{cases}$

On remarque que la première ligne est égale à deux fois la seconde. Par conséquent, le système (S_3) est équivalent à

$$x-y=-1.$$

L'ensemble des couples (x,y) vérifiant cette relation est $\{(y-1,y)\mid y\in\mathbb{R}\}$. Le système (S_3) a donc une infinité de solutions.

Conclusion : D'après ces exemples, pour les systèmes 2×2, au moins trois cas de figure peuvent se présenter : ou bien le système a une seule solution, ou bien il n'en a pas du tout, ou bien il en a une infinité. Nous allons voir dans ce chapitre que même pour des systèmes linéaires plus généraux, il n'y a pas d'autre scénario possible.

Notation : Dans tout ce chapitre, les systèmes considérés sont à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$. Dans un souci de simplification des notations, nous adopterons la convention que le symbole $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. On rappelle que $\mathbb{K}^n$ désigne l'ensemble des n-uplets d'éléments de $\mathbb{K}$, c'est-à-dire l'ensemble des $(u_1,\cdots,u_n)$ avec chaque u_i appartenant à $\mathbb{K}$. Ainsi $\mathbb{R}^2$ est l'ensemble des couples de réels, $\mathbb{C}^3$, l'ensemble des triplets de nombres complexes, etc.

1.2 Définitions

Définition 1.2.1. *Soit $m\in\mathbb{N}^*$ et $n\in\mathbb{N}^*$. On appelle* **système linéaire** *de m équations à n inconnues à coefficients dans $\mathbb{K}$ (ou encore système $m\times n$) tout système d'équations du type*

$$(S)\qquad \left\{\begin{array}{l} a_{11}x_1+\cdots+a_{1n}x_n=b_1\\ \dots\dots\dots\dots\dots\dots\dots\dots\\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m\end{array}\right.$$

avec $a_{ij}\in\mathbb{K}$ pour $1\leq i\leq m$ et $1\leq j\leq n$, et $b_i\in\mathbb{K}$.

Les a_{ij} sont appelés **coefficients** *du système, les x_i,* **inconnues** *et les b_i,* **second membre***.*

On appelle **matrice** *associée au système (S) le tableau de nombres*

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

On dit que A est une matrice à m lignes, n colonnes et à coefficients dans $\mathbb{K}$.

Exemple : La matrice du système (S_1) de la page précédente est $\begin{pmatrix}2 & 3\\ 1 & -1\end{pmatrix}$.

Notation : Lorsqu'il n'y a pas d'ambiguïté sur l'appellation des inconnues, on écrit le système (S) sous la forme condensée suivante :

$$\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1\\ a_{21} & a_{22} & \cdots & a_{2n} & b_2\\ \vdots & & & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m\end{array}$$

Ainsi le système (S_1) peut se récrire :

$$\begin{array}{cc|c} 2 & 3 & 8\\ 1 & -1 & -1\end{array}$$

Définition 1.2.2. *Si $b_1=\cdots=b_m=0$, le système est dit* **homogène***. Pour un système linéaire général (S) de m équations à n inconnues, le système*

$$(S')\qquad \left\{\begin{array}{l} a_{11}x_1+\cdots+a_{1n}x_n=0\\ \dots\dots\dots\dots\dots\dots\dots\dots\\ a_{m1}x_1+\cdots+a_{mn}x_n=0\end{array}\right.$$

est appelé **système homogène associé** *à (S).*

Exemple : Le système homogène associé à (S_1) est le système (S_1') : $\begin{cases} 2x+3y=0 \\ x-y=0 \end{cases}$

Définition 1.2.3. *Soit (S) un système linéaire $m \times n$. On appelle solution de (S) tout n-uplet $(u_1, \cdots, u_n)$ de $\mathbb{K}^n$ tel que*

$$\begin{cases} a_{11}u_1 + \cdots + a_{1n}u_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}u_1 + \cdots + a_{mn}u_n = b_m. \end{cases}$$

Définition 1.2.4. *Deux systèmes (S_1) et (S_2) sont dits équivalents s'ils ont le même ensemble de solutions, c'est-à-dire si toute solution de (S_1) est solution de (S_2) et vice versa.*

Exemple : Les systèmes

$$\begin{cases} x_1 = 1 \\ x_1 - x_2 = 2 \end{cases} \quad \text{et} \quad \begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}$$

sont équivalents car ont tous les deux pour unique solution le couple $(1, -1)$.

Définition 1.2.5. *On dit qu'un système linéaire $m \times n$ est* **carré** *si $m = n$. On dit qu'un système carré est* **triangulaire** *si l'on a*

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i,j) \quad \textit{tel que} \quad i < j \quad \textit{(système triangulaire inférieur)}$$

ou bien

$$a_{ij} = 0 \quad \textit{pour tout couple} \quad (i,j) \quad \textit{tel que} \quad i > j \quad \textit{(système triangulaire supérieur).}$$

Un système triangulaire est dit **à diagonale non nulle** *s'il est triangulaire et si tous les* **termes diagonaux** *a_{ii} sont non nuls.*

Exemple : Le système suivant est triangulaire supérieur à diagonale non nulle :

$$\begin{cases} x_1 + 5x_2 + x_4 = 1 \\ \quad x_2 + x_3 = -5 \\ \quad\quad x_3 - 5x_4 = 0 \\ \quad\quad\quad x_4 = -1 \end{cases}$$

Définition 1.2.6. *On dit qu'un système $m \times n$ est* **échelonné** *s'il est de la forme :*

$$\begin{cases} a_{1j_1}x_{j_1} + \dots\dots\dots\dots\dots\dots\dots + a_{1n}x_n = b_1 \\ \quad a_{2j_2}x_{j_2} + \dots\dots\dots + a_{2n}x_n = b_2 \\ \quad\quad a_{3j_3}x_{j_3} + \dots\dots + a_{3n}x_n = b_3 \\ \quad\quad\quad \dots\dots\dots\dots\dots\dots\dots \\ \quad\quad\quad a_{kj_k}x_{j_k} + \dots + a_{kn}x_n = b_k \\ \quad\quad\quad\quad 0 = b_{k+1} \\ \quad\quad\quad\quad \cdots \\ \quad\quad\quad\quad 0 = b_m \end{cases}$$

où $1 \leq j_1 < j_2 \cdots < j_k \leq n$ avec $1 \leq k \leq m$, et les coefficients a_{mj_m} pour $1 \leq m \leq k$ sont tous non nuls.

Les k premières équations sont appelées **équations principales** *et les inconnues $x_{j_1}, \cdots, x_{j_k}$ sont appelées* **inconnues principales**.

Remarque : Tout système triangulaire supérieur à diagonale non nulle est échelonné : on a $k = n$ et $j_i = i$ pour tout $i \in \{1, \cdots, n\}$.

Exemple : Le système 4×5 suivant est échelonné :

$$\left\{\begin{array}{ccccccccccc} 2x_1 & + & 3x_2 & & & & & + & x_5 & = & 1 \\ & & x_2 & + & x_3 & + & x_4 & - & x_5 & = & 4 \\ & & & & & & & & x_5 & = & 0 \\ & & & & & & & & 0 & = & -3 \end{array}\right.$$

On a $k = 3$, $j_1 = 1$, $j_2 = 2$ et $j_3 = 5$. Les trois premières équations sont les équations principales, et x_1, x_2 et x_5 sont les inconnues principales.

1.3 Résolution des systèmes échelonnés

Nous allons voir que la résolution de tels systèmes est particulièrement aisée. Commençons par traiter le cas particulier des systèmes triangulaires.

1.3.1 Systèmes triangulaires à diagonale non nulle

Proposition 1.3.1. *Soit (S) un système triangulaire supérieur à diagonale non nulle. Alors (S) a une unique solution obtenue par la* **méthode de la remontée** *:*

$$x_n = \frac{b_n}{a_{nn}}, \quad x_{n-1} = \frac{b_{n-1} - a_{n-1n}x_n}{a_{n-1n-1}}, \ \cdots\cdots\cdots, \ x_1 = \frac{b_1 - a_{12}x_2 - \cdots - a_{1n}x_n}{a_{11}}.$$

Preuve : Le système (S) est du type

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & \cdots & + & a_{1n-1}x_{n-1} & + & a_{1n}x_n & = & b_1 \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & a_{n-1n-1}x_{n-1} & + & a_{n-1n}x_n & = & b_{n-1} \\ & & & & & & a_{nn}x_n & = & b_n. \end{array}\right.$$

Comme $a_{nn} \neq 0$, la dernière équation permet de calculer x_n. Comme $a_{n-1n-1} \neq 0$, l'avant-dernière équation donne $x_{n-1} = \frac{b_{n-1} - a_{n-1n}x_n}{a_{n-1n-1}}$. Puis, de proche en proche, comme $a_{kk} \neq 0$,

$$x_k = \frac{b_k - a_{kk+1}x_{k+1} - \cdots - a_{kn}x_n}{a_{kk}}.$$

Comme $x_{k+1}, \cdots, x_n$ ont déjà été calculés, la formule ci-dessus détermine x_k. ■

Exercice : Montrer que les systèmes triangulaires inférieurs peuvent être résolus de façon analogue par la *méthode de la descente*.

1.3.2 Systèmes échelonnés

Considérons un système échelonné général $m \times n$:

$$(S) \qquad \left\{\begin{array}{ccccccccc} a_{1j_1}x_{j_1} & + & \cdots\cdots & \cdots & \cdots & + & a_{1n}x_n & = & b_1 \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & a_{kj_k}x_{j_k} & + & \cdots & + & a_{kn}x_n & = & b_k \\ & & & & & & 0 & = & b_{k+1} \\ & & & & & & 0 & = & b_m. \end{array}\right.$$

1er cas : L'un des b_j avec $j \in \{k+1, \cdots, m\}$ est non nul. Alors (S) n'a pas de solution.

2ème cas : $b_{k+1} = \cdots = b_m = 0$. Alors (S) est équivalent au système (Σ) suivant :

$$(\Sigma) \qquad \left\{ \begin{array}{ccccccc} a_{1j_1}x_{j_1} & + & \dots\dots\dots\dots\dots\dots & + & a_{1n}x_n & = & b_1 \\ & & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & \\ & & a_{kj_k}x_{j_k} \quad + \quad \cdots & + & a_{kn}x_n & = & b_k. \end{array} \right.$$

Ce nouveau système se résout facilement par la méthode de la remontée en considérant les inconnues non principales (x_j avec $j \neq j_i$ pour tout i) comme des paramètres libres. Si $k = n$, le système (Σ) est tout simplement un système triangulaire supérieur à diagonale non nulle, et la proposition 1.3.1 s'applique. Sinon, on doit avoir $k < n$, et on obtient une infinité de solutions $(x_1, \cdots, x_n)$ données par les formules suivantes :

$$\begin{array}{lcl} x_{j_k} & = & \frac{b_k - a_{kj_k+1}x_{j_k+1} - \cdots - a_{kn}x_n}{a_{kj_k}}, \\ \dots\dots\dots\dots & & \dots\dots\dots , \\ x_{j_1} & = & \frac{b_1 - a_{1j_1+1}x_{j_1+1} - \cdots - a_{1n}x_n}{a_{1j_1}}, \end{array}$$

avec x_j pour $j \neq j_i$ choisi arbitrairement dans $\mathbb{K}$.

Exemple : Soit α un paramètre réel. On veut résoudre dans $\mathbb{R}^5$ le système suivant :

$$\left. \begin{array}{ccccr|c} 2 & 3 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & \alpha \end{array} \right.$$

1er cas : $\alpha \neq 0$. Le système n'a pas de solution.
2ème cas : $\alpha = 0$. Le système est équivalent à

$$\left. \begin{array}{ccccr|c} 2 & 3 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{array} \right.$$

Les inconnues principales sont x_1, x_2 et x_5. Les deux autres inconnues x_3 et x_4 sont des paramètres libres. Par la méthode de la remontée, on trouve :

$$\left\{ \begin{array}{l} x_5 = 6, \\ x_2 = 3 - x_4 + x_5 = 9 - x_4, \\ x_1 = \frac{1 - x_5 - 3x_2}{2} = -16 + \frac{3}{2}x_4. \end{array} \right.$$

L'ensemble des solutions de (S) est

$$\mathcal{E} = \left\{ \left(-16 + \frac{3}{2}x_4, 9 - x_4, x_3, x_4, 6 \right) \,\middle|\, x_3 \in \mathbb{R}, x_4 \in \mathbb{R} \right\}.$$

1.4 Méthode du pivot de Gauss

La méthode du pivot de Gauss consiste à transformer un système (S) en un système échelonné équivalent à l'aide de *transformations élémentaires* .

Les transformations élémentaires sont de trois types :
($T1$) Échange de deux lignes du système,
($T2$) Multiplication d'une ligne par un scalaire non nul,
($T3$) Ajout à une ligne d'un multiple d'une autre ligne.

L'importance accordée aux transformations élémentaires est justifiée par le résultat suivant :

Proposition 1.4.1. *Deux systèmes (S) et $(\widetilde{S})$ se déduisant l'un de l'autre par une succession de transformations élémentaires sont équivalents.*

Remarque : Autrement dit, faire des transformations élémentaires ne change pas l'ensemble des solutions d'un système linéaire.

Preuve : Notons $(a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ les coefficients du système (S). Il suffit de vérifier qu'effectuer une transformation élémentaire sur (S) ne modifie pas l'ensemble des solutions.

- Pour $(T1)$ (i.e permutations de deux lignes), c'est évident.
- Pour $(T2)$ c'est également clair : si $(x_1, \cdots, x_n)$ est solution, alors on a pour tout $\alpha \neq 0$ et $i \in \{1, \cdots, m\}$,

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i \iff \alpha a_{i1}x_1 + \cdots + \alpha a_{in}x_n = \alpha b_i.$$

- Reste à vérifier pour $(T3)$. Supposons que l'on ajoute $\alpha(L_{i_1})$ à (L_{i_0}) (avec $i_0 \neq i_1$). Soit $(\widetilde{a}_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ les coefficients du système $(\widetilde{S})$ obtenu après avoir ajouté $\alpha(L_{i_1})$ à la ligne (L_{i_0}). Notons $\mathcal{E}$ (resp. $\widetilde{\mathcal{E}}$) l'ensemble des solutions de (S) (resp. $(\widetilde{S})$). On a

$$\text{(1.1)} \qquad \begin{cases} \widetilde{a}_{ij} = a_{ij} \quad \text{si} \quad i \neq i_0, \\ \widetilde{a}_{i_0j} = a_{i_0j} + \alpha a_{i_1j}. \end{cases}$$

Soit $(u_1, \cdots, u_n)$ une solution de (S). On a par définition

$$\text{(1.2)} \qquad \forall i \in \{1, \cdots, m\},\ a_{i1}u_1 + \cdots + a_{in}u_n = b_i.$$

Comme $\widetilde{a}_{ij} = a_{ij}$ pour $i \neq i_0$, le n-uplet $(u_1, \cdots, u_n)$ satisfait les lignes de $(\widetilde{S})$ distinctes de i_0. En ajoutant α fois l'égalité (1.2) avec $i = i_1$ à l'égalité (1.2) avec $i = i_0$, on trouve

$$(a_{i_01} + \alpha a_{i_11})u_1 + \cdots + (a_{i_0n} + \alpha a_{i_1n})u_n = b_{i_0} + \alpha b_{i_1}$$

qui est exactement la ligne i_0 de $(\widetilde{S})$.
Donc $(u_1, \cdots, u_n)$ est solution de $(\widetilde{S})$. D'où $\mathcal{E} \subset \widetilde{\mathcal{E}}$.
Pour montrer l'inclusion réciproque, il suffit de remarquer que l'on passe de $(\widetilde{S})$ à (S) en retranchant $\alpha(L_{i_1})$ à (L_{i_0}). On reprend alors le raisonnement précédent en échangeant les rôles de (S) et de $(\widetilde{S})$, et en remplaçant α par $-\alpha$.

■

Attention : Lorsque l'on fait des opérations élémentaires, il ne faut pas oublier de les effectuer aussi sur le second membre.

Exemple : Mettre le système suivant sous forme échelonnée :

$$(S) \ : \ \begin{array}{cccc|c} 0 & 4 & 0 & 4 & 2 \\ 2 & 3 & 1 & 0 & 1 \\ 2 & 0 & -1 & 0 & 0 \end{array}$$

$$(S) \iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 \\ 0 & 4 & 0 & 4 & 2 \end{array}\right) \quad \text{(on a échangé } (L_1) \text{ et } (L_3)),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & \frac{1}{2} \end{array}\right) \quad \left(\text{on a multiplié } (L_3) \text{ par } \tfrac{1}{4}\right),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & \frac{1}{2} \end{array}\right) \quad \text{(on a retranché } (L_1) \text{ à } (L_2)),$$

$$\iff \left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 \\ 0 & 0 & -\frac{2}{3} & 1 & \frac{1}{6} \end{array}\right) \quad \left(\text{on a retranché } (\tfrac{1}{3}L_2) \text{ à } (L_3)\right).$$

Le système obtenu après cette succession de transformations élémentaires est *échelonné* (les inconnues principales sont x_1, x_2 et x_3). On peut donc le résoudre par la méthode de la remontée. La proposition 1.4.1 assure que ce nouveau système est équivalent au système initial.

Cette propriété remarquable n'est pas propre à l'exemple étudié : on a le résultat suivant :

Théorème 1.4.2. *Tout système linéaire est équivalent à un système échelonné.*

La preuve de ce résultat repose sur l'algorithme du **pivot de Gauss** qui est une méthode *itérative* permettant de transformer n'importe quel système linéaire en un système échelonné équivalent après un nombre fini de transformations élémentaires.

Algorithme du pivot de Gauss : Soit (S) un système de taille $m \times n$ et de matrice A.

Première itération :

Premier cas : La matrice A est nulle. L'algorithme est alors terminé.

Deuxième cas : $A \neq 0$. Soit j_1 l'indice de la première colonne non nulle.

1*ère étape :* Par permutation de lignes on se ramène au cas où $a_{1j_1} \neq 0$. Le système (S) est donc équivalent à un système de matrice

$$m \text{ lignes} \left\{ \overbrace{\left(\begin{array}{ccc|cc} 0 & \cdots & 0 & a_{1j_1} & * \\ \vdots & & \vdots & * & * \\ 0 & \cdots & 0 & * & * \end{array}\right)}^{n \text{ colonnes}} \right. .$$

2*ème étape :* On veut faire apparaître des 0 dans la colonne j_1 sous le coefficient a_{1j_1}. Pour cela, on retranche $\frac{a_{ij_1}}{a_{1j_1}}(L_1)$ à la ligne (L_i) $(i \geq 2)$. Le système obtenu après ces $m-1$ opérations élémentaires a pour matrice

$$\left(\begin{array}{ccc|c|c} 0 & \cdots & 0 & a_{1j_1} & * \\ \vdots & & \vdots & 0 & * \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & 0 & * \end{array}\right).$$

Itération suivante : On applique la première itération au système $(m-1) \times n$ constitué par les lignes 2 à m du système obtenu à la fin de la première étape.

Fin de l'algorithme : L'algorithme s'arrête au bout d'au plus $m-1$ itérations ou lorsque le sous-système obtenu a toutes ses lignes nulles.

Remarque : Pour un système $m \times n$, l'itération type nécessite environ mn additions, soustractions, multiplications ou divisions. L'algorithme du pivot donne donc un système échelonné au bout d'environ $m^2 n$ opérations (n^3 opérations si le système est carré). Son côté automatique le rend facilement exécutable par un ordinateur. Dans les cas pratiques, la méthode du pivot de Gauss n'est pas forcément la plus rapide. **Il n'est pas interdit de réfléchir avant d'appliquer aveuglement l'algorithme !**

1.5 Structure de l'ensemble des solutions d'un système linéaire

1.5.1 Un exemple

Cherchons à résoudre le système $(S)\ :\ \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 2 & 3 & 1 & 0 & 2 \\ 1 & 0 & 1 & 1 & 4 \end{array}$

On applique l'algorithme du pivot de Gauss :

$$(S) \iff \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & -1 & 1 & 2 & 4 \end{array} \qquad \text{(on a retranché} \quad 2(L_1) \quad \text{à} \quad (L_2), \quad \text{et} \quad (L_1) \quad \text{à} \quad (L_3))\,,$$

$$\iff \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 & 2 \\ 0 & 0 & 2 & 4 & 6 \end{array} \qquad \text{(on a ajouté} \quad (L_2) \quad \text{à} \quad (L_3))\,.$$

Ce dernier système est échelonné et a pour inconnues principales x_1, x_2 et x_3. On a

$$(S) \iff \begin{cases} x_3 = 3 - 2x_4, \\ x_2 = 2 - 2x_4 - x_3 = -1, \\ x_1 = x_4 - x_2 = x_4 + 1. \end{cases}$$

La solution générale du système s'écrit donc

$$\begin{cases} x_1 = x_1^0 + \lambda x_1' \\ x_2 = x_2^0 + \lambda x_2' \\ x_3 = x_3^0 + \lambda x_3' \\ x_4 = x_4^0 + \lambda x_4' \end{cases} \quad \text{avec} \quad \begin{cases} x_1^0 = 1 \\ x_2^0 = -1 \\ x_3^0 = 3 \\ x_4^0 = 0 \end{cases} \quad \text{et} \quad \begin{cases} x_1' = 1 \\ x_2' = 0 \\ x_3' = -2 \\ x_4' = 1 \end{cases}$$

et λ est un paramètre réel quelconque.

On remarque que $(x_1^0, x_2^0, x_3^0, x_4^0)$ est une solution de (S) alors que $(\lambda x_1', \lambda x_2', \lambda x_3', \lambda x_4')$ est la solution générale du système homogène (S') associé à (S). Autrement dit la solution générale de (S) est somme d'une solution particulière de (S) et de la solution générale de (S').

Nous allons voir que cette propriété n'est pas propre à cet exemple : l'ensemble des solutions d'un système linéaire, s'il n'est pas vide, peut toujours s'écrire comme la somme d'une solution particulière et de l'ensemble des solutions du système homogène associé.

1.5.2 Cas des systèmes homogènes

Proposition 1.5.1. *Soit (S') un système homogène $m \times n$ et $\mathcal{E}'$ l'ensemble de ses solutions. Alors*

i) $\mathcal{E}'$ contient $(0, \cdots, 0)$ (et n'est donc pas vide),

ii) $\mathcal{E}'$ est stable par addition : si $(x_1, \cdots, x_n) \in \mathcal{E}'$ et $(y_1, \cdots, y_n) \in \mathcal{E}'$ alors $(x_1+y_1, \cdots, x_n+y_n) \in \mathcal{E}'$, *iii)* *$\mathcal{E}'$ est stable par multiplication par un scalaire : si $(x_1, \cdots, x_n) \in \mathcal{E}'$ et $\lambda \in \mathbb{K}$ alors $(\lambda x_1, \cdots, \lambda x_n) \in \mathcal{E}'$.*

On dit que $\mathcal{E}'$ est un **sous-espace vectoriel** *de $\mathbb{K}^n$.*

Preuve : Le second membre d'un système homogène est une colonne de 0. Il est donc immédiat que $(0, \cdots, 0) \in \mathcal{E}'$.

Supposons que $(x_1, \cdots, x_n)$ et $(y_1, \cdots, y_n)$ soient solutions de (S'). Alors pour tout $i \in \{1, \cdots, m\}$, on a

$$a_{i1}x_1 + \cdots + a_{in}x_n = 0 \quad \text{et} \quad a_{i1}y_1 + \cdots + a_{in}y_n = 0.$$

Donc, en sommant les deux égalités,

$$a_{i1}(x_1 + y_1) + \cdots + a_{in}(x_n + y_n) = 0.$$

En conséquence, $(x_1 + y_1, \cdots, x_n + y_n) \in \mathcal{E}'$.

Il est également clair que pour tout $\lambda \in \mathbb{K}$, on a $a_{i1}\lambda x_1 + \cdots + a_{in}\lambda x_n = 0$. Les points *ii)* et *iii)* sont donc démontrés. ■

1.5.3 Cas général

Proposition 1.5.2. *Soit (S) un système linéaire et (S') le système linéaire homogène associé. Notons $\mathcal{E}$ et $\mathcal{E}'$ leurs ensembles de solutions respectifs. Supposons de plus que $\mathcal{E}$ ne soit pas vide et donnons-nous $(x_1^0, \cdots, x_n^0)$ une solution particulière de (S). Alors*

$$(x_1, \cdots, x_n) \in \mathcal{E} \iff \exists (x'_1, \cdots, x'_n) \in \mathcal{E}' \quad \textit{tel que} \quad (x_1, \cdots, x_n) = (x_1^0 + x'_1, \cdots, x_n^0 + x'_n).$$

Autrement dit, si $\mathcal{E}$ n'est pas vide alors $\mathcal{E}$ est la somme des solutions de $\mathcal{E}'$ et d'une solution particulière de $\mathcal{E}$. On dit que $\mathcal{E}$ est un **sous-espace affine** de $\mathbb{K}^n$.

Preuve : $\Longrightarrow$ Soit $(x_1^0, \cdots, x_n^0)$ une solution particulière de (S). Alors on a pour tout $i \in \{1, \cdots, m\}$,

$$a_{i1}x_1^0 + \cdots + a_{in}x_n^0 = b_i. \tag{1.3}$$

Par ailleurs, $(x_1, \cdots, x_n) \in \mathcal{E}$ si et seulement si pour tout $i \in \{1, \cdots, m\}$,

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i.$$

En soustrayant à (1.3) cette deuxième relation, on trouve

$$a_{i1}(x_1 - x_1^0) + \cdots + a_{in}(x_n - x_n^0) = 0,$$

ce qui signifie que $(x'_1, \cdots, x'_n) \stackrel{\text{déf}}{=} (x_1 - x_1^0, \cdots, x_n - x_n^0)$ est solution du système homogène associé à (S).

$\Longleftarrow$ Supposons que $(x'_1, \cdots, x'_n) \in \mathcal{E}'$. Alors on a $a_{i1}x'_1 + \cdots + a_{in}x'_n = 0$ pour tout $i \in \{1, \cdots, m\}$. En ajoutant l'égalité (1.3), on trouve

$$a_{i1}(x_1^0 + x'_1) + \cdots + (x_n^0 + x'_n) = b_i$$

pour tout $i \in \{1, \cdots, m\}$, et donc $(x_1^0 + x'_1, \cdots, x_n^0 + x'_n)$ est solution de (S). ■

Chapitre 2

Vecteurs

Le but de ce chapitre est de généraliser la notion de vecteur du plan ou de l'espace vue au lycée, à la dimension n.

2.1 Rappels sur les vecteurs du plan

- Approche "géométrique" : un **vecteur du plan** est un "segment orienté". Plus exactement, c'est une classe d'équivalence de segments orientés pour la relation d'équipolence.
- Approche "analytique" : un **vecteur du plan** est caractérisé par ses **composantes** x et y suivant une base de vecteurs $(\vec{\imath}, \vec{\jmath})$ fixée une fois pour toutes.

Dans ce chapitre, nous allons privilégier l'approche *analytique*. De ce fait, un vecteur $\vec{u}$ du plan sera pour nous un couple de réels :

$$\vec{u} = (u_1, u_2) \quad \text{avec} \quad u_1 \ \text{ et } \ u_2 \ \text{ réels,}$$

et on adoptera la même notation $\mathbb{R}^2$ pour désigner l'ensemble des vecteurs du plan ou l'ensemble des couples de réels.

Rappelons quelques opérations classiques sur les vecteurs du plan :

- **addition** : si $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$, le vecteur somme $\vec{u} + \vec{v}$ est égal au vecteur [1] de composantes $(u_1 + v_1, u_2 + v_2)$ que l'on peut noter $\overrightarrow{u + v}$. Le vecteur de composantes nulles est appelé **vecteur nul** et noté $\vec{0}$. Enfin, si les vecteurs $\vec{u}$ et $\vec{v}$ sont tels que $\vec{u} + \vec{v} = \vec{0}$, on note $\vec{v} = -\vec{u}$.
- **multiplication par un réel** : pour tout réel λ et vecteur $\vec{u} = (u_1, u_2)$, on définit le vecteur [2] $\lambda\vec{u} = \overrightarrow{\lambda u} = (\lambda u_1, \lambda u_2)$.

On peut combiner les deux opérations précédentes et obtenir par exemple les relations :

(i) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$,

(ii) $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$,

et bien d'autres encore que nous verrons plus en détails dans la partie suivante.

Lorsqu'il existe un réel λ tel que $\vec{u} = \lambda\vec{v}$, ou $\vec{v} = \lambda\vec{u}$, on dit que les vecteurs $\vec{u}$ et $\vec{v}$ sont **colinéaires**. Dans le cas contraire, on dit qu'ils sont **linéairement indépendants**. Deux vecteurs linéairement indépendants forment une **base** du plan. Nous reviendrons plus précisément sur la signification de ces notions dans les sections suivantes.

---

1. Du fait que la somme de deux vecteurs du plan redonne un vecteur du plan, on dit que l'addition des vecteurs est une **loi interne**.

2. Du fait que la multiplication est effectuée avec un réel et non pas un vecteur, on dit que la multiplication par un réel est une **loi externe**.

Si la multiplication d'un vecteur du plan par un réel redonne un vecteur, on ne peut en revanche pas multiplier deux vecteurs du plan entre eux pour former un troisième vecteur du plan. Il existe cependant une opération qui "ressemble" à un produit de deux vecteurs mais qui donne un réel : il s'agit du **produit scalaire**. Le produit scalaire d'un couple $(\vec{u}, \vec{v})$ de vecteurs du plan est noté $\vec{u} \cdot \vec{v}$ et défini par :

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2.$$

Le produit scalaire permet de définir la notion d'**orthogonalité** : deux vecteurs $\vec{u}$ et $\vec{v}$ sont orthogonaux si

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 = 0.$$

Enfin, à partir du produit scalaire, on peut définir la **norme** $\|\vec{u}\|$ d'un vecteur $\vec{u}$:

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}.$$

Le lecteur n'ignore sans doute pas que toutes ces notions peuvent s'adapter à la dimension 3, c'est-à-dire aux vecteurs de l'espace. Le but de ce chapitre est d'étudier le cas de la dimension n avec n entier supérieur ou égal à 1. Nous verrons en sus que l'on peut sans effort supplémentaire autoriser les vecteurs à avoir des composantes complexes !

2.2 Vecteurs en dimension n

Comme la construction que nous allons faire s'applique indistinctement au cas réel ou au cas complexe, nous adoptons désormais la notation $\mathbb{K}$ pour désigner $\mathbb{R}$ ou $\mathbb{C}$. Qu'ils soient réels ou complexes, les éléments de $\mathbb{K}$ sont appelés **scalaires**.

Définition 2.2.1. *On appelle* **vecteur à** n **composantes** *tout n-uplet $(x_1, \cdots, x_n)$ d'éléments de $\mathbb{K}$. L'ensemble des vecteurs à n composantes (appelés simplement vecteurs en l'absence d'ambiguïté) est noté $\mathbb{K}^n$, et l'on pose $\vec{x} = (x_1, \cdots, x_n)$. Les scalaires x_i sont appelés* **composantes** *du vecteur $\vec{x}$.*

Remarques : 1. Pour éviter tout risque de confusion entre vecteurs et scalaires, nous avons choisi de noter les vecteurs avec des flèches. Il faut savoir cependant que cette convention est loin d'être universelle. La plupart des ouvrages d'algèbre linéaire notent les vecteurs sans flèche. Du reste, les flèches disparaîtront dans les cours de l'an prochain.

2. Deux vecteurs $\vec{x}$ et $\vec{y}$ sont égaux s'ils ont le même nombre de composantes et si $x_i = y_i$ pour tout i.

On munit l'ensemble $\mathbb{K}^n$ d'une **addition** interne " $+$ " définie pour tous vecteurs $\vec{x}$ et $\vec{y}$ par

$$\vec{x} + \vec{y} \overset{\text{déf}}{=} \overrightarrow{x+y} = (x_1 + y_1, \cdots, x_n + y_n),$$

et d'une **multiplication externe** définie pour tout scalaire $\lambda \in \mathbb{K}$ et vecteur $\vec{x} \in \mathbb{K}^n$ par

$$\lambda \vec{x} \overset{\text{déf}}{=} \overrightarrow{\lambda x} = (\lambda x_1, \cdots, \lambda x_n).$$

Pour $\vec{x}$ et $\vec{y}$ vecteurs de $\mathbb{K}^n$, on adopte aussi la notation $-\vec{x} \overset{\text{déf}}{=} (-x_1, \cdots, -x_n)$ et $\vec{x} - \vec{y} \overset{\text{déf}}{=} (x_1 - y_1, \cdots, x_n - y_n)$. Le vecteur de $\mathbb{K}^n$ dont toutes les composantes sont nulles est noté $\vec{0}$.

Le lecteur pourra vérifier par lui-même la proposition suivante.

Proposition 2.2.2. *Pour tous vecteurs $\vec{x}$, $\vec{y}$ et $\vec{z}$ de $\mathbb{K}^n$, on a les relations suivantes :*

(i) $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$,

(ii) $\vec{x} + \vec{0} = \vec{0} + \vec{x} = \vec{x}$,

(iii) $\vec{x} + (-\vec{x}) = (-\vec{x}) + \vec{x} = \vec{0}$,

(iv) $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

Ceux qui ont suivi le module *Arithmétique et groupes* du premier semestre auront reconnu en $(\mathbb{K}^n, +)$ un **groupe commutatif**.

Mais l'ensemble $\mathbb{K}^n$ possède des propriétés supplémentaires relatives à la multiplication externe. En effet, pour tous scalaires λ et μ, et pour tous vecteurs $\vec{x}$ et $\vec{y}$, on a

(v) $\lambda(\vec{x}+\vec{y}) = \lambda\vec{x} + \lambda\vec{y}$,

(vi) $(\lambda+\mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$,

(vii) $\lambda(\mu\vec{x}) = (\lambda\mu)\vec{x}$,

(viii) $(1)\vec{x} = \vec{x}$.

On dit que $\mathbb{K}^n$ est un **espace vectoriel** sur $\mathbb{K}$.

Remarque : Par convention $\mathbb{K}^0$ est l'espace vectoriel “trivial” contenant un seul élément noté $\vec{0}$. C'est le plus “petit”de tous les espaces vectoriels sur $\mathbb{K}$.

L'étude détaillée des espaces vectoriels fera l'objet d'un module du prochain semestre. Nous nous limitons cette année à l'étude des sous-espaces vectoriels de $\mathbb{K}^n$ définis ci-dessous :

Définition 2.2.3. *Soit $X \subset \mathbb{K}^n$. On dit que X est un* **sous-espace vectoriel**[3] *de $\mathbb{K}^n$ si*

(i) X n'est pas vide,

(ii) pour tout couple $(\vec{x}, \vec{y})$ de vecteurs de X, on a $\vec{x}+\vec{y} \in X$,

(iii) pour tout vecteur $\vec{x}$ de X et scalaire λ, on a $\lambda\vec{x} \in X$.

Exemple : D'après le chapitre précédent, l'ensemble des solutions d'un système linéaire homogène de m équations à n inconnues est un sous-espace vectoriel de $\mathbb{K}^n$.

Pour prouver qu'un sous-ensemble X de $\mathbb{K}^n$ est un s.e.v de $\mathbb{K}^n$, on fait généralement appel à la proposition suivante :

Proposition 2.2.4. *$X \subset \mathbb{K}^n$ est un sous-espace vectoriel de $\mathbb{K}^n$ si et seulement si*

(i) X contient $\vec{0}$,

(ii) $\forall\lambda \in \mathbb{K}, \forall\mu \in \mathbb{K}, \forall\vec{x} \in X, \forall\vec{y} \in X, \ \lambda\vec{x} + \mu\vec{y} \in X$.

Preuve : La proposition 2.2.2 assure que tout sous-espace vectoriel vérifie les deux propriétés ci-dessus.

Réciproquement, considérons un sous-ensemble X de $\mathbb{K}^n$ vérifiant les propriétés (i) et (ii). L'ensemble X n'est pas vide car contient $\vec{0}$. De plus, d'après *(ii)*, pour tout couple $(\vec{x}, \vec{y}) \in X \times X$, on a

$$\vec{x} + \vec{y} = 1\,\vec{x} + 1\,\vec{y} \in X,$$

et pour tout scalaire λ, on a

$$\lambda\,\vec{x} = \lambda\,\vec{x} + 1\,\vec{0} \in X.$$

■

Remarque : Pour simplifier, nous avons occulté l'approche géométrique. Comme dans le plan, un vecteur de $\mathbb{R}^n$ peut être vu comme un segment orienté de l'espace affine (c'est-à-dire de l'ensemble des points) à n dimensions. Nous reviendrons sur cet aspect à la fin du cours.

3. ou s.e.v en abrégé

2.3 Familles de vecteurs

2.3.1 Combinaisons linéaires

Introduisons tout d'abord la notion de **famille de vecteurs**.

Définition 2.3.1. *Une* **famille de vecteurs** *de $\mathbb{K}^n$ est une "collection" de vecteurs $\vec{x}_i$ où l'indice i décrit un ensemble non vide I donné et chaque $\vec{x}_i$ est un vecteur de $\mathbb{K}^n$. On utilise la notation $(\vec{x}_i)_{i\in I}$ pour désigner cette famille. Dans le cas très fréquent où $I = \{i_1, \cdots, i_k\}$, on utilisera également la notation $(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$.*

- *Si $J \subset I$, on dit que $(\vec{x}_j)_{j\in J}$ est une* **sous-famille** *de $(\vec{x}_i)_{i\in I}$.*
- *Si $I \subset K$, on dit que $(\vec{x}_k)_{k\in K}$ est une* **sur-famille** *de $(\vec{x}_i)_{i\in I}$.*

Attention : Contrairement aux ensembles de vecteurs, une famille de vecteurs peut contenir *plusieurs fois* le même vecteur.

Remarque 2.3.2. *Il est commode d'étendre la définition ci-dessus au cas où $I = \emptyset$. Par convention, la famille indexée par l'ensemble vide est $\emptyset$. On l'appelle* **famille vide** *de $\mathbb{K}^n$. Bien évidemment, la famille vide est sous-famille de toute famille de vecteurs.*

Dans la suite du cours, on se limite à des familles *finies* de vecteurs. Le plus souvent, ces familles seront indexées par $I = \{1, \cdots, k\}$ et notées $(\vec{x}_1, \cdots, \vec{x}_k)$. On prendra garde à ne pas confondre la famille $(\vec{x}_1, \cdots, \vec{x}_k)$ avec les composantes $(x_1, \cdots, x_k)$ d'un vecteur de $\mathbb{K}^k$.

Définition 2.3.3. *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$. On dit que $\vec{y} \in \mathbb{K}^n$ est* **combinaison linéaire** *de la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ s'il existe un p-uplet $(\lambda_1, \cdots, \lambda_p)$ d'éléments de $\mathbb{K}$ tel que*

$$\vec{y} = \lambda_1\vec{x}_1 + \cdots + \lambda_p\vec{x}_p.$$

Convention : Toute combinaison linéaire de zéro vecteur (i.e de la famille vide) est égale au vecteur nul.

Proposition 2.3.4. *Tout sous-espace vectoriel de $\mathbb{K}^n$ est stable par combinaison linéaire d'un nombre fini de ses vecteurs.*

Preuve : La proposition 2.2.4 montre la stabilité par combinaison linéaire de *deux* vecteurs. Une récurrence élémentaire donne le cas général. ■

2.3.2 Sous-espaces vectoriels engendrés

Proposition 2.3.5. *Soit $(\vec{u}_1, \cdots, \vec{u}_k)$ une famille de vecteurs de $\mathbb{K}^n$. L'ensemble de toutes les combinaisons linéaires des $\vec{u}_i$ est un sous-espace vectoriel de $\mathbb{K}^n$. On l'appelle* **sous-espace vectoriel engendré** *par la famille $(\vec{u}_1, \cdots, \vec{u}_k)$ et on le note $\mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$. C'est le plus petit sous-espace vectoriel contenant tous les vecteurs de la famille $(\vec{u}_1, \cdots, \vec{u}_k)$.*

Preuve : Pour vérifier que $\mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ est un sous-espace vectoriel, on va appliquer la proposition 2.2.4.

- $\mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ contient $\vec{0}$.
- Si $\vec{x}$ et $\vec{y}$ sont deux éléments de $\mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ alors on peut trouver deux k-uplets $(\lambda_1, \cdots, \lambda_k)$ et $(\mu_1, \cdots, \mu_k)$ tels que

$$\vec{x} = \sum_{i=1}^{k} \lambda_i\vec{u}_i \quad \text{et} \quad \vec{y} = \sum_{i=1}^{k} \mu_i\vec{u}_i.$$

Pour tout couple (λ, μ) de $\mathbb{K}^2$, on a donc

$$\lambda\vec{x} + \mu\vec{y} = \sum_{i=1}^{k}(\lambda\lambda_i + \mu\mu_i)\vec{u}_i.$$

Donc $\lambda\vec{x} + \mu\vec{y}$ est bien combinaison linéaire de $(\vec{u}_1, \cdots, \vec{u}_k)$.
Enfin, si X est un sous-espace vectoriel contenant chacun des vecteurs $\vec{u}_i$, il contient toute combinaison linéaire de la famille $(\vec{u}_1, \cdots, \vec{u}_k)$ (cf Prop. 2.3.4) donc $\mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$. ■

Convention : Le sous-espace vectoriel engendré par la famille vide de $\mathbb{K}^n$ est $\{\vec{0}\}$.

Nous laissons au lecteur le soin d'établir le résultat suivant :

Proposition 2.3.6. *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$, et $(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$ une sous-famille de $(\vec{x}_1, \cdots, \vec{x}_p)$. Alors on a*

$$\mathrm{Vect}\,(\vec{x}_{i_1}, \cdots, \vec{x}_{i_k}) \subset \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p).$$

Pour déterminer le sous-espace vectoriel engendré par une famille de vecteurs, on fait souvent appel à la proposition suivante :

Proposition 2.3.7. *Le sous-espace vectoriel engendré par une famille de vecteurs donnée est invariant par les opérations suivantes :*
$(T1)$ *Permutation de deux vecteurs,*
$(T2)$ *Multiplication d'un vecteur par un scalaire non nul,*
$(T3)$ *Ajout à l'un des vecteurs d'une combinaison linéaire des autres vecteurs.*

Preuve : Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs. L'invariance de $\mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$ par les transformations $(T1)$ et $(T2)$ est évidente.

Pour (T3), il suffit de considérer le cas où l'on ajoute un seul vecteur, le cas général suit par récurrence. Quitte à changer l'ordre des vecteurs (ce qui ne change pas les sous-espaces vectoriels engendrés), il suffit de prouver par exemple que pour tout $\alpha \in \mathbb{K}$, on a

$$\mathrm{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p) = \mathrm{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p).$$

Soit $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p)$. Alors il existe $(\lambda_1, \cdots, \lambda_p) \in \mathbb{K}^p$ tel que

$$\vec{y} = \lambda_1(\vec{x}_1 + \alpha\vec{x}_2) + \lambda_2\vec{x}_2 + \cdots + \lambda_p\vec{x}_p.$$

On a donc

$$\vec{y} = \lambda_1\vec{x}_1 + (\lambda_2 + \alpha\lambda_1)\vec{x}_2 + \cdots + \lambda_p\vec{x}_p.$$

et, par conséquent, $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p)$. On a donc montré que

$$\mathrm{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p) \subset \mathrm{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p).$$

Pour montrer l'inclusion réciproque, on considère $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p)$. Il existe donc $(\mu_1, \cdots, \mu_p) \in \mathbb{K}^p$ tel que $\vec{y} = \mu_1\vec{x}_1 + \mu_2\vec{x}_2 + \cdots + \mu_p\vec{x}_p$, ce qui peut se récrire

$$\vec{y} = \mu_1(\vec{x}_1 + \alpha\vec{x}_2) + (\mu_2 - \alpha\mu_1)\vec{x}_2 + \mu_2\vec{x}_3 + \cdots + \mu_p\vec{x}_p.$$

On a donc bien $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1 + \alpha\vec{x}_2, \vec{x}_2, \cdots, \vec{x}_p)$. ■

2.3.3 Familles génératrices, familles liées, familles libres

Définition 2.3.8. *Soit X un sous-espace vectoriel de $\mathbb{K}^n$. On dit qu'une famille de vecteurs $(\vec{x}_1, \cdots, \vec{x}_k)$ de X est* **génératrice** *si tout élément de X est combinaison linéaire de $(\vec{x}_1, \cdots, \vec{x}_k)$.*

Remarque : En d'autres termes, dire que $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice de X signifie que $X = \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$.

Dans la première partie de ce chapitre, nous avons rappelé la définition de colinéarité de deux vecteurs du plan, et de sa négation, l'indépendance linéaire. Dans cette section, nous généralisons ces deux notions au cas d'une famille de k vecteurs de $\mathbb{K}^n$.

Définition 2.3.9. *Soit $(\vec{u}_1, \cdots, \vec{u}_k)$ une famille de $\mathbb{K}^n$.*
- *On dit que $(\vec{u}_1, \cdots, \vec{u}_k)$ est* **liée** *(ou linéairement dépendante) s'il existe un k-uplet $(\lambda_1, \cdots, \lambda_k) \neq (0, \cdots, 0)$ tel que $\displaystyle\sum_{i=1}^{k} \lambda_i \vec{u}_i = \vec{0}$.*
- *On dit que $(\vec{u}_1, \cdots, \vec{u}_k)$ est* **libre** *(ou linéairement indépendante) si elle n'est pas liée.*

Exemples :
- Toute famille contenant le vecteur nul est liée.
- Une famille de deux vecteurs $(\vec{x}_1, \vec{x}_2)$ est liée si et seulement si $\vec{x}_1$ et $\vec{x}_2$ sont **colinéaires**, i.e il existe $\lambda \in \mathbb{K}$ tel que
$$\vec{x}_1 = \lambda \vec{x}_2 \quad \text{ou} \quad \vec{x}_2 = \lambda \vec{x}_1.$$
- Par convention, la famille vide est libre.

Proposition 2.3.10. *Si $(\vec{u}_1, \cdots, \vec{u}_k)$ est libre et $\vec{x} \in \mathrm{Vect}\,(\vec{u}_1, \cdots, \vec{u}_k)$ alors il existe un unique k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}$ tel que $\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{u}_i$.*

Autrement dit,
$$\left(\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{u}_i = \sum_{i=1}^{k} \mu_i \vec{u}_i\right) \implies \left(\lambda_1 = \mu_1, \cdots, \lambda_k = \mu_k\right).$$

Preuve : L'égalité $\sum_{i=1}^{k} \lambda_i \vec{u}_i = \sum_{i=1}^{k} \mu_i \vec{u}_i$ entraîne $\sum_{i=1}^{k} (\lambda_i - \mu_i) \vec{u}_i = \vec{0}$, et donc, puisque $(\vec{u}_1, \cdots, \vec{u}_k)$ est libre, $\lambda_i - \mu_i = 0$ pour tout $i \in \{1, \cdots, k\}$. ■

Proposition 2.3.11. *Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille de vecteurs de $\mathbb{K}^n$, et $\vec{y} \in \mathbb{K}^n$.*

(i) Si $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$ alors la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée.

(ii) Si la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée et si de plus $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre alors $\vec{y}$ appartient à $\mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$.

Preuve : Si $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$ alors il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}$ tel que $\vec{y} = \sum_{i=1}^{k} \lambda_k \vec{x}_k$. On a donc $\sum_{i=1}^{k} \lambda_k \vec{x}_k - \vec{y} = \vec{0}$. Le $(k+1)$-uplet $(\lambda_1, \cdots, \lambda_k, -1)$ n'est pas identiquement nul donc la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée.

Réciproquement, supposons que $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ soit liée et que $(\vec{x}_1, \cdots, \vec{x}_k)$ soit libre. Alors il existe un $(k+1)$-uplet non identiquement nul $(\lambda_1, \cdots, \lambda_k, \lambda)$ de $\mathbb{K}^{k+1}$ tel que

$$\sum_{i=1}^{k} \lambda_i \vec{x}_i + \lambda \vec{y} = \vec{0}. \tag{2.1}$$

On peut de plus affirmer que $\lambda \neq 0$. En effet, si λ était nul alors (2.1) entraînerait que $\sum_{i=1}^{k} \lambda_i \vec{x}_i = \vec{0}$. Mais $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre, donc $(\lambda_1, \cdots, \lambda_k) = (0, \cdots, 0)$, puis $(\lambda_1, \cdots, \lambda_k, \lambda) = (0, \cdots, 0)$, ce qui est contraire à l'hypothèse faite.

Donc λ n'est pas nul, et on peut écrire d'après (2.1),

$$\vec{y} = -\sum_{i=1}^{k} \frac{\lambda_i}{\lambda} \vec{x}_i.$$

Autrement dit, $\vec{y} \in \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k)$. ∎

Proposition 2.3.12. *Une famille $(\vec{x}_1, \cdots, \vec{x}_k)$ est liée si et seulement si elle contient un vecteur qui est combinaison linéaire des autres vecteurs.*

Preuve : $\Longrightarrow$ Supposons que $(\vec{x}_1, \cdots, \vec{x}_k)$ soit liée. Alors il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ non nul tel que $\sum_{i=1}^{k} \lambda_i \vec{x}_i = \vec{0}$. Comme le k-uplet n'est pas nul, l'un des λ_i (disons λ_k pour fixer les idées) n'est pas nul et l'on a donc

$$\vec{x}_k = -\sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda_k} \vec{x}_i,$$

et donc $\vec{x}_k$ est combinaison linéaire des autres vecteurs de la famille.

$\Longleftarrow$ Supposons par exemple que $\vec{x}_k$ soit combinaison linéaire de $(\vec{x}_1, \cdots, \vec{x}_{k-1})$. Alors $\vec{x}_k \in \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_{k-1})$ et la proposition 2.3.11 montre que la famille $(\vec{x}_1, \cdots, \vec{x}_k)$ est liée. ∎

Pour déterminer si une famille de vecteurs est libre ou liée, on a souvent recours à la proposition suivante (laissée au lecteur en guise d'exercice) :

Proposition 2.3.13. *(i) Toute sous-famille d'une famille libre est libre.*

(ii) Toute sur-famille d'une famille liée est liée.

(iii) Toute sur-famille d'une famille génératrice est génératrice.

(iv) Une sous-famille d'une famille non génératrice n'est pas génératrice non plus.

2.3.4 Bases

Définition 2.3.14. *Soit X un s.e.v de $\mathbb{K}^n$. On dit que la famille de vecteurs $(\vec{u}_1, \cdots, \vec{u}_k)$ est une* **base** *de X si elle est à la fois libre, et génératrice de X.*

Tout vecteur $\vec{x}$ de X se décompose alors de manière unique en $\vec{x} = \sum_{i=1}^{k} x_i \vec{u}_i$. Le k-uplet $(x_1, \cdots, x_k)$ s'appelle **coordonnées** *de $\vec{x}$ par rapport à la base $(\vec{u}_1, \cdots, \vec{u}_k)$.*

Exemples :

1. Dans $\mathbb{R}^3$ (ou $\mathbb{C}^3$), la famille constituée des vecteurs $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$ et $\vec{e}_3 = (0, 0, 1)$ est une base. En effet, il est clair que $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est libre. De plus, tout vecteur $\vec{x} = (x_1, x_2, x_3)$ peut s'écrire

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

La base $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est appelée **base canonique** de $\mathbb{R}^3$ (ou $\mathbb{C}^3$).

2. En revanche, la famille $(\vec{e}_1, \vec{e}_2)$ n'est pas une base de $\mathbb{R}^3$. En effet, toute combinaison linéaire de $\vec{e}_1$ et de $\vec{e}_2$ a sa troisième composante nulle. Le vecteur $\vec{e}_3$ n'est donc pas dans $\mathrm{Vect}\,(\vec{e}_1, \vec{e}_2)$.

3. Définissons $\vec{u} = (1, 1, 1)$. La famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ est génératrice de $\mathbb{R}^3$ puisque contient la famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ qui est déjà génératrice. Mais $\vec{e}_1 + \vec{e}_2 + \vec{e}_3 - \vec{u} = \vec{0}$. Donc $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ est liée et $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{u})$ n'est pas une base.

4. Plus généralement, la famille $(\vec{e}_1, \cdots, \vec{e}_n)$ de vecteurs de $\mathbb{K}^n$ définie par

$$\vec{e}_i \stackrel{\text{déf}}{=} (0, \cdots, 0, \underbrace{1}_{\text{position } i}, 0, \cdots, 0)$$

est une base de $\mathbb{K}^n$. On l'appelle **base canonique** de $\mathbb{K}^n$.

Théorème 2.3.15. *Soit X un sous-espace vectoriel de $\mathbb{K}^n$, et $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille de vecteurs de X. Les trois propriétés suivantes sont équivalentes :*

i) $(\vec{x}_1, \cdots, \vec{x}_k)$ *est une famille libre maximale (i.e si on ajoute un ou plusieurs vecteurs à la famille, on obtient une famille liée),*

ii) $(\vec{x}_1, \cdots, \vec{x}_k)$ *est une famille génératrice minimale (i.e si on retire un ou plusieurs vecteurs à la famille, la famille obtenue n'est plus génératrice de X),*

iii) $(\vec{x}_1, \cdots, \vec{x}_k)$ *est une base de X.*

Preuve : En vertu de la proposition 2.3.13, il suffit de traiter les cas où l'on ajoute ou retire un seul vecteur.

i) $\Rightarrow$ *iii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une famille libre maximale. Montrons que cette famille est aussi génératrice.
Soit $\vec{y} \in X$. Alors par hypothèse, la famille $(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée. La proposition 2.3.11 assure donc que $\vec{y} \in \text{Vect}(\vec{x}_1, \cdots, \vec{x}_k)$.

iii) $\Rightarrow$ *i)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une base de X. Par définition, cette famille est donc *libre*. De plus, si $\vec{y} \in X$ alors $\vec{y} \in \text{Vect}(\vec{x}_1, \cdots, \vec{x}_k)$ puisque $(\vec{x}_1, \cdots, \vec{x}_k)$ est aussi génératrice. D'après la proposition 2.3.11, on conclut donc que $\text{Vect}(\vec{x}_1, \cdots, \vec{x}_k, \vec{y})$ est liée. Donc $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre *maximale*.

ii) $\Rightarrow$ *iii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ génératrice minimale. Supposons *par l'absurde* que cette famille ne soit pas une base. Alors elle est liée, c'est-à-dire que l'un de ses vecteurs –disons $\vec{x}_k$ pour fixer les idées– est combinaison linéaire des autres :

$$\exists(\alpha_1, \cdots, \alpha_{k-1}) \in \mathbb{K}^{k-1},\ \vec{x}_k = \sum_{i=1}^{k-1} \alpha_i \vec{x}_i. \tag{2.2}$$

Soit maintenant $\vec{y} \in X$ arbitraire. Comme $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice, il existe un k-uplet $(\lambda_1, \cdots, \lambda_k)$ d'éléments de $\mathbb{K}^k$ tel que

$$\vec{y} = \sum_{i=1}^{k} \lambda_i \vec{x}_i.$$

En tenant compte de (2.2), on trouve

$$\vec{y} = \sum_{i=1}^{k-1} (\lambda_i + \alpha_i \lambda_k) \vec{x}_i.$$

En conséquence $(\vec{x}_1, \cdots, \vec{x}_{k-1})$ est génératrice, ce qui contredit l'hypothèse "$(\vec{x}_1, \cdots, \vec{x}_k)$ génératrice minimale".

iii) $\Rightarrow$ *ii)* : Soit $(\vec{x}_1, \cdots, \vec{x}_k)$ une base de X. Alors $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice. Retirons un vecteur à cette famille, par exemple $\vec{x}_k$. La proposition 2.3.11 montre que $\vec{x}_k$ ne peut être engendré par la famille $(\vec{x}_1, \cdots, \vec{x}_{k-1})$ car alors $(\vec{x}_1, \cdots, \vec{x}_k)$ serait liée. En conséquence, $(\vec{x}_1, \cdots, \vec{x}_k)$ est bien génératrice minimale. ■

Partant d'une famille donnée $(\vec{x}_1, \cdots, \vec{x}_p)$ de X, on peut se demander s'il est possible de construire une base de X ayant pour p premiers vecteurs $(\vec{x}_1, \cdots, \vec{x}_p)$. Une condition nécessaire est clairement que $(\vec{x}_1, \cdots, \vec{x}_p)$ soit libre. Le théorème suivant, dit de la base incomplète, montre que cette condition est aussi suffisante.

Théorème de la base incomplète. *Soit X un sous-espace vectoriel de $\mathbb{K}^n$ non réduit à $\{\vec{0}\}$, $m \in \mathbb{N}^*$ et $p \in \{0, \cdots, m\}$. Soit $(\vec{x}_1, \cdots, \vec{x}_m)$ une famille génératrice de X telle que $(\vec{x}_1, \cdots, \vec{x}_p)$ soit libre*[4]*. Alors, parmi les vecteurs $\vec{x}_{p+1}, \ldots \vec{x}_n$, on peut choisir des vecteurs $\vec{x}_{i_1}, \ldots, \vec{x}_{i_k}$ avec $p < i_1 < \cdots < i_k \leq m$ tels que $(\vec{x}_1, \cdots, \vec{x}_m, \vec{x}_{i_1}, \cdots, \vec{x}_{i_k})$ soit une base de X.*

Preuve : Soit $(\vec{x}_1, \cdots, \vec{x}_m)$ une famille génératrice de X dont les p premiers vecteurs constituent une famille libre. Considérons l'ensemble

$$\mathcal{E} = \Big\{(\vec{x}_1, \cdots, \vec{x}_p, \vec{x}_{j_1}, \cdots, \vec{x}_{j_\ell}) \mid 0 \leq \ell \leq m-p,\ p < j_1 < \cdots < j_\ell \leq m \text{ et } (\vec{x}_1, \cdots, \vec{x}_{j_\ell}) \text{ libre}\Big\}.$$

Soit $\mathcal{C}$ l'ensemble des cardinaux des familles de $\mathcal{E}$.

Par construction, l'ensemble $\mathcal{E}$ contient $(\vec{x}_1, \cdots, \vec{x}_p)$. Donc $\mathcal{C}$ est un sous-ensemble non vide de $\mathbb{N}$ ayant visiblement m comme majorant. En conséquence, $\mathcal{C}$ admet un élément maximal que l'on peut toujours noter $p+k$. (On a donc bien $0 \leq k \leq m-p$).

Soit $(\vec{x}_1, \cdots, \vec{x}_p, \vec{x}_{j_1}, \cdots, \vec{x}_{j_\ell})$ une famille de $\mathcal{E}$ de cardinal maximal. Cette famille est libre par construction. Mais elle est aussi maximale. En effet, vu sa définition, elle engendre tous les vecteurs de $(\vec{x}_1, \cdots, \vec{x}_m)$ donc X lui-même puisque $X = \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_m)$. La proposition 2.3.15 permet alors de conclure que $(\vec{x}_1, \cdots, \vec{x}_p, \vec{x}_{j_1}, \cdots, \vec{x}_{j_\ell})$ est une base. ■

2.4 Vecteurs et systèmes linéaires

Il est parfois commode de noter les composantes d'un vecteur $\vec{x}$ de $\mathbb{K}^n$ en colonne :

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{au lieu de} \quad \vec{x} = (x_1, \cdots, x_n).$$

On parle de **vecteur colonne** pour désigner l'écriture en colonne, et de **vecteur ligne** pour désigner l'écriture en ligne.

L'écriture en colonne permet d'établir une correspondance étroite entre systèmes linéaires et familles de vecteurs. En particulier, résoudre un système linéaire est équivalent à montrer qu'un certain vecteur est combinaison linéaire d'une famille de vecteurs donnés.

En effet, soit le système linéaire $m \times n$ suivant :

$$(S) \qquad \left\{ \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array} \right.$$

avec $a_{ij} \in \mathbb{K}$ pour $1 \leq i \leq m$ et $1 \leq j \leq n$, et $b_i \in \mathbb{K}$.

Définissons les $n+1$ vecteurs colonnes de $\mathbb{K}^m$ suivants :

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots\ldots\ldots\ldots, \vec{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \quad \text{et} \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

4. Avec la convention que $(\vec{x}_1, \cdots, \vec{x}_p)$ est la famille vide si $p = 0$.

Alors résoudre (S) revient à trouver un n-uplet $(x_1, \cdots, x_n)$ de $\mathbb{K}^n$ tel que

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}.$$

Autrement dit, on a le résultat suivant :

Théorème 2.4.1. *Le système (S) admet une solution si et seulement si* $\vec{b} \in \mathrm{Vect}\,(\vec{a}_1, \cdots, \vec{a}_n)$.

Le système homogène (S') associé à (S) peut s'écrire sous la forme

$$x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{0}$$

et l'on obtient donc le résultat important suivant :

Théorème 2.4.2. *Le système* (S') *admet une solution non nulle si et seulement si la famille* $(\vec{a}_1, \cdots, \vec{a}_n)$ *est liée.*

En utilisant la proposition sur la structure de l'ensemble des solutions d'un système linéaire, on en déduit que

Théorème 2.4.3. *Le système (S) n'admet qu'une seule solution si et seulement si* $(\vec{a}_1, \cdots, \vec{a}_n)$ *est libre et* $\vec{b} \in \mathrm{Vect}\,(\vec{a}_1, \cdots, \vec{a}_n)$.

Donnons une application fondamentale de cette correspondance entre systèmes linéaires et familles de vecteurs.

Théorème 2.4.4. *Dans* $\mathbb{K}^n$ *une famille comportant* $n+1$ *vecteurs est toujours liée.*

Preuve : On fait une récurrence sur la dimension n, l'hypothèse de récurrence étant :

$(\mathcal{P}_n)$ Dans $\mathbb{K}^n$, toute famille de $n+1$ vecteurs est liée.

• **Montrons que** $(\mathcal{P}_1)$ **est vraie :** Si α et β sont deux éléments de $\mathbb{K}^1$, c'est-à-dire de $\mathbb{K}$. Il est clair qu'il existe $(\lambda, \mu) \in \mathbb{K}^2$ non nul tel que $\lambda\alpha + \mu\beta = 0$: si $\alpha \neq 0$, on peut choisir par exemple $\lambda = -\beta/\alpha$ et $\mu = 1$; si $\alpha = 0$, le couple $(1, 0)$ fait l'affaire.

• **Supposons** $(\mathcal{P}_\mathbf{n})$ **vraie et démontrons** $(\mathcal{P}_\mathbf{n+1})$ **:** Soit donc $(\vec{x}_1, \cdots, \vec{x}_{n+2})$ une famille de $n+2$ vecteurs de $\mathbb{K}^{n+1}$. Notons ces vecteurs en colonne :

$$\vec{x}_i = \begin{pmatrix} x_i^1 \\ \vdots \\ x_i^{n+1} \end{pmatrix} \quad \text{pour} \quad i = 1, \cdots, n+2.$$

D'après la proposition précédente, $(\vec{x}_1, \cdots, \vec{x}_{n+2})$ est liée si et seulement si le système linéaire (S) suivant admet une solution $(\lambda_1, \cdots, \lambda_{n+2})$ non nulle :

$$(S) \qquad \begin{cases} x_1^1\lambda_1 + \cdots + x_{n+2}^1\lambda_{n+2} = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_1^{n+1}\lambda_1 + \cdots + x_{n+2}^{n+1}\lambda_{n+2} = 0. \end{cases}$$

1er cas : $x_1^{n+1} = \cdots = x_{n+2}^{n+1} = 0$.

Dans ce cas, la dernière ligne du système est vérifiée par n'importe quel $(n+2)$-uplet $(\lambda_1, \cdots, \lambda_{n+2})$. Si l'on pose $\vec{x}'_i \overset{\text{déf}}{=} (x_i^1, \cdots, x_i^n)$, la famille $(\vec{x}'_1, \cdots, \vec{x}'_{n+1})$ est une famille de $n+1$ vecteurs de $\mathbb{K}^n$. En vertu de l'hypothèse de récurrence, elle est liée. Donc $(\vec{x}'_1, \cdots, \vec{x}'_{n+2})$ aussi (cf prop. 2.3.13). On peut donc trouver $(\lambda_1, \cdots, \lambda_{n+2}) \in \mathbb{K}^{n+2}$ non nul tel que les n premières lignes de (S) soient aussi satisfaites.

2ème cas : Les coefficients de la dernière ligne de (S) ne sont pas tous nuls.

Quitte à changer l'ordre des vecteurs de la famille et à multiplier la dernière ligne de (S) par un scalaire non nul, on peut supposer que $x_{n+2}^{n+1} = 1$. On a donc

$$(S) \iff \begin{cases} x_1^1 \lambda_1 + \cdots + x_{n+1}^1 \lambda_{n+1} + x_{n+2}^1 \lambda_{n+2} = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_1^n \lambda_1 + \cdots + x_{n+1}^n \lambda_{n+1} + x_{n+2}^n \lambda_{n+2} = 0, \\ \lambda_{n+2} = -\sum_{i=1}^{n+1} x_i^{n+1} \lambda_i. \end{cases}$$

Pour $i \in \{1, \cdots, n+1\}$, définissons le vecteur $\vec{x}'_i$ de $\mathbb{K}^n$ par

$$\vec{x}'_i \stackrel{\text{déf}}{=} (x_i^1 - x_{n+2}^1 x_i^{n+1}, \cdots, x_i^n - x_{n+2}^n x_i^{n+1}).$$

Le système (S) donne $\sum_{i=1}^{n+1} \lambda_i \vec{x}'_i = \vec{0}$. D'après $(\mathcal{P}_n)$, la famille $(\vec{x}'_1, \cdots, \vec{x}'_{n+1})$ est liée. Donc il existe $(\lambda_1, \cdots, \lambda_{n+1})$ non nul tel que $\sum_{i=1}^{n+1} \lambda_i \vec{x}'_i = \vec{0}$. En définissant λ_{n+2} conformément à la dernière ligne du système ci-dessus, on obtient $\sum_{i=1}^{n+2} \lambda_i \vec{x}_i = \vec{0}$.

■

2.5 Dimension

Une preuve similaire à celle du théorème ci-dessus permet de montrer plus généralement que si X est un sous-espace vectoriel de $\mathbb{K}^n$ engendré par une famille de p vecteurs alors toute famille de vecteurs de X comportant au moins $p+1$ éléments est liée. Cela va nous permettre d'introduire la notion de **dimension** d'un sous-espace vectoriel.

Proposition 2.5.1. *Soit X un sous-espace vectoriel de $\mathbb{K}^n$ non réduit à $\{\vec{0}\}$. Alors X admet une base composée d'au plus n vecteurs. De plus, toutes les bases de X comportent le même nombre d'éléments. Ce nombre est appelé* **dimension** *de X. On le note* $\dim X$.

Preuve : La preuve de l'existence d'une base se fait par récurrence limitée. Par hypothèse, X contient un vecteur $\vec{x}_1 \neq 0$. On choisit alors un autre vecteur $\vec{x}_2$ de X tel que $(\vec{x}_1, \vec{x}_2)$ soit une base. Si un tel vecteur n'existe pas, $(\vec{x}_1)$ est une famille libre maximale et donc une base (cf th. 2.3.15).

Plus généralement, supposons connue une famille libre $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ de X. Deux cas peuvent se présenter. Ou bien $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ est *libre maximale* auquel cas le th. 2.3.15 assure que $(\vec{x}_1, \cdots, \vec{x}_{j-1})$ est une base de X, ou bien cette famille libre *n'est pas maximale*, auquel cas on peut trouver $\vec{x}_j \in X$ tel que $(\vec{x}_1, \cdots, \vec{x}_j)$ soit libre.

Enfin, en vertu du théorème 2.4.4, le procédé de construction s'arrête au plus tard après l'obtention d'une famille libre à n éléments.

Reste à vérifier que toutes les bases ont le même nombre d'éléments. Considérons deux bases $(\vec{x}_1, \cdots, \vec{x}_k)$ et $(\vec{y}_1, \cdots, \vec{y}_\ell)$ de X. On a

$$X = \text{Vect}\,(\vec{x}_1, \cdots, \vec{x}_k) = \text{Vect}\,(\vec{y}_1, \cdots, \vec{y}_\ell).$$

La première égalité montre que X est engendré par k vecteurs. D'après le théorème 2.4.4, toute famille de $k+1$ vecteurs de X est donc liée. Comme $(\vec{y}_1, \cdots, \vec{y}_\ell)$ est libre, on a donc $\ell \leq k$. En échangeant les rôles des deux familles, on obtient $k \leq \ell$. ■

Remarque : Soit X un sous-espace vectoriel de dimension k. Pour montrer qu'une famille $(\vec{x}_1, \cdots, \vec{x}_k)$ de vecteurs de X est une base, il suffit d'établir que $(\vec{x}_1, \cdots, \vec{x}_k)$ est génératrice ou que $(\vec{x}_1, \cdots, \vec{x}_k)$ est libre.

Remarque 2.5.2. *(i) Par convention, le sous-espace vectoriel $\{\vec{0}\}$ a pour dimension 0.*

(ii) La base canonique de $\mathbb{K}^n$ comporte n éléments. Donc $\dim \mathbb{K}^n = n$. On en déduit qu'une famille de n vecteurs de $\mathbb{K}^n$ est une base si et seulement si elle est libre <u>ou</u> génératrice.

(iii) Les sous-espaces vectoriels de dimension 1 sont appelés **droites vectorielles** *ou droites.*

(iv) Les sous-espaces de dimension $n-1$ de $\mathbb{K}^n$ sont appelés **hyperplans**. *Lorsque $n = 3$, on parle plutôt de* **plan**. *Enfin, si $n = 2$, les hyperplans sont des droites.*

Exemples :

- **Équation d'une droite de $\mathbb{R}^2$:**
  Soit $(\alpha, \beta) \in \mathbb{R}^2 \backslash (0,0)$. La droite (vectorielle) engendrée par le vecteur $\vec{u} = (\alpha, \beta)$ de $\mathbb{R}^2$ est l'ensemble des vecteurs $\vec{v} = (x, y)$ tels que $\beta x - \alpha y = 0$.
  Réciproquement, si $(a, b) \neq (0,0)$, l'ensemble des vecteurs $\vec{v} = (x, y)$ de $\mathbb{R}^2$ tels que $ax + by = 0$ est la droite (vectorielle) de $\mathbb{R}^2$ orthogonale au vecteur $\vec{u} = (a, b)$.
- **Équation d'un plan de $\mathbb{R}^3$:**
  Soit (a, b, c) un triplet non nul (i.e. distinct de $(0,0,0)$) de $\mathbb{R}^3$. L'ensemble des vecteurs $\vec{v} = (x, y, z)$ de $\mathbb{R}^3$ tels que
  $$ax + by + cz = 0$$
  est un plan (vectoriel) de $\mathbb{R}^3$. C'est le plan orthogonal au vecteur $\vec{u} = (a, b, c)$.
- **Équation d'un hyperplan de $\mathbb{R}^n$:**
  Plus généralement, si $(a_1, \cdots, a_n)$ est un n-uplet non nul de $\mathbb{R}^n$, alors l'ensemble des vecteurs $\vec{v} = (x_1, \cdots, x_n)$ tels que
  $$a_1 x_1 + \cdots + a_n x_n = 0$$
  est un hyperplan de $\mathbb{R}^n$.

2.6 Produit scalaire et orthogonalité

Dans toute cette partie, nous supposons pour simplifier que $\mathbb{K} = \mathbb{R}$, et l'on s'intéresse donc uniquement à des vecteurs de $\mathbb{R}^n$.

Définition 2.6.1. *Soit $(\vec{x}, \vec{y})$ un couple de vecteurs de $\mathbb{R}^n$. Le nombre*
$$\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i$$
est appellé **produit scalaire** *de $\vec{x}$ et de $\vec{y}$.*

Remarque : L'espace $\mathbb{R}^n$ muni de l'addition vectorielle, de la multiplication par un scalaire et du produit scalaire est appelé **espace euclidien**.

Proposition 2.6.2. *Pour tous vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$ et réel λ, on a*

(i) $(\vec{x} + \vec{y}) \cdot \vec{z} = \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}$, **(distributivité)**

(ii) $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$, **(commutativité)**

(iii) $(\lambda \vec{x}) \cdot \vec{y} = \lambda (\vec{x} \cdot \vec{y})$, **(homogénéité)**

(iv) $\vec{x} \cdot \vec{x} \geq 0$ *et* $\vec{x} \cdot \vec{x} = 0$ *si et seulement si* $\vec{x} = 0$ **(définie positivité)**.

Preuve : La première propriété résulte de la distributivité de la multiplication sur l'addition dans $\mathbb{R}$. La deuxième est une conséquence de la commutativité de la multiplication dans $\mathbb{R}$. La troisième propriété est évidente. Finalement, on remarque que

$$\vec{x} \cdot \vec{x} = \sum_{i=1}^{n} x_i^2$$

ce qui donne immédiatement *(iv)*. ■

Définition 2.6.3. *Pour tout vecteur $\vec{x}$ de $\mathbb{R}^n$, le réel positif*

$$\|\vec{x}\| \stackrel{\text{déf}}{=} \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

est appelé **norme** *de $\vec{x}$.*

Identité remarquable : Pour tout couple $(\vec{x}, \vec{y})$ de vecteurs de $\mathbb{R}^n$, on a

$$\boxed{\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + 2\vec{x}\cdot\vec{y} + \|\vec{y}\|^2.}$$

Preuve : Par définition de la norme de $\vec{x}+\vec{y}$, on a

$$\|\vec{x}+\vec{y}\|^2 = \sum_{i=1}^{n} (x_i+y_i)^2.$$

Mais on a $(x_i+y_i)^2 = x_i^2 + 2x_iy_i + y_i^2$ pour tout $i \in \{1, \cdots, n\}$, donc

$$\|\vec{x}+\vec{y}\|^2 = \underbrace{\sum_{i=1}^{n} x_i^2}_{\|\vec{x}\|^2} + 2\underbrace{\sum_{i=1}^{n} x_iy_i}_{\vec{x}\cdot\vec{y}} + \underbrace{\sum_{i=1}^{n} y_i^2}_{\|\vec{y}\|^2}.$$

■

Inégalité de Cauchy-Schwarz : pour tout couple $(\vec{x}, \vec{y})$ de vecteurs de $\mathbb{R}^n$, on a

$$\boxed{\left|\vec{x}\cdot\vec{y}\right| \leq \|\vec{x}\|\,\|\vec{y}\|.}$$

Preuve : Le cas $\vec{y} = \vec{0}$ est évident. Supposons donc que $\vec{y} \neq 0$. On pose alors

$$\forall \lambda \in \mathbb{R},\ f(\lambda) \stackrel{\text{déf}}{=} \|\vec{x}+\lambda\vec{y}\|^2.$$

D'après l'identité remarquable ci-dessus, on a

$$f(\lambda) = \|\vec{y}\|^2\lambda^2 + 2\lambda\vec{x}\cdot\vec{y} + \|\vec{x}\|^2.$$

La fonction f est positive. C'est par ailleurs un polynôme de degré 2 en λ. Le discriminant réduit correspondant doit donc être négatif ou nul. Autrement dit, on doit avoir

$$|\vec{x}\cdot\vec{y}|^2 - \|\vec{x}\|^2\|\vec{y}\|^2 \leq 0.$$

Cela entraîne l'inégalité de Cauchy-Schwarz. ■

Inégalité triangulaire : Pour tout couple $(\vec{x}, \vec{y})$ de vecteurs de $\mathbb{R}^n$, on a

$$\boxed{\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.}$$

Preuve : En élevant au carré et en appliquant l'identité remarquable, on voit que l'inégalité à prouver est équivalente à

$$\|\vec{x}\|^2 + 2\vec{x}\cdot\vec{y} + \|\vec{y}\|^2 \leq \|\vec{x}\|^2 + 2\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2.$$

Cette deuxième inégalité découle trivialement de l'inégalité de Cauchy-Schwarz. ∎

Deuxième inégalité triangulaire : Pour tout couple $(\vec{x}, \vec{y})$ de vecteurs de $\mathbb{R}^n$, on a

$$\boxed{\big|\|\vec{x}\| - \|\vec{y}\|\big| \leq \|\vec{x}-\vec{y}\|.}$$

Preuve : D'après l'inégalité triangulaire appliquée aux vecteurs $\vec{x}$ et $\vec{y}-\vec{x}$, on a

$$\|\vec{y}\| = \|\vec{x} + (\vec{y}-\vec{x})\| \leq \|\vec{x}\| + \|\vec{y}-\vec{x}\|.$$

D'après l'inégalité triangulaire appliquée aux vecteurs $\vec{y}$ et $\vec{x}-\vec{y}$, on a

$$\|\vec{x}\| = \|\vec{y} + (\vec{x}-\vec{y})\| \leq \|\vec{y}\| + \|\vec{x}-\vec{y}\|.$$

En combinant ces deux inégalités, on obtient le résultat souhaité. ∎

Définition 2.6.4. *On dit que deux vecteurs $\vec{x}$ et $\vec{y}$ de $\mathbb{R}^n$ sont* **orthogonaux** *si $\vec{x}\cdot\vec{y} = 0$. On note $\vec{x}\perp\vec{y}$.*
On dit qu'une famille $(\vec{x}_1, \cdots, \vec{x}_p)$ de vecteurs de $\mathbb{R}^n$ est **orthogonale** *si $\vec{x}_i\cdot\vec{x}_j = 0$ pour tout couple d'indices i et j tel que $i \neq j$.*
On dit que $(\vec{x}_1, \cdots, \vec{x}_p)$ est **orthonormale** *si elle est orthogonale et si de plus $\|\vec{x}_i\| = 1$ pour tout $i \in \{1, \cdots, n\}$.*

Remarque : Le vecteur nul est orthogonal à tous les vecteurs.

Théorème de Pythagore. Si $\vec{x}$ et $\vec{y}$ sont deux vecteurs orthogonaux de $\mathbb{R}^n$ alors on a

$$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

Plus généralement, si $(\vec{x}_1, \cdots, \vec{x}_p)$ est une famille orthogonale alors on a

$$\Big\|\sum_{i=1}^{p} \vec{x}_i\Big\|^2 = \sum_{i=1}^{p} \|\vec{x}_i\|^2.$$

Preuve : La première égalité est une conséquence triviale de l'identité remarquable de la page précédente et de la définition de l'orthogonalité. Quant à la deuxième égalité, elle se montre par récurrence à partir de la première. ∎

Proposition 2.6.5. *Une famille orthogonale de vecteurs non nuls est toujours libre.*

Preuve : Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille orthogonale de vecteurs non nuls de $\mathbb{R}^n$, et $(\lambda_1, \cdots, \lambda_p)$ un p-uplet de $\mathbb{R}^n$ tel que

$$\sum_{i=1}^{p} \lambda_i \vec{x}_i = 0.$$

Fixons $j \in \{1, \cdots, n\}$ et prenons le produit scalaire de l'égalité ci-dessus avec $\vec{x}_j$. On obtient $\lambda_j\|\vec{x}_j\|^2 = 0$. Mais $\vec{x}_j \neq \vec{0}$ donc $\|\vec{x}_j\| \neq 0$. Cela assure que $\lambda_j = 0$. ∎

Remarque : En particulier, une famille orthonormale de n vecteurs de $\mathbb{R}^n$ est libre. C'est donc une base. On dit que c'est une **base orthonormale** de $\mathbb{R}^n$.

Exemples :

(i) En toute dimension, la base canonique de $\mathbb{R}^n$ est orthonormale.

(ii) Dans $\mathbb{R}^3$, définissons $\vec{u} = \frac{1}{\sqrt{3}}(1,1,1)$, $\vec{v} = \frac{1}{\sqrt{2}}(1,0,-1)$ et $\vec{w} = \frac{1}{\sqrt{6}}(1,-2,1)$. La famille $(\vec{u}, \vec{v}, \vec{w})$ est une base orthonormale de $\mathbb{R}^3$.

Chapitre 3

Matrices

Dans le premier chapitre, nous avons défini les matrices associées aux systèmes linéaires comme des "tableaux de nombres" contenant tous les coefficients du système.

L'objet central de ce chapitre est l'étude des matrices. Même si cette étude peut paraître "gratuite" et des plus théoriques, il faut savoir que la théorie des matrices trouve des applications dans la plupart des disciplines scientifiques : mathématiques bien sûr, mais aussi informatique, statistiques, économie, physique, ...

Dans ce chapitre, on présente les fondements de la théorie des matrices, vues comme "tableaux de nombres". Comme d'habitude, la lettre $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

3.1 Espaces de matrices

3.1.1 Définitions

Commençons par donner la définition des matrices.

Définition 3.1.1. *On appelle* **matrice à** m **lignes et** n **colonnes** *à coefficients dans* $\mathbb{K}$ *un tableau à* m *lignes et* n *colonnes constitué de scalaires de* $\mathbb{K}$ *:*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{3.1}$$

Le scalaire $a_{ij} \in \mathbb{K}$ *se trouve à l'intersection de la* i*-ième ligne et de la* j*-ième colonne ; il est appelé le* ij*-ième* **coefficient** *de la matrice.*

On note $\mathcal{M}_{m,n}(\mathbb{K})$ *l'ensemble de toutes les matrices à* m *lignes et* n *colonnes. Les matrices de* $\mathcal{M}_{m,n}(\mathbb{K})$ *sont aussi appelées* **matrices** $m \times n$ *(lire* m *fois* n*).*

Exemples : La matrice $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ est une matrice 2×3. La matrice $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ est une matrice 3×1. La matrice $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ est une matrice 1×3.

Notations : Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. On note a_{ij} ou encore $(A)_{ij}$ le ij-ième coefficient de A. Pour désigner A, on peut aussi utiliser aussi la notation $(a_{ij})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ ou bien (a_{ij}).

Définition 3.1.2. *On dit que deux matrices* A *et* B *sont égales si elles ont même taille* $m \times n$ *et si pour tous les indices tels que* $1 \leq i \leq m$ *et* $1 \leq j \leq n$*, on a* $(A)_{ij} = (B)_{ij}$*.*

Si A est une matrice $1 \times n$, on dit que A est une **matrice ligne**. Si A est matrice $m \times 1$, on dit que A est une **matrice colonne**.

3.1.2 Opérations linéaires

Un vecteur ligne peut être considéré comme une matrice ligne, et un vecteur colonne, comme une matrice colonne. On sait calculer la somme de deux vecteurs et le produit d'un vecteur par un scalaire. De la même façon, on va définir la somme de deux matrices et le produit d'une matrice par un scalaire.

Définition 3.1.3. *Soient A et B deux matrices $m \times n$. Alors la somme $A+B$ de ces deux matrices est la matrice $m \times n$ définie par les formules*

$$\forall\ 1 \leq i \leq m,\ 1 \leq j \leq n, \qquad (A+B)_{ij} = (A)_{ij} + (B)_{ij}.$$

Exemple :

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

Définition 3.1.4. *Soit A une matrice $m \times n$ et soit λ un scalaire. Alors le produit de λ et A est la matrice λA de taille $m \times n$ définie par les formules*

$$\forall\ 1 \leq i \leq m,\ 1 \leq j \leq n, \qquad (\lambda A)_{ij} = \lambda (A)_{ij}.$$

Exemple :

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

Pour toutes matrices A et B de $\mathcal{M}_{m,n}(\mathbb{K})$, on définit :

- la **matrice opposée** $-A$ par $-A \overset{\text{déf}}{=} (-1)A$,
- $A - B \overset{\text{déf}}{=} A + (-B)$,
- la **matrice nulle** $\mathbb{O}$: c'est la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ dont tous les coefficients sont nuls.

L'opération d'addition vérifie les propriétés suivantes :

$$A + (B + C) = (A + B) + C, \quad A + \mathbb{O} = A, \quad A + (-A) = \mathbb{O}, \quad A + B = B + A.$$

Cela signifie que l'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ est un groupe commutatif par rapport à l'addition.
De plus, on a

$$\lambda(A + B) = \lambda A + \lambda B, \quad (\lambda + \mu)A = \lambda A + \mu A, \quad \lambda(\mu A) = (\lambda\mu)A, \quad 1\,A = A.$$

Cela signifie que l'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ est un espace vectoriel sur $\mathbb{K}$.

3.1.3 Transposition

Définition 3.1.5. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$. On appelle* **matrice transposée** *de A la matrice*[1] A^T *de $\mathcal{M}_{n,m}(\mathbb{R})$ définie par les formules*

$$\forall\ 1 \leq i \leq n,\ 1 \leq j \leq m, \qquad (A^T)_{ij} = (A)_{ji}.$$

Exemple : Soit $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$. Alors, $A^T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

1. La notation ${}^tA$ est utilisée par certains auteurs.

Lemme 3.1.6. *Soit* $\alpha, \beta \in \mathbb{R}$ *et* $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$. *Alors,*

$$(\alpha A + \beta B)^T = \alpha A^T + \beta B^T. \tag{3.2}$$

Preuve : Soit $1 \leq i \leq n$ et $1 \leq j \leq m$. On a

$$(\,(\alpha A+\beta B)^T)_{ij} = (\alpha A+\beta B)_{ji} = (\alpha A)_{ji} + (\beta B)_{ji} = \alpha (A)_{ji} + \beta (B)_{ji} = \alpha (A^T)_{ij} + \beta (B^T)_{ij}.$$

Cela implique (3.2). ■

Remarque : Le lemme précédent signifie que l'opération de transposition est **linéaire**.

Lemme 3.1.7. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{R})$. *Alors,*

$$(A^T)^T = A. \tag{3.3}$$

Preuve : Soit $1 \leq i \leq n$ et $1 \leq j \leq m$. On a

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

Cela implique (3.3). ■

Remarque : Ce lemme signifie que l'opération de transposition est **involutive**.

D'après la définition de la transposition, si $A \in \mathcal{M}_{n,n}(\mathbb{R})$ alors la matrice A^T appartient aussi à $\mathcal{M}_{n,n}(\mathbb{R})$. On peut donc comparer A et A^T, ce qui motive la définition suivante.

Définition 3.1.8. *On dit qu'une matrice* A *de* $\mathcal{M}_{n,n}(\mathbb{R})$ *est* **symétrique** *si l'on a* $A^T = A$.

Exemple : La matrice $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ est symétrique.

3.2 Produit de matrices

3.2.1 Définition

Dans cette partie, nous allons définir – quand c'est possible – une multiplication entre matrices. Nous allons voir que cette opération est naturellement liée à l'étude des système linéaires.

Définition 3.2.1. *Soit* A *une matrice* $m \times n$ *et soit* B *une matrice* $n \times p$. *Alors le produit* AB *des matrices* A *et* B *est une matrice* $m \times p$ *dont les éléments sont donnés par les formules :*

$$\forall\ 1 \leq i \leq m,\ 1 \leq j \leq p, \qquad (AB)_{ij} = \sum_{k=1}^{n} (A)_{ik}\,(B)_{kj}. \tag{3.4}$$

Attention : On peut calculer le produit AB si et seulement si le nombre de colonnes de A est égal au nombre de lignes de B !

Exemple très important. Soit a une matrice $1 \times n$ et soit b une matrice $n \times 1$,

$$a = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Le nombre des colonnes de a est égal au nombre de lignes de b. Donc, le produit ab est bien défini. Ce produit est une matrice 1×1, c'est-à-dire un scalaire!
Un calcul simple montre que

$$ab = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n. \tag{3.5}$$

Si l'on suppose de plus que a et b sont à coefficients réels, et si l'on note $\vec{a} = (a_1, \cdots, a_n)$ et $\vec{b} = (b_1, \cdots, b_n)$, on voit que le nombre réel ab est le produit scalaire des vecteurs $\vec{a}$ et $\vec{b}$.

Remarque très importante. Soit A et B deux matrices comme dans la définition 3.2.1. Soit A_i la i-ième ligne de A. C'est une matrice $1 \times n$. Soit B^j la j-ième colonne de B. C'est une matrice $n \times 1$.
La formule (3.4) montre que le ij-ième coefficient de la matrice AB est égal à $A_i\, B^j$. On arrive donc à la règle mnémotechnique :

Le ij-ième coefficient de AB est égal au produit de la i-ième ligne de A par la j-ième colonne de B.

Encore une remarque très importante. Soit A et B deux matrices telles que $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$. En général,

$$AB \neq BA,$$

et ce pour plusieurs raisons :

1. Le produit BA peut ne pas être défini.

   En effet, BA est défini si et seulement si $m = p$. En conséquence, si $m \neq p$, AB est défini mais pas BA.

2. AB et BA peuvent être de tailles différentes.

   En effet, supposons $m = p$ de telle sorte que AB et BA soient définies. Dans ce cas, la matrice AB est une matrice $m \times m$, et la matrice BA est une matrice $n \times n$. Donc, si $m \neq n$, AB et BA sont des matrices de tailles différentes!

3. Finalement, même dans le cas où $m = n = p$, AB n'est pas forcément égal à BA. Donnons un exemple. Soit

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

   Dans ce cas, $n = m = l = 2$, et les deux produits AB et BA sont bien définis. De plus, les deux matrices AB et BA sont 2×2. Mais, on a :

   $$A\,B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbb{O} \quad \text{et} \quad B\,A = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix} \neq \mathbb{O}.$$

3.2.2 Multiplication de matrices et systèmes linéaires

Soit A une matrice $m \times n$ et soit b une matrice $m \times 1$. On cherche une matrice colonne x de taille $n \times 1$ telle que $Ax = b$.

$$\text{Si l'on note} \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad \text{et} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

alors l'équation $Ax = b$ se récrit

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{3.6}$$

En calculant le produit Ax, on voit que l'équation (3.6) est équivalente au système linéaire :

$$(S) \qquad \left\{ \begin{array}{r} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{array} \right.$$

Donc, l'équation $Ax = b$ est équivalente au système (S).

3.2.3 Propriétés de la multiplication de matrices

Proposition 3.2.2. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$ *et* $C \in \mathcal{M}_{p,q}(\mathbb{K})$. *Alors,*

$$(AB)C = A(BC). \tag{3.7}$$

On dit que la multiplication des matrices est **associative**.

Preuve : Soient i et ℓ deux entiers tels que $1 \leq i \leq m$ et $1 \leq \ell \leq q$. Alors,

$$\begin{aligned} ((AB)\,C)_{i\ell} &= \sum_{k=1}^{p} (AB)_{ik} (C)_{k\ell} = \\ &= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} (A)_{ij} (B)_{jk} \right) (C)_{k\ell} = \\ &= \sum_{j=1}^{n} (A)_{ij} \left(\sum_{k=1}^{p} (B)_{jk}\, (C)_{k\ell} \right) = \\ &= \sum_{j=1}^{n} A_{ij} (BC)_{j\ell} = \\ &= (\, A(BC)\,)_{i\ell}. \end{aligned}$$

Cela implique (3.7). ■

Proposition 3.2.3. *Soit* $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$ *et* $C \in \mathcal{M}_{n,p}(\mathbb{K})$. *Alors,*

$$(A + B)C = AC + BC. \tag{3.8}$$

Soit $C \in \mathcal{M}_{m,n}(\mathbb{K})$ *et* $A, B \in \mathcal{M}_{n,p}(\mathbb{K})$. *Alors,*

$$C(A + B) = CA + CB. \tag{3.9}$$

On dit que la multiplication des matrices est **distributive**.

Preuve : Démontrons (3.8). Soient i et k deux entiers tels que $1 \leq i \leq m$ et $1 \leq k \leq p$. Alors,

$$\begin{aligned}
((A+B)\,C)_{ik} &= \sum_{j=1}^{n} (A+B)_{ij}(C)_{jk} = \\
&= \sum_{j=1}^{n} \left((A)_{ij} + (B)_{ij}\right)\,(C)_{jk} = \\
&= \left(\sum_{j=1}^{n} (A)_{ij}(C)_{jk}\right) + \left(\sum_{j=1}^{n} (B)_{ij}(C)_{jk}\right) = \\
&= (AC)_{ik} + (BC)_{ik}.
\end{aligned}$$

Cela implique (3.8). On démontre (3.9) de même façon. ■

On laisse au lecteur la preuve du

Lemme 3.2.4. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$ *et* $\lambda \in \mathbb{K}$. *Alors,*

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

Enfin, on vérifie le

Lemme 3.2.5. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *et* $B \in \mathcal{M}_{n,p}(\mathbb{R})$. *Alors* $(AB)^T = B^T\,A^T$.

Preuve : Soient i et k deux entiers tels que $1 \leq i \leq p$ et $1 \leq k \leq m$. Alors,

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

Cela achève la preuve du lemme. ■

3.3 Matrices carrées

3.3.1 Définitions

Si A est une matrice $n \times n$, on dit que A est une **matrice carrée** de taille n. On note $\mathcal{M}_n(\mathbb{K})$ l'ensemble de toutes les matrices carées de taille n.

Les matrices carrées de la forme

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} \tag{3.10}$$

sont appelées **matrices triangulaires supérieures** ; les matrices carrées de la forme

$$\begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} \tag{3.11}$$

sont appelées **matrices triangulaires inférieures**.

Les matrices carrées de la forme

$$\begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix} \tag{3.12}$$

sont appelées **matrices diagonales**. Les matrices diagonales sont aussi notées

$$\mathrm{diag}(a_{11}, a_{22}, \dots a_{nn}).$$

La matrice $\mathrm{diag}(1, 1, 1, \dots 1)$ est appelée **matrice identité** et notée I_n. On note δ_{ij} le ij-ième coefficient de I_n :

$$\delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j, \\ 0 & \text{sinon.} \end{cases} \tag{3.13}$$

Le coefficient δ_{ij} est appelé **symbole de Kronecker**.

3.3.2 Algèbre des matrices carrées

Rappelons que $\mathcal{M}_n(\mathbb{K})$ est un espace vectoriel par rapport aux opérations de multiplication par un scalaire et d'addition de matrices. La définition de produit de matrices implique de plus le résultat suivant :

Lemme 3.3.1. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. Alors le produit AB est bien défini et appartient à $\mathcal{M}_n(\mathbb{K})$.*

On a aussi le résultat suivant :

Lemme 3.3.2. *Pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, on a*

$$AI_n = I_n A = A, \tag{3.14}$$

où $I_n \in \mathcal{M}_n(\mathbb{K})$ est la matrice identité.

Preuve : Montrons que $AI_n = A$. Soient i et j deux entiers tels que $1 \leq i, j \leq n$. Alors, en utilisant (3.13), on obtient

$$(AI_n)_{ij} = \sum_{k=1}^{n} (A)_{ik} \delta_{kj} = (A)_{ij} \delta_{jj} = (A)_{ij}.$$

On démontre l'égalité $I_n A = A$ de la même façon. ■

Définition 3.3.3. *On appelle* **algèbre** *un espace vectoriel $\mathcal{A}$ muni d'une loi de multiplication interne telle que*

(i) si $A, B \in \mathcal{A}$, alors $AB \in \mathcal{A}$;
(ii) si $A, B, C \in \mathcal{A}$, alors $(A+B)C = AC + BC$ et $C(A+B) = CA + CB$;
(iii) si $A, B, C \in \mathcal{A}$, alors $(AB)C = A(BC)$;
(iv) si $A, B \in \mathcal{A}$ et si λ est un scalaire, alors $\lambda(AB) = (\lambda A)B = A(\lambda B)$;
(v) il existe $I \in \mathcal{A}$ tel que pour tout $A \in \mathcal{A}$, on ait $AI = IA = A$.

Les deux derniers lemmes et les résultats de la section 3.2.3 montrent que

$$\boxed{\mathcal{M}_n(\mathbb{K}) \quad \text{est une algèbre.}}$$

Enfin, on rappelle que, en général, même pour les matrices carrées, $AB \neq BA$. Cela signifie que $\mathcal{M}_n(\mathbb{K})$ est une **algèbre non commutative**.

3.4 Rang

À toute matrice, on peut associer un nombre entier qui s'appelle *le rang*, et qui est étroitement lié à la notion de dimension vue dans le chapitre précédent. Le but de cette partie est de définir le rang et d'en donner quelques propriétés fondamentales.

3.4.1 Rang d'une famille de vecteurs

Avant de définir le rang d'une matrice, nous devons définir celui d'une famille de vecteurs. C'est l'objet de cette section.

Définition 3.4.1. *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$. On appelle* **rang de la famille** *$(\vec{x}_1, \cdots, \vec{x}_p)$, noté $\mathrm{rg}\,(\vec{x}_1, \cdots, \vec{x}_p)$ la dimension du sous-espace vectoriel* $\mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$.

Remarque : Autrement dit, le rang de la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ est le nombre (commun) d'éléments des bases de $\mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$.

Proposition 3.4.2. *Soit $(\vec{x}_1, \cdots, \vec{x}_p)$ une famille de vecteurs de $\mathbb{K}^n$.*

i) On a toujours $\mathrm{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) \leq p$.

ii) On a $\mathrm{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) \leq n$ avec égalité si et seulement si $(\vec{x}_1, \cdots, \vec{x}_p)$ engendre $\mathbb{K}^n$.

iii) On a $\mathrm{rg}\,(\vec{x}_1, \cdots, \vec{x}_p) = p$ si et seulement si la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ est libre.

Preuve : Notons $X = \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_p)$.

(i) D'après le théorème de la base incomplète, il existe une sous-famille de $(\vec{x}_1, \cdots, \vec{x}_p)$ qui est une base de X. Cette famille a au plus p éléments.

(ii) Comme $X \subset \mathbb{K}^n$, les familles de vecteurs de X comptant au moins $n+1$ éléments sont toutes liées. Donc $\dim X \leq n$ avec égalité si $X = \mathbb{K}^n$ puisque $\dim \mathbb{K}^n = n$.

Maintenant, si $\dim X = n$, on peut, d'après le théorème de la base incomplète et la définition de la dimension, extraire une sous-famille à n éléments de $(\vec{x}_1, \cdots, \vec{x}_p)$ qui soit une base de X. Cette base de X est une famille libre de $\mathbb{K}^n$. C'est en fait une famille libre *maximale* (en vertu du théorème 2.4.4). Donc c'est une base de $\mathbb{K}^n$. Par conséquent, $X = \mathbb{K}^n$.

(iii) Par définition de X, la famille $(\vec{x}_1, \cdots, \vec{x}_p)$ est génératrice. Si de plus $(\vec{x}_1, \cdots, \vec{x}_p)$ est libre alors $(\vec{x}_1, \cdots, \vec{x}_p)$ est une base de X, et donc $\dim X = p$.

Si au contraire $(\vec{x}_1, \cdots, \vec{x}_p)$ est liée, alors l'un des vecteurs, par exemple $\vec{x}_p$, est combinaison linéaire des autres. Donc $X = \mathrm{Vect}\,(\vec{x}_1, \cdots, \vec{x}_{p-1})$, et donc d'après *i)*, $\dim X \leq p-1$.

■

Pour déterminer le rang d'une famille de vecteurs, on a souvent recours au résultat suivant :

Proposition 3.4.3. *Le rang d'une famille de vecteurs est invariant par les transformations élémentaires* $(T1)$, $(T2)$ *et* $(T3)$ *définies dans la proposition* 2.3.7*.*

Preuve : On sait que le sous-espace vectoriel engendré par une famille de vecteurs donnée, est invariant par les transformations élémentaires (cf proposition 2.3.7). Le résultat est donc immédiat. ■

3.4.2 Rang d'une matrice

Définition 3.4.4. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$*. On appelle rang de la matrice* A *(noté* $\operatorname{rg} A$*) le rang de la famille de vecteurs constituée des* m *vecteurs lignes de* A*. Autrement dit, si l'on note* $\vec{a}_i$ *la* i*-ème ligne de* A*, on a*

$$\operatorname{rg} A = \operatorname{rg}(\vec{a}_1, \cdots, \vec{a}_m).$$

Par analogie avec les systèmes échelonnés, on définit les matrices échelonnées comme suit :

Définition 3.4.5. *On dit que la matrice* A *de* $\mathcal{M}_{m,n}(\mathbb{K})$ *est* **échelonnée** *si elle est de la forme*

$$A = \begin{pmatrix} 0 & a_{1j_1} & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & * \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & a_{kj_k} & * \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

avec $1 \leq j_1 < \cdots < j_k \leq n$ *et* $a_{ij_i} \neq 0$ *pour* $i \in \{1, \cdots, k\}$*.*

Remarque : Un système linéaire est échelonné si et seulement si sa matrice l'est.

Le rang d'une matrice échelonnée est très facile à calculer :

Proposition 3.4.6. *Soit* A *une matrice échelonnée du type décrit dans la définition précédente. Alors on a* $\operatorname{rg} A = k$*.*

Preuve : Comme les $m-k$ dernières lignes de A sont nulles, le rang de A est égal à celui de ses k premières lignes. On vérifie alors facilement que la famille constituée par les k premières lignes est libre. En effet, du fait que $a_{ij_i} \neq 0$ pour tout $i \in \{1, \cdots, k\}$, l'unique solution du système linéaire

$$\begin{cases} a_{1j_1}\lambda_1 = 0 \\ a_{1j_2}\lambda_1 + a_{2j_2}\lambda_2 = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1j_k}\lambda_1 + a_{2j_k}\lambda_2 + \cdots + a_{kj_k}\lambda_k = 0 \end{cases}$$

est $(\lambda_1, \cdots, \lambda_k) = (0, \cdots, 0)$.

La proposition 3.4.2 permet de conclure que $\operatorname{rg} A = k$. ■

3.4.3 Calcul du rang d'une famille de vecteurs par pivot de Gauss

Nous avons vu dans le premier chapitre que l'algorithme du pivot de Gauss permettait de transformer n'importe quel système linéaire en un système échelonné équivalent.

Si l'on considère maintenant une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$, le même algorithme permet de transformer A en une matrice échelonnée B. Comme l'algorithme consiste à effectuer une succession de transformations élémentaires, la proposition 3.4.3 assure que $\operatorname{rg} A = \operatorname{rg} B$, et le rang de B peut être facilement calculé grâce à la proposition 3.4.6.

Ces considérations nous suggèrent une méthode systématique pour calculer le rang d'une famille de vecteurs ou d'une matrice :

Comment calculer le rang d'une famille $(\vec{x}_1, \cdots, \vec{x}_m)$ de vecteurs de $\mathbb{K}^n$

1ère étape. On dispose les m vecteurs en ligne. Si l'on note $\vec{x}_i = (x_{i1}, \cdots, x_{in})$, on obtient la matrice à p lignes et n colonnes suivante :

$$A = \begin{pmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_m \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}.$$

2ème étape. On applique aux lignes de la matrice A des transformations élémentaires $(T1)$, $(T2)$ ou $(T3)$ afin d'obtenir une matrice B <u>échelonnée</u> et de même rang que A. Pour ce faire, la méthode du pivot de Gauss est tout indiquée.

Conclusion. Le **rang de** A est égal au **nombre de lignes non nulles de** B.

Exemple : Calculer le rang de la famille $(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$ composée des vecteurs de $\mathbb{R}^5$ suivants :

$$\vec{V}_1 = (1,0,0,2,-1), \quad \vec{V}_2 = (0,1,-2,1,0), \quad \vec{V}_3 = (0,-1,2,1,-1), \quad \vec{V}_4 = (0,0,0,2,-1).$$

On écrit la matrice A de lignes $\vec{V}_1$, $\vec{V}_2$, $\vec{V}_3$ et $\vec{V}_4$:

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

puis on applique l'algorithme du pivot de Gauss afin de transformer A en une matrice échelonnée. En ajoutant (L_2) à (L_3), puis en retranchant (L_3) à (L_4), on trouve

$$\operatorname{rg} A = \operatorname{rg} \begin{pmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 1 & 0 & 0 & 2 & \text{-1} \\ 0 & 1 & \text{-2} & 1 & 0 \\ 0 & 0 & 0 & 2 & \text{-1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Cette dernière matrice est échelonnée avec trois lignes non nulles. Donc

$$\boxed{\operatorname{rg}(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4) = \operatorname{rg} A = 3.}$$

En examinant soigneusement les transformations élémentaires ayant permis de transformer A en une matrice échelonnée, on obtient comme renseignement supplémentaire le fait que $\vec{V}_4 = \vec{V}_2 + \vec{V}_3$.

Comme le rang de la famille $(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$ vaut 3, cela permet de conclure que $(\vec{V}_1, \vec{V}_2, \vec{V}_3)$ est une base de $\operatorname{Vect}(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$.

Remarque très importante : Dans le chapitre suivant, on démontrera que le rang d'une matrice est égal à celui de sa transposée : $\operatorname{rg} A^T = \operatorname{rg} A$.
Cette propriété a deux conséquences très importantes :
- pour calculer le rang d'une famille de vecteurs, on peut, si on le souhaite, disposer les vecteurs en colonne au lieu de les disposer en ligne, puis calculer le rang de la matrice ainsi obtenue,
- pour calculer le rang d'une matrice, on peut effectuer les transformations élémentaires sur les colonnes au lieu de les faire sur les lignes, voire alterner les deux types d'opérations.

3.5 Matrices inversibles

3.5.1 Définitions

Soit $b \in \mathcal{M}_{n,1}(\mathbb{K})$ et $A \in \mathcal{M}_n(\mathbb{K})$. Considérons l'equation $Ax = b$ pour $x \in \mathcal{M}_{n,1}(\mathbb{K})$. Nous avons vu que cette équation est équivalente à un système linéaire.
S'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $BA = I_n$, alors

$$Ax = b \quad \Longrightarrow \quad BAx = Bb \quad \Longrightarrow \quad x = Bb.$$

Donc si l'équation $Ax = b$ admet effectivement une solution, elle est nécessairement donnée par la formule $x = Bb$. Réciproquement, si de plus $AB = I_n$ alors

$$x = Bb \quad \Longrightarrow \quad Ax = ABb \quad \Longrightarrow \quad Ax = b,$$

donc $x = Bb$ est bien une solution de l'équation $Ax = b$.

Ces observations motivent les définitions qui suivent.

Définition 3.5.1. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. S'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que*

$$BA = AB = I_n,$$

on dit que A est **inversible**. *La matrice B est appelée* **inverse** *de A. On la note A^{-1}.*

Remarque 3.5.2. *D'après les considérations précédentes, lorsque A est inversible, l'équation $Ax = b$ admet une unique solution : $x = A^{-1}b$.*

1) La matrice I_n est inversible et $I_n^{-1} = I_n$. En effet, $I_n I_n = I_n$.

2) La matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est inversible. On a $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ (exercice : le vérifier).

3) La matrice $\mathbb{O}$ n'est pas inversible. En effet, si $\mathbb{O}$ était inversible, on aurait

$$I_n = \mathbb{O}\,\mathbb{O}^{-1} = \mathbb{O}$$

car pour toute matrice A, on a $\mathbb{O}\,A = \mathbb{O}$.

4) La matrice $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ n'est pas inversible. En effet, on vérifie facilement que $BB = \mathbb{O}$. Si l'on suppose (par l'absurde) que B^{-1} existe alors on a

$$B = B\,I_n = B\,(\,BB^{-1}\,) = (\,BB\,)\,B^{-1} = \mathbb{O}\,B^{-1} = \mathbb{O}.$$

Mais, $B \neq \mathbb{O}$!

5) Plus généralement pour les matrices 2×2, on dispose du critère et de la formule donnés dans la proposition ci-dessous :

Proposition 3.5.3. *Soit* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. *Alors*

(i) A *est inversible si et seulement si* $ad - bc \neq 0$.

(ii) Si $ad - bc \neq 0$ *alors* $A^{-1} = \dfrac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Preuve : Remarquons tout d'abord que

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ad-bc & bd-bd \\ -ac+ac & -bc+ad \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

En conséquence, si $ad - bc \neq 0$, en divisant l'égalité ci-dessus par $ad - bc$, on constate que A est inversible et que A^{-1} est bien donnée par l'expression de la proposition.

Pour la preuve du fait que A inversible implique $ad - bc \neq 0$, le lecteur devra patienter jusqu'au chapitre suivant. ■

Remarque très importante : Si A est une matrice $n \times m$ avec $n \neq m$, on ne peut pas définir l'inverse de A. En fait, les produits AB et BA sont bien définis si seulement si B est une matrice $m \times n$. Mais, si $m \neq n$, les matrices AB et BA sont de tailles différentes.
On pourrait songer à étendre légèrement la définition de matrice inverse et chercher une matrice B de taille $m \times n$ telle que

$$AB = I_n \in \mathcal{M}_n(\mathbb{K}), \quad \text{et} \quad BA = I_m \in \mathcal{M}_m(\mathbb{K}).$$

Mais si $n \neq m$, on peut démontrer qu'une telle matrice n'existe jamais...

3.5.2 Propriétés des matrices inverses

Lemme 3.5.4. *Soit* $A \in \mathcal{M}_n(\mathbb{K})$ *une matrice inversible. Alors, la matrice inverse de* A *est unique.*

Preuve : Soit $B, C \in \mathcal{M}_n(\mathbb{K})$ deux matrices inverses de A. Alors, on a

$$B = B\, I_n = B\,(AC) = (BA)\,C = I_n\, C = C.$$

■

En pratique on utilise souvent la proposition suivante qui sera démontrée dans le chapitre suivant.

Proposition 3.5.5. *Soit* $A \in \mathcal{M}_n(\mathbb{K})$. *Les trois propriétés suivantes sont équivalentes :*

(a) la matrice A *est inversible,*
(b) il existe $B \in \mathcal{M}_n(\mathbb{K})$ *telle que*[2] $AB = I_n$,
(c) il existe $C \in \mathcal{M}_n(\mathbb{K})$ *telle que*[3] $CA = I_n$.

Remarque 3.5.6.

- *Dans le cas (b),* $A^{-1} = B$. *En effet,* $A^{-1} = A^{-1}(AB) = (A^{-1}A)\,B = B$.
- *Dans le cas (c),* $A^{-1} = C$. *En effet,* $A^{-1} = (CA)\,A^{-1} = C\,(AA^{-1}) = C$.

Enfin, vérifions la

Proposition 3.5.7. *On a :*

(i) Si $A, B \in \mathcal{M}_n(\mathbb{K})$ *sont inversibles, la matrice* AB *est inversible et* $(AB)^{-1} = B^{-1}A^{-1}$;

2. On dit alors que A est **inversible à droite**
3. On dit alors que A est **inversible à gauche**

(ii) Si $A \in \mathcal{M}_n(\mathbb{K})$ *est inversible, la matrice* A^{-1} *est inversible et* $(A^{-1})^{-1} = A$ *;*
(iii) Si $A \in \mathcal{M}_n(\mathbb{R})$ *est inversible, la matrice* A^T *est inversible et* $(A^T)^{-1} = (A^{-1})^T$.

Preuve : Si A et B sont inversibles, on a

$$(B^{-1}A^{-1})\,(AB) = B^{-1}\,(A^{-1}A)\,B = B^{-1}B = I_n;$$

Un calcul analogue donne $(AB)(B^{-1}A^{-1}) = I_n$. Cela achève la preuve de la propriété *(i)*. La deuxième propriété est évidente. Le calcul

$$(B^{-1})^T\,B^T = (BB^{-1})^T = I_n^T = I_n \quad \text{et} \quad B^T(B^{-1})^T = (B^{-1}B)^T = I_n^T = I_n$$

implique la troisième propriété. ■

3.5.3 Calcul de l'inverse par la méthode du pivot de Gauss

Pour vérifier qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible et pour calculer la matrice inverse on utilise souvent la méthode du pivot de Gauss. Dans cette partie, nous décrivons l'algorithme et considérons deux exemples.

Algorithme d'inversion des matrices

Soit A une matrice $n \times n$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Première étape. Pour $b \in \mathbb{K}^n$ arbitraire, on considère le système linéaire

$$\text{(3.15)} \qquad \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{array} \right.$$

À l'aide de la méthode du pivot de Gauss, on transforme (3.15) en un système échelonné. Un système échelonné $n \times n$ est toujours triangulaire supérieur. Donc, on peut écrire le résultat sous la forme

$$\text{(3.16)} \quad \left\{ \begin{array}{r} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n = \widetilde{b}_1 \\ c_{22}x_2 + \cdots + c_{2n}x_n = \widetilde{b}_2 \\ \dots\dots\dots\dots\dots \\ c_{nn}x_n = \widetilde{b}_n \end{array} \right. \quad \text{où} \quad \begin{array}{l} \widetilde{b}_1 = d_{11}b_1 + d_{12}b_2 + \cdots + d_{1n}b_n, \\ \widetilde{b}_2 = d_{22}b_1 + d_{22}b_2 + \cdots + d_{2n}b_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \widetilde{b}_n = d_{n1}b_1 + d_{n2}b_2 + \cdots + d_{nn}b_n. \end{array}$$

Par construction, les coefficients c_{ij} et d_{ij} ne dépendent que de la matrice A.

Deuxième étape. Du fait que le système (3.16) est échelonné, deux cas peuvent se présenter :
- *Premier cas :* $c_{nn} = 0$. Alors l'analyse faite dans la section 1.3 du chapitre 1 montre que le système (3.15) a, ou bien pas de solution du tout, ou bien une infinité de solutions, suivant la valeur de b. D'après la remarque 3.5.2, la matrice A n'est donc pas inversible, et **on peut arrêter les calculs...**
- *Deuxième cas :* $c_{nn} \neq 0$. Comme le système est échelonné, les autres coefficients diagonaux c_{ii} ne sont pas nuls non plus. En appliquant la méthode de la remontée, on en déduit que, qu'elle que soit la valeur de b, le système (3.15) a une unique solution, et l'on peut passer à l'étape suivante de l'algorithme.

Troisième étape. On suppose ici que $c_{nn} \neq 0$. On applique alors au système (3.16) la méthode de la remontée. On obtient

$$\begin{array}{c} x_1 = r_{11}b_1 + r_{12}b_2 + \cdots + r_{1n}b_n, \\ x_2 = r_{21}b_1 + r_{22}b_2 + \cdots + r_{2n}b_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n = r_{n1}b_1 + r_{n2}b_2 + \cdots + r_{nn}b_n. \end{array}$$

Les coefficients r_{ij} ne dépendent que de la matrice A. On définit alors

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{pmatrix}.$$

Reste à vérifier que $AR = I_n$. En vertu de la remarque 3.5.6, cela impliquera que A est inversible, et que $A^{-1} = R$.

Afin de montrer que $AR = I_n$, on note

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Alors, on a

$$\forall b \in \mathcal{M}_{n,1}(\mathbb{K}), \quad Ax = b \quad \text{et} \quad x = Rb.$$

Donc, $\forall b \in \mathcal{M}_{n,1}(\mathbb{K})$, on obtient $(AR)b = b$. On fixe j dans $\{1, \cdots, n\}$ et on choisit b tel que

$$b_j = 1 \quad \text{et} \quad b_i = 0 \quad \text{si} \quad i \neq j$$

de telle sorte que $(AR)b$ soit égale à la j-ième colonne de la matrice AR.

L'égalité $(AR)b = b$ implique alors que la j-ième colonne de AR est égale à la j-ième colonne de I_n. Comme j est arbitraire, on peut conclure que $AR = I_n$. ■

Remarque 3.5.8. *Notons au passage que l'algorithme proposé a permis d'établir que la matrice A est inversible si et seulement si le coefficient c_{nn} du système échelonné* (3.16) *n'est pas nul.*

Exemples

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

• Première étape : On considère le système :

$$\left\{ \begin{array}{rcl} 1x_1 + 0x_2 + 1x_3 & = & b_1 \\ 1x_1 + 1x_2 + 1x_3 & = & b_2 \\ 0x_1 + 1x_2 + 1x_3 & = & b_3 \end{array} \right. \iff \begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 1 & 1 & 1 & b_2 \\ 0 & 1 & 1 & b_3 \end{array}$$

En utilisant la méthode du pivot de Gauss, on obtient

$$\begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 1 & 1 & 1 & b_2 \\ 0 & 1 & 1 & b_3 \end{array} \iff \begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 0 & 1 & 0 & -b_1 + b_2 \\ 0 & 1 & 1 & b_3 \end{array} \iff \begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 0 & 1 & 0 & -b_1 + b_2 \\ 0 & 0 & 1 & b_1 - b_2 + b_3 \end{array}$$

• Deuxième étape : La matrice obtenue est triangulaire supérieure à diagonale non nulle. Le lemme 3.5.8 implique donc que la matrice A est inversible.

• Troisième étape : en utilisant la méthode de la remontée, on obtient

$$\begin{cases} x_1 = b_2 - b_3, \\ x_2 = -b_1 + b_2, \\ x_3 = b_1 - b_2 + b_3. \end{cases}$$

Donc,

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

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

• Première étape : on considère le système : $\begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 1 & 1 & 1 & b_2 \\ 2 & 1 & 2 & b_3 \end{array}$

En utilisant la méthode du pivot de Gauss, on obtient

$$\begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 1 & 1 & 1 & b_2 \\ 2 & 1 & 2 & b_3 \end{array} \iff \begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 0 & 1 & 0 & -b_1 + b_2 \\ 0 & 1 & 0 & b_3 \end{array} \iff \begin{array}{ccc|c} 1 & 0 & 1 & b_1 \\ 0 & 1 & 0 & -b_1 + b_2 \\ 0 & 0 & 0 & b_1 - b_2 + b_3 \end{array}$$

• Deuxième étape : La matrice obtenue est triangulaire supérieure avec un 0 sur la diagonale. La remarque 3.5.8 permet de conclure que la matrice A n'est pas inversible.

3.6 Interprétation matricielle de la méthode du pivot de Gauss

Dans tout ce qui suit, $n \geq 1$ et toutes les matrices appartiennent à $\mathcal{M}_n(\mathbb{K})$.

3.6.1 Matrices élémentaires

• Pour $(k, l) \in \{1, \cdots, n\}^2$ tels que $k \neq l$, on note A_{kl} la matrice dont tous les coefficients sont nuls sauf les coefficients diagonaux qui ne sont pas d'indice kk ou ll, et les kl-ième et lk-ième coefficients, qui sont égaux à 1.

Exemple : Matrice A_{23} dans le cas $n = 4$:

$$A_{23} = \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

• Pour $j \in \{1, \cdots, n\}$ et $\lambda \in \mathbb{K} \setminus \{0\}$, on note $B_j(\lambda)$ la matrice diagonale dont tous les coefficients diagonaux sont égaux à 1 sauf le jj-ième coefficient qui vaut λ.

Exemple : Matrice $B_2(\lambda)$ dans le cas $n = 4$:

$$B_2(\lambda) = \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

• Pour $(k, l) \in \{1, \cdots, n\}^2$ tels que $k \neq l$, et $\alpha \in \mathbb{K}$, on note $C_{k,l}(\alpha)$ la matrice qui coïncide avec la matrice identité au kl-ième coefficient près qui vaut α.

Exemple : Matrice $C_{14}(\alpha)$ dans le cas $n = 4$:

$$C_{14}(\alpha) = \begin{pmatrix} \mathbf{1} & 0 & 0 & \alpha \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

Les matrices A_{kl}, $B_j(\lambda)$ et $C_{kl}(\alpha)$ définies ci-dessus sont appelées **matrices élémentaires**.

On laisse au lecteur la vérification du lemme suivant :

Lemme 3.6.1. *Soit M une matrice $n \times n$. Alors,*

(i) faire l'opération $M \to A_{kl}M$ revient à échanger les lignes k et l de M (transformation de type (T1)),

(ii) faire l'opération $M \to B_j(\lambda)M$ revient à multiplier par λ la j-ième ligne de M (transformation de type (T2)),

(iii) faire l'opération $M \to C_{kl}(\alpha)M$ revient à ajouter α fois la l-ième ligne à la k-ième ligne de M (transformation de type(T3)).

Lemme 3.6.2. *Toutes les matrices élémentaires sont inversibles ; on a*

$$(A_{kl})^{-1} = A_{kl}, \quad (B_j(\lambda))^{-1} = B_j(1/\lambda), \quad (C_{kl}(\alpha))^{-1} = C_{kl}(-\alpha).$$

Preuve : On démontre ce lemme en utilisant les relations entre les matrices élémentaires et les transformations élémentaires établies dans le lemme 3.6.1. Par exemple, on a

$$B_j(1/\lambda)\, B_j(\lambda) I_n = I_n. \tag{3.17}$$

En effet, en multipliant une matrice par $B_j(\lambda)$, on multiplie sa j-ième ligne par λ et, en multipliant une matrice par $B_j(1/\lambda)$, on divise sa j-ième ligne par λ. Donc, en multipliant une matrice par $B_j(1/\lambda)\, B_j(\lambda)$, on ne change rien. En échangeant les rôles de λ et de $1/\lambda$, on obtient de même $B_j(\lambda)\, B_j(1/\lambda) = I_n$. Donc $B_j(\lambda)$ est inversible, d'inverse $B_j(\lambda^{-1})$.

La preuve de l'inversibilité des autres matrices élémentaires est similaire. ■

3.6.2 Interprétation matricielle de la méthode de Gauss

On rappelle que le pivot de Gauss est une méthode itérative permettant à transformer n'importe quel système linéaire en un système échelonné par une succession finie de transformations élémentaires. En pratique, on peut se limiter aux transformations de type (T1) ou (T3). On rappelle aussi que tout système linéaire $n \times n$ est équivalent à une équation de la forme $Mx = b$ avec $M \in \mathcal{M}_n(\mathbb{K})$ et $b \in \mathcal{M}_{n,1}(\mathbb{K})$, l'inconnue x étant une matrice colonne : $x \in \mathcal{M}_{n,1}(\mathbb{K})$.

Le lemme 3.6.1 montre qu'effectuer une transformation élémentaire sur un système linéaire revient à multiplier à gauche la matrice du système par une matrice élémentaire.

On en déduit le résultat suivant :

Théorème 3.6.3. *Soit M une matrice $n \times n$. Il existe une famille finie E_1, E_2 ... E_L de matrices élémentaires $n \times n$ de type A_{kl} ou $C_{kl}(\alpha)$ telles que la matrice*

$$U = E_L\, E_{L-1} \ldots E_2\, E_1 M \tag{3.18}$$

soit une matrice $n \times n$ échelonnée.

Comme le produit de matrices inversibles est inversible et comme toutes les matrices élémentaires sont inversibles, on constate que M et U sont inversibles simultanément.

Par ailleurs, une matrice $n \times n$ échelonnée est triangulaire supérieure et on montrera dans le chapitre suivant qu'une matrice triangulaire supérieure est inversible si et seulement si tous ses coefficients diagonaux sont non nuls.

Toutes ces observations conduisent au

Lemme 3.6.4. *Sous les hypothèses du théorème* 3.6.3, *la matrice M est inversible si et seulement si la matrice U est une matrice triangulaire supérieure à coefficients diagonaux tous non nuls.*

3.7 Quelques compléments (hors-programme)

Dans cette partie, nous donnons quelques notions supplémentaires sur les matrices afin de satisfaire la curiosité du lecteur voulant aller un peu plus loin dans la théorie. Toutes ces notions sont hors-programme cette année. Elles seront vues plus en détails dans les cours d'algèbre linéaire du semestre prochain.

3.7.1 Dimension

En début de chapitre, nous avons établi que l'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ était un espace vectoriel. En généralisant légèrement les définitions du chapitre précédent (réservées cette année aux familles de vecteurs de $\mathbb{K}^n$), on peut se demander si l'espace vectoriel $\mathcal{M}_{m,n}(\mathbb{K})$ possède des bases et, dans l'affirmative, déterminer sa dimension (c'est-à-dire le nombre d'éléments de ses bases).

Proposition 3.7.1. *La dimension de l'espace $\mathcal{M}_{m,n}(\mathbb{K})$ est égale à mn.*

Preuve : Pour chaque couple (i,j) tel que $1 \leq i \leq m$ et $1 \leq j \leq n$, notons $M(i,j)$ la matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ dont tous les coefficients sont nuls sauf le ij-ième qui vaut 1. Ces matrices forment une famille libre. De plus, quelle que soit la matrice A, on a

$$A = \sum_{1\leq i\leq m} \sum_{1\leq j\leq n} (A)_{ij} M(i,j).$$

Cela montre que cette famille libre est génératrice. Donc c'est une base de $\mathcal{M}_{m,n}(\mathbb{K})$. La dimension de $\mathcal{M}_{m,n}(\mathbb{K})$ est égale au nombre des éléments de la base, c'est-à-dire au nombre des matrices $M(i,j)$, donc à mn. ■

3.7.2 Matrices conjuguées

Définition 3.7.2. *Soit $A \in \mathcal{M}_{m,n}(\mathbb{C})$. On appelle* **matrice conjuguée** *de A la matrice A^* de $\mathcal{M}_{n,m}(\mathbb{C})$ définie par les formules*

$$\forall\ 1 \leq k \leq n,\ 1 \leq \ell \leq m, \qquad (A^*)_{k\ell} = \overline{(A)_{\ell k}},$$

où la barre désigne la conjugaison complexe.

Exemple : Soit $A = \begin{pmatrix} i & 0 \\ i & 1 \end{pmatrix}$. Alors $A^* = \begin{pmatrix} -i & -i \\ 0 & 1 \end{pmatrix}$.

L'opération de conjugaison a des propriétés similaires à celles de l'opération de transposition [4]. Le lecteur pourra vérifier par lui-même les deux lemmes suivants.

4. Du reste, restreinte aux matrices réelles, elle coïncide avec la transposition.

Lemme 3.7.3. *Soit* $\alpha, \beta \in \mathbb{C}$ *et* $A, B \in \mathcal{M}_{m,n}(\mathbb{C})$. *Alors,*

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*. \tag{3.19}$$

Remarque : Ce lemme signifie que l'opération de conjugaison est **anti-linéaire**.

Lemme 3.7.4. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{C})$. *Alors,*

$$(A^*)^* = A. \tag{3.20}$$

Remarque : Ce lemme signifie que l'opération de conjugaison est **involutive**.

Une adaptation facile de la preuve de la proposition (3.2.5) donne le résultat suivant :

Proposition 3.7.5. *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{C})$ *et* $B \in \mathcal{M}_{n,p}(\mathbb{C})$. *Alors,*

$$(AB)^* = B^* A^*.$$

Si de plus A *est inversible alors* A^* *aussi et l'on a* $(A^*)^{-1} = (A^{-1})^*$.

De la définition 3.7.2, on déduit que la conjuguée d'une matrice carrée de taille $n \times n$ est aussi une matrice carrée de taille $n \times n$. Cela motive la définition suivante (à mettre en parallèle avec la définition de matrice symétrique) :

Définition 3.7.6. *Soit* $A \in \mathcal{M}_n(\mathbb{C})$. *On dit que* A *est une* **matrice hermitienne** *si* $A^* = A$.

Exemple : La matrice $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ est hermitienne.

3.7.3 Trace

Définition 3.7.7. *Soit* $A \in \mathcal{M}_n(\mathbb{K})$. *Le nombre*

$$\operatorname{Tr} A = \sum_{j=1}^{n} (A)_{jj}$$

est appelé **trace** *de la matrice* A.

Étudions les propriétés de la trace.

Lemme 3.7.8. *On a*

1) $\forall\ A, B \in \mathcal{M}_n(\mathbb{K})$, $\qquad \operatorname{Tr}(A+B) = \operatorname{Tr} A + \operatorname{Tr} B$;

2) $\forall\ A \in \mathcal{M}_n(\mathbb{K})$ *et* $\forall\ \lambda \in \mathbb{K}$, $\qquad \operatorname{Tr}(\lambda A) = \lambda \operatorname{Tr} A$;

3) $\forall\ A, B \in \mathcal{M}_n(\mathbb{K})$, $\qquad \operatorname{Tr}(AB) = \operatorname{Tr}(BA)$;

4) $\forall\ A \in \mathcal{M}_n(\mathbb{R})$, $\qquad \operatorname{Tr}(A^T) = \operatorname{Tr} A$;

5) $\forall\ A \in \mathcal{M}_n(\mathbb{C})$, $\qquad \operatorname{Tr}(A^*) = \overline{\operatorname{Tr} A}$.

Preuve : Vérifions la troisième propriété. On a

$$\operatorname{Tr}(AB) = \sum_{j=1}^{n} (AB)_{jj} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} (A)_{ji} (B)_{ij} \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} (B)_{ij} (A)_{ji} \right) = \sum_{i=1}^{n} (BA)_{ii} = \operatorname{Tr}(BA).$$

On laisse au lecteur la démonstration des autres propriétés. ■

Bibliographie

[1] J.-M. Arnaudiès et H. Fraysse : Cours de mathématiques 1, Algèbre, **Dunod**.

[2] J.-M. Arnaudiès et J. Lelong-Ferrand : Cours de mathématiques, tome 1 (algèbre), **Dunod**.

[3] T. Cuesta : Algèbre 1, MIAS1, Polycopié de l'Université Paris XII.

[4] C. Deschamps et A. Warusfel : Mathématiques. Cours et exercices corrigés, **Dunod**.

[5] J. Dixmier : Cours de mathématiques du premier cycle, **Gauthier-Villars**.

[6] S. Lipschutz : Algèbre linéaire, cours et problèmes, **Série Schaum**.

[7] F. Liret et D. Martinais : Mathématiques pour le DEUG, Algèbre, 1ère année, **Dunod**.

[8] J.-M. Monier : Algèbre et géométrie, 1ère année, **Dunod**.

[9] E. Ramis, C. Deschamps et J. Odoux : Cours de mathématiques spéciales, algèbre, Vol. 1, **Masson**.