Algèbre linéaire

Peip Post PACES

2M001

2014-2015

Myriam Comte

POLYTECH
PARIS-UPMC

Table des matières

Chapitre 1

Rappels sur les nombres complexes

1.1 Introduction

On considère l'équation

$$x^3 + ax^2 + bx + c = 0,$$

qui après le changement de variable $x = y - \frac{a}{3}$ devient

$$y^3 + py + q = 0.$$

Dans la première moitié du 16e siècle le mathématicien Cardan donna une formule explicite pour une solution de cette équation

$$y = \Big(-\frac{q}{2} + \sqrt{\Big(\frac{q}{2}\Big)^2 + \Big(\frac{p}{3}\Big)^3}\Big)^{\frac{1}{3}} + \Big(-\frac{q}{2} - \sqrt{\Big(\frac{q}{2}\Big)^2 + \Big(\frac{p}{3}\Big)^3}\Big)^{\frac{1}{3}} \tag{1.1}$$

à laquelle on donna son nom. Cependant il s'interrogea et Bombelli par la suite sur les solutions de l'équation

$$x^3 = 15x + 4$$

dont 4 est solution. Mais si on remplace dans la formule (1.1), on se retrouve avec la quantité $\sqrt{-121}$, racine carrée d'un nombre négatif ! Surmontant sa répulsion, Bombelli note a plus b (meno di memo) , que l'on écrit $a + ib$ aujourd'hui et on calcule comme pour des nombres usuels en remplaçant i^2 par -1 quand on le rencontre. Ce nombre $\sqrt{-1}$ qui n'existe pas, sera appelé imaginaire, d'où la notation i. Dès lors, par exemple,

$$(2 + i)^3 = 2 + 11i$$
$$(2 - i)^3 = 2 - 11i$$

d'où, en additionnant les racines cubiques,

$$4 = (2 + 11i)^{\frac{1}{3}} + (2 - 11i)^{\frac{1}{3}},$$

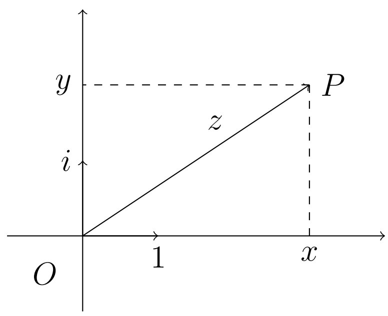
et on retrouve la formule (1.1). Les nombres complexes étaient nés. Par la suite Gauss prouva qu'un polynôme complexe non nul de degré n avait n racines comptées avec leur ordre de multiplicité à la fin du 18e siècle. Dès le 18e siècle, les nombres complexes s'étaient introduits dans le calcul intégral, et grâce aux développements en séries, Euler montra que ces nombres complexes permettent un lien entre la fonction exponentielle e^x et les fonctions trigonométriques sinus et cosinus :

$$e^{ix} = \cos x + i \sin x.$$

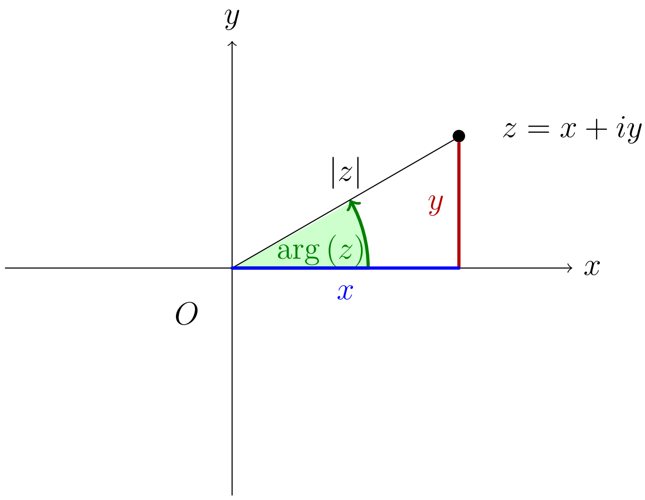
Il en déduisit la formule $e^{i\pi} = -1$. À la fin du 18e siècle et au début du 19e siècle, ces nombres furent interprétés géométriquement : ainsi la multiplication par i s'interprète comme une rotation d'un quart de tour dans le sens trigonométrique. Notons enfin qu'en physique, et notamment en électricité, la lettre i est réservée à l'intensité du courant, si bien que les physiciens préfèrent noter $j^2 = -1$. Malheureusement, la tradition mathématique donne à la lettre j une toute autre signification $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$...

1.2 Interprétation géométrique

On sait bien que le plan rapporté à un repère cartésien orthonormé est naturellement identifié à l'espace $\mathbb{R}^2$, au sens où, une fois un repère choisi, tout point P du plan est caractérisé par le couple $(x, y) \in \mathbb{R}^2$ de ses coordonnées cartésiennes et que réciproquement, tout couple de réels représente un unique point du plan dans ce repère. On obtient donc une interprétation géométrique de $\mathbb{C}$ en associant à tout nombre complexe $z = x + iy$ le point P du plan de coordonnées (x, y) (on a choisi un repère cartésien orthonormé une fois pour toutes). Cette interprétation explique pourquoi on parle parfois de plan complexe pour C. Le complexe z s'appelle l'affixe de P. On représente aussi le complexe z par le vecteur partant de l'origine et ayant P comme extrémité.



Définition 1.2.1. *On appelle module d'un nombre complexe z la longueur du segment OP. On le note $|z|$. On appelle argument d'un nombre complexe $z \neq 0$ l'angle orienté (défini modulo 2π) entre le segment OP et le demi axe positif des abscisses. On le note* $\arg(z)$.



Notons que l'on a $|0| = 0$, mais que l'argument de 0 n'est pas défini, puisque dans ce cas $P = O$. En fait, 0 est le seul nombre complexe de module nul.

Définition 1.2.2. *Soit $z = x + iy$ un nombre complexe. Le nombre x est appelé partie réelle de z et y sa partie imaginaire. On note*

$$x = \Re(z) \quad \textit{et} \quad y = \Im(z).$$

Proposition 1.2.3. *On a*

$$|z| = \sqrt{\Re(z)^2 + \Im(z)^2}$$

et

$$\begin{cases} \arg(z) = \arctan\left(\dfrac{\Im(z)}{\Re(z)}\right) \mod 2\pi \quad \textit{si} \quad \Re(z) > 0 \\ \arg(z) = \arctan\left(\dfrac{\Im(z)}{\Re(z)}\right) + \pi \mod 2\pi \quad \textit{si} \quad \Re(z) < 0 \\ \arg(z) = \dfrac{\pi}{2} \mod 2\pi \quad \textit{si} \quad \Re(z) = 0 \quad \textit{et} \quad \Im(z) > 0 \\ \arg(z) = -\dfrac{\pi}{2} \mod 2\pi \quad \textit{si} \quad \Re(z) = 0 \quad \textit{et} \quad \Im(z) < 0. \end{cases}$$

Réciproquement si on pose $\rho = |z|$ et $\theta = \arg(z)$, on a

$$\Re(z) = \rho\cos\theta$$
$$\Im(z) = \rho\sin\theta.$$

démonstration 1.2.4. *Le segment OP étant l'hypothénuse d'un triangle rectangle, d'après le théorème de Pythagore, on voit que*

$$|z| = \sqrt{\Re(z)^2 + \Im(z)^2}.$$

Pour l'argument, on voit que si $\arg z \neq \frac{\pi}{2} + k\pi$*, c'est-à-dire quand* $\Re(z) \neq 0$ *on a* $tg(\arg(z)) = \frac{\Im(z)}{\Re(z)}$ *(*tg *est la notation traditionnelle française pour la tangente, et non pas tan !). On en déduit les valeurs de l'argument à l'aide des fonctions circulaires inverses (rappelons que la fonction arctg prend ses valeurs dans l'intervalle* $]-\frac{\pi}{2}, \frac{\pi}{2}[$*. La formule en sens inverse découle directement de la définition des sinus et cosinus d'un angle. Ainsi, on a par exemple* $|i| = 1$ *et* $\arg(i) = \frac{\pi}{2}$*. On note aussi que si z est réel, son module est égal à sa valeur absolue : le module est le prolongement naturel de la valeur absolue de* $\mathbb{R}$ *à* $\mathbb{C}$*. C'est d'ailleurs la raison pour laquelle on utilise la même notation pour les deux notions.*

Remarque 1.2.5. *Le passage de* $(\Re(z), \Im(z))$ *à* $(|z|, \arg(z))$ *correspond du point de vue géométrique au passage des coordonnées cartésiennes aux coordonnées polaires. Le couple module-argument donne donc une autre représentation du plan complexe.*

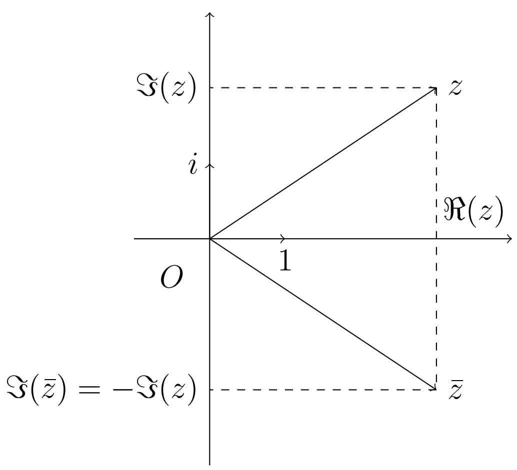
Définition 1.2.6. *Soit* $z = x + iy$ *un nombre complexe. On définit son nombre conjugué par*

$$\bar{z} = x - iy.$$

Ce qui donne

$$\Re(\bar{z}) = \Re(z) \quad \text{et} \quad \Im(\bar{z}) = -\Im(z).$$

Du point de vue géométrique, la conjugaison complexe est une symétrie orthogonale par rapport à l'axe réel.



Les nombres complexes de module 1 décrivent le cercle unité lorsque l'argument décrit un intervalle de longueur 2π.

Définition 1.2.7. *Pour tout* $\theta \in \mathbb{R}$*, on pose*

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{1.2}$$

C'est un nombre complexe de module 1*. On a*

$$e^{i(\theta+\theta')} = e^{i\theta} \times e^{i\theta'}. \tag{1.3}$$

De plus

$$\arg\left(e^{i\theta}\right)=\theta \quad \textit{et} \quad \frac{1}{e^{i\theta}}=\overline{e^{i\theta}}=e^{-i\theta}.$$

Remarque 1.2.8. *L'exponentielle complexe apparaît ici simplement comme une notation commode. Mais c'est en réalité une vraie exponentielle, c'est-à-dire qu'on peut démontrer que c'est nombre* e *élevé à la puissance* $i\theta$*. Notons le cas particulier de* $\theta=\pi$ *qui donne la célèbre formule d'Euler* $\mathrm{e}^{i\pi}=-1$.

On déduit de la définition les formules de De Moivre.

Proposition 1.2.9. *On a*

$$\cos\theta=\frac{e^{i\theta}+e^{-i\theta}}{2}$$
$$\sin\theta=\frac{e^{i\theta}-e^{-i\theta}}{2i}$$

Ces formules servent (entre autres) à linéariser les puissances des fonctions circulaires à l'aide de la formule du binôme. En effet, on a

$$(e^{i\theta})^n=e^{in\theta}, \quad \text{pour tout} \quad n \quad \text{entier}. \tag{1.4}$$

1.3 Polynômes

Un polynôme à une indéterminée à coefficients complexes est une expression de la forme

$$P(X)=a_0+a_1X+a_2X^2+\ldots+a_nX^n \tag{1.5}$$

avec $a_i\in\mathbb{C}$. Si $a_n\neq 0$, alors l'entier n est le degré de P. Par convention, le degré du polynôme nul est $-\infty$. La lettre X désigne l'indéterminée. On peut lui donner un sens mathématique précis, mais pour ce qui nous concerne, on peut tout aussi bien y penser comme à une variable. En fait, un polynôme à coefficients complexes définit une application de $\mathbb{C}$ dans $\mathbb{C}$ par

$$z\to P(z)=a_0+a_1z+a_2z^2+\ldots+a_nz^n. \tag{1.6}$$

L'ensemble de tous les polynômes à une indéterminée à coefficients complexes est noté $\mathbb{C}[X]$. On peut additionner et multiplier les polynômes entre eux de façon naturelle. Tout ce que l'on a dit plus haut se spécialise au cas réel, en remplaçant systématiquement $\mathbb{C}$ par $\mathbb{R}$. En particulier, tout polynôme à coefficients réels définit une fonction polynômiale de $\mathbb{R}$ dans $\mathbb{R}$, mais aussi de $\mathbb{C}$ dans $\mathbb{C}$. On dit qu'un nombre $\zeta\in\mathbb{C}$ est une racine d'un polynôme P si $P(\zeta)=0$. Un problème que l'on doit souvent résoudre est de trouver les racines des polynômes. On admettra le résultat suivant.

Proposition 1.3.1. *Si ζ est une racine de P, alors P est divisible par $(X-\zeta)$, c'est-à-dire qu'il existe un autre polynôme Q tel que*

$$P(X) = (X-\zeta)Q(X). \tag{1.7}$$

Corollaire 1.3.2. *Un polynôme de degré $n \geq 0$ admet au plus n racines.*

démonstration 1.3.3. *En effet, s'il en avait $n+1$, $\zeta_1, \zeta_2, \cdot, \zeta_{n+1}$, alors on aurait*

$$P(X) = (X-\zeta_1)(X-\zeta_2)\dots(X-\zeta_{n+1})Q(X)$$

donc le degré de P serait supérieur à $n+1$, qui est le degré du produit des $n+1$ premiers termes.

Le théorème suivant qui donne le nombre de racines d'un polynôme est célèbre, c'est le théorème de D'Alembert que l'on admettra car sa démonstration est loin d'être élémentaire.

Théorème 1.3.4. *Tout polynôme de $\mathbb{C}[X]$ de degré $n \geq 0$ admet exactement n racines complexes (distinctes ou confondues). En particulier, si $n \geq 1$, il se factorise entièrement en facteurs du premier degré*

$$P(X) = a_n(X-\zeta_1)(X-\zeta_2)\dots(X-\zeta_n).$$

Nous allons rappeler les résultats pour les polynômes à coefficients réels de degré 2.

Théorème 1.3.5. *Soit $P(X) = aX^2 + bX + c$ où $a \neq 0, b, c$ sont des éléments de $\mathbb{R}$. On appelle discriminant le nombre $\Delta = b^2 - 4ac$. Si $\Delta > 0$ alors P admet deux racines réelles distinctes*

$$\zeta_1 = \frac{-b-\sqrt{\Delta}}{2a} \quad \textit{et} \quad \zeta_2 = \frac{-b+\sqrt{\Delta}}{2a}.$$

Si $\Delta = 0$ alors P admet une racine double

$$\zeta_1 = \frac{-b}{2a}.$$

Si $\Delta < 0$ alors P admet deux racines complexes conjuguées

$$\zeta_1 = \frac{-b-i\sqrt{-\Delta}}{2a} \quad \textit{et} \quad \zeta_2 = \frac{-b+i\sqrt{-\Delta}}{2a}.$$

démonstration 1.3.6. *Il suffit de remarquer qu'on peut écrire P sous la forme*

$$P(X) = a\Big((X+\frac{b}{2a})^2 - \frac{b^2-4ac}{4a^2}\Big).$$

Chapitre 2

Matrices

Dans tout ce chapitre, on notera K pour désigner $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Définitions

Définition 2.1.1. *Une matrice est une application de $\{1, ..., p\} \times \{1, ..., n\}$ dans K assimilée à un tableau rectangulaire de n colonnes et p lignes. On note $M_{p,n}$ l'ensemble des matrices à p lignes et n colonnes.*

$$A = \begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{p,1} & \cdot & \cdot & a_{p,n} \end{pmatrix} = \begin{pmatrix} a_{i,j} \end{pmatrix}_{\{1 \leq i \leq p, 1 \leq j \leq n\}}$$

Le terme $a_{i,j}$ désigne ainsi le terme de la matrice situé sur la i-ème ligne et la j-ème colonne.

Définition 2.1.2. *On dit que la matrice est carrée si $n = p$ et on dit qu'elle est d'ordre n. On note $M_n(K)$ l'ensemble des matrices carrées à coefficients dans K.*

Définition 2.1.3. *Une matrice carrée d'ordre n est triangulaire supérieure si pour tout $i \in \{1, \cdot, n\}$, et tout $1 \leq j \leq i \leq n$, $a_{i,j} = 0$.*

Une matrice carrée d'ordre n est triangulaire inférieure si pour tout $i \in \{1, \cdot, n\}$, et tout $1 \leq i \leq j \leq n$, $a_{i,j} = 0$.

Une matrice carrée d'ordre n est diagonale si $a_{i,j} = 0$ pour $i \neq j$.

Cela signifie que pour une matrice triangulaire supérieure tous les termes situés sous la diagonale sont nuls et que pour une matrice triangulaire inférieure tous les termes situés au-dessus de la diagonale sont nuls. Pour une matrice diagonale, seuls les termes situés sur la diagonale sont non nuls.

Remarque 2.1.4. *On note I_n la matrice diagonale $n \times n$ dont les termes diagonaux sont égaux à* 1.

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Définition 2.1.5. *Soit*

$$A = \left(\; a_{i,j} \; \right) \in M_{p,n}.$$

On définit la transposée de A que l'on note A^T la matrice

$$A^T = \left(\; a_{j,i} \; \right) \in M_{n,p}.$$

La transposition échange les lignes et les colonnes. On utilise aussi la notation ${}^tA$.

2.2 Opérations sur les matrices

On définit l'addition de deux matrices de $M_{p,n}$ de la façon suivante :

Définition 2.2.1. *Soient*

$$A = \left(\; a_{ij} \; \right)_{\{1 \leq i \leq p, 1 \leq j \leq n\}} \text{ et } \quad B = \left(\; b_{ij} \; \right)_{\{1 \leq i \leq p, 1 \leq j \leq n\}},$$

deux éléments de $M_{p,n}$. On définit la matrice $C \in M_{p,n}$ de leur somme par

$$C = \left(\; c_{ij} \; \right)_{\{1 \leq i \leq p, 1 \leq j \leq n\}}, \text{ où } \quad c_{ij} = a_{ij} + b_{ij}.$$

De même, on définit la multiplication d'une matrice A par $\lambda \in K$,

Définition 2.2.2. *Soient*

$$A = \left(\; a_{ij} \; \right)_{\{1 \leq i \leq p, 1 \leq j \leq n\}} \text{ et } \quad \lambda \in K.$$

On définit la multiplication de la matrice A par $\lambda \in K$ par

$$\lambda A = \left(\; \lambda a_{ij} \; \right)_{\{1 \leq i \leq p, 1 \leq j \leq n\}}.$$

Peut-on définir le produit de deux matrices ? On ne peut faire le produit de la matrice A par la matrice B que si le nombre de colonnes de A est égal au nombre de lignes de B.

Définition 2.2.3. *On peut définir le produit de deux matrices $A \in M_{q,p}$ et $B \in M_{p,n}$ par $C \in M_{q,n}$ avec*

$$c_{i,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k}.$$

Exemples 2.2.4.

$$\textit{Si } A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix} \quad \textit{et} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}$$

le produit AB est bien défini, c'est un élément de $M_{2,2}$ et vaut

$$AB = \begin{pmatrix} 9 & -2 \\ -1 & -1 \end{pmatrix}$$

ATTENTION : le produit de deux matrices n'est pas commutatif, i.e. on n'a pas en général $AB = BA$.

Calculer

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

et

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

Théorème 2.2.5. *Supposons que*

$$A, A' \in M_{n,p}, \quad B, B' \in M_{p,q} \quad \textit{et} \quad C \in M_{q,r}.$$

Alors

i) $(A + A')B = AB + A'B$ et $A(B + B') = AB + AB'$,

ii) Pour tout $\lambda \in K$, $\lambda(AB) = (\lambda A)B = A(\lambda B)$,

iii) $(AB)C = A(BC)$ (associativité).

démonstration 2.2.6. *Notons tout d'abord que tous les produits sont bien définis.*

i) On applique la définition (2.2.3) : le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $(A + A')B$ est égal à

$$\sum_{j=1}^{p} (a_{i,j} + a'_{i,j}) b_{j,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k} + \sum_{j=1}^{p} a'_{i,j} b_{j,k}$$

et le terme à droite de l'égalité est le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $AB + A'B$. On procède de même pour l'autre égalité.

ii) On utilise à nouveau la définition (2.2.3) : le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $\lambda(AB)$ est égal à

$$\lambda(\sum_{j=1}^{p} a_{i,j} b_{j,k}) = \sum_{j=1}^{p} (\lambda a_{i,j}) b_{j,k} = \sum_{j=1}^{p} a_{i,j} (\lambda b_{j,k}).$$

iii) Utilisons toujours la définition (2.2.3), le terme situé sur la i-ème ligne et la j-ème colonne de la matrice AB est égal à

$$\sum_{l=1}^{p} a_{i,l}b_{l,j}.$$

Le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $(AB)C$ est donc égal à

$$\sum_{j=1}^{q}(\sum_{l=1}^{p} a_{i,l}b_{l,j})c_{j,k}$$

On inverse alors l'ordre des deux sommes

$$\sum_{j=1}^{q}(\sum_{l=1}^{p} a_{i,l}b_{l,j})c_{j,k} = \sum_{l=1}^{p}\sum_{j=1}^{q} a_{i,l}b_{l,j}c_{j,k} = \sum_{l=1}^{p} a_{i,l}(\sum_{j=1}^{q} b_{l,j}c_{j,k})$$

Le terme

$$\sum_{j=1}^{q} b_{l,j}c_{j,k}$$

est terme situé sur la l-ème ligne et la k-ème colonne de la matrice BC, le terme

$$\sum_{l=1}^{p} a_{i,l}(\sum_{j=1}^{q} b_{l,j}c_{j,k})$$

est donc bien le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $A(BC)$.

Proposition 2.2.7. *Soit $A \in M_{n,p}$ alors*

$$AI_p = I_nA = A.$$

démonstration 2.2.8. *Découle de façon immédiate de la définition (2.2.3).*

Remarque 2.2.9. *Attention aux mauvais réflexes :*

i) Si $AB = AC$, on ne peut pas simplifier par A pour dire que $B = C$: c'est faux en général.

ii) Si $AB = 0$, on ne peut pas dire que $A = 0$ ou $B = 0$: c'est faux en général.

Proposition 2.2.10. *Si le produit AB est bien défini, il en est de même du produit B^TA^T et on a $(AB)^T = B^TA^T$.*

démonstration 2.2.11. *Si $A \in M_{q,p}$ et $B \in M_{p,n}$ alors $AB \in M_{q,n}$. On a aussi $B^T \in M_{n,p}$ et $A^T \in M_{p,q}$ donc B^TA^T est bien défini et c'est un élément de $M_{n,q}$. Par* *définition, le terme situé sur la i-ème ligne et la k-ème colonne de la matrice B^TA^T est égal à*

$$\sum_{j=1}^{p} b_{j,i}a_{k,j} = \sum_{j=1}^{p} a_{k,j}b_{j,i},$$

qui est le terme situé sur la k-ème ligne et la i-ème colonne de la matrice AB et donc le terme situé sur la i-ème ligne et la k-ème colonne de la matrice $(AB)^T$.

Définition 2.2.12. *Une matrice $A \in M_n(K)$ est inversible si et seulement s'il existe une matrice $B \in M_n(K)$ telle que $AB = BA = I_n$. On note alors $B = A^{-1}$.*

Proposition 2.2.13. *Si A est inversible, alors son inverse est unique.*

démonstration 2.2.14. *On va supposer que A admet deux inverses B_1 et B_2, ce qui signifie que $AB_1 = B_1A = I_n$ et $AB_2 = B_2A = I_n$*

$$AB_1 = B_1A = I_n \quad et \quad AB_2 = B_2A = I_n.$$

On a alors

$$B_2(AB_1) = B_2 = (B_2A)B_1$$

par associativité, et donc

$$B_2 = B_1.$$

2.3 Systèmes linéaires

Les systèmes linéaires interviennent dans de nombreux contextes d'applications de l'algèbre linéaire : sciences de l'ingénieur, météorologie, économie, mais aussi codes de transmission d'information et cryptographie.

2.3.1 Définitions

On ne va traiter que le cas des systèmes linéaires faisant intervenir seulement des nombres réels. Mais tout ce que l'on dit dans ce cas reste valable pour des systèmes linéaires faisant intervenir des nombres complexes. Soit $n \in \mathbb{N}^*$ un entier naturel supérieur à 1. Une équation linéaire à n inconnues $x_1, x_2, ..., x_n$ est une équation de la forme

$$a_1x_1 + a_2x_2 + \ + a_nx_n = b,$$

où $a_1, a_2, ..., a_n$ et b sont des nombres réels donnés. Soit $m \in \mathbb{N}^*$ un autre entier naturel supérieur à 1.

Définition 2.3.1. *Un système de m équations linéaires à n inconnues, ou système linéaire, est une liste de m équations linéaires.*

On écrit usuellement de tels systèmes en m lignes placées les unes sous les autres.

Exemples 2.3.2. *Le système*

$$\begin{cases} x_1 + 2x_2 - x_3 = 1 \\ 2x_1 - 3x_2 = 4 \end{cases}$$

est un système linéaire de deux équations à trois inconnues.

La forme générale d'un système linéaire de m équations à n inconnues, ou encore système $m\times$n, est la suivante

$$\left\{ \begin{array}{ccccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,j}x_j & + & \dots & + & a_{1,n}x_n = b_1(\text{équation 1}) \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,j}x_j & + & \dots & + & a_{2,n}x_n = b_2(\text{équation 2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{i,1}x_1 & + & a_{i,2}x_2 & + & \dots & + & a_{i,j}x_j & + & \dots & + & a_{i,n}x_n = b_i(\text{équation i}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,j}x_j & + & \dots & + & a_{m,n}x_n = b_m(\text{équation m}) \end{array} \right.$$

Les nombres $a_{i,j}$, pour $i = 1, ..., m$, $j = 1, ..., n$, sont les coefficients du système. Ce sont des données. Les nombres b_i, $i = 1, ..., m$, constituent le second membre du système et sont également des données. Il convient de bien observer comment on a rangé le système en lignes (une ligne par équation) numérotées de 1 à m par l'indice i, et en colonnes : les termes correspondant à une même inconnue x_j sont alignés verticalement les uns sous les autres. L'indice j varie de 1 à n. Il y a donc n colonnes à gauche des signes d'égalité, plus une colonne supplémentaire à droite pour le second membre. La notation avec double indice $a_{i,j}$ correspond à ce rangement : le premier indice (ici i) est le numéro de ligne et le second indice (ici j) est le numéro de colonne. Il est extrêmement important de toujours respecter cette convention. Dans l'exemple ci-dessus, on a $m = 2$ (nombre d'équations = nombre de lignes), $n = 3$ (nombre d'inconnues = nombre de colonnes à gauche du signe =) et $a_{1,1} = 1$, $a_{1,2} = 2$, $a_{1,3} = -1$, $a_{2,1} = 2$, $a_{2,2} = -3$, $a_{2,3} = 0$, $b_1 = 1$ et $b_2 = 4$.

Définition 2.3.3. *Une solution du système linéaire est une liste de n nombres réels* $(s_1, s_2, ..., s_n)$ *(un n-uplet) tels que si l'on substitue* s_1 *pour* x_1, s_2 *pour* x_2*, etc., dans le système linéaire, on obtient une égalité. L'ensemble des solutions du système est l'ensemble de tous ces n-uplets.*

Ainsi, (2,0,1) est une solution du système linéaire de l'exemple ci-dessus. Lorsqu'on cherche à déterminer l'ensemble des solutions d'un système linéaire on dit que l'on veut résoudre ce système linéaire.

Définition 2.3.4. *On dit que deux systèmes linéaires sont équivalents s'ils ont le même ensemble de solutions.*

À partir de là, le jeu pour résoudre un système linéaire donné consistera à le transformer en un système équivalent dont la résolution sera plus simple que celle du système de départ.

Remarque 2.3.5. *Deux systèmes équivalents ont toujours le même nombre d'inconnues. En revanche, le nombre d'équations peut être différent. Dans ce dernier cas, on peut toujours ajouter au système avec le moins d'équations le nombre manquant à l'aide d'équations triviales* $0 \times x_1 + 0 \times x_2 + \ + 0 \times x_n = 0$*, lesquelles ne modifient pas l'ensemble des solutions !*

2.3.2 Systèmes échelonnés réduits

Pour résoudre un système linéaire, on va l'écrire sous forme matricielle, à savoir $AX = B$, où

$$A = \left(\ a_{ij} \ \right)_{\{1 \leq i \leq m, 1 \leq j \leq n\}} \quad X = \begin{pmatrix} x_1 \\ . \\ . \\ x_n \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_1 \\ . \\ . \\ b_m \end{pmatrix}.$$

Et on va le ramener à ce qu'on appelle un système échelonné réduit qui lui soit équivalent.

Définition 2.3.6. *Une matrice* A *est dite échelonnée si et seulement si elle a les deux propriétés suivantes*

1. *Si une ligne est entièrement nulle, toutes les lignes situées en dessous sont également entièrement nulles.*
2. *Dans chaque ligne non entièrement nulle (à partir de la deuxième), le premier coefficient non nul en comptant à partir de la gauche est situé strictement à droite du premier coefficient non nul de la ligne précédente.*

   On dit qu'une matrice est échelonnée réduite si et seulement elle a en plus les deux propriétés suivantes
3. *Le premier coefficient non nul d'une ligne en comptant à partir de la gauche vaut 1.*
4. *Et c'est le seul élément non nul de sa colonne.*

Remarque 2.3.7. *Grace à 1), on voit que 2) a un sens : si une ligne contient un élément non nul, alors la ligne précédente contient aussi un élément non nul, sinon cela contredirait 1). Par ailleurs, de 2) et de 1), on déduit que tous les coefficients situés dans la même colonne qu'un tel premier élément non nul d'une ligne et en dessous de cet élément, sont nuls.*

Définition 2.3.8. *Soit U une matrice échelonnée réduite. Les positions de pivot de U sont les emplacements (au sens du couple (numéro de ligne, numéro de colonne)) des coefficients valant 1 du point 3) de la définition 2.3.6.*

Exemples 2.3.9. *La matrice*

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.1}$$

est échelonnée mais n'est pas échelonnée réduite. La matrice

$$B = \begin{pmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \tag{2.2}$$

est échelonnée réduite et ses positions de pivot sont $(1,1)$, $(2,2)$ *et* $(3,4)$. *On reconnaît (à l'oeil) les matrices échelonnées à la disposition caractéristique des zéros en escalier descendant du haut à gauche vers le bas à droite.*

Le coefficient 1 situé en position $(3,5)$ n'est pas un pivot car il n'est pas le premier élément non nul de sa ligne. Dans une matrice échelonnée réduite, on appelle colonnes de pivot les colonnes qui contiennent une position de pivot et lignes de pivot les lignes qui contiennent une position de pivot. D'après le point 3) de la définition 2.3.6, on voit qu'il y a au plus une position de pivot par ligne, et d'après le point 4), au plus une position de pivot par colonne. Par conséquent, le nombre de colonnes de pivot est égal au nombre de lignes de pivot, tous deux étant égaux au nombre de positions de pivot. Cette observation banale jouera un role important plus loin dans les questions de dimension d'un espace vectoriel. Les positions de pivot permettent d'introduire une classification des inconnues.

Définition 2.3.10. *Les inconnues correspondant à une colonne de pivot sont appelées inconnues ou variables essentielles. Les autres sont appelées inconnues ou variables libres.*

Remarquons qu'un système échelonné a toujours au moins une variable essentielle, mais qu'il n'a pas forcément de variables libres. Pour résoudre les systèmes linéaires, nous avons besoin d'introduire une dernière définition.

Définition 2.3.11. *Soit le système linéaire* $AX = B$, *où*

$$A = \begin{pmatrix} a_{ij} \end{pmatrix}_{\{1\leq i\leq m, 1\leq j\leq n\}} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \textit{et} \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

On appelle la matrice augmentée du système, la matrice $\tilde{A}$ définie par

$$\tilde{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ a_{i,1} & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,n} & b_i \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,j} & \dots & a_{m,n} & b_m \end{pmatrix}.$$

On ajouté à la matrice A une colonne formée par le vecteur B.

Nous pouvons maintenant résoudre les systèmes échelonnés réduits dans tous les cas.

Théorème 2.3.12. *Un système échelonné réduit est compatible si et seulement si sa matrice augmentée ne contient aucune ligne de la forme*

$$\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix}$$

avec $b \neq 0$. Dans ce cas, on obtient une description paramétrique de l'ensemble des solutions en exprimant les variables essentielles en fonction du second membre et des variables libres.

démonstration 2.3.13. *Supposons que la matrice augmentée du système contienne une ligne de la forme*

$$\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix}$$

avec $b \neq 0$. Cette ligne correspond à l'équation linéaire

$$0 \times x_1 + 0 \times x_2 + \quad + 0 \times x_n = b,$$

laquelle n'a évidemment aucune solution. Le système est par conséquent incompatible, l'ensemble des solutions noté S est vide, $S = \emptyset$.

Dans le cas où aucune ligne n'est de cette forme, alors on peut visiblement résoudre. En effet, les éventuelles lignes nulles donnent des équations de la forme

$$0 \times x_1 + 0 \times x_2 + \quad + 0 \times x_n = 0,$$

qui sont toujours satisfaites. De plus, chaque ligne non nulle réécrite sous forme d'équation prend la forme

$$x_{i,l} + B_l(x_{libres}) = b_l,$$

où $x_{i,l}$ est la l-ème variable essentielle (qui n'apparaît que dans cette équation située à la ligne l), $B_l(x_{libres})$ est une somme composée de coefficients du système multipliés par les variables libres (désignées collectivement par x_{libres} mais en fait, seules celles *situées à droite de $x_{i,l}$ interviennent) s'il y a des variables libres, $B_l(x_{libres}) = 0$ s'il n'y en a pas, et b_l est la l-ème ligne du second membre. Par conséquent,*

$$x_{i,l} = -B_l(x_{libres}) + b_l,$$

fournit une représentation paramétrique de l'ensemble des solutions, les variables libres parcourant indépendamment $\mathbb{R}$.

On a donc établit le résultat suivant

Corollaire 2.3.14. *Dans le cas d'un système échelonné réduit $m \times n$ on a*
a) Soit il n'y a aucune solution s'il y a une ligne de la forme

$$\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix} \quad \textit{avec} \quad b \neq 0.$$

b) Soit il y a une solution unique s'il n'y a pas de telle ligne ni de variables libres.
c) Soit il y a une infinité de solutions s'il n'y a pas de telle ligne mais qu'il existe des variables libres.

2.3.3 Algorithme de Gauss

À partir de maintenant, la stratégie pour résoudre un système général sera de se ramener à un système échelonné réduit qui lui soit équivalent. On va pour cela raisonner uniquement sur les matrices échelonnées réduites, sans référence particulière aux systèmes linéaires, et introduire un algorithme travaillant sur les matrices à cet effet. Un algorithme est une description précise d'une suite d'opérations à effectuer, dans un ordre donné, qui aboutit au bout d'un nombre fini d'étapes si possible connu à l'avance au résultat voulu. Il est bien préférable de se laisser guider par une méthode stricte dont l'application garantit un nombre minimal de calculs (en général) que de s'essayer de résoudre par rapport à une variable, puis remplacer dans les autres équations ... C'est une source d'erreur de calcul. De plus, dans la pratique, les systèmes à résoudre sont énormes (des milliers, voire des millions d'équations et d'inconnues) et qu'il n'est pas question d'effectuer les calculs à la main. Ce sont des ordinateurs qui s'en chargent, et ces derniers ont besoin de programmes, lesquels sont la traduction de tel ou tel algorithme.

L'algorithme de Gauss (ou encore du pivot de Gauss) est fondé sur les notions suivantes.

Définition 2.3.15. *On appelle opérations élémentaires sur les lignes les trois opérations suivantes :*

i) Échanger deux lignes (échange).

ii) Multiplier une ligne par une constante non nulle (homothétie).

iii) Remplacer une ligne par elle-meme plus un multiple d'une autre ligne (substitution).

Il faut bien remarquer qu'en effectuant une opération élémentaire, on ne mélange jamais les colonnes. Ce que contient une colonne après l'opération ne dépend que de ce qu'elle contenait avant l'opération.

Définition 2.3.16. *On dit que deux matrices A et B de même taille $m \times n$ sont équivalentes si B se déduit de A par une suite finie d'opérations élémentaires. Dans ce cas, on note $A \sim B$.*

Proposition 2.3.17. *Si les matrices augmentées de deux systèmes linéaires sont équivalentes, alors les systèmes linéaires sont équivalents.*

démonstration 2.3.18. *Il suffit de le vérifier sur les opérations élémentaires, l'équivalence des systèmes se propageant visiblement de proche en proche à chacune d'entre elles. Soient donc deux systèmes linéaires dont les matrices augmentées*

$$A = \left(\begin{array}{c} a_{ij} \end{array} \right)_{\{1 \leq i \leq m, 1 \leq j \leq n+1\}} \quad \textit{et} \quad A' = \left(\begin{array}{c} a'_{ij} \end{array} \right)_{\{1 \leq i \leq m, 1 \leq j \leq n+1\}}$$

(on note $a_{i,n+1}$ la colonne correspondant au second membre pour simplifier la notation), diffèrent par une opération élémentaire. Notons S_A l'ensemble des solutions du système associé à A et $S_{A'}$ l'ensemble des solutions du système associé à A'. Il faut distinguer suivant les trois cas possibles.

Le cas de l'échange est clair : on intervertit l'ordre de deux équations ce qui ne change pas l'ensemble des solutions.

Le cas de l'homothétie : $a'_{i,j} = \lambda a_{i,j}$ avec $\lambda \neq 0$ pour un certain i et tout $j = 1, ..., n+1$. Soit $(s_1, s_2, ..., s_n) \in S_A$. Ce n-uplet vérifie en particulier l'équation numéro i

$$a_{i,1}s_1 + a_{i,2}s_2 + \quad + a_{i,n}s_n = a_{i,n+1}.$$

Multipliant les deux membres par λ, on voit que

$$a'_{i,1}s_1 + a'_{i,2}s_2 + \quad + a'_{i,n}s_n = a'_{i,n+1}.$$

et comme les autres équations du système associé à A' sont les mêmes que celles de A, on en déduit que $(s_1, s_2, ..., s_n) \in S_{A'}$. En d'autres termes, on vient de montrer que $S_A \subset S_{A'}$. Inversant les rôles de A et A', on en déduit que $S_{A'} \subset S_A$, d'où finalement $S_A = S_{A'}$, les deux systèmes sont équivalents.

Le cas de la substitution est très semblable :

$$a'_{i,j} = a_{i,j} + \lambda a_{k,j}$$

pour un certain i, un certain k et tout $j = 1, ..., n+1$. Soit $(s_1, s_2, ..., s_n) \in S_A$. Ce n-uplet vérifie en particulier les équations numéros i et k

$$a_{i,1}s_1 + a_{i,2}s_2 + \quad + a_{i,n}s_n = a_{i,n+1}$$
$$a_{k,1}s_1 + a_{k,2}s_2 + \quad + a_{k,n}s_n = a_{k,n+1}$$

d'où en multipliant la deuxième égalité par λ et en additionnant

$$a_{i,1}s_1 + a_{i,2}s_2 + \quad + a_{i,n}s_n + \lambda(a_{k,1}s_1 + a_{k,2}s_2 + \quad + a_{k,n}s_n) = a_{i,n+1} + \lambda a_{k,n+1}$$

On factorise le membre de gauche

$$(a_{i,1} + \lambda a_{k,1})s_1 + (a_{i,2} + \lambda a_{k,2})s_2 + \ldots + (a_{i,n} + \lambda a_{k,n})s_n = a_{i,n+1} + \lambda a_{k,n+1},$$

qui n'est autre que

$$a'_{i,1}s_1 + a'_{i,2}s_2 + \quad + a'_{i,n}s_n = a'_{i,n+1}.$$

Les autres équations n'étant pas modifiées, on en déduit comme précédemment que $S_A \subset S_{A'}$, puis que $S_A = S_{A'}$. Les deux systèmes sont équivalents.

Les opérations élémentaires appliquées aux matrices augmentées produisant des systèmes équivalents entre eux, on va s'en servir pour se ramener à un système échelonné réduit. Le théorème fondamental est le suivant.

Théorème 2.3.19. *Toute matrice A est équivalente à une unique matrice échelonnée réduite U.*

démonstration 2.3.20. *Ce théorème est en deux parties, une partie d'existence (il existe U échelonnée réduite équivalente à A) et une partie unicité (c'est la seule matrice échelonnée réduite). Commençons par l'existence, laquelle se démontre grâce à l'algorithme de Gauss proprement dit. L'idée générale de l'algorithme de Gauss consiste à utiliser des substitutions de lignes pour placer des zéros là où il faut de façon à créer d'abord une forme échelonnée, puis une forme échelonnée réduite. Soit A une matrice $m \times n$ quelconque.*

Algorithme de Gauss

1) Passage à une forme échelonnée.

Étape 1 : Choix du pivot.

On commence par inspecter la première colonne. Soit elle ne contient que des zéros, auquel cas on passe directement à l'étape 3, soit elle contient au moins un terme non nul. On choisit alors un tel terme, que l'on appelle le pivot. Si c'est le terme $a_{1,1}$ on passe directement à l'étape 2, si c'est un terme $a_{i,1}$ avec $i \neq 1$, on échange les lignes 1 *et i et on passe à l'étape 2. Au terme de l'étape 1, on a obtenu une matrice de la forme*

$$A = \begin{pmatrix} 0 & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{m,2} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix},$$

dans le premier cas, ou bien

$$A \sim \begin{pmatrix} a'_{1,1} & a'_{1,2} & \dots & a'_{1,j} & \dots & a'_{1,n} \\ a'_{2,1} & a'_{2,2} & \dots & a'_{2,j} & \dots & a'_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{i,1} & a'_{i,2} & \dots & a'_{i,j} & \dots & a'_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{m,1} & a'_{m,2} & \dots & a'_{m,j} & \dots & a'_{m,n} \end{pmatrix},$$

avec $a'_{1,1} \neq 0$ dans le second cas.

Étape 2 : Élimination.

On ne touche plus à la ligne 1, et on se sert du pivot $a'_{1,1}$ pour éliminer tous les termes $a'_{i,1}$, pour $i \geq 2$. Pour cela, il suffit de remplacer la ligne i par elle-même moins $\frac{a'_{i,1}}{a'_{1,1}} \times$ la ligne 1, ceci pour $i = 2, ..., m$.

De cette façon, on fait apparaître des 0 en début de chaque ligne, à partir de la deuxième.

Au terme de l'étape 2, on a obtenu une matrice de la forme

$$A \sim \begin{pmatrix} a'_{1,1} & a'_{1,2} & \dots & a'_{1,j} & \dots & a'_{1,n} \\ 0 & a''_{2,2} & \dots & a''_{2,j} & \dots & a''_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{i,2} & \dots & a''_{i,j} & \dots & a''_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{m,2} & \dots & a''_{m,j} & \dots & a''_{m,n} \end{pmatrix}.$$

Étape 3 : Boucle.

Au début de l'étape 3, on a obtenu dans tous les cas de figure une matrice de la *forme*

$$A \sim \begin{pmatrix} \tilde{a}_{1,1} & a^1_{1,2} & \dots & \tilde{a}_{1,j} & \dots & \tilde{a}_{1,n} \\ 0 & \tilde{a}_{2,2} & \dots & \tilde{a}_{2,j} & \dots & \tilde{a}_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & \tilde{a}_{i,2} & \dots & \tilde{a}_{i,j} & \dots & \tilde{a}_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & \tilde{a}_{m,2} & \dots & \tilde{a}_{m,j} & \dots & \tilde{a}_{m,n} \end{pmatrix},$$

dont la première colonne est bien celle d'une matrice échelonnée. On va donc conserver cette première colonne. Si $\tilde{a}_{1,1} \neq 0$*, on conserve aussi la première ligne, et l'on va boucler sur l'étape 1 en l'appliquant à la sous-matrice* $(m-1) \times (n-1)$ *qui reste*

$$\begin{pmatrix} \tilde{a}_{2,2} & \dots & \tilde{a}_{2,j} & \dots & \tilde{a}_{2,n} \\ \vdots & & \vdots & & \vdots \\ \tilde{a}_{i,2} & \dots & \tilde{a}_{i,j} & \dots & \tilde{a}_{i,n} \\ \vdots & & \vdots & & \vdots \\ \tilde{a}_{m,2} & \dots & \tilde{a}_{m,j} & \dots & \tilde{a}_{m,n} \end{pmatrix}.$$

Si $\tilde{a}_{1,1} = 0$*, on boucle sur l'étape 1 en l'appliquant à la sous-matrice* $m \times (n-1)$

$$\begin{pmatrix} \tilde{a}_{1,2} & \dots & \tilde{a}_{1,j} & \dots & \tilde{a}_{1,n} \\ \tilde{a}_{2,2} & \dots & \tilde{a}_{2,j} & \dots & \tilde{a}_{2,n} \\ \vdots & & \vdots & & \vdots \\ \tilde{a}_{i,2} & \dots & \tilde{a}_{i,j} & \dots & \tilde{a}_{i,n} \\ \vdots & & \vdots & & \vdots \\ \tilde{a}_{m,2} & \dots & \tilde{a}_{m,j} & \dots & \tilde{a}_{m,n} \end{pmatrix}.$$

Au terme de cette deuxième itération de la boucle, on aura obtenu une matrice de la forme

$$A \sim \begin{pmatrix} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \dots & \tilde{a}_{1,j} & \dots & \tilde{a}_{1,n} \\ 0 & \tilde{\tilde{a}}_{2,2} & \dots & \tilde{\tilde{a}}_{2,j} & \dots & \tilde{\tilde{a}}_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \tilde{\tilde{a}}_{i,j} & \dots & \tilde{\tilde{a}}_{i,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \tilde{\tilde{a}}_{m,j} & \dots & \tilde{\tilde{a}}_{m,n} \end{pmatrix},$$

et ainsi de suite. Comme chaque itération de la boucle travaille sur une matrice qui a une colonne de moins que la précédente, il est clair qu'au bout d'au plus $n-1$ *itérations de la boucle, on aura construit une matrice échelonnée équivalente à la matrice de départ.*

2) Passage à une forme échelonnée réduite.

Étape 1 : Homothéties. On repère le premier élément non nul de chaque ligne non nulle, et on multiplie cette ligne par l'inverse de cet élément. Ceci crée une matrice échelonnée avec des 1 *en position de pivot.*

Étape 2 : Élimination. On élimine les termes situés au dessus des positions de pivot comme précédemment, en procédant à partir du bas à droite de la matrice. Ceci ne modifie pas la structure échelonnée de la matrice en raison de la disposition des zéros dont on part. Cette étape requiert en général beaucoup moins de calculs que l'élimination de la première partie de l'algorithme, car les pivots valent 1 *et il y a peu de termes à modifier.*

Fin de l'algorithme de Gauss

Passons maintenant à la partie unicité du théorème. Elle repose sur l'observation simple suivante. Comme les opérations élémentaires sur les lignes ne mélangent pas les colonnes, si l'on a deux matrices équivalentes, et si l'on supprime dans ces deux matrices la même colonne, alors les matrices obtenues sont encore équivalentes.

Nous allons raisonner par l'absurde. Supposons donc qu'il existe une matrice A *et deux matrices échelonnées réduites* U *et* U' *telles que* $A \sim U$ *et* $A \sim U'$ *avec* $U \neq U'$. *Par transitivité et symétrie, on en déduit que* $U \sim U'$. *Comme* $U \neq U'$, *il existe une première colonne en comptant à partir de la gauche qui diffère d'au moins un coefficient entre les deux matrices. On supprime toutes les colonnes se trouvant à droite de cette colonne, ainsi que celles qui ne sont pas des colonnes de pivot à gauche dans l'une ou l'autre matrice (ce sont les mêmes puisque ces colonnes des deux matrices sont égales par définition). Il en résulte deux matrices équivalentes* $\tilde{U}$ *et* $\tilde{U}'$, *d'après la remarque ci-dessus. Par construction, ces matrices ne diffèrent que par leur dernière colonne. De plus, elles ont les formes suivantes.*

Cas 1 : la première colonne différente n'est pas une colonne de pivot, on a alors

$$\tilde{U} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & u_1 \\ 0 & 1 & \dots & 0 & 0 & u_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & u_{k-1} \\ 0 & 0 & \dots & 0 & 1 & u_k \\ 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix},$$

pour un certain $k \leq m$.

Cas 2 : la première colonne différente est une colonne de pivot, on a alors

$$\tilde{U} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

De même pour $\tilde{U}'$, on obtient

Cas 1 : la première colonne différente n'est pas une colonne de pivot, on a alors

$$\tilde{U}' = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & u'_1 \\ 0 & 1 & \dots & 0 & 0 & u'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & u'_{k-1} \\ 0 & 0 & \dots & 0 & 1 & u'_k \\ 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix},$$

pour un certain $k \leq m$.

Cas 2 : la première colonne différente est une colonne de pivot, on a alors

$$\tilde{U}' = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

On voit que $\tilde{U}$ et $\tilde{U}'$ sont encore échelonnées réduites. Interprétons maintenant $\tilde{U}$ et $\tilde{U}'$ comme étant les matrices augmentées de deux systèmes linéaires. Comme ces matrices sont équivalentes, les systèmes linéaires en question ont le même ensemble de solutions. Si le premier système est incompatible, ce qui est le cas 2 pour $\tilde{U}$, alors le deuxième système est aussi incompatible, ce qui est aussi le cas 2 pour $\tilde{U}'$. Mais alors, $\tilde{U} = \tilde{U}'$ ce qui est impossible car on est parti de l'hypothèse $\tilde{U} \neq \tilde{U}'$. Les deux *systèmes sont donc compatibles. Or ce sont des systèmes sans variable libre et on a donc* $S_{\tilde{U}} = (u_1, u_2, ..., u_k) = S_{\tilde{U}'} = (u'_1, u'_2, ..., u'_k)$*. On en déduit dans ce cas aussi que* $\tilde{U} = \tilde{U}'$*, ce qui est également impossible. Contradiction. Ceci montre que* U *et* U' *ne peuvent être différentes, d'où l'unicité.*

Remarque 2.3.21. *Si une matrice donnée n'est équivalente qu'à une seule matrice échelonnée réduite, elle est par contre équivalente à une infinité de matrices échelonnées.*

Exemples 2.3.22. *On considère la matrice*

$$A = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 2 & 6 & 1 & 2 \\ 3 & 3 & 2 & 10 \end{pmatrix}, \tag{2.3}$$

Passage à une forme échelonnée.

Première itération de la boucle, étape 1. Le choix du pivot est tout fait, on garde $a_{1,1} = 1$.

Première itération de la boucle, étape 2. On remplace la ligne 2 par elle-même moins $2\times$ *la ligne 1 et la ligne 3 par elle-même moins* $3\times$ *la ligne 1. On obtient*

$$A \sim \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & -6 & 2 & 10 \end{pmatrix}, \tag{2.4}$$

On échange les lignes 3 et 2, on obtient

$$A \sim \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & -6 & 2 & 10 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \tag{2.5}$$

Cette matrice est échelonnée ($m - 1 = 3 - 1 = 2$ *itérations maximum).*

Passage à une forme échelonnée réduite.

Étape 1, homothéties. On multiplie la ligne 1 par 1*, la ligne 2 par* $\frac{-1}{6}$ *et la ligne 3 par* 1 *et l'on obtient*

$$A \sim \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 1 & \frac{-1}{3} & \frac{-5}{3} \\ 0 & 0 & 1 & 2 \end{pmatrix}, \tag{2.6}$$

Étape 2, première itération. On ne touche plus à la ligne 3 et on remplace la ligne 2 par elle-même moins $\frac{-1}{3}\times$ *la ligne 3. On obtient*

$$A \sim \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \tag{2.7}$$

Étape 2, deuxième itération. On ne touche plus à la ligne 2 et on remplace la ligne 1 par elle-même moins $3\times$ la ligne 2. On obtient

$$A \sim \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \tag{2.8}$$

qui est bien échelonnée réduite.

Le théorème 2.3.19 permet d'étendre un certain nombre de définitions aux matrices quelconques.

Définition 2.3.23. *Soit A une matrice quelconque et U l'unique matrice échelonnée réduite qui lui est équivalente. Les positions, colonnes et lignes de pivot de A sont les positions, colonnes et lignes de pivot de U. Si A est la matrice augmentée d'un système linéaire, alors les inconnues correspondant à une colonne de pivot sont appelées inconnues ou variables essentielles. Les autres sont appelées inconnues ou variables libres.*

Il faut faire attention que les positions de pivot ne sont en général pas apparentes sur la matrice A. Il faut effectivement calculer la matrice U, ou au moins une matrice échelonnée équivalente à A pour les déterminer. Ainsi, dans l'exemple 2.3.22, on voit trois positions de pivot : $(1,1)$, $(2,2)$ et $(3,3)$ sur la matrice échelonnée réduite que l'on ne pouvait pas deviner sur la matrice A elle-même.

En regroupant tous les résultats précédents, on obtient la discussion générale de la résolution des systèmes linéaires.

Théorème 2.3.24. *Un système linéaire est compatible si et seulement si la matrice échelonnée réduite équivalente à sa matrice augmentée ne contient aucune ligne de la forme*

$$\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix} \quad \textit{avec} \quad b \neq 0.$$

Dans ce cas, on obtient une description paramétrique de l'ensemble des solutions en exprimant les variables essentielles en fonction du second membre et des variables libres.

De même,

Corollaire 2.3.25. *Soit un système linéaire $m \times n$ quelconque, A sa matrice augmentée et U l'unique matrice échelonnée réduite équivalente à A. On a*

a) Soit il n'y a aucune solution si U contient une ligne de la forme

$$\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix} \quad \textit{avec} \quad b \neq 0.$$

b) Soit il y a une solution unique si U ne contient aucune telle ligne et qu'il n'y a pas de variables libres.

c) Soit il y a une infinité de solutions si U ne contient aucune telle ligne mais qu'il existe des variables libres.

Remarque 2.3.26. *On n'a décrit qu'un seul algorithme de résolution, l'algorithme de Gauss. Or cet algorithme est bien insuffisant pour résoudre numériquement, c'est-à-dire sur ordinateur, les énormes systèmes linéaires rencontrés dans la pratique. L'analyse numérique matricielle est l'étude d'algorithmes qui généralisent celui de Gauss, ou qui sont de nature totalement différente, dans le but de résoudre effectivement et efficacement de tels systèmes. C'est un vaste champ de recherche toujours très actif de nos jours.*

Chapitre 3

Espaces vectoriels

L'ensemble des vecteurs du plan ou de l'espace est muni d'une loi de composition interne (à savoir la somme de deux vecteurs) et d'une loi de composition externe (à savoir la multiplication d'un vecteur par un scalaire). De plus, ces deux lois satisfont un certains nombres de propriétés. Les espaces vectoriels généralisent cette situation. Dans tout ce chapitre, on considérera un espace défini sur $\mathbb{R}$ ou $\mathbb{C}$ que l'on notera K.

3.1 Définitions

Définition 3.1.1. *$(E, +, \cdot)$ est un espace vectoriel sur K si et seulement si*

1) $(E, +)$ est un groupe commutatif c'est-à-dire qu'il vérifie les quatre propriétés suivantes :

a) Associativité , c'est-à-dire que pour tout élément u, v et w de E

$$(u + v) + w = u + (v + w)$$

b) Il existe un élément neutre , c'est-à-dire qu'il existe un élément de E, noté 0_E, vérifiant :

$$\forall v \in E, v + 0_E = 0_E + v = v$$

Cet élément neutre est unique.

c) Tout élément de E admet un symétrique, c'est-à-dire qu'il existe un élément v' de E tel que

$$v + v' = v' + v = 0_E$$

Cet élément v' de E est noté $-v$.

d) Commutativité, c'est-à-dire que pour tout élément u et v de E,

$$u + v = v + u.$$

2) On a pour tous x,y dans E, et tous λ, μ dans K,

$$(\lambda + \mu)x = \lambda x + \mu x \tag{3.1}$$
$$\lambda(x + y) = \lambda x + \lambda y \tag{3.2}$$
$$\lambda(\mu x) = (\lambda\mu)x \tag{3.3}$$
$$1 \cdot x = x \tag{3.4}$$

Dans la suite le signe $\cdot$ sera omis.

Remarque 3.1.2. *1. Au lieu de K-espace vectoriel, on dit aussi espace vectoriel sur K.*

2. *Les élements de K sont appelés des scalaires et les éléments de E des vecteurs.*
3. *La loi de composition interne sur E (notée usuellement +) est appelée couramment l'addition et $v + v'$ est appelée somme des vecteurs v et v'. La loi de composition externe sur E est appelée couramment multiplication par un scalaire. La multiplication du vecteur v par le scalaire λ sera notée λv.*
4. *0_E est l'élément neutre de la loi interne de E et est appelé vecteur nul. Il ne doit pas être confondu avec l'élément 0 de K. Lorsqu'il n'y aura pas de risque de confusion, 0_E sera aussi noté 0.*

Exemples 3.1.3 (Espaces vectoriels).
1) $\mathcal{F}([a,b],\mathbb{R})$ l'ensemble des fonctions définies de $[a,b]$ dans $\mathbb{R}$.
2) $\mathbb{R}$, $\mathbb{C}$.
3) $M_{p,n}$ est un espace vectoriel.
4) $K[X] = \left\{ P(X) = \sum_{k=0}^{n} a_k X^k \mid n \in \mathbb{N}, \forall k,\ a_k \in K \right\}$ les polynômes à coefficients dans K

Définition 3.1.4. Espace vectoriel produit :
Soient E et F deux espaces vectoriels. On définit l'addition et la multiplication pour tous x, x' dans E, tous y, y' dans F et tout λ dans K,

$$\lambda(x, y) = (\lambda x, \lambda y) \tag{3.5}$$
$$(x, y) + (x', y') = (x + x', y + y') \tag{3.6}$$

Alors $E \times F$ muni de ces deux opérations est un espace vectoriel.

Exemples 3.1.5.
$\mathbb{R}^n$, $\mathbb{C}^n$.

3.2 Sous-espaces vectoriels

Dans la pratique, afin d'éviter d'avoir huit propriétés à vérifier, on cherchera un espace vectoriel qui contient l'ensemble étudié et on introduira la notion de sous-espace vectoriel.

Définition 3.2.1. $(F, +, \cdot)$ *est un sous-espace vectoriel de* $(E, +, \cdot)$ *si et seulement si*

$$\begin{array}{rc} i) & F \quad \text{est stable par addition et multiplication} \\ ii) & F \quad \text{est non vide} \end{array}$$

On utilisera l'abréviation sev pour sous-espace vectoriel. Notons qu'un sev n'est jamais vide, puisqu'il contient toujours au moins le vecteur nul.

Remarque 3.2.2. *L'addition et la multiplication par un scalaire définissent donc des opérations sur tout sev* F*. De plus, si* $u \in F$*, alors* $-u = (-1)u \in F$ *par i). Un sev* F *muni des deux opérations satisfait donc les huit propriétés algébriques de base. C'est donc également un espace vectoriel sur* K*.*

Critère :

$$F \text{ sev de } E \iff F \neq \emptyset \text{ et } \forall x, y \in F, \quad \forall \lambda, \mu \in K, \quad \lambda x + \mu y \in F$$

Définition 3.2.3. Combinaison linéaire *Soient* E *un* K*-espace vectoriel, et soient* $u_1, \ldots, u_n$ *des éléments de* E*. On dit qu'un élément* u *est* combinaison linéaire *des* $(u_k)_{k=1,\ldots,n}$*, s'il existe des scalaires* $\lambda_1, \ldots, \lambda_n$ *dans* K *tels que*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n = \sum_{k=1}^{n} \lambda_k u_k.$$

Remarque 3.2.4. *Un sev est donc non vide et stable par combinaison linéaire.*

Exemples 3.2.5.

1. E *et* $\{0\}$ *sont des sev de* E*.*
2. $C^0([a, b], \mathbb{R}) \subset \mathcal{F}([a, b], \mathbb{R})$
3. *L'ensemble* $F = \{(x_1, x_2) \in \mathbb{R}^2; x_2 = 0\}$ *est un sev de* $\mathbb{R}^2$*.*
4. *L'ensemble* $F = \{(x_1, x_2) \in \mathbb{R}^2; x_1 = 1\}$ *n'est pas un sev de* $\mathbb{R}^2$*.*

Proposition 3.2.6. *Si* U *et* V *sont deux sev alors* $U \cap V$ *aussi. De même une intersection finie de sev est un sev.*

démonstration 3.2.7.

i) $0 \in U$ *et* $0 \in V$ *donc* $0 \in U \cap V$*.*

ii) Soient $x, y \in U \cap V$, alors $\forall \lambda, \mu \in K$, on a $\lambda x + \mu y \in U$ car $x, y \in U$ et U est un sev. De même comme $x, y \in V$, $\lambda x + \mu y \in V$ car V est un sev. On a donc $\lambda x + \mu y \in U \cap V$.

Remarque 3.2.8. *Généralement la réunion de deux sev n'est pas un sev, sauf si l'un est inclus dans l'autre.*

Définition 3.2.9. *Soient U et V sont deux sev de E. On appelle somme de U et V, l'ensemble*

$$U + V = \{u + v; u \in U \text{ et } v \in V\}.$$

C'est un sev de E.

De la même façon, si $U_1, ..., U_n$ sont des sev de E,

$$U_1 + ... + U_n = \{u_1 + ...u_n; u_1 \in U_1, ..., u_n \in U_n\}$$

est un sev de E.

démonstration 3.2.10.

i) $0 \in U_i$ pour tout $i = 1, \ldots, n$. Et $0 + 0 + \ldots + 0 = 0 \in U_1 + ... + U_n$.

ii) Soient x et y deux éléments de $U_1, ..., U_n$. Il existe donc x_i, y_i dans U_i pour tout i tels que

$$x = x_1 + x_2 + \ldots + x_n,$$
$$y = y_1 + y_2 + \ldots + y_n.$$

On multiplie la première équation par λ, la seconde par μ et on additionne membre à membre

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1) + (\lambda x_2 + \mu y_2) + \ + (\lambda x_k + \mu y_k).$$

Comme U_i est un sev, on a $\lambda x_i + \mu y_i \in U_i$ pour tout i, d'où

$$\lambda x + \mu y \in U_1 + ... + U_n.$$

Définition 3.2.11. *Soient $U_1, ..., U_n$ des sev de E, on dit que E est somme directe si $\forall x \in E \quad \exists!(x_1, ..., x_n) \in U_1 \times ... \times U_n; x = x_1 + ... + x_n$. On note $E = U_1 \bigoplus ... \bigoplus U_n$*

Proposition 3.2.12.

$$E = U_1 \bigoplus ... \bigoplus U_n \Longleftrightarrow \begin{cases} E = U_1 + ... + U_n \\ u_1 + ... + u_n = 0 \Longleftrightarrow u_i = 0 \quad \forall i \in \{1, ..., n\} \end{cases}$$

démonstration 3.2.13. *La condition nécessaire est évidente, puisque c'est l'unicité de la décomposition du vecteur nul. Pour la condition suffisante, on suppose que*

$$u_1 + ... + u_n = 0 \Longleftrightarrow u_i = 0 \quad \forall i \in \{1, ..., n\}$$

Soit $y \in E$ et considérons deux décompositions de y,

$$y = y_1 + y_2 + \ + y_n,$$
$$y = y'_1 + y'_2 + \ + y'_n$$

Soustrayant membre à membre, il vient

$$0 = (y_1 - y'_1) + (y_2 - y'_2) + \ \ldots + (y_n - y'_n),$$

avec $y_i - y'_i \in U_i$ puisque U_i est un sev. Par notre hypothèse, on en déduit que $y_i - y'_i = 0$ pour tout i, donc la somme est bien directe.

On a un cas particulier important, celui de deux sev en somme directe.

Proposition 3.2.14.

$$E = U_1 \bigoplus U_2 \iff \begin{cases} E = U_1 + U_2 \\ U_1 \cap U_2 = \{0\} \end{cases}$$

On dit que U_1 et U_2 sont supplémentaires.

démonstration 3.2.15. *Condition nécessaire : si $E = U_1 \bigoplus U_2$ alors en particulier $E = U_1 + U_2$. De plus, prenant $x \in U_1 \cap U_2$, on voit que $0 = x + (-x)$ avec $x \in U_1$ et $-x \in U_2$. Par conséquent, $x = 0$.*

Condition suffisante : si $E = U_1 + U_2$ et $U_1 \cap U_2 = \{0\}$, prenons $x_1 \in U_1$ et $x_2 \in U_2$ tels que $x_1 + x_2 = 0$. On en déduit que $x_2 = -x_1 \in U_1$, d'où $x_2 \in U_1 \cap U_2$, d'où $x_2 = x_1 = 0$.

Remarque 3.2.16. *Attention, cette proposition est fausse à partir de trois sev. On peut parfaitement avoir $U_1 \cap U_2 \cap U_3 = \{0\}$ sans que la somme $U_1 + U_2 + U_3$ ne soit directe.*

Exemples 3.2.17.

1) Dans $\mathbb{R}^2$: $\mathbb{R}^2 = \mathbb{R}e_1 \bigoplus \mathbb{R}e_2$, où $e_1 = (1, 0)$ et $e_2 = (0, 1)$.

2) $\mathcal{F}([-a, a], \mathbb{R}) = \mathcal{P} \bigoplus \mathcal{I}$ où $\mathcal{P}$ est l'ensemble des fonctions réelles paires définies sur $[-a, a]$, et $\mathcal{I}$ est l'ensemble des fonctions réelles impaires définies sur $[-a, a]$.

Définition 3.2.18 (Sous-espace engendré)**.** *Soit $A \subset E$. Le sev A' engendré par A est l'intersection des sev contenant A.*

Exemples 3.2.19.

1) $A = \{u\}$, alors $A' = \mathbb{R}u$ ou $A' = \mathbb{C}u$.

2) $A = \{u_1, ..., u_n\}$, alors $A' = Ku_1 + ... + Ku_n = \{\lambda_1 u_1 + ... + \lambda_n u_n; (\lambda_1, ..., \lambda_n) \in K^n\}$

3) $A = \{u_i; i \in I\}$, alors $A' = \{\sum_{j \in J} \lambda_j u_j;$ où $J \subset I$, J finie , $\lambda_j \in K\}$

démonstration 3.2.20. *On appelle F l'ensemble des combinaisons linéaires des vecteurs $(u_1, \ldots, u_n)$. Cet ensemble est non vide, car il contient la combinaison linéaire particulière $0 \times u_1 + \ldots + 0 \times u_n$ qui vaut 0_E. On peut également vérifier que $u_1, \ldots, u_n$ appartiennent à F, en effet pour tout k compris entre 1 et n, u_k est combinaison linéaire de $u_1, \ldots, u_n$ (il suffit de considérer la combinaison linéaire où tous les coefficients sont nuls sauf le kième qui vaut 1). Il s'agit maintenant de prouver que F est stable par combinaison linéaire de deux vecteurs. Soient u et v deux vecteurs de F et deux scalaires α et β . Comme u est élément de F, il existe des scalaires $\lambda_1, \ldots, \lambda_n$ tels que*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n.$$

De même, v étant élément de F , il existe des scalaires $\mu_1, \ldots, \mu_n$ tels que

$$v = \mu_1 u_1 + \cdots + \mu_n u_n$$

En utilisant les règles de calcul dans un espace vectoriel, on obtient :

$$\alpha u + \beta v = (\alpha\lambda_1 + \beta\mu_1)u_1 + \ldots + (\alpha\lambda_n + \beta\mu_n)u_n$$

C'est une combinaison linéaire des vecteurs $u_1, \ldots, u_n$ donc un élément de F. Si G est un sous-espace vectoriel contenant $u_1, \ldots, u_n$ alors il est stable par combinaison linéaire ; il contient donc toute combinaison linéaire des vecteurs $u_1, \ldots, u_n$. Par conséquent F est inclus dans G : F est le plus petit sous-espace (au sens de l'inclusion) contenant $u_1, \ldots, u_n$.

Notation *Le sous-espace engendré par $u_1, \ldots, u_n$ et est noté $vect(u_1, \ldots, u_n)$. On a donc*

$$u \in vect(u_1, \ldots, u_n) \Leftrightarrow \exists(\lambda_1, \ldots, \lambda_n) \in K^n | u = \lambda_1 u_1 + \cdots + \lambda_n u_n. \tag{3.7}$$

Définition 3.2.21. Système générateur

$\{u_i; i \in I\}$ est un système générateur de E si et seulement si E est le sous-espace engendré par l'ensemble $\{u_i; i \in I\}$.

Définition 3.2.22. Famille libre

$\{u_i; i \in I\}$ est une famille libre de E si et seulement si

$$\forall J \subset I, J \text{ finie }, \sum_{j \in J} \lambda_j u_j = 0 \Longrightarrow \lambda_j = 0 \quad \forall j \in J.$$

Si une famille n'est pas libre, elle est dite liée.

On emploiera indifféremment le terme de famille ou de système.

Proposition 3.2.23.

(u, v) liée $\iff$

$$\begin{cases} u = 0 \\ \text{ou} \\ v \text{ est colinéaire à } u \text{ i. e.} \exists \lambda \in K \text{ tel que } v = \lambda u \end{cases}$$

démonstration 3.2.24. *Supposons d'abord la famille liée. Il existe donc α, β non tous les deux nuls tels que*

$$\alpha u + \beta v = 0.$$

Si $u = 0$, c'est terminé. Si $u \neq 0$, alors $\beta \neq 0$ car sinon on aurait aussi $\alpha = 0$, ce qui est contraire à l'hypoythèse. Comme $\beta \neq 0$,on peut diviser par β, ce qui donne

$$v = \frac{-\alpha}{\beta} u.$$

et v est un multiple de u. Réciproquement, s'il existe un scalaire λ tel que $v = \lambda u$, alors

$$\lambda u + v = 0,$$

et donc (u, v) est liée.

Remarque 3.2.25. *Toute famille extraite d'une famille libre est libre. Toute famille qui contient une famille liée est liée.*

Proposition 3.2.26. Ajout d'un vecteur à une partie libre

Soient E un espace vectoriel et $(u_1, \ldots, u_n)$ une partie libre de E. Si v est un vecteur de E tel que $(u_1, \ldots, u_n, v)$ soit une partie liée de E, alors le vecteur v est combinaison linéaire des vecteurs $u_1, \ldots, u_n$.

démonstration 3.2.27. *Les vecteurs $u_1, \ldots, u_n, v$ sont lin'eairement dépendants. Il existe donc des scalaires $\lambda_1, \ldots, \lambda_n, \beta$ non tous nuls tels que*

$$\lambda_1 u_1 + \cdots + \lambda_n u_n + \beta v = 0.$$

Si $\beta = 0$, l'égalité devient

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0.$$

avec $\lambda_1, \ldots, \lambda_n$ non tous nuls. Ce qui contredirait le fait que la famille $(u_1, \ldots, u_n)$ est libre.

Donc β est non nul, on peut donc diviser par β et on obtient

$$v = -\frac{\lambda_1}{\beta} u_1 - \ldots - \frac{\lambda_n}{\beta} u_n,$$

Ce qui signifie que v est combinaison linéaire des vecteurs $u_1, \ldots, u_n$.

Remarque 3.2.28. *Le même raisonnement montre que si une partie $(u_1, \ldots, u_n)$ est liée alors l'un des vecteurs de cette famille est combinaison linéaire des autres.*

Définition 3.2.29. Base

Une base est une famille libre et génératrice.

Proposition 3.2.30. $\{u_i; i \in I\}$ *est une base de* E *si et seulement si*

$$\forall u \in E \quad \exists! J \subset I, J \text{ finie et } \exists!(\lambda_j)_{j\in J}, \text{tels que } u = \sum_{j\in J} \lambda_j u_j.$$

Exemples 3.2.31.

1) Dans $\mathbb{R}^3$ *ou* $\mathbb{C}^3$*, considéré comme un espace vectoriel sur* $\mathbb{R}$*,* (e_1, e_2, e_3) *où* $e_1 = (1, 0, 0)$*,* $e_2 = (0, 1, 0)$ *et* $e_3 = (0, 0, 1)$ *est une base appelée base canonique. Plus généralement dans* $\mathbb{R}^n$ *on note* $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ *le vecteur dont toutes les coordonnées sont nulles sauf la ième. Alors* $(e_i)_{i=1,\ldots,n}$ *est la base canonique de* $\mathbb{R}^n$*.*

2) Dans $\mathbb{R}^3$ *ou* $\mathbb{C}^3$*,* (e_1, e_2) *n'est pas une base.*

3) Dans $\mathbb{R}^3$*,* (e_1, e_2, e_3, u) *où* $u = (1, 1, 1)$ *est génératrice et liée.*

4) Dans $K[X]$ *la famille* $(X^n)_{n\in\mathbb{N}}$ *est une base.*

On remarque que dans le dernier exemple, la base a un nombre infini d'éléments, ce qui n'est pas le cas des exemples précédents. Dans tous les cas, on a

Théorème 3.2.32. *Tout espace vectoriel admet une base.*
Tout système libre peut être complété en une base.
De tout système générateur, on peut extraire une base.

3.3 Espaces vectoriels de dimension finie

On suppose que E admet une famille finie de générateurs.

Théorème 3.3.1 (de la base extraite)**.** *Soit* F *un sev de* E *engendré par une famille finie de vecteurs* S*. Alors il existe un sous-ensemble de* S *qui est une base de* F*.*

démonstration 3.3.2. *Soit* $S = \{v_1, v_2, ..., v_k\}$ *la famille génératrice en question. Soit elle est libre et c'est la base cherchée, soit elle est liée. Si elle est liée et si* $Card(S) = 1$*, i.e. le nombre d'éléments de* S*, alors c'est que* $v_1 = 0$ *et donc* $F = \{0\}$*. Dans ce cas,* $S' = \emptyset$ *est la base cherchée. Si* $Card(S) \geq 2$ *alors il existe un indice* p *tel que* v_p *soit combinaison linéaire des autres vecteurs de* S*, d'après la remarque 3.2.28. On pose alors* $S' = S \setminus \{v_p\}$ *(on enlève* v_p *à* S*). Montrons que* S' *est encore une famille génératrice. Par le choix de l'indice* p*, on peut écrire*

$$v_p = \mu_1 v_1 + \ldots + \mu_{p-1} v_{p-1} + \mu_{p+1} v_{p+1} + \ldots + \mu_k v_k$$

(formule à modifier de façon évidente si $p = 1$ *ou* $p = k$*). Comme* S *est génératrice, pour tout* $x \in F$*, il existe des scalaires* λ_i *tels que*

$$x = \lambda_1 v_1 + \ldots + \lambda_{p-1} v_{p-1} + \lambda_{p+1} v_{p+1} + \ldots + \lambda_k v_k$$

Remplaçant v_p *par son expression en fonction des vecteurs de* S'*, on obtient*

$$x = (\lambda_1 + \lambda_p \mu_1) v_1 + \ldots + (\lambda_{p-1} + \lambda_p \mu_{p-1}) v_{p-1} + (\lambda_{p+1} + \lambda_p \mu_{p+1}) v_{p+1} + \ldots + (\lambda_k + \lambda_p \mu_k) v_k$$

d'où S' génératrice. De plus, $Card(S') = Card(S) - 1$.

On recommence alors le raisonnement : soit S' est libre et on a terminé, soit elle est liée et on peut lui enlever un vecteur en conservant le caractère générateur. Le processus s'arrête en au plus k itérations, car il ne reste alors plus aucun vecteur à enlever, et il ne peut s'arrêter que sur une famille libre. Celle-ci est donc la base extraite de S cherchée.

Théorème 3.3.3. *(de la base incomplète) Soit F un sev de E et S une famille libre de vecteurs de F. Alors il existe une base B de F qui contient S.*

démonstration 3.3.4. *Soit S la famille libre en question. Soit cette famille engendre F et on a terminé, soit elle n'engendre pas F. On va supposer que E est engendré par $(u_1, \ldots, u_d)$ et que $S = \{v_1, \ldots, v_r\}$ On va substituer les vecteurs v_i aux vecteurs u_i. On peut écrire*

$$v_1 = \mu_{1,1}u_1 + \ldots + \mu_{1,d}u_d,$$

avec l'un des coefficients non nul. Quitte à renuméroter les vecteurs, on peut supposer que $\mu_1 \neq 0$. On substitue alors u_1 à v_1. L'ensemble $\{v_1, u_2, \ldots, u_d\}$ est-il encore générateur ? Oui car

$$u_1 = \frac{1}{\mu_{1,1}}v_1 - \ldots - \frac{\mu_{1,d}}{\mu_{1,1}}u_d$$

c'est donc une combinaison linéaire de $v_1, u_2, \ldots, u_d$. Supposons que l'on a substitué $u_1, \ldots, u_k$ à $v_1, \ldots, v_k$ pour $k \leq \min\{r-1, d-1\}$. On a

$$v_{k+1} = \lambda_{k+1,1}v_1 + \ldots + \lambda_{k+1,k}v_k + \mu_{k+1,k+1}u_{k+1} + \ldots + \mu_{k+1,d}u_d,$$

L'un des coefficients $\mu_{k+1,l}$ n'est pas nul pour un un $l \in \{k+1, \ldots, d\}$ car sinon la partie $(v_1, \ldots, v_{k+1})$ serait liée. On peut alors substituer v_{k+1} à u_l comme précédemment et la famille ainsi obtenue est toujours génératrice. on a de même obtenu que $r \leq d$ car sinon l'un des u_i serait combinaison linéaire des autres éléments de S, ce qui contredirait le fait que S est libre. Pour conclure, on applique la proposition 3.2.26 à S.

On a ainsi aussi démontré

Théorème 3.3.5. *Supposons que E admette une base de cardinal égal à n.*

Si L est libre alors card $L \leq n$. S'il y a égalité alors L est une base.

Si G est générateur alors card $G \geq n$. S'il y a égalité alors G est une base.

On en déduit un théorème fondamental.

Théorème 3.3.6. *(de la dimension) Supposons que E admette une famille finie de générateurs. Toutes les bases d'un sev F de E ont le même nombre d'éléments.*

démonstration 3.3.7. *Soit B une base de F avec $CardB = k$ et B' une autre base de F avec $CardB' = k'$. Comme la deuxième base est une famille libre, on déduit du Théorème 3.3.5 que $k' \leq k$. Inversant les rôles de B et B', il vient aussi $k \leq k'$.*

On peut donc maintenant poser raisonnablement la définition suivante.

Définition 3.3.8. *La dimension d'un sous-espace vectoriel F de E est le nombre d'éléments de chacune de ses bases,*

$$dimF = CardB, \quad \textit{pour toute base} \quad B \quad \textit{de} \quad F.$$

Remarque 3.3.9.

la dimension de l'ensemble $\{0\}$ est égal au cardinal de $\emptyset$ soit 0.

Les sous-espaces de dimension 1 sont appelés droites vectorielles ou droites.

Les sous-espaces de dimension $n-1$ dans un espace de dimension n sont appelés hyperplans. Lorsque $n=3$, on parle plutôt de plans. Si $n=2$, les hyperplans sont les droites vectorielles !

Théorème 3.3.10. Théorème de la base incomplète *Soit E un espace vectoriel de dimension n. Soit L un système libre de r vecteurs. Il existe des vecteurs $\{u_1, ..., u_{n-r}\}$ tels que $L \cup \{u_1, ..., u_{n-r}\}$ est une base de E.*

La démonstration repose sur les théorèmes 3.2.26 et 3.3.5.

Corollaire 3.3.11.

$\forall x \neq 0$, il existe une base dont le premier vecteur est x.

Tout sous-espace admet un supplémentaire.

Remarque 3.3.12. *Comme $\emptyset$ est une famille libre de F pour tout sev F, le théorème de la base incomplète permet de conclure que tout sev F admet une base B et que $CardB \leq dimE$.*

Théorème 3.3.13. *Soient E un espace vectoriel de dimension finie, U_1 et U_2 deux sev de E tels que $U_1 \subset U_2$. Alors*

$$dimU_1 \leq dimU_2 \leq dimE.$$

démonstration 3.3.14. *. D'après la remarque 3.3.12, U_1 admet une base, qui est donc aussi une famille libre de U_2. Donc son cardinal est inférieur à $dimU_2$. Comme $U_2 \subset E$, la même remarque implique que $dimU_2 \leq dimE$.*

Corollaire 3.3.15. *Si de plus $dimU_1 = dimU_2$, alors $U_1 = U_2$.*

démonstration 3.3.16. *Soit B_1 une base de U_1. On a donc $U_1 = vectB_1$. Mais c'est aussi une famille libre de U_2 de $dimU_2$ éléments. Donc, c'est aussi une base de U_2, ce qui implique que $vectB_1 = U_2$. Ceci s'applique en particulier au cas $U_2 = E$.*

Exemples 3.3.17.

1) Dans $\mathbb{R}^2$, la droite engendrée par $(\alpha,\beta) \neq (0,0)$ est l'ensemble des vecteurs (x,y) tels que $\alpha y - \beta x = 0$. Réciproquement , si $(a,b) \neq (0,0)$, l'ensemble des vecteurs (x,y) de $\mathbb{R}^2$ tels que $ax + by = 0$ est la droite vectorielle engendrée par $(b,-a)$.

2) Dans $\mathbb{R}^3$, si $(a,b,c) \neq (0,0,0)$, l'ensemble des vecteurs (x,y,z) tels que $ax + by + cz = 0$ est un plan vectoriel de $\mathbb{R}^3$.

3) Dans $\mathbb{R}^n$, si $(a_1,...,a_n) \neq (0,...,0)$ alors $a_1x_1 + ... + a_nx_n = 0$ est l'équation d'un hyperplan.

4) Dans $\mathbb{R}^3$, la droite engendrée par $(\alpha,\beta,\gamma) \neq (0,0,0)$ a pour équation : $\left\{\begin{array}{l}\beta x - \alpha y = 0\\ \gamma y - \beta x = 0\end{array}\right.$ *intersection de deux plans.*

Voyons maintenant les liens existant entre bases et sommes directes de sous espaces vectoriels.

Théorème 3.3.18. *Soit E un espace vectoriel de dimension finie et B une base de E. On se donne une partition $B_1, B_2, \ldots, B_k$ de B, c'est-à-dire un découpage de B en parties disjointes*

$$B_i \cap B_j = \emptyset \quad pour \quad i \neq j \quad et \quad B = \bigcup_{i=1}^{k} B_i.$$

Alors on a

$$E = vectB_1 \bigoplus \ldots \bigoplus vectB_k$$

Réciproquement, si

$$E = U_1 \bigoplus \ldots \bigoplus U_k$$

et B_i est une base de U_i, alors $B = \cup_{i=1}^{k} B_i$ est une base de E.

démonstration 3.3.19. *On commence par la première partie du théorème. Comme tout vecteur x de E est combinaison linéaire des vecteurs de B, famille génératrice, en regroupant ensemble les termes correspondant à chaque B_i, on écrit ainsi x comme une somme de vecteurs appartenant à $vectB_i$, pour $i = 1,...,k$, soit*

$$E = vectB_1 + ... + vectB_k$$

Il reste à montrer que la somme est directe. Pour cela on décompose le vecteur nul

$$0 = x_1 + x_2 + \quad + x_k \quad avec \quad x_i \in vectB_i.$$

Chaque x_i est combinaison linéaire des vecteurs de B_i avec certains coefficients, donc la somme ci-dessus est en fait une combinaison linéaire des vecteurs de B. Comme chaque *vecteur de B_i n'apparaît qu'à un seul endroit dans cette somme, puisque $B_i \cap B_j = \emptyset$ pour $i \neq j$, les coefficients de la combinaison linéaire globale sont simplement la réunion des coefficients des combinaisons linéaires pour chaque x_i. Or la famille B est libre, donc ces coefficients sont tous nuls. Par conséquent, $x_i = 0$ pour tout i et la somme est bien directe.*

Réciproquement, pour la deuxième partie du théorème, dire que

$$E = U_1 + \ldots + U_k$$

est clairement dire que B est une famille génératrice, et dire que la somme est directe, c'est-à-dire l'unicité de la décomposition du vecteur nul implique immédiatement que B est libre.

Corollaire 3.3.20.
$$E = U_1 \bigoplus U_2 \Longleftrightarrow \begin{cases} E = U_1 + U_2 \\ \dim E = \dim U_1 + \dim U_2 \end{cases}$$

démonstration 3.3.21. *On compte les vecteurs de la base grâce au théorème 3.3.18.*

On peut regrouper tous les résultats sur les supplémentaires dans le théorème suivant :

Théorème 3.3.22.
$$E = U_1 \bigoplus U_2 \Longleftrightarrow \begin{cases} E = U_1 + U_2 \\ U_1 \cap U_2 = \{0\} \end{cases} \Longleftrightarrow \begin{cases} E = U_1 + U_2 \\ \dim E = \dim U_1 + \dim U_2 \end{cases}$$

$$\Longleftrightarrow \begin{cases} U_1 \cap U_2 = \{0\} \\ \dim E = \dim U_1 + \dim U_2 \end{cases}$$

Corollaire 3.3.23. *Soit E un espace vectoriel de dimension finie. Si U_1 est un sous-espace vectoriel de E, alors U_1 admet au moins un supplémentaire.*

démonstration 3.3.24. *Soit B_1 une base de U_1. C'est une famille libre de E, donc d'après le théorème de la base incomplète, on peut lui ajouter un ensemble fini de vecteurs B_2 de telle sorte que $B_1 \cup B_2$ soit une base de E. Le supplémentaire cherché n'est autre que $U_2 = vect B_2$.*

Proposition 3.3.25. *On a la formule plus générale*

$$dim(U_1 + U_2) = dim U_1 + dim U_2 - dim(U_1 \cap U_2).$$

démonstration 3.3.26. *Nous admettrons ce résultat qui utilise le même type d'arguments que précédemment, (commencer avec une base de $U_1 \cap U_2$).*

Corollaire 3.3.27. *L'intersection de deux hyperplans de E est soit un hyperplan, soit un sev de dimension $dim E-2$. En particulier, l'intersection de deux droites de $\mathbb{R}^2$ est soit une droite, soit $\{0\}$, et l'intersection de deux plans de $\mathbb{R}^3$ est soit un plan, soit une droite.*

démonstration 3.3.28. *Soient U_1 et U_2 deux hyperplans de E, c'est-à-dire deux sev de dimension $n-1$ si $dim E = n$. De la formule précédente, on déduit que*

$$dim(U_1 \cap U_2) = dim U_1 + dim U_2 - dim(U_1 + U_2) = 2n - 2 - dim(U_1 + U_2).$$

Comme $U_1 \subset U_1 + U_2 \subset E$, on a

$$n - 1 = dim U_1 \leq dim(U_1 + U_2) \leq dim E = n.$$

On en déduit un encadrement de la dimension de $U_1 \cap U_2$:

$$n - 2 \leq dim(U_1 \cap U_2) \leq n - 1.$$

Il y a donc deux possibilités, soit $dim(U_1 \cap U_2) = n-1$, ce qui n'est possible que si $U_1 \cap U_2 = U_1 = U_2$ (les deux hyperplans sont confondus), soit $dim(U_1 \cap U_2) = n-2$. Pour $n = 2$, on a donc soit $dim(U_1 \cap U_2) = 1$ c'est une droite, soit $dim(U_1 \cap U_2) = 0$, les deux droites sont supplémentaires. Pour $n = 3$, on a donc soit $dim(U_1 \cap U_2) = 2$ c'est un plan, soit $dim(U_1 \cap U_2) = 1$ c'est une droite.

Théorème 3.3.29. $E = U_1 \bigoplus ... \bigoplus U_n \Longleftrightarrow \begin{cases} E = U_1 + ... + U_n \\ \dim E = \displaystyle\sum_{i=1}^{n} \dim U_i \end{cases}$

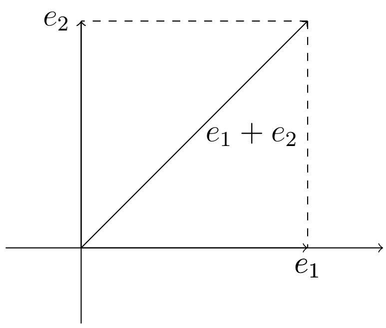
démonstration 3.3.30. *On compte les vecteurs de la base grâce au théorème 3.3.18.*

ATTENTION

$$\begin{cases} E = U_1 + ... + U_n \\ U_i \cap U_j = \{0\} \quad \text{pour} \quad i \neq i \end{cases} \not\Longrightarrow E = U_1 \bigoplus ... \bigoplus U_n$$

Contre exemple :

On se place dans $\mathbb{R}^2$ et on considère $U_1 = \mathbb{R}e_1$, $U_2 = \mathbb{R}e_2$ et $U_3 = \mathbb{R}(e_1+e_2)$ les droites engendrées respectivement par e_1, e_2 et e_1+e_2. On a bien $U_1 \cap U_2 = \{0\}$, $U_2 \cap U_3 = \{0\}$ et $U_1 \cap U_3 = \{0\}$. Cependant U_1, U_2 et U_3 ne sont pas en somme directe car $0 = e_1 - (e_1 + e_2) + e_2$.



Chapitre 4

Applications linéaires

Nous allons maintenant nous intéresser aux applications entre sev qui conservent l'addition et la multiplication par un scalaire.

4.1 Définitions

Définition 4.1.1. *Soient E et F deux espaces vectoriels et f une fonction définie sur E à valeurs dans F.*

$$f \text{ est linéaire } \iff \forall x, y \in E, \forall \lambda, \mu \in K, f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

Exemples 4.1.2.

1)

$$\begin{aligned} f : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto x + y \end{aligned}$$

2)

$$\begin{aligned} f : \quad \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ x &\mapsto (2x, -x) \end{aligned}$$

3)

$$\begin{aligned} f : \quad \mathbb{R}[X] &\longrightarrow \mathbb{R} \\ P &\mapsto P(0) \end{aligned}$$

4) L'application Id qui à chaque élément associe lui-même est linéaire.

Définition 4.1.3. *L'ensemble de toutes les applications linéaires de E dans F est noté $L(E, F)$. Si $E = F$, on le note plus simplement $L(E)$.*

Remarque 4.1.4. *On constate que $f(0) = 0$. En effet soit $\lambda \neq 1 \in K$. On a alors*

$$f(0) = f(\lambda \times 0) = \lambda f(0),$$

par linéarité. On en déduit donc que $f(0) = 0$ puisque $\lambda \neq 1$.

Introduisons maintenant quelques points de vocabulaire, ainsi que quelques propriétés associées.

On rappelle d'abord les notions de théorie des ensembles suivantes.

Une application f d'un ensemble X dans un ensemble Y est dite injective si deux éléments disjoints de X ont des images disjointes par f , c'est-à-dire

$$x \neq y \Rightarrow f(x) \neq f(y),$$

ou encore par contraposition,

$$f(x) = f(y) \Rightarrow x = y.$$

Elle est surjective si tout élément de Y est l'image d'au moins un élément de X par f , c'est-à-dire

$$\forall y \in Y \exists x \in X; y = f(x).$$

Elle est bijective si elle est simultanément injective et surjective. Dans ce cas, elle admet une application réciproque notée

$$f^{-1} : Y \to X$$

telle que

$$\forall x \in X, f^{-1}(f(x)) = x \quad \text{et} \quad \forall y \in Y, f(f^{-1}(y)) = y.$$

Si f est une application linéaire bijective de E dans F, on dit que c'est un isomorphisme entre E et F. Dans ce cas, l'application réciproque f^{-1} est également linéaire de F dans E, et c'est donc aussi un isomorphisme.

démonstration 4.1.5. *Soit $f : E \to F$ un isomorphisme. L'application réciproque est déterminée par*

$$f^{-1} : F \to E; f^{-1}(y) = x \Longleftrightarrow y = f(x).$$

Calculons $f^{-1}(\lambda y_1 + \mu y_2)$. Ce vecteur de E est déterminé par

$$f^{-1}(\lambda y_1 + \mu y_2) = x \Longleftrightarrow \lambda y_1 + \mu y_2 = f(x).$$

Par ailleurs, on a aussi

$$f^{-1}(y_1) = x_1 \Longleftrightarrow y_1 = f(x_1) \quad \textit{et} \quad f^{-1}(y_2) = x_2 \Longleftrightarrow y_2 = f(x_2).$$

En combinant ces deux dernires relations, il vient

$$\lambda y_1 + \mu y_2 = \lambda f(x_1) + \mu f(x_2) = f(\lambda x_1 + \mu x_2)$$

par linéarité de f *. En comparant avec la caractérisation ci-dessus, on obtient*

$$f^{-1}(\lambda y_1 + \mu y_2) = \lambda x_1 + \mu x_2 = \lambda f^{-1}(y_1) + \mu f^{-1}(y_2),$$

et f^{-1} *est bien linéaire.*

Proposition 4.1.6. *Si* $f : E \to F$ *est linéaire et* $g : F \to G$ *est linéaire, alors* $g \circ f$ *est linéaire.*

démonstration 4.1.7. *On a, pour tous* $x, y \in E$ *et tout* $\lambda, \mu \in K$*,*

$$\begin{aligned} g \circ f(\lambda x + \mu y) &= g(f(\lambda x + \mu y)) = g(\lambda f(x) + \mu f(y)) \\ &= \lambda g(f(x)) + \mu g(f(y)) = \lambda g \circ f(x) + \mu g \circ f(y), \end{aligned}$$

par linéarité de f *puis de* g*, donc* $g \circ f$ *est linéaire.*

4.2 Image, noyau, rang

On associe à chaque application linéaire deux sous-espaces vectoriels fondamentaux.

Définition 4.2.1. *Soit* $f \in L(E, F)$*. On appelle noyau de* f *l'ensemble*

$$Kerf = \{x \in E; f(x) = 0\} = f^{-1}(\{0\}). \tag{4.1}$$

On appelle image de f *l'ensemble*

$$Imf = \{y \in F; \exists x \in E, f(x) = y\} = f(E). \tag{4.2}$$

Les expressions de droite sont respectivement l'image réciproque et l'image directe d'un ensemble. L'image réciproque d'une partie A d'un ensemble Y par une application $f : X \to Y$ n'est autre que l'ensemble des éléments de X dont l'image par f appartient à A,

$$f^{-1}(A) = \{x \in X; f(x) \in A\}.$$

Il s'agit d'un concept purement ensembliste que l'on utilise ici dans le contexte plus spécifique de l'algèbre linéaire.

Proposition 4.2.2. *Le noyau de* f *est un sev de* E *et l'image de* f *est un sev de* F*.*

démonstration 4.2.3. *On va montrer que ces ensembles contiennent le vecteur nul de leur espace respectif et sont stables par combinaison linéaire. Comme* $f(0_E) = 0_F$ *(on distingue exceptionnellement les vecteurs nuls), on voit que* $0_E \in Kerf$ *et* $0_F \in Imf$. *Soient* $x_1, x_2 \in Kerf$ *et* $\lambda_1, \lambda_2 \in K$. *On a donc* $f(x_1) = f(x_2) = 0$. *Par conséquent,*

$$f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2) = 0,$$

d'où

$$\lambda_1 x_1 + \lambda_2 x_2 \in Kerf$$

Soient $y_1, y_2 \in Imf$ *et* $\lambda_1, \lambda_2 \in K$. *Il existe donc* $x_1, x_2 \in E$ *tels que* $y_1 = f(x_1)$ *et* $y_2 = f(x_2)$. *Par conséquent,*

$$\lambda_1 y_1 + \lambda_2 y_2 = \lambda_1 f(x_1) + \lambda_2 f(x_2) = f(\lambda_1 x_1 + \lambda_2 x_2)$$

d'où $\lambda_1 y_1 + \lambda_2 y_2 \in Imf$.

On a le théorème très important suivant

Théorème 4.2.4. *Soit* $f \in L(E, F)$. *Alors*

f *est linéaire et injective* $\Longleftrightarrow Kerf = \{0\}$.

f *est linéaire et surjective* $\Longleftrightarrow Imf = F$.

démonstration 4.2.5. *La deuxième propriété est une simple reformulation de la surjectivité.*

Soit f *une application linéaire injective. Comme* $f(0) = 0$, *c'est donc que* $0 \in E$ *est l'unique vecteur dont l'image est* $0 \in F$, *d'où* $Kerf = \{0\}$. *Réciproquement, supposons que* $Kerf = \{0\}$ *et donnons nous deux vecteurs* x_1 *et* x_2 *de* E *qui ont la même image par* f , $f(x_1) = f(x_2)$. *Passant le second membre au premier membre, il vient*

$$0 = f(x_1) - f(x_2) = f(x_1 - x_2)$$

par linéarité. On en déduit que

$$x_1 - x_2 \in Kerf,$$

donc $x_1 = x_2$ *et f est injective.*

Remarque 4.2.6. *Pour démontrer qu'une application linéaire est injective, on utilisera TOUJOURS la caractérisation par le noyau et non pas la définition générale de l'injectivité.*

On va maintenant étudier l'action des applications linéaires sur les bases. Tout d'abord une première définition :

Définition 4.2.7. *Soit* $f \in L(E, F)$. *On définit son rang par*

$$rgf = dimImf. \tag{4.3}$$

Proposition 4.2.8. *Si* $(v_1, v_2, \dots, v_m)$ *est une base de* E*, alors*

$$Imf = vect(f(v_1), f(v_2), ..., f(v_m)) \quad et \quad rgf \leq dimE = m. \tag{4.4}$$

démonstration 4.2.9. *Comme* $(v_1, v_2, \dots, v_m)$ *est une base de* E*, pour tout* $y \in Imf$ *, il existe* $x = \lambda_1 v_1 + \dots + \lambda_m v_m \in E$ *tel que* $y = f(x)$.

Par linéarité de f *, il vient*

$$y = f(\lambda_1 v_1 + \dots + \lambda_m v_m) = \lambda_1 f(v_1) + \dots + \lambda_m f(v_m) \in vect(f(v_1), f(v_2), ..., f(v_m)).$$

Donc

$$Imf \subset vect(f(v_1), f(v_2), ..., f(v_m)).$$

Réciproquement, si $y \in vect(f(v_1), f(v_2), ..., f(v_m))$*, alors il existe des scalaires* λ_j *tels que*

$$y = \lambda_1 f(v_1) + \dots + \lambda_m f(v_m).$$

Posant

$$x = \lambda_1 v_1 + \dots + \lambda_m v_m \in E,$$

on voit que $y = f(x)$*, donc* $y \in Imf$ *. Finalement,* $Imf = vect(f(v_1), f(v_2), ..., f(v_m))$. *La famille de* m *vecteurs* $(f(v_1), f(v_2), ..., f(v_m))$ *est une famille génératrice de* Imf. *Par conséquent* $dimImf \leq m = dimE$.

Proposition 4.2.10. *Soit* $f \in L(E, F)$*, alors*

i) $rgf \leq min(dimE, dimF)$.

ii) L'application f *est surjective si et seulement si* $rgf = dimF$.

ii) L'application f *est injective si et seulement si* $rgf = dimE$.

démonstration 4.2.11.

i) On a déjà vu que $rgf \leq dimE$. *Par ailleurs, comme* Imf *est un sev de* F*, on évidemment aussi* $rgf \leq dimF$.

ii) f *est surjective si et seulement si* $Imf = F$ *si et seulement si* $dimImf = dimF$ *(cf. Corollaire 3.3.15).*

iii) Si $rgf = dimE$*, alors* $vect(f(v_1), f(v_2), ..., f(v_{dimE}))$ *est une famille génératrice de* Imf *dont le nombre d'éléments est égal à la dimension de* Imf *. C'est donc une base et en particulier une famille libre. Or, si* $x \in Kerf$ *, alors on a*

$$x = \sum_{i=1}^{dimE} \lambda_i v_i$$

et

$$0 = f(x) = f(\sum_{i=1}^{dimE} \lambda_i v_i) = \sum_{i=1}^{dimE} \lambda_i f(v_i).$$

On en déduit que $\lambda_1 = \lambda_2 = \ldots = \lambda_{dimE} = 0$*, d'où* $x = 0$*. Par conséquent,* $Kerf = \{0\}$ *et* f *est injective. Réciproquement, si* f *est injective, alors* $Kerf = \{0\}$*, donc en remontant le raisonnement précédent, on en déduit que la famille* $(f(v_1), f(v_2), ..., f(v_m))$ *est libre, donc c'est une base de* Imf *. Par conséquent,* $rgf = dimE$*.*

Corollaire 4.2.12. *Deux sev sont isomorphes, i.e. il existe un isomorphisme entre eux, si et seulement si leur dimension sont égales.*

Corollaire 4.2.13. *Si* f *est bijective, l'image d'une base est une base.*

On déduit que si $\{u_i; 1 \leq i \leq n\}$ est une base de E, f est entièrement déterminée par $\{f(u_i); 1 \leq i \leq n\}$.

Quelles sont les applications linéaires de $\mathbb{R}$ dans $\mathbb{R}$?

Proposition 4.2.14. *Soient* E *et* F *deux espaces vectoriels et* $f : E \to F$ *linéaire. Si* $\dim E < +\infty$*, alors*

$$dimE = dimImf + dimKerf.$$

démonstration 4.2.15. *Si* f *est injective, en désignant par* $(v_1, v_2, \ldots, v_m)$ *une base de* E*, nous avons vu que la famille à* m *éléments* $(f(v_1), f(v_2), ..., f(v_m))$ *est une famille libre de* F *donc une famille libre de* Imf*. De plus,* $(f(v_1), f(v_2), ..., f(v_m))$ *est une partie génératrice de* Imf*. Donc* $(f(v_1), f(v_2), ..., f(v_m))$ *est une base de* Imf*). La dimension de* Imf *est donc égale à* m *qui est la dimension de* E *et le théorème du rang est vrai.*

Si f *n'est pas injective, le noyau de* f *est un sous espace de* E *de dimension* p *avec* $1 \leq p \leq m$*. Soit* $(v_1, v_2, \ldots, v_p)$ *une base de* $Kerf$*. D'après le théorème de la base incomplète, il existe* $m - p$ *vecteurs* $v_{p+1}, \ldots, v_m$ *de* E *tels que* $(v_1, v_2, \ldots, v_m)$ *soit une base de* E*. Alors* Imf *est engendré par les vecteurs* $f(v_1), f(v_2), ..., f(v_m)$*. Mais, comme pour tout* i *compris entre* 1 *et* p *on a* $f(v_i) = 0$*,* Imf *est engendrée par les vecteurs* $f(v_{p+1}), ..., f(v_m)$*. Montrons que ces vecteurs sont linéairement indépendants. Soient* $\alpha_{p+1}, \ldots, \alpha_m$ *des scalaires tels que*

$$\alpha_{p+1} f(v_{p+1}) + \ldots + \alpha_m f(v_m) = 0.$$

Puisque f est linéaire cette égalité équivaut à l'égalité

$$f(\alpha_{p+1} v_{p+1} + \ldots + \alpha_m v_m) = 0$$

qui prouve que le vecteur $\alpha_{p+1} v_{p+1} + \ldots + \alpha_m v_m$ *appartient au noyau de* f *. Il existe donc des scalaires* $\lambda_1, \ldots, \lambda_p$ *tels que*

$$\alpha_{p+1} v_{p+1} + \ldots + \alpha_m v_m = \lambda_1 v_1 + \ldots + \lambda_p v_p.$$

Comme $(v_1, v_2, \ldots, v_m)$ *est une base de* E*, les vecteurs* $v_1, v_2, \ldots, v_m$ *sont linéairement indépendants et par conséquent :*

$$\forall i \in \{1, \ldots, p\}, \lambda_i = 0 \quad \textit{et} \quad \forall j \in \{p+1, \ldots, m\}, \alpha_j = 0.$$

Les vecteurs $f(v_{p+1}), ..., f(v_m)$ *définissent donc une base de* Imf *. Le sous espace vectoriel* Imf *est donc de dimension* $m - p$*. Ce qui achève la démonstration.*

Théorème 4.2.16. *L(E, F)={ applications linéaires de E dans F } est un espace vectoriel. Si* $\dim E < +\infty$ *et* $\dim F < +\infty$ *alors*

$$\dim L(E, F) = \dim E \times \dim F$$

démonstration 4.2.17. *L'ensemble des applications de E à valeurs dans F est un espace vectoriel, il suffit donc de montrer que $L(E, F)$ est un sev de ce dernier. L'application nulle qui envoie tout élément de E sur 0 est dans $L(E, F)$, donc $L(E, F) \neq \emptyset$.*

Soient f et g deux éléments de $L(E, F)$ et $\lambda, \mu \in K$. Alors pour tous $x, y \in E$ et $\alpha, \beta \in K$ on a

$$(\lambda f + \mu g)(\alpha x + \beta y) = (\lambda f)(\alpha x + \beta y) + (\mu g)(\alpha x + \beta y)$$

par linéarité de f et g on obtient

$$(\lambda f + \mu g)(\alpha x + \beta y) = \lambda(\alpha f(x) + \beta f(y)) + \mu(\alpha g(x) + \beta g(y))$$

et en regroupant les termes en α et β,

$$(\lambda f + \mu g)(\alpha x + \beta y) = \alpha(\lambda f(x) + \mu g(x)) + \beta(\lambda f(y) + \mu g(y))$$

soit encore

$$(\lambda f + \mu g)(\alpha x + \beta y) = \alpha(\lambda f + \mu g)(x) + \beta(\lambda f + \mu g)(y)$$

ce qui donne la linéarité de $\lambda f + \mu g$. On a ainsi démontré que $L(E, F)$ est espace vectoriel.

Soient $(e_1, \ldots, e_n)$ une base de E et $(f_1, \ldots, f_p)$ une base de F. Soit $x \in E$, alors il s'écrit sous la forme

$$x = \sum_{i=1}^{n} x_i e_i \quad \textit{avec} \quad x_i \in K \quad \textit{pour} \quad i = 1, \ldots, n.$$

Soit $u \in L(E, F)$, alors

$$u(x) = \sum_{i=1}^{n} x_i u(e_i).$$

D'autre part, comme $(f_1, \ldots, f_p)$ est une base de F, il existe $a_{j,i} \in K$ pour tout $i = 1, \ldots, n$ et pour $j = 1, \ldots, p$, tels que

$$u(e_i) = \sum_{j=1}^{p} a_{j,i} f_j.$$

Notons $u_{i,j}$ l'application linéaire de E dans F telle que

$$u_{i,j}(e_j) = f_i \quad \textit{et} \quad u_{i,j}(e_k) = 0 \quad \textit{si} \quad k \neq j.$$

Alors

$$u(e_i) = \sum_{j=1}^{p} a_{j,i} u_{j,i}(e_i).$$

D'où

$$u(x) = \sum_{i=1}^{n} x_i (\sum_{j=1}^{p} a_{j,i} u_{j,i}(e_i)) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{j,i} x_i u_{j,i}(e_i)$$

par linéarité de $u_{i,j}$, on a

$$u(x) = \sum_{i=1}^{n} x_i (\sum_{j=1}^{p} a_{j,i} u_{j,i}(e_i)) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{j,i} u_{j,i}(x_i e_i).$$

Par définition de $u_{i,j}$, on peut encore l'écrire sous la forme

$$u(x) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{j,i} u_{j,i}(\sum_{k=1}^{n} x_k e_k).$$

soit

$$u(x) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{j,i} u_{j,i}(x).$$

On a donc que

$$\{u_{j,i} \quad \text{pour} \quad i = 1, \ldots, n \quad \text{et pour} \quad j = 1, \ldots, p\}$$

est un système générateur.

Est-il libre ? Si

$$\sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_{j,i} u_{j,i} = 0$$

avec $\lambda_{j,i} \in K$ pour $i = 1, \ldots, n$ et pour $j = 1, \ldots, p$, alors pour tout $k = 1, \ldots, n$, on a

$$\sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_{j,i} u_{j,i}(e_k) = 0,$$

soit encore

$$\sum_{j=1}^{p} \lambda_{j,k} f_j = 0,$$

et comme $(f_1, \ldots, f_p)$ une base de F, on a

$$\lambda_{j,k} = 0 \quad \text{pour tout} \quad j = 1, \ldots, p \quad \text{et} \quad k = 1, \ldots, n.$$

Donc $\{u_{j,i} \quad$ pour $\quad i = 1, \ldots, n \quad$ et pour $\quad j = 1, \ldots, p\}$ est libre et c'est une base de $L(E, F)$. La dimension de $L(E, F)$ est donc bien $n \times p$.

4.3 Matrice d'une application linéaire

Soient E et F deux espaces vectoriels de dimension n et p.

4.3.1 Définitions

Définition 4.3.1. Matrice d'une application linéaire
Soit

$$\begin{array}{rcl} f: & E \longrightarrow & F \\ & (e_j)_{1\leq j\leq n} & (f_i)_{1\leq i\leq p} \end{array}$$

une application linéaire. Alors $M_f = (K_1...K_n)$ *où* K_j *est le vecteur des coordonnées de* $f(e_j)$ *dans la base* $(f_i)_{1\leq i\leq n}$ *est la matrice associée à l'application linéaire* f.

Théorème 4.3.2. *L'application*

$$\begin{array}{c} L(E,F) \longrightarrow M_{p,n} \\ f \mapsto M_f \end{array}$$

est une bijection.

démonstration 4.3.3. *Découle de la définition.*

Pour étudier f on peut étudier M_f et inversement. Comment trouver M_f ?

Exemples 4.3.4.
1)

$$\begin{array}{rc} f: & \mathbb{R}^2 \longrightarrow \mathbb{R} \\ & (x,y) \mapsto x+y \\ & M_f = \begin{pmatrix} 2 & 2 \end{pmatrix}, \end{array}$$

2)

$$\begin{array}{rc} f: & \mathbb{R} \longrightarrow \mathbb{R}^2 \\ & x \mapsto (2x,-x) \\ & M_f = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \end{array}$$

3)

$$\begin{array}{rc} f: & \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ & (x,y) \mapsto (x+y, y-2x) \\ & M_f = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}, \end{array}$$

Soient donc deux espaces vectoriels E et F de dimension finie et B et B' leur base respective.

Théorème 4.3.5. *Soient $f, g \in L(E, F)$ et $\lambda \in K$, M la matrice de f dans les bases B et B' , et N celle de g. Alors la matrice de $f + g$ dans ces mêmes bases est $M + N$ et celle de λf est λM.*

démonstration 4.3.6. *Immédiat en utilisant les définitions et la linéarité des composantes dans une base.*

4.3.2 Produit de matrices

Soient E, F et G trois espaces vectoriels de dimension finie et de bases $(e_k)_{1\leq k\leq n}$, $(f_j)_{1\leq j\leq p}$ et $(g_i)_{1\leq i\leq q}$. Soient $\varphi : E \longrightarrow F$ et $\psi : F \longrightarrow G$ linéaires.Si $M_\psi = \begin{pmatrix} a_{ij} \end{pmatrix}_{\{1\leq i\leq q, 1\leq j\leq p\}}$ et $M_\varphi = \begin{pmatrix} b_{jk} \end{pmatrix}_{\{1\leq j\leq p, 1\leq k\leq n\}}$. On a

$$\begin{aligned}
\psi o\varphi(e_k) = \psi(\sum_{j=1}^{p} b_{jk} f_j) &= \sum_{j=1}^{p} b_{jk}\psi(f_j) \\
&= \sum_{j=1}^{p} b_{jk}(\sum_{i=1}^{q} a_{ij} g_i) \\
&= \sum_{i=1}^{q}(\sum_{j=1}^{p} a_{ij} b_{jk}) g_i \\
&= \sum_{i=1}^{q} c_{ik} g_i
\end{aligned}$$

Alors $M_{\psi o\varphi} = \begin{pmatrix} c_{ik} \end{pmatrix}_{\{1\leq i\leq q, 1\leq k\leq n\}}$.

Définition 4.3.7. *On définit $M_\psi \times M_\varphi = M_{\psi o\varphi}$.*

Remarque 4.3.8. *On retrouve alors la définition du produit de deux matrices $A \in M_{q,p}$ et $B \in M_{p,n}$ vue dans la section Matrices, qui n'est que la traduction dans des bases de la composition d'applications linéaires.*

Proposition 4.3.9. *On a $f : E \to E$ bijective si et seulement si sa matrice A dans une base est inversible. Dans ce cas, la matrice de f^{-1} dans la même base est A^{-1} .*

démonstration 4.3.10. *On a que f est bijective si et seulement s'il existe $f^{-1} \in L(E)$ telle que*

$$f \circ f^{-1} = f^{-1} \circ f = Id.$$

Soit B une base de E et A la matrice de f dans cette base et A' celle de f^{-1} dans la même base. La matrice de l'application identité Id dans une base étant la matrice identité I_n , on en déduit que $AA' = A'A = I_n$, et réciproquement.

4.3.3 Changement de base

On a vu que la définition d'une matrice dépendait des bases dans lesquelles on se trouve. Soient $(e_i)_{1\leq i\leq n}$ et $(e'_i)_{1\leq i\leq n}$ deux bases de E. On considère la matrice dont les colonnes sont les coordonnées des vecteurs e'_i, $P = (e'_1 ... e'_n)$ i.e.

$$P = \left(\; p_{ij} \; \right)_{\{1\leq i\leq n, 1\leq j\leq n\}}$$

avec

$$e'_j = \sum_{i=1}^{n} p_{ij} e_i.$$

C′ est la matrice de passage de la base $(e_i)_{1\leq i\leq n}$ à la base $(e'_i)_{1\leq i\leq n}$. Si x est un vecteur de E de coordonnées X dans la base $(e_i)_{1\leq i\leq n}$ et de coordonnées X' dans la base $(e'_i)_{1\leq i\leq n}$, alors

$$X = PX'.$$

Proposition 4.3.11. *P est inversible.*

démonstration 4.3.12. *On peut même donner son inverse qui est la matrice de passage de la base $(e'_i)_{1\leq i\leq n}$ à la base $(e_i)_{1\leq i\leq n}$.*

$$P^{-1} = \left(\; q_{ij} \; \right)_{\{1\leq i\leq n, 1\leq j\leq n\}}$$

avec

$$e_j = \sum_{i=1}^{n} q_{ij} e'_i.$$

Exercice

On considère le vecteur $xe_1 + ye_2$ de $\mathbb{R}^2$. Quelles sont ses coordonnées dans la base (u, v) où $u = e_1 + e_2$ et $v = -e_1 + e_2$?

Problème de changement de base

On considère

$$\begin{array}{lll} E: & (e_i)_{1\leq i\leq n} & \longrightarrow (e'_i)_{1\leq i\leq n} \text{ avec } P \text{ matrice de passage} \\ F: & (f_j)_{1\leq j\leq p} & \longrightarrow (f'_j)_{1\leq j\leq p} \text{ avec } Q \text{ matrice de passage} \end{array}$$

Soit $\varphi : E \longrightarrow F$ linéaire dont la matrice dans les bases $(e_i)_{1\leq i\leq n}$ et $(f_j)_{1\leq j\leq p}$ est A. Quelle est la matrice B de φ dans les bases $(e'_i)_{1\leq i\leq n}$ et $(f'_j)_{1\leq j\leq p}$? On a $X = PX'$, $Y = QY'$ et $Y = AX$. On peut donc écrire $QY' = APX'$ et donc

$$B = Q^{-1}AP.$$

Exemples 4.3.13.

$$\begin{aligned} f: \quad E &\longrightarrow E \\ e_1 &\mapsto e_1 \\ e_2 &\mapsto -e_2 \end{aligned}$$

Définition 4.3.14. *Soit $(e_i)_{1\leq i\leq n}$ une famille de vecteurs. Le rang de cette famille est la dimension du sev engendré par cette famille.*

Soit $A = (K_1...K_n)$ une matrice de $M_{p,n}$. Le rang de A est le rang de la famille $\{K_1, ..., K_n\}$.

Proposition 4.3.15. *$A \in M_n(K)$ est inversible $\Longleftrightarrow rgA = n$.*

4.3.4 Interprétation matricielle de la méthode de Gauss

On va voir que l'algorithme de Gauss de réduction d'une matrice $m \times n$ à la forme échelonnée réduite s'interprète en termes de produits matriciels.

Définition 4.3.16. *On appelle matrice élémentaire toute matrice qui résulte de l'application d'une opération élémentaire sur les lignes à la matrice identité I_m.*

Exemples 4.3.17. *Dans le cas 3×3*

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

L'échange des lignes 1 et 3 donne la matrice élémentaire

$$E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

Le remplacement de la ligne 2 par elle-même plus 2 fois la ligne 1 donne la matrice élémentaire

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

La multiplication de la ligne 3 par 5 donne la matrice élémentaire

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix},$$

et ainsi de suite.

L'interprétation matricielle de la méthode de Gauss est fondée sur la remarque suivante.

Proposition 4.3.18. *Soit A une matrice $m \times n$ et E une matrice élémentaire. La matrice EA est celle qui résulte de l'application de la même opération élémentaire à la matrice A.*

démonstration 4.3.19. *Écrivons la matrice A comme une ligne de n vecteurs colonnes de $\mathbb{R}^m$,*

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}.$$

Par définition du produit matriciel, il vient donc

$$EA = \begin{pmatrix} Ea_1 & Ea_2 & \dots & Ea_n \end{pmatrix}.$$

Il suffit par conséquent de vérifier quel est l'effet de la multiplication par E sur un seul vecteur-colonne $x \in \mathbb{R}^m$. Soit $(e_1, e_2, ..., e_m)$ la base canonique de $\mathbb{R}^m$. Ces vecteurs sont les vecteurs-colonne de la matrice I_m, donc par définition d'une matrice élémentaire, Ee_i est le vecteur obtenu par l'opération élémentaire considérée appliquée au vecteur e_i. Or les opérations élémentaires sur les lignes définissent clairement des applications linéaires de $\mathbb{R}^m$ dans $\mathbb{R}^m$. Comme tout vecteur $x \in \mathbb{R}^m$ est combinaison linéaire des e_i , Ex n'est autre que le vecteur obtenu par l'opération élémentaire considérée appliquée au vecteur x.

Théorème 4.3.20. *Soit $A \in M_{m,n}$ et $U \in M_{m,n}$ l'unique matrice échelonnée réduite qui lui est équivalente. Alors il existe une unique matrice M inversible telle que*

$$U = MA \Longleftrightarrow A = M^{-1}U.$$

démonstration 4.3.21. *D'après la proposition précédente, chaque étape de l'algorithme de Gauss s'interprète matriciellement comme la multiplication à gauche de la matrice obtenue à l'étape précédente par une matrice élémentaire. Ainsi on a*

Premère étape : $A_1 = E_1A$.

Deuxième étape : $A_2 = E_2A_1 = E_2(E_1A) = (E_2E_1)A$.

Par récurrence, à la fin de l'algorithme, on a

pème étape : $A_p = E_pA_{p-1} = (E_pE_{p-1} \dots E_2E_1)A$.

On pose donc $M = E_pE_{p-1} \dots E_2E_1$. Comme chacune des opérations élémentaires est bijective, chaque matrice élémentaire E_k est inversible, d'où M est inversible.

Remarque 4.3.22. *Si la matrice M n'est pas très facile à calculer, on peut montrer que la matrice $M^{-1} = E_1^{-1}E_2^{-1} \dots E_{p-1}^{-1}E_p^{-1}$ est en fait un sous-produit gratuit de l'algorithme de Gauss, que l'on obtient sans calcul supplémentaire. Cette remarque est intéressante, car la formule $A = M^{-1}U$ est ce que l'on appelle une factorisation de A, et les factorisations d'une matrice ont de multiples applications.*

Remarque 4.3.23. *Dans le cas où* A *est une matrice carrée inversible, on a* $U = I_m$*, donc* $M = A^{-1}$*. On retrouve donc le calcul de* A^{-1} *par la méthode de Gauss en utilisant* $\tilde{A} = (AI_m)$*. En effet,* $\tilde{U} = M\tilde{A} = (MAMI_m) = (I_mA^{-1})$*.*

Chapitre 5

Déterminants

Dans ce chapitre nous ne considérerons que des matrices carrées.

5.1 Définitions

Dans le cas de $\mathbb{R}^2$, on considère deux vecteurs non nuls $X = (x_1, x_2)$ et $Y = (y_1, y_2)$ qui sont colinéaires. Il existe donc $\lambda \in \mathbb{R}^*$ tel que

$$X = \lambda Y,$$

ce qui peut aussi s'écrire sous la forme

$$\begin{aligned} x_1 &= \lambda y_1 \\ x_2 &= \lambda y_2. \end{aligned}$$

Multiplions la première ligne par y_2 et la seconde par y_1 et soustrayons les deux égalités ainsi obtenues, on a alors

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1 = 0.$$

Nous allons généraliser ce calcul ùn nombre fini de vecteurs. Pour cela, nous allons introduire la notion de déterminant.

Définition 5.1.1. *Le déterminant est l'unique application définie de*

$$\begin{aligned} M_n(\mathbb{R}) &\longrightarrow \mathbb{R} \\ A &\mapsto \det A \end{aligned}$$

qui vérifie

1)

$$\det\begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i,1}+a'_{i,1} & \cdot & \cdot & a_{i,n}+a'_{i,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{pmatrix} = \det\begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i,1} & \cdot & \cdot & a_{i,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{pmatrix} + \det\begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a'_{i,1} & \cdot & \cdot & a'_{i,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{pmatrix}$$

et

$$\det\begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ \lambda a_{i,1} & \cdot & \cdot & \lambda a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{pmatrix} = \lambda \det\begin{pmatrix} a_{1,1} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i,1} & \cdot & \cdot & a_{i,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{pmatrix}$$

2) Si A a deux lignes identiques, $\det A = 0$.

3) $\det I_n = 1$.

Notations : Si $A = \left(\; a_{ij} \; \right)_{\{1\leq i\leq n, 1\leq j\leq n\}}$ alors $\det A = \left| \; a_{ij} \; \right|$

Proposition 5.1.2. *1) Si on échange deux lignes dans une matrice, la valeur du déterminant est multipliée par -1.*

2) Si une ligne de A est nulle, $\det A = 0$.

3) Toutes les propriétés énoncées pour les lignes sont valables pour les colonnes.

ATTENTION

$$\det(A+B) \neq \det A + \det B$$

en général, comme le montre l'exemple suivant :

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

démonstration 5.1.3.

1) On considère la matrice

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

et la matrice où on a inversé la ligne i avec la ligne j

$$A' = \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

Introduisons une troisième matrice

$$\tilde{A} = \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} + a_{j,1} & \dots & a_{i,k} + a_{j,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} + a_{i,1} & \dots & a_{j,k} + a_{i,k} & \dots & a_{j,n} + a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

Comme cette matrice admet deux lignes identiques, son déterminant est nul.

Calculons ce déterminant en utilisant la propriété 1)

$$\det \tilde{A} = \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} + a_{i,1} & \dots & a_{j,k} + a_{i,k} & \dots & a_{j,n} + a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}$$
$$+ \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} + a_{i,1} & \dots & a_{j,k} + a_{i,k} & \dots & a_{j,n} + a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

En utilisant toujours la même propriété,

$$\det \tilde{A} = \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}$$

$$+ \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}$$

$$+ \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}$$

$$+ \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j,1} & \dots & a_{j,k} & \dots & a_{j,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

Ce qui nous donne

$$0 = \det \tilde{A} = \det A + 0 + 0 + \det A',$$

donc

$$\det A = -\det A'.$$

2) Il suffit de prendre $\lambda = 0$ dans la propriété 2).

Proposition 5.1.4. *Soit A une matrice $n \times n$ et A' la matrice obtenue en ajoutant à une colonne de A une combinaison linéaire des autres colonnes de A. Alors on a*

$$\det A = \det A'.$$

démonstration 5.1.5. *Soit*

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,k} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} & \dots & a_{n,n} \end{pmatrix}.$$

et soient λ_j pour $j = 1, \dots, n, j \neq k$. On pose

$$A' = \begin{pmatrix} a_{1,1} & \dots & a_{1,k} + \sum\limits_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j a_{1,j} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,k} + \sum\limits_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j a_{i,j} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,k} + \sum\limits_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j a_{n,j} & \dots & a_{n,n} \end{pmatrix}.$$

Par linéarité par rapport à la colonne k, on en déduit

$$\det A' = \det A + \sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j \det \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,j} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,n} \end{pmatrix},$$

et chacun des déterminants sous le signe de sommation est nul car les colonnes j et k sont égales.

Corollaire 5.1.6. *Si une colonne de A est combinaison linéaire des autres colonnes alors* $\det A = 0$.

démonstration 5.1.7. *En effet, on soustrait à cette colonne la combinaison linéaire en question, ce qui ne modifie pas le déterminant. La matrice obtenue a une colonne nulle, et par linéarité par rapport à cette colonne, le déterminant est nul.*

5.2 Calculs

Dans le cas de la dimension 2,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Dans le cas de la dimension 3,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix},$$

ou

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Exemples 5.2.1.

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

Proposition 5.2.2. *Si A est diagonale i. e. $a_{ij} = 0$ si $i \neq j$, alors* $\det A = \prod_{i=1}^{n} a_{ii}$. *Si A est triangulaire supérieure alors* $\det A = \prod_{i=1}^{n} a_{ii}$.

démonstration 5.2.3. *On traite le cas des matrices triangulaires supérieures, le cas des matrices triangulaires inférieures est identique. Soit donc*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{pmatrix},$$

Par linéarité par rapport à la première colonne, on a

$$\det A = a_{1,1} \det \begin{pmatrix} 1 & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{pmatrix}.$$

On ajoute à chaque colonne $j \geq 2$ le vecteur $-a_{1,j} \times$ la colonne 1. Ceci ne modifie pas le déterminant d'après la section précédente. Il vient donc

$$\det A = a_{1,1} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{pmatrix}.$$

Par linéarité par rapport à la deuxième colonne, on a

$$\det A = a_{1,1}a_{2,2} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{i,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{pmatrix},$$

et l'on continue ainsi jusqu'à avoir parcouru toutes les colonnes de la matrice. Au bout de n étapes, on a obtenu

$$\det A = a_{1,1}a_{2,2} \dots a_{n,n} \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = a_{1,1}a_{2,2} \dots a_{n,n} \det I_n,$$

d'où le résultat par la propriété 3). On procède de façon similaire pour les matrices triangulaires inférieures, et les matrices diagonales sont à la fois supérieure et inférieure.

Corollaire 5.2.4. *Soit E une matrice élémentaire de la méthode de Gauss.*

i) Si E est la matrice d'un remplacement de ligne, alors $\det E = 1$.

ii) Si E est la matrice d'un échange de lignes, alors $\det E = -1$.

iii) Si E est la matrice d'une multiplication d'une ligne par λ, alors $\det E = \lambda$.
Dans tous les cas, ce déterminant est non nul.

démonstration 5.2.5.

i) Dans ce cas, E est triangulaire inférieure ou supérieure avec des 1 sur la diagonale.

ii) Dans ce cas, E est aussi obtenue en échangeant les colonnes i et j de la matrice I_n .

iii) Matrice diagonale, tous les éléments diagonaux valent 1 sauf un qui vaut λ.

5.3 Propriétés

Théorème 5.3.1. *A est inversible si et seulement si* $\det A \neq 0$.

démonstration 5.3.2. *On procède en deux temps. Tout d'abord, si A n'est pas inversible, alors il existe une relation de dépendance linéaire entre ses colonnes, c'est-à-dire qu'au moins une colonne est combinaison linéaire des autres, donc* $\det A = 0$.

Réciproquement, si A est inversible, alors A^T *est aussi inversible, donc l'algorithme de Gauss appliqué à* A^T *fournit comme matrice échelonnée réduite équivalente à la matrice* I_n *. Il existe donc des matrices élémentaires* E_j *telles que*

$$(E_p E_{p-1} \dots E_2 E_1)A^T = I_n \Longleftrightarrow A(E_1^T E_2^T \dots E_{p-1}^T E_p^T) = I_n.$$

puisque $I_n^T = I_n$*. Or, multiplier une matrice à gauche par une matrice élémentaire effectue l'opération élémentaire correspondante sur les lignes de la matrice, donc par transposition, multiplier une matrice à droite par la transposée d'une matrice élémen -taire effectue l'opération élémentaire correspondante sur les colonnes. Dans le cas i), on ne modifie pas le déterminant, puisqu'on ajoute à une colonne un multiple d'une autre. Dans le cas ii), on multiplie le déterminant par* -1 *puisqu'on échange deux colonnes. Dans le cas iii), on multiplie le déterminant par* λ *par linéarité par rapport à la colonne multipliée par* λ*. On en déduit que*

$$\det(AE_1^T) = \lambda_1 \det A,$$

puis que

$$\det(AE_1^T E_2^T) = \lambda_2 \det(AE_1^T) = \lambda_1 \lambda_2 \det A,$$

puis par récurrence que

$$1 = \det I_n = \lambda_1 \dots \lambda_2 \lambda_p \det A,$$

où λ_i *vaut 1,* -1 *ou le facteur d'homothétie* λ *suivant la nature de l'opération élémentaire effectuée à l'étape i, c'est-à-dire* $\lambda_i = \det E_i$ *. Dans tous les cas, on voit que*

$$\det A \neq 0,$$

ce qui conclut la démonstration du Théorème.

On en déduit le résultat suivant :

Corollaire 5.3.3. *Soit un système de vecteurs* $\{V_1, ..., V_N\}$ *de coordonnées* $\{v_1, ..., v_n\}$ *dans une base* $\{e_1, ..., e_n\}$*. Alors*

$$\{V_1, ..., V_N\} \textit{est une base} \Longleftrightarrow \det_{(e_i)}(v_1 ... v_n) \neq 0.$$

De la démonstration du théorème précédent, on déduit le résultat suivant :

Théorème 5.3.4. $\det(AB) = \det A \det B$.

démonstration 5.3.5. *Supposons d'abord que B soit inversible. Soient E_i les matrices élémentaires et λ_i les scalaires non nuls qui leur sont associés tels que*

$$B(E_1^T E_2^T \dots E_{p-1}^T E_p^T) = I_n \quad \textit{et} \quad \det B = \frac{1}{\lambda_1 \dots \lambda_2 \lambda_p}$$

Posons $C = AB$. Par le même raisonnement que précédemment, il vient

$$\det C(E_1^T E_2^T \dots E_{p-1}^T E_p^T) = \lambda_1 \dots \lambda_2 \lambda_p \det C$$

Or

$$C(E_1^T E_2^T \dots E_{p-1}^T E_p^T) = AB(E_1^T E_2^T \dots E_{p-1}^T E_p^T) = AI_n = A,$$

d'où le résultat dans ce cas.

Si B n'est pas inversible, alors $\det B = 0$ d'une part. D'autre part, $rang B < n$ et comme $rang(AB) \leq rang B$, on voit que AB n'est pas inversible non plus, d'òu $\det(AB) = 0 = \det A \det B$ également dans ce cas.

Corollaire 5.3.6. *Si A est inversible alors*

$$\det(A^{-1}) = \frac{1}{\det A}. \tag{5.1}$$

démonstration 5.3.7. *En effet $AA^{-1} = I_n$, donc $\det A \det(A^{-1}) = \det I_n = 1$.*

Corollaire 5.3.8. *On a $\det(A^T) = \det A$.*

démonstration 5.3.9. *Soit $A \in M_n$. Par l'algorithme de Gauss, on a une factorisation $A = M^{-1}U$ avec U échelonnée réduite, donc en particulier triangulaire supérieure et*

$$M^{-1} = E_1^{-1} E_2^{-1} \dots E_{p-1}^{-1} E_p^{-1}$$

Par conséquent, en transposant on a aussi $A^T = U^T (M^{-1})^T$ avec

$$(M^{-1})^T = (E_p^{-1})^T (E_{p-1}^{-1})^T \dots (E_2^{-1})^T (E_1^{-1})^T.$$

Utilisant la multiplicativité du déterminant, on en déduit

$$\det A = \frac{\det U}{\det E_1 \det E_2 \dots \det E_{p-1} \det E_p}$$
$$\det A^T = \frac{\det U^T}{\det(E_1^T) \det(E_2^T) \dots \det(E_{p-1}^T) \det(E_p^T)}.$$

Or U est triangulaire supérieure, son déterminant est le produit de ses termes diagonaux. Par conséquent, U^T est triangulaire inférieure et son déterminant est le produit de ces mêmes termes diagonaux, c'est-à-dire $\det U = \det(U^T)$. De même, les matrices E_i sont soit triangulaires (substitution), soit symétriques c'est-à-dire égales à leur transposée (échange de lignes et homothétie). Par conséquent, $\det E_i = \det E_i^T$ aussi, d'où le résultat.

Le résultat suivant donne le comportement du déterminant lorsque l'on fait un changement de base :

Corollaire 5.3.10. *Si* $A' = P^{-1}AP$ *alors* $\det A = \det A'$.

démonstration 5.3.11. *On a*

$$\det A' = \det(P^{-1}) \det A \det P = \det A$$

d'après 5.3.6.

Remarque 5.3.12. *ATTENTION !*

$$\det(\lambda A) = \det((\lambda I_n)A) = \det(\lambda I_n) \det A = \lambda^n \det A.$$

5.4 Applications

On considère un système de n équations à n inconnues. On peut l'écrire sous la forme $AX = Y$, où

$$A = \left(\; a_{ij} \; \right)_{\{1\leq i\leq n, 1\leq j\leq n\}} \quad X = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}.$$

Proposition 5.4.1. *Un système linéaire admet une unique solution si et seulement si* $\det A \neq 0$. *Dans ce cas l'unique solution est donnée par les formules de Cramer :*

$$x_k = \frac{\det \begin{pmatrix} a_{11} & \cdot & a_{1k-1} & y_1 & a_{1k+1} & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ a_{n1} & \cdot & a_{nk-1} & y_n & a_{nk+1} & \cdot & a_{nn} \end{pmatrix}}{\det A}$$

démonstration 5.4.2. *Écrivons*

$$A = (a_1 a_2 \dots a_i \dots a_n) \quad \textit{et} \quad I_n = (e_1 e_2 \dots e_i \dots e_n)$$

comme des lignes de vecteurs, où e_i *sont les vecteurs de la base canonique de* $\mathbb{R}^n$. *Introduisons la matrice*

$$I_{n,k(x)} = (e_1 e_2 \dots x \dots e_n)$$

dans laquelle on a remplacé e_k *par* x. *Il vient*

$$AI_{n,k(x)} = (Ae_1 Ae_2 \dots Ax \dots Ae_n) = \begin{pmatrix} a_{11} & \cdot & a_{1k-1} & y_1 & a_{1k+1} & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ a_{n1} & \cdot & a_{nk-1} & y_n & a_{nk+1} & \cdot & a_{nn} \end{pmatrix}$$

Par conséquent,

$$\det A \det I_{n,k(x)} = \det \begin{pmatrix} a_{11} & \cdot & a_{1k-1} & y_1 & a_{1k+1} & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ a_{n1} & \cdot & a_{nk-1} & y_n & a_{nk+1} & \cdot & a_{nn} \end{pmatrix}$$

formule qui reste vraie même si A n'est pas inversible, du moment que x est solution du système linéaire. Or en développant $\det I_{n,k(x)}$ par rapport à sa première colonne ou en utilisant la linéarité par rapport à la colonne k, on voit aisément que

$$\det I_{n,k(x)} = x_k.$$

Exemples 5.4.3. *Soit le système*

$$\begin{cases} x + y = 1 \\ 2x - 3y = 4 \end{cases}$$

Le déterminant est égal à

$$\begin{vmatrix} 1 & 1 \\ 2 & -3 \end{vmatrix} = -5,$$

et on a

$$x = \frac{\begin{vmatrix} 1 & 1 \\ 4 & -3 \end{vmatrix}}{-5} = \frac{7}{5},$$

et

$$y = \frac{\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix}}{-5} = \frac{-2}{5}.$$

Remarque 5.4.4. *Sauf pour $n = 2$, et $n = 3$, ou pour des matrices très particulières, ce n'est pas la bonne façon de résoudre un système linéaire. On a tout intérêt à utiliser la méthode de Gauss dès que $n \geq 3$ (à condition de faire bien attention à ne pas diviser par 0 quand les coefficients dépendent de paramètres). En revanche, les formules de Cramer sont intéressantes pour la théorie.*

5.5 Interprétation géométrique des déterminants

On a une interprétation géométrique de $\mathbb{R}^n$ pour $n = 1, 2, 3$. En effet, en dimension 2, les déterminants sont liés aux questions de surface et en dimension 3 aux questions de volume. En dimension 2, deux vecteurs v_1, v_2 définissent un parallélogramme, alors qu'en dimension 3, trois vecteurs v_1, v_2, v_3 définissent un parallélépipède. On prendra comme unité de surface la surface du carré unité dont les cotés sont les vecteurs de la base canonique, et comme unité de volume, le volume du cube unité construit de la même façon en dimension 3.

Proposition 5.5.1. *La surface du parallélogramme est donnée par*

$$|\det(v_1 v_2)|.$$

Le volume du parallélépipède est donné par

$$|\det(v_1 v_2 v_3)|.$$

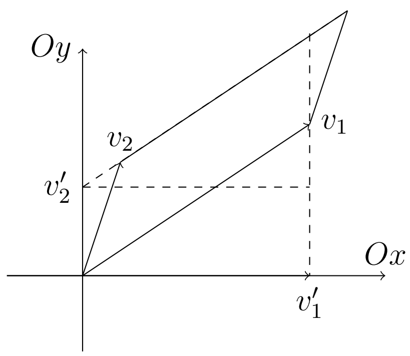
démonstration 5.5.2. *Traitons le cas n = 2. Le résultat est vrai si*

$$(v_1 v_2) = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

En effet, dans ce cas on a affaire à un rectangle de côtés $|a|$ et $|b|$, donc de surface $|ad|$, alors que le déterminant de la matrice vaut ad. Supposons que (v_1, v_2) est une famille libre. Notons

$$(v_1 v_2) = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.$$

Si $a_{1,1} \neq 0$, alors $v_2' = v_2 - \frac{a_{1,2}}{a_{1,1}} v_1$, est un multiple de e_2, c'est-à-dire que sa première composante est nulle. L'opération ne change ni le déterminant, ni la surface du parallélogramme. Comme la famille de départ était libre, $v_2' \neq 0$ et ce vecteur a une deuxième composante $a_{2,2}'$ non nulle. On pose alors $v_1' = v_1 - \frac{a_{2,1}}{a_{2,2}'} v_2'$, ce qui produit un vecteur multiple de e_1. L'opération ne change ni le déterminant ni la surface des parallélogrammes. On est donc ramené au premier cas d'un rectangle aux côtés parallèle aux axes, pour lequel le résultat est déjà acquis.



Les diverses opérations ci-dessus ne modifient pas les surfaces.

Si $a_{1,1} = 0$, alors $a_{1,2} \neq 0$ puisque la famille est libre, et on échange les rôles de v_1 et v_2. Enfin, si la famille est liée, alors le déterminant vaut 0. Dans ce cas, le parall élogramme est réduit à un segment et est donc de surface nulle.

Le cas tridimensionnel se traite de façon analogue.

Chapitre 6

Valeurs propres et vecteurs propres

6.1 Définitions

Définition 6.1.1. *Soit f une fonction de $L(E)$.*

$$\lambda \text{ est une valeur propre de } f \iff Ker(f - \lambda Id) \neq \{0\}.$$

On appelle vecteur propre de f tout $x \in E$, $x \neq 0$, tel que $f(x) = \lambda x$.
Pour toute valeur propre λ, on appelle espace propre noté

$$E_\lambda(f) = \{x \in E; f(x) = \lambda x\} = Ker(f - \lambda Id).$$

Exemples 6.1.2.
1)

$$\begin{aligned} \mathbb{R}[X] &\longrightarrow \mathbb{R}[X] \\ P &\mapsto Q, \end{aligned}$$

où $Q(X) = XP(X)$ n'a pas de valeur propre.
2)

$$\begin{aligned} \mathbb{R}[X] &\longrightarrow \mathbb{R}[X] \\ P &\mapsto P' \end{aligned}$$

a une seule valeur propre, c'est 0 et $E_0 = \{constantes\}$.

De la définition découle les équivalence suivantes :

Proposition 6.1.3. *Soit E est un espace de dimension finie.*

$$\begin{aligned} \lambda \text{ est une valeur propre } &\iff f - \lambda Id \text{ est non injective} \\ &\iff Ker(f - \lambda Id) \neq \{0\} \\ &\iff det(A - \lambda I) = 0. \end{aligned}$$

Dans toute la suite on se place dans un espace de dimension finie.

Définition 6.1.4. $P_A(X) = det(A - XI) \in K[X]$ *est le polynôme caractéristique de* f.

Remarque 6.1.5. *La matrice* A *dépend de la base dans laquelle on définit* f. *Mais d'après le résultat 5.3.10, le polynôme caractéristique sera le même quelque soit la base dans laquelle on travaille.*

Remarque 6.1.6. *On voit en développant par rapport à la première colonne que le degré de* P_A *est égal à* n *et que le terme de plus haut degré de* $P_A(X)$ *est* $(-1)^n X^n$.

Corollaire 6.1.7.

$$\lambda \text{ est une valeur propre } \iff \lambda \text{ est racine } de P_A.$$

On est ainsi ramené pour la recherche des valeurs propres à un problème classique des mathématiques, d'une importance historique majeure, trouver les racines d'un polynôme de degré n comme on l'a vu au premier chapitre . Comme un tel polynôme a au plus n racines, on peut tout de suite en déduire le corollaire suivant.

Corollaire 6.1.8. *Toute matrice* $n \times n$ *admet au plus* n *valeurs propres.*

L'existence de valeurs propres va donc dépendre de K.

Corollaire 6.1.9. *Si* $K = \mathbb{C}$, *toute matrice admet au moins une valeur propre.*

démonstration 6.1.10. *C'est le théorème de d'Alembert qui affirme que tout polynôme à coefficients complexes admet au moins une racine dans* $\mathbb{C}$.

Corollaire 6.1.11. *Les matrices* A *et* A^T *admettent les mêmes valeurs propres.*

démonstration 6.1.12. *En effet,*

$$P_A^T(X) = \det(A^T - XI_n) = \det((A - XI_n)^T) = \det(A - XI_n) = P_A(X).$$

6.2 Propriétés des espaces propres en dimension finie

Soient λ une valeur propre, α l'ordre de multiplicité de λ dans le polynôme caractéristique de f, $\alpha \geq 1$ et $\beta = dim E_\lambda$.

Proposition 6.2.1. $1 \leq \beta \leq \alpha$, *les inégalités peuvent être strictes.*

démonstration 6.2.2. *On considère une base de E_λ et on la complète en une base de E. Dans la base ainsi constituée la matrice de f s'écrit*

$$\left(\begin{array}{cccc|c} \lambda & & & & \\ & \lambda & 0 & & A' \\ & & \lambda & & \\ 0 & & & \lambda & \\ \hline & 0 & & & A'' \end{array}\right)$$

Donc, en utilisant le résultat 5.3.10

$$P_A(X) = \left(\begin{array}{cccc|c} \lambda - X & & & & \\ & \lambda - X & & 0 & A' \\ & & \lambda - X & & \\ 0 & & & \lambda - X & \\ \hline & 0 & & & A'' - XI_{n-\beta} \end{array}\right).$$

On en déduit que

$$P_A(X) = (\lambda - X)^\beta \det(A'' - XI_{n-\beta}),$$

et de ce fait $\beta \leq \alpha$.

Exercices

1)

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

2)

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

Proposition 6.2.3. *La somme des espaces propres est directe.*

Remarque 6.2.4. *On peut avoir $\sum_\oplus E_\lambda = \{0\}$ ou $\sum_\oplus E_\lambda \subset_{\neq} E$ ou $\sum_\oplus E_\lambda = E$.*

démonstration 6.2.5. *Soient $\lambda_1, \ldots, \lambda_p \in K$ les valeurs propres distinctes de A et soient $E_{\lambda_1}, ..., E_{\lambda_p}$ les sous-espaces propres associés dans E. On raisonne par récurrence sur le nombre de sous-espaces propres. Un seul sous-espace propre est trivialement en somme directe, car la seule décomposition du vecteur nul est le vecteur nul.*

Supposons le résultat acquis pour $k-1$ sous-espaces propres avec $k \leq p$. Donnons-nous $x_i \in E_{\lambda_i}$, $i = 1, ..., k$ tels que $x_1 + x_2 + \ + x_k = 0$.

Multiplions cette relation par λ_k. Il vient

$$\lambda_k x_1 + \lambda_k x_2 + \; + \lambda_k x_k = 0.$$

Multiplions également cette égalité par la matrice A. Il vient

$$Ax_1 + Ax_2 + \; + Ax_k = \lambda_1 x_1 + \lambda_2 x_2 + \; + \lambda_k x_k = 0.$$

Soustrayons les deux égalités obtenues membre à membre. On obtient

$$(\lambda_k - \lambda_1)x_1 + (\lambda_k - \lambda_2)x_2 + \; + (\lambda_k - \lambda_{k-1})x_{k-1} = 0.$$

Posant $y_i = (\lambda_k - \lambda_i)x_i \in E_{\lambda_i}$, $i = 1, \dots, k-1$, on a obtenu une décomposition du vecteur nul sur $k-1$ sous-espaces propres. Par l'hypothèse de récurrence, on obtient $y_i = 0$ pour $i = 1, \dots, k-1$. Or on a pris des valeurs propres distinctes, donc $\lambda_k - \lambda_i \neq 0$. Par conséquent, $x_i = 0$ pour $i = 1, \dots, k-1$. Reportant ceci dans la première relation, on en déduit finalement que $x_k = 0$, et les sous-espaces propres sont bien en somme directe.

6.2.1 Diagonalisation

Définition 6.2.6. *On dit que $f \in L(E)$ est diagonalisable s'il existe une base $(b_i)_{i \in I}$ de E et $\lambda_i \in \mathbb{C} \quad \forall i \in I$ telle que*

$$f(b_i) = \lambda_i b_i \quad \forall i \in I.$$

Si $(b_i)_{1 \leq i \leq n}$ est une base de vecteurs propres la matrice de f dans cette base s'écrit

$$D = \begin{pmatrix} \lambda_1 & \cdot & \cdot & 0 \\ \cdot & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & \lambda_n \end{pmatrix}.$$

Si P est la matrice de passage de la base $(e_i)_{1 \leq i \leq n}$ à la base $(b_i)_{1 \leq i \leq n}$ et A la matrice de f dans la base $(e_i)_{1 \leq i \leq n}$, alors

$$D = P^{-1}AP$$

ce qui est équivalent à

$$A = PDP^{-1}.$$

Cela va nous permettre d'introduire la notion de matrice diagonalisable.

Définition 6.2.7. *On dit qu'une matrice $A \in M_n$ est diagonalisable s'il existe une matrice inversible $P \in M_n$ et une matrice diagonale $D \in M_n$ telle que $A = P^{-1}DP$.*

Proposition 6.2.8. *A est diagonalisable si et seulement s'il existe une base de E constituée de vecteurs propres.*

démonstration 6.2.9. *Supposons qu'il existe $B = \{u_1, u_2, \ldots, u_n\}$ base de E telle que $Au_i = \lambda_i u_i$, alors la matrice D de l'application linéaire $x \to Ax$ dans la base B est diagonale, ses éléments diagonaux étant les valeurs propres λ_i. Si P désigne la matrice de passage de la base canonique dans B, on a bien que $D = P^{-1}AP$. Réciproquement, supposons qu'il existe une matrice inversible $P \in M_n$ et une matrice diagonale $D \in M_n$ telle que $A = P^{-1}DP$. Interprétons P comme la matrice de passage de la base canonique à une autre base $B = \{u_1, u_2, \ldots, u_n\}$ (les composantes de u_i dans la base canonique sont simplement les colonnes de P). Alors on voit que*

$$Au_i = P^{-1}DPu_i = P^{-1}De_i = P^{-1}(\lambda_i e_i) = \lambda_i P^{-1}e_i = \lambda_i u_i,$$

donc $Au_i = \lambda_i u_i$, et les u_i forment une base de vecteurs propres de A.

Corollaire 6.2.10. *f est diagonalisable si et seulement si $E = \sum_{\oplus} E_\lambda$.*

démonstration 6.2.11. *En effet, la réunion des bases des sous-espaces propres forme alors une base de E constituée de vecteurs propres de A.*

Corollaire 6.2.12. *Si A est diagonalisable alors son polynôme caractéristique est scindé, i.e. on peut le factoriser en polynômes du premier degré.*

démonstration 6.2.13. *Si A est diagonalisable alors il existe une matrice inversible $P \in M_n$ et une matrice diagonale $D \in M_n$ telle que $A = P^{-1}DP$. On alors $P_A(X) = \det(D - XI_n) = \prod_{i=1}^n (\lambda_i - X)$.*

Corollaire 6.2.14. *Soit A une matrice. Si son polynôme caractéristique est scindé et si toutes les racines de P_A sont simples, alors A est diagonalisable.*

démonstration 6.2.15. *On applique les résultats de 6.2.14 et 6.2.1. En effet dans ce cas la dimension des espaces propres est égal à 1, la somme des dimension est donc égal à n, et $E = \sum_{\oplus} E_\lambda$.*

Exemples 6.2.16.

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

6.2.2 Applications

- **Systèmes linéaires**
On considère un système de n équations à n inconnues. On peut l'écrire sous la forme $AX = Y$, où

$$A = \left(\begin{array}{c} a_{ij} \end{array} \right)_{\{1 \leq i \leq n, 1 \leq j \leq n\}} \quad X = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}.$$

Si $A = PDP^{-1}$ le système devient

$$PDP^{-1}X = Y \Longleftrightarrow DP^{-1}X = P^{-1}Y.$$

On pose $Z = P^{-1}X$ et $V = P^{-1}Y$, on obtient alors le système $DZ = V$, très simple à résoudre. On trouve X par $X = PZ$.
- **Systèmes différentiels**
On veut résoudre le système différentiel $\frac{dX}{dt} = AX$. Si $A = PDP^{-1}$ le système devient

$$\frac{dX}{dt} = PDP^{-1}X \Longleftrightarrow P^{-1}\frac{dX}{dt} = DP^{-1}X.$$

On pose $Z = P^{-1}X$ alors $\frac{dZ}{dt} = P^{-1}\frac{dX}{dt}$ car la dérivation est linéaire et le système devient $\frac{dZ}{dt} = DZ$, qui se résout en exponentielle. Et la solution X est égale à PZ.
- **Puissance d'une matrice**
Proposition 6.2.17. *Si $A = PDP^{-1}$ alors $A^m = PD^mP^{-1}$. Plus généralement, si Q est un polynôme à une indéterminée, et on a*

$$D = \begin{pmatrix} \lambda_1 & \cdot & \cdot & 0 \\ \cdot & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & \lambda_n \end{pmatrix},$$

alors

$$Q(A) = P \begin{pmatrix} Q(\lambda_1) & \cdot & \cdot & 0 \\ \cdot & Q(\lambda_2) & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & Q(\lambda_n) \end{pmatrix} P^{-1}.$$

démonstration 6.2.18. *On a*

$$A^m = (PDP^{-1})(PDP^{-1})\dots(PDP^{-1}) = PD(P^{-1}P)D\dots(P^{-1}P)DP^{-1} = PD^mP^{-1}.$$

Il est évident que

$$D^m = \begin{pmatrix} \lambda_1^m & \cdot & \cdot & 0 \\ \cdot & \lambda_2^m & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \cdot & \cdot & \lambda_n^m \end{pmatrix},$$

d'où le résultat. Le cas du polynôme Q en découle immédiatement.