Notes de cours
L1 — MATH120

Hervé Le Dret

11 mai 2004

Chapitre 1

Rappels sur les nombres complexes

Dans ces notes de cours, on travaillera essentiellement à l'aide de nombres réels, dont les propriétés algébriques sont supposées acquises, mais parfois également avec les nombres complexes.

1.1 Définition et premières propriétés

On adopte une présentation algébrique des nombres complexes. Pour cela, on va munir l'ensemble $\mathbb{R}^2$ de deux lois de composition internes, respectivement appelées addition et multiplication, qui sont définies de la façon suivante.

– Pour tout couple $((x,y),(x',y'))$ de $\mathbb{R}^2 \times \mathbb{R}^2$, on définit la somme par

$$(x,y) + (x',y') = (x+x', y+y'),$$

– et le produit par

$$(x,y) \times (x',y') = (xx' - yy', xy' + yx').$$

L'ensemble $\mathbb{R}^2$ muni de ces deux lois est appelé *ensemble des nombres complexes*, ou encore *plan complexe* (pour une raison géométrique que l'on détaillera un peu plus loin et qui correspond à l'interprétation qui en est donnée dans l'enseignement secondaire), et on le note $\mathbb{C}$.

Par convention, il est commode d'identifier les nombres complexes de la forme $(x,0)$ avec le nombre réel x, si bien que l'on pose

$$\forall x \in \mathbb{R}, \quad (x,0) = x.$$

En particulier, on notera $(1,0) = 1$ et $(0,0) = 0$. On introduit enfin la notation suivante

$$i = (0,1).$$

Théorème 1 *Avec les notations précédentes, on a*

$$(x,y) = x + iy,$$

et l'on a

$$i^2 = -1.$$

Démonstration. En effet, soit $z = (x,y) \in \mathbb{C}$. Alors z peut encore s'écrire, en utilisant les définitions des opérations internes

$$z = (x+0, 0+y) = (x,0) + (0,y) = (x,0) + (0,1) \times (y,0) = x + iy.$$

De plus, $i^2 = (0,1) \times (0,1) = (-1,0) = -1$. □

Ces notations sont extrêmement pratiques car elles simplifient les calculs.

Proposition 1 *Les calculs algébriques élémentaires sur les nombres complexes s'effectuent exactement comme les calculs sur les nombres réels en remplaçant* i^2 *par* -1.

Démonstration. Soit $z = (x,y) = x + iy$ et $z' = (x',y') = x' + iy'$ deux nombres complexes. Considérons le cas de l'addition. On a

$$\begin{aligned} z + z' = (x,y) + (x',y') &= (x+x', y+y') = (x+x') + i(y+y') \\ &= x + x' + iy + iy' = (x+iy) + (x'+iy'). \end{aligned}$$

De même pour le produit

$$\begin{aligned} zz' = (x,y) \times (x',y') &= (xx' - yy', xy' + yx') \\ &= (xx' - yy') + i(xy' + yx') \\ &= xx' + i(xy' + yx') + i^2 yy' \\ &= (x+iy)(x'+iy'), \end{aligned}$$

si bien qu'à partir de maintenant, on peut sans regret abandonner les définitions utilisant les couples de réels. □

Définition 1 *Soit* $z = x + iy$ *un nombre complexe. Le réel* x *s'appelle la* partie réelle *de* z *et le réel* y *sa* partie imaginaire. *On les note*

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

Remarque 1 On a donc $z = \Re(z) + i\Im(z)$. Notons bien que la partie réelle et la partie imaginaire d'un nombre complexe sont toutes les deux des nombres réels. On peut réécrire les définitions des opérations internes sous la forme suivante

$$\begin{aligned} \Re(z+z') &= \Re(z) + \Re(z') &\text{ et }\quad \Im(z+z') &= \Im(z) + \Im(z'), \\ \Re(zz') &= \Re(z)\Re(z') - \Im(z)\Im(z') &\text{ et }\quad \Im(zz') &= \Re(z)\Im(z') + \Im(z)\Re(z'). \end{aligned}$$

Les nombres complexes de la forme iy avec $y \in \mathbb{R}$ sont appelés des *nombres imaginaires purs*. Le carré d'un nombre imaginaire pur est un réel négatif, $(iy)^2 = -y^2$ et réciproquement i est une racine carrée de -1 (l'autre étant $-i$). À l'époque où ces nombres ont été inventés, à la Renaissance, cette propriété était jugée tellement choquante que ces nombres ont été qualifiés d'« imaginaires », c'est-à-dire qu'on ne croyait pas vraiment à leur existence. Depuis le temps, on s'est habitué à l'idée...

Notons que les puissances successives du nombre i sont de la forme $i^2 = -1$, $i^3 = i^2 \times i = -i$, $i^4 = (i^2)^2 = (-1)^2 = 1$ et plus généralement $i^{2p} = (-1)^p$ pour les puissances paires et $i^{2p+1} = (-1)^p i$ pour les puissances impaires.

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 $(0,1) = j$. Malheureusement, la tradition mathématique donne à la lettre j une toute autre signification $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$. Ce n'est qu'une question de contexte, et naturellement nous suivrons ici la tradition mathématique. $\square$

Théorème 2 *Les propriétés algébriques des opérations sur les complexes sont les mêmes que celles des opérations sur les réels. En particulier, pour tout $z = x + iy \in \mathbb{C}$, on a*

$$-z = -x - iy \text{ et si } z \neq 0, \frac{1}{z} = \frac{x - iy}{x^2 + y^2}.$$

Démonstration. Dans cette démonstration, on notera systématiquement $x = \Re(z)$, $y = \Im(z)$, $x' = \Re(z')$, etc.

Il y a toute une liste de propriétés à vérifier.

– L'addition est associative

$$\begin{aligned}(z+z')+z'' &= \big((x+x')+i(y+y')\big)+x''+iy'' \\ &= (x+x'+x'')+i(y+y'+y'') \\ &= (x+iy)+\big((x'+x'')+i(y'+y'')\big) \\ &= z+(z'+z''),\end{aligned}$$

pour tous z, z', z'' dans $\mathbb{C}$. On n'a donc pas à se préoccuper des parenthèses dans une somme de complexes.

– Elle admet un élément neutre, qui n'est autre que $(0,0) = 0$, car

$$z + 0 = x + 0 + i(y+0) = x + iy = z = 0 + z,$$

pour tout $z \in \mathbb{C}$.

– Tout $z \in \mathbb{C}$ admet un opposé pour l'addition, qui n'est autre que $-x - iy$, car

$$z + (-x - iy) = x - x + i(y - y) = 0 = (-x - iy) + z.$$

On note l'opposé de z par $-z = -x - iy$.

– L'addition est commutative

$$\begin{aligned} z+z' &= (x+x')+i(y+y') \\ &= (x'+x)+i(y'+y) \\ &= z'+z, \end{aligned}$$

pour tous z, z' dans $\mathbb{C}$. On n'a donc pas à se préoccuper de l'ordre des termes dans une somme de complexes.

Ces quatre propriétés algébriques font de $(\mathbb{C}, +)$ ce qu'on appelle un *groupe commutatif* ou *groupe abélien*.

Pour ce qui concerne maintenant la multiplication, on a les propriétés suivantes.

– La multiplication est associative

$$\begin{aligned} [zz']z'' &= [(x+iy)(x'+iy')](x''+iy'') \\ &= \big[(xx'-yy')+i(xy'+yx')\big](x''+iy'') \\ &= (xx'-yy')x''-(xy'+yx')y''+i\big((xx'-yy')y''+(xy'+yx')x''\big) \\ &= xx'x''-xy'y''-(yy'x''+yx'y'')+i(xx'y''+xy'x''+yx'x''-yy'y'') \\ &= x(x'x''-y'y'')-y(y'x''+x'y'')+i\big(x(x'y''+y'x'')+y(x'x''-y'y'')\big) \\ &= (x+iy)\big[(x'x''-y'y'')+i(y'x''+x'y'')\big] \\ &= (x+iy)[(x'+iy')(x''+iy'')] \\ &= z[z'z''], \end{aligned}$$

pour tous z, z', z'' dans $\mathbb{C}$. On n'a donc pas à se préoccuper des parenthèses dans un produit de complexes.

– Elle est distributive par rapport à l'addition

$$\begin{aligned} z(z'+z'') &= (x+iy)(x'+x'+i(y'+y'')) \\ &= x(x'+x'')-y(y'+y'')+i\big(x(y'+y'')+y(x'+x'')\big) \\ &= xx'-yy'+i(xy'+yx')+xx''-yy''+i(xy''+yx'') \\ &= zz'+zz'', \end{aligned}$$

pour tous z, z', z'' dans $\mathbb{C}$. De même, $(z+z')z'' = zz''+z'z''$. Ceci permet de développer et factoriser les expressions algébriques utilisant des nombres complexes.

Ajoutées aux quatre précédentes, ces deux propriétés algébriques supplémentaires font de $(\mathbb{C}, +, \times)$ ce que l'on appelle en algèbre un *anneau*. Mais ce n'est pas tout, car

– La multiplication est commutative

$$\begin{aligned} zz' &= (xx'-yy')+i(xy'+yx') \\ &= (x'x-y'y)+i(x'y+y'x) \\ &= z'z, \end{aligned}$$

pour tous z, z' dans $\mathbb{C}$. On n'a donc pas à se préoccuper de l'ordre des termes dans un produit de complexes.

– Elle admet un élément neutre, qui n'est autre que $(1,0) = 1$, car

$$z \times 1 = x \times 1 - y \times 0 + i(y \times 1 + x \times 0) = x + iy = z = 1 \times z,$$

pour tout $z \in \mathbb{C}$.

Ajoutées aux six précédentes, ces deux propriétés algébriques supplémentaires font de $(\mathbb{C}, +, \times)$ ce que l'on appelle en algèbre un *anneau unitaire commutatif*. Mais ce n'est (encore) pas tout, car

– Tout $z \in \mathbb{C}$, $z \neq 0$ admet un inverse pour la multiplication. En effet, $z \neq 0$ si et seulement si $x \neq 0$ ou $y \neq 0$, c'est-à-dire, si et seulement si $x^2 + y^2 > 0$. Le nombre réel $x^2 + y^2$ n'est pas nul, donc il a un inverse $\frac{1}{x^2+y^2}$ pour la multiplication dans $\mathbb{R}$. On note alors que

$$(x+iy)(x-iy)\frac{1}{x^2+y^2} = [(x^2 - (-y^2)) + i \times 0]\frac{1}{x^2+y^2} = 1.$$

On note l'inverse de $z \neq 0$ par $z^{-1} = \frac{1}{z} = \frac{x-iy}{x^2+y^2}$.

Ajoutées aux huit précédentes, cette dernière propriété algébrique supplémentaire fait de $(\mathbb{C}, +, \times)$ ce que l'on appelle en algèbre un *corps commutatif*. □

Remarque 2 L'ensemble des nombres complexes privé de 0 est noté

$$\mathbb{C}^* = \mathbb{C} \setminus \{0\}.$$

On remarque que les propriétés algébriques 5, 7, 8 et 9 font de $(\mathbb{C}^*, \times)$ également un groupe commutatif. □

L'ensemble $\mathbb{R}$ des nombres réels muni de l'addition et de la multiplication usuelles possède également ces neuf propriétés algébriques : c'est aussi un corps commutatif (il existe bien d'autres corps commutatifs en algèbre). En fait, à cause de l'identification $(x,0) = x$ pour tout x réel, on considère d'ordinaire que $\mathbb{R}$ est inclus dans $\mathbb{C}$. En effet, l'addition et la multiplication de $\mathbb{C}$ prolongent celles de $\mathbb{R}$, manifestement. On dit que $\mathbb{R}$ est un sous-corps de $\mathbb{C}$ ou que $\mathbb{C}$ est une extension de corps de $\mathbb{R}$.

Ce qu'il faut retenir de cette discussion, c'est que les calculs algébriques usuels (développements, factorisations, divisions par des termes non nuls) *s'effectuent exactement de la même façon dans* $\mathbb{C}$ *que dans* $\mathbb{R}$ (ou d'ailleurs que dans tout autre corps commutatif). En particulier, les identités remarquables bien connues sur $\mathbb{R}$ restent valables dans $\mathbb{C}$. Ainsi, on a toujours, pour tous $a, b \in \mathbb{C}$,

$$\begin{aligned} a^2 - b^2 &= (a-b)(a+b) \text{ et plus généralement} \\ a^n - b^n &= (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}), \end{aligned}$$

dont une variante amusante est $a^2+b^2=(a-ib)(a+ib)$ obtenue en remplaçant b par ib, la formule du binôme

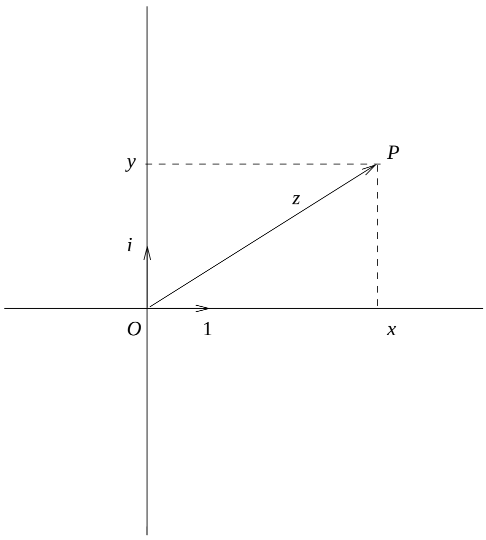
$$(a+b)^n=\sum_{k=0}^{n} C_n^k a^{n-k} b^k$$

où $C_n^k=\frac{n!}{k!(n-k)!}$ est le nombre de combinaisons de k objets parmi n objets, etc., etc.

1.2 Interprétation géométrique, module, argument, trigonométrie

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.

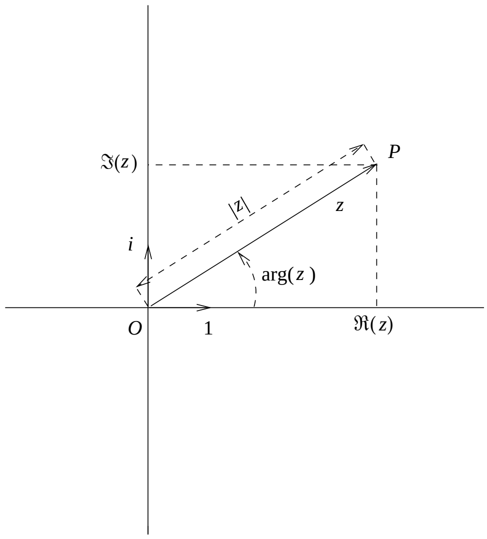
On obtient donc une interprétation géométrique de $\mathbb{C}$ en associant à tout nombre complexe $z=x+iy$ le point P du plan de coordonnées (x,y) (on a choisi un repère cartésien orthonormé une fois pour toutes). Cette interprétation explique pourquoi on parle parfois de plan complexe pour $\mathbb{C}$. 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 (c'est-à-dire pour le moment une flèche de O à P. Nous reviendrons plus loin sur la notion de vecteur en général).



Le plan complexe

Dans cette représentation, les nombres réels sont situés sur l'axe des abscisses, appelé dès lors *axe réel*, et les nombres imaginaires purs sur l'axe des ordonnées, appelé *axe imaginaire*.

Définition 2 *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* $\arg(z)$.



Module et argument

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.

Proposition 2 *On a*

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

et

$$\arg(z) = \mathrm{arctg}\left(\frac{\Im(z)}{\Re(z)}\right) \mod 2\pi$$

quand $\Re(z) > 0$ *et*

$$\arg(z) = \mathrm{arctg}\left(\frac{\Im(z)}{\Re(z)}\right) + \pi \mod 2\pi$$

quand $\Re(z) < 0$. *Enfin, quand* $\Re(z) = 0$, *on a* $\arg(z) = \frac{\pi}{2} \mod 2\pi$ *si* $\Im(z) > 0$ *et* $\arg(z) = -\frac{\pi}{2} \mod 2\pi$ *si* $\Im(z) < 0$.

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

$$\begin{cases} \Re(z) &= \rho\cos\theta, \\ \Im(z) &= \rho\sin\theta. \end{cases}$$

Démonstration. 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 $\operatorname{tg}(\arg(z)) = \frac{\Im(z)}{\Re(z)}$ (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 pourquoi on utilise la même notation pour les deux notions.

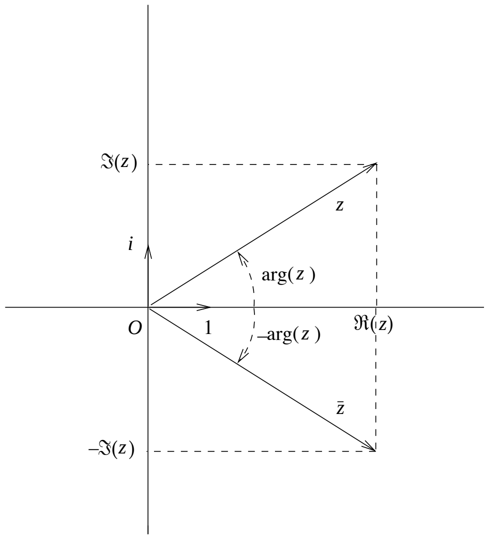
Remarque 3 Le passage $(\Re(z), \Im(z)) \leftrightarrow (|z|, \arg(z))$ correspond du point de vue géométrique au passage des coordonnées cartésiennes aux coordonnées polaires. Le couple module-argument donne donc une autre représentation du plan complexe. □

Introduisons maintenant une notion très importante, la *conjugaison complexe*.

Définition 3 *Pour tout $z = x + iy$ avec x, y réels, on définit son* complexe conjugué *par*

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

On peut également dire que $\Re(\overline{z}) = \Re(z)$ et $\Im(\overline{z}) = -\Im(z)$. Du point de vue géométrique, la conjugaison complexe se traduit simplement par une symétrie orthogonale par rapport à l'axe réel.



La conjugaison complexe

Ainsi, par exemple $\overline{1} = 1$ et $\overline{i} = -i$. Quelques propriétés sont plus ou moins évidentes.

Proposition 3 *Pour tout $z \in \mathbb{C}$, on a*

$$\overline{\overline{z}} = z,\ |\overline{z}| = |z|,\ \arg(\overline{z}) = -\arg(z),\ \Re(z) = \frac{z+\overline{z}}{2} \ \textit{ et } \ \Im(z) = \frac{z-\overline{z}}{2i}.$$

D'autres le sont marginalement moins.

Proposition 4 *Pour tous $z, z' \in \mathbb{C}$, on a*

$$\overline{z+z'} = \overline{z} + \overline{z'} \quad \textit{et} \quad \overline{zz'} = \overline{z}\,\overline{z'}.$$

Par ailleurs

$$z\overline{z} = |z|^2, \quad |zz'| = |z||z'|,$$

et si $z \neq 0$ on a

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

Enfin, un nombre complexe z est réel si et seulement si $\overline{z} = z$ et il est imaginaire pur si et seulement si $\overline{z} = -z$.

Démonstration. Montrons seulement les points non évidents. Soient $z = x + iy$ et $z' = x' + iy'$ avec les conventions habituelles. On a donc $zz' = (xx' - yy') + i(xy' + yx')$, d'où

$$\overline{zz'} = (xx' - yy') - i(xy' + yx') = (x - iy)(x' - iy') = \overline{z}\,\overline{z'}.$$

On a aussi
$$z\overline{z} = (x+iy)(x-iy) = x^2+y^2 = |z|^2,$$
par l'identité remarquable déjà notée plus haut. Du coup, $|zz'|^2 = zz'\overline{zz'} = zz'\overline{z}\overline{z'} = |z|^2|z'|^2$. Également, on en déduit que $z\frac{\overline{z}}{|z|^2} = 1$, d'où $z^{-1} = \frac{\overline{z}}{|z|^2}$ quand z est non nul.

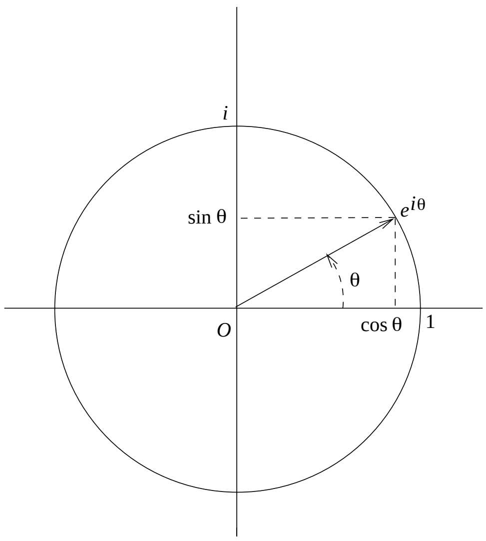
Enfin, z est réel si et seulement si sa partie imaginaire est nulle, donc si et seulement si $\frac{z-\overline{z}}{2i} = 0$ ou encore si et seulement si $z = \overline{z}$. De même, z est imaginaire pur si et seulement si sa partie réelle est nulle donc si et seulement si $\frac{z+\overline{z}}{2} = 0$ ou encore $z = -\overline{z}$. □

Ainsi on a $\frac{1}{i} = -i$. Remarquons que si $z = x+iy$ avec x, y réels, alors on a bien
$$\overline{z} = \overline{x+iy} = \overline{x}+\overline{i}\overline{y} = x-iy$$
puisque $\overline{x} = x$ et $\overline{y} = y$.

Les nombres complexes de module 1 jouent un rôle particulier.

Définition 4 *On note*
$$S^1 = \{z \in \mathbb{C}; |z| = 1\}.$$

Du point de vue géométrique, l'ensemble des nombres complexes de module 1 est représenté par le *cercle unité*, c'est-à-dire le *cercle trigonométrique*.



Le cercle unité

En fait, il est clair que l'on a

Proposition 5 *On a*

$$S^1 = \{z \in \mathbb{C}, \exists \theta \in \mathbb{R}; z = \cos\theta + i \sin\theta\},$$

avec $\theta = \arg(z) \mod 2\pi$.

Si l'on spécialise à ce cas les résultats précédents, on voit en particulier que

Proposition 6 *Pour tous* $z, z' \in S^1$*, on a* $zz' \in S^1$ *et* $z^{-1} = \bar{z} \in S^1$.

Démonstration. En effet, $|zz'| = |z||z'| = 1$ et $z^{-1} = \bar{z}/1$. □

Ceci signifie que la multiplication est une loi de composition interne sur S^1 et que $(S^1, \times)$ est un groupe commutatif.

Proposition 7 *Pour tous* $z, z' \in S^1$*, on a*

$$\arg(zz') = \arg(z) + \arg(z') \mod 2\pi.$$

Démonstration. Soit $\theta = \arg(z)$ et $\theta' = \arg(z')$. On a donc

$$z = \cos\theta + i \sin\theta \quad \text{et} \quad z' = \cos\theta' + i \sin\theta'.$$

On calcule le produit

$$zz' = \cos\theta\cos\theta' - \sin\theta\sin\theta' + i(\cos\theta\sin\theta' + \sin\theta\cos\theta'),$$

d'où par les formules d'addition des angles en trigonométrie

$$zz' = \cos(\theta + \theta') + i \sin(\theta + \theta'),$$

d'où le résultat. □

Cette formule justifie d'introduire l'exponentielle complexe.

Définition 5 *Pour tout* $\theta \in \mathbb{R}$*, on pose*

$$e^{i\theta} = \cos\theta + i \sin\theta \in S^1,$$

qui satisfait la relation fonctionnelle de l'exponentielle

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

De plus, $\theta = \arg(e^{i\theta})$ *et*

$$\frac{1}{e^{i\theta}} = \overline{e^{i\theta}} = e^{-i\theta}.$$

Remarque 4 L'exponentielle complexe apparaît ici simplement comme une notation commode. Mais c'est en réalité une « vraie » exponentielle, c'est-à-dire qu'il s'agit bien du nombre e élevé à la puissance $i\theta$, comme on le montre à un niveau un peu plus élevé. Notons le cas particulier de $\theta = \pi$ qui donne la célèbre formule d'Euler $e^{i\pi} = -1$.

On a en fait défini ainsi une application de $\mathbb{R}$ dans S^1, surjective et périodique de période 2π ($e^{i\theta} = e^{i(\theta+2\pi)}$ pour tout θ). Comme elle transforme l'addition en multiplication, on dit que c'est un homomorphisme du groupe $(\mathbb{R},+)$ dans le groupe $(S^1,\times)$. Il s'agit en fait de l'« enroulement » naturel de la droite réelle sur le cercle unité. □

On déduit immédiatement de la définition les formules de de Moivre.

Proposition 8 *On a*

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2} \quad \textit{et} \quad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

Ces formules servent (entre autres) à linéariser les puissances des fonctions circulaires à l'aide de la formule du binôme. En effet, on déduit immédiatement de la propriété d'homomorphisme que $(e^{i\theta})^n = e^{in\theta}$ pour tout n entier. Donnons un exemple.

$$\begin{aligned}
\sin^3\theta &= \left(\tfrac{e^{i\theta}-e^{-i\theta}}{2i}\right)^3 \\
&= \left(\tfrac{1}{2i}\right)^3 [e^{i3\theta} - 3e^{i2\theta}e^{-i\theta} + 3e^{i\theta}e^{-i2\theta} - e^{-i3\theta}] \\
&= \left(\tfrac{1}{2i}\right)^2 \tfrac{e^{i3\theta}-e^{-i3\theta}}{2i} - 3\left(\tfrac{1}{2i}\right)^2 \tfrac{e^{i\theta}-e^{-i\theta}}{2i} \\
&= -\tfrac{1}{4}\sin 3\theta + \tfrac{3}{4}\sin\theta.
\end{aligned}$$

Proposition 9 *Pour tout $z \in \mathbb{C}$, il existe $u \in S^1$ tel que $z = |z|u$. Si $z \neq 0$, alors u est unique et on a*

$$z = |z|e^{i\arg(z)}.$$

Démonstration. En effet, si $z = 0$, alors $|z| = 0$ et on peut écrire $z = |z| \times 1$ avec $1 \in S^1$, par exemple.

Si $z \neq 0$, alors $|z| > 0$, donc $u = z/|z|$ est bien défini. De plus $|u| = |z|/|z| = 1$, donc $u \in S^1$. Enfin, si $u' \in S^1$ est tel que $z = |z|u'$, divisant par $|z|$ on en déduit que $u' = u$, c'est-à-dire l'unicité.

Pour identifier u, il suffit de remarquer que deux complexes non nuls z et z' tels que $z = tz'$ avec $t > 0$ ont même argument, ce qui est géométriquement évident. Ici on a donc $\arg(z) = \arg(u)$ et bien sûr $u = e^{i\arg(u)}$. □

Corollaire 3 *Pour tous* $z, z' \in \mathbb{C}^*$,

$$zz' = |z||z'|e^{i(\arg(z)+\arg(z'))}.$$

En particulier,

$$\arg(zz') = \arg(z) + \arg(z') \mod 2\pi.$$

Démonstration. On a déjà établi ce résultat dans S^1, on va l'étendre à $\mathbb{C}^*$. On pose $u = z/|z|$ et $u' = z'/|z'|$. On a déjà noté que $\arg(u) = \arg(z)$ et $\arg(u') = \arg(z')$. De plus, on sait que $\arg(uu') = \arg(u) + \arg(u') \mod 2\pi$. Or, clairement $zz' = |z||z'|uu'$, d'où le résultat puisque $|z||z'| > 0$. □

Remarque 5 Soit $z' \in \mathbb{C}^*$ fixé. On définit une application $\mathbb{C} \to \mathbb{C}$, $z \mapsto zz'$. D'après ce qui précède, cette application a pour effet de multiplier le module de z par celui de z', c'est-à-dire d'effectuer une *homothétie* du plan de rapport $|z'|$, puis d'augmenter l'argument de z par celui de z', c'est-à-dire d'effectuer une *rotation* du plan d'angle $\arg(z')$. La composition de ces deux applications s'appelle une *similitude* (de centre O). Toute les similitudes de centre O sont représentées par la multiplication par un complexe donné. □

1.3 Polynômes et équations polynomiales

Les équations polynomiales jouent un grand rôle dans l'étude de $\mathbb{C}$. De façon un peu informelle, un *polynôme à une indéterminée* à coefficients complexes est une expression de la forme

$$P(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$$

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 coefficient complexe définit une application de $\mathbb{C}$ dans $\mathbb{C}$ par

$$z \mapsto P(z) = a_0 + a_1z + a_2z^2 + \cdots + a_nz^n,$$

et l'on peut sans danger confondre, dans ce cas précis, le polynôme et la fonction polynomiale associée. 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 polynomiale 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 $P(\zeta) = 0$, c'est-à-dire est solution de cette équation polynomiale. Un problème de première importance est de trouver les racines des polynômes. On admettra le résultat suivant.

Proposition 10 *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*

$$P(X) = (X - \zeta)Q(X).$$

Corollaire 4 *Un polynôme de degré $n \geq 0$ admet au plus n racines.*

Démonstration. En effet, s'il en avait $n+1$, $\zeta_1, \zeta_2, \dots, \zeta_{n+1}$, alors on aurait

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

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

En fait, le théorème suivant est célèbre, c'est le théorème de d'Alembert que l'on admettra car sa démonstration est loin d'être élémentaire.

Théorème 5 *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*

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

Ce théorème s'applique en particulier aux polynômes de $\mathbb{R}[X]$ de degré n, qui ont donc n racines complexes. Par contre, ces racines ne sont pas forcément réelles. Par exemple, les deux racines de $P(X) = X^2 + 1$ sont $\pm i$.

Proposition 11 *Soit P un polynôme de $\mathbb{R}[X]$ et ζ une racine de P. Alors, $\bar{\zeta}$ est aussi une racine de P.*

Démonstration. En effet, si l'on prend le complexe conjugué de l'équation $P(\zeta) = 0$, on obtient

$$0 = \overline{a_0 + a_1\zeta + \cdots + a_n\zeta^n} = a_0 + a_1\bar{\zeta} + \cdots + a_n\bar{\zeta}^n$$

puisque $\bar{a}_k = a_k$. $\square$

Les racines d'un polynôme réel sont donc de deux sortes. Soit elles sont réelles, auquel cas elles sont égales à leur conjugué, soit elles ne sont pas réelles, et dans ce cas elles sont groupées par paires avec leur conjugué.

En remarquant que

$$(X-\zeta)(X-\overline{\zeta}) = X^2 - 2\Re\zeta X + |\zeta|^2 = (X-\Re\zeta)^2 + \Im\zeta^2,$$

on obtient une factorisation réelle des polynômes réels en termes du premier degré, correspondant aux racines réelles, et en termes du second degré sans racine réelle, de la forme ci-dessus, correspondant aux racines à partie imaginaire non nulle.

1.4 Racines *n*-èmes de l'unité

On va en considérer quelques équations polynomiales très particulières.

Définition 6 *Soit $n \geq 1$ un nombre entier. On dit qu'un nombre complexe ω est une racine n-ème de l'unité si $\omega^n = 1$.*

Pour $n = 1$, il n'y a manifestement qu'une racine 1-ème de l'unité $\omega_0 = 1$. Pour $n = 2$, il y a deux racines carrées $\omega_0 = 1$ et $\omega_1 = -1$. En effet, on cherche les solutions de l'équation $\omega^2 - 1 = 0$. Or par les identités remarquables, $\omega^2 - 1 = (\omega - 1)(\omega + 1)$ et un produit de deux complexes est nul si et seulement au moins un des deux facteurs est nul. Généralisons.

Proposition 12 *Un nombre complexe ω est une racine n-ème de l'unité si et seulement si c'est une racine de l'équation polynomiale de degré n*

$$z^n - 1 = 0.$$

Remarquons que cette équation a toujours la solution triviale $\omega_0 = 1$ et que par les identités remarquables, les autres racines sont solution de l'équation polynomiale

$$z^{n-1} + z^{n-2} + \cdots + z + 1 = 0.$$

Remarquons tout d'abord que toutes ces racines sont forcément de module 1.

Proposition 13 *Soit ω une racine n-ème de l'unité. Alors $\omega \in S^1$.*

Démonstration. En effet, soit $\omega \in \mathbb{C}$ tel que $\omega^n = 1$. Prenant les modules, on en déduit que $|\omega^n| = |\omega|^n = 1$. Or $|\omega|$ est un nombre réel positif, et sur $\mathbb{R}$, 1 n'a qu'une seule racine n-ème positive, à savoir $1 = |\omega|$. □

Toute racine n-ème de l'unité ω est donc de la forme $\omega = e^{i\theta}$ avec $\theta = \arg(\omega)$.

Proposition 14 *Il y a exactement n racines n-èmes de l'unité, données par*

$$\omega_k = e^{i\frac{k2\pi}{n}}, k = 0, \ldots, n-1.$$

Démonstration. En effet, une telle racine doit satisfaire $e^{in\theta} = 1$, c'est-à-dire

$$\cos(n\theta) + i\sin(n\theta) = 1,$$

soit encore

$$\cos(n\theta) = 1 \quad \text{et} \quad \sin(n\theta) = 0.$$

On en déduit que $n\theta = k2\pi$, pour k entier, ou $\theta = \frac{k2\pi}{n}$. Les n nombres $\omega_k = e^{i\frac{k2\pi}{n}}$, $k = 0, \ldots, n-1$ sont donc des racines n-ème de l'unité. Elles sont distinctes, puisque placées régulièrement sur le cercle unité à des angles d'intervalle $2\pi/n$. Enfin, la 2π-périodicité de l'exponentielle complexe montre que les autres valeurs de k entières redonnent les mêmes exponentielles complexes. □

Pour $n = 2$, on retrouve bien les 2 racines carrées $\omega_0 = e^0 = 1$ et $\omega_1 = e^{i\pi} = -1$.

Pour $n = 3$, on trouve 3 racines cubiques de l'unité $\omega_0 = 1$, $\omega_1 = e^{i2\pi/3}$ et $\omega_2 = e^{i4\pi/3}$. En particulier

$$\omega_1 = \cos(2\pi/3) + i\sin(2\pi/3) = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = j$$

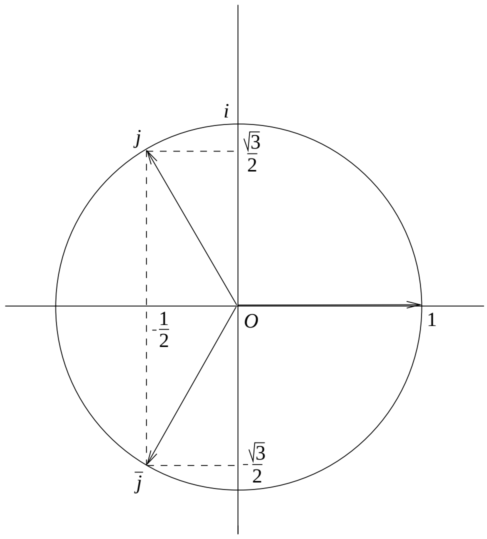
au sens des mathématiciens. Le nombre j est donc une racine cubique de l'unité, la troisième étant $\omega_2 = \bar{j}$. Ces deux nombres sont les racines de l'équation du second degré

$$z^2 + z + 1 = 0,$$

dont le discriminant est $\Delta = 1 - 4 = -3$, d'où $j = \frac{-1+\sqrt{-3}}{2} = \frac{-1+i\sqrt{3}}{2}$ et $\bar{j} = \frac{-1-\sqrt{-3}}{2} = \frac{-1-i\sqrt{3}}{2}$.

Pour $n = 4$, on trouve quatre racines quatrièmes de l'unité $\omega_0 = 1$, $\omega_1 = i$, $\omega_2 = -1$ et $\omega_3 = -i$.

Et ainsi de suite.



Les trois racines cubiques de l'unité

Chapitre 2

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). Pour ce qui concerne les mathématiques, ils forment la base calculatoire de l'algèbre linéaire. Ils permettent également de traiter une bonne partie de la théorie de l'algèbre linéaire en dimension finie.

2.1 Définition

On ne va traiter que le cas des systèmes linéaires faisant intervenir seulement des nombres réels. On s'apercevra plus tard que 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* $x_1, x_2, \ldots, x_n$ est une équation de la forme

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

où $a_1, a_2, \ldots, a_n$ et b sont des nombres réels donnés.

Soit $m \in \mathbb{N}^*$ un autre entier naturel supérieur à 1.

Définition 7 *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.

Exemple 1 Voici un système de 2 équations à 3 inconnues

$$\begin{cases} 2x_1 & -x_2 & +\frac{3}{2}x_3 & = & 8, \\ x_1 & & -4x_3 & = & -7. \end{cases}$$

On aurait pu l'écrire tout aussi bien

$$\left\{\begin{array}{rrrcr} 2x_1 & -x_2 & +\frac{3}{2}x_3 & = & 8, \\ x_1 & +0\times x_2 & -4x_3 & = & -7. \end{array}\right.$$

□

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

$$\left\{\begin{array}{cccccccl} a_{11}x_1 & +a_{12}x_2 & +a_{13}x_3 & + & \cdots & +a_{1n}x_n & = & b_1\,(\leftarrow \text{équation n}^\circ\ 1) \\ a_{21}x_1 & +a_{22}x_2 & +a_{23}x_3 & + & \cdots & +a_{2n}x_n & = & b_2\,(\leftarrow \text{équation n}^\circ\ 2) \\ \vdots & \vdots & \vdots & & & \vdots & = & \vdots \\ a_{i1}x_1 & +a_{i2}x_2 & +a_{i3}x_3 & + & \cdots & +a_{in}x_n & = & b_i\,(\leftarrow \text{équation n}^\circ\ i) \\ \vdots & \vdots & \vdots & & & \vdots & = & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +a_{m3}x_3 & + & \cdots & +a_{mn}x_n & = & b_m\,(\leftarrow \text{équation n}^\circ\ m) \end{array}\right.$$

Les nombres a_{ij}, $i=1,\ldots,m$, $j=1,\ldots,n$, sont les *coefficients* du système. Ce sont des données. Les nombres b_i, $i=1,\ldots,m$, 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_{ij} 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 1, on a $m=2$ (nombre d'équations = nombre de lignes), $n=3$ (nombre d'inconnues = nombre de colonnes à gauche du signe =) et $a_{11}=2$, $a_{12}=-1$, $a_{13}=3/2$, $a_{21}=1$, $a_{22}=0$, $a_{23}=-4$, $b_1=8$ et $b_2=-7$.

Définition 8 *Une solution du système linéaire est une liste de n nombres réels* $(s_1,s_2,\ldots,s_n)$ *(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, $(5,13/2,3)$ est une solution du système linéaire de l'exemple 1. En règle générale, on s'attache à déterminer l'ensemble des solutions d'un système linéaire. C'est ce que l'on appelle *résoudre* le système linéaire. Ceci amène à poser la définition suivante.

Définition 9 *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. Nous verrons plus loin comment procéder de façon systématique pour arriver à ce but.

Remarque 6 Deux systèmes équivalents ont toujours visiblement le même nombre d'inconnues. Par contre, ils n'ont pas forcément le même nombre d'équations. Dans ce dernier cas, on peut toujours ajouter au système avec le moins d'équations le nombre manquant à l'aide d'équations triviales

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

lesquelles ne modifient clairement pas l'ensemble des solutions. □

Exemple 2 Résolution dans le cas d'un système 2×2. Considérons le système suivant

$$\left\{ \begin{array}{rrcr} x_1 & -2x_2 & = & -1, \\ -x_1 & +3x_2 & = & 3. \end{array} \right.$$

Si x_1 et x_2 désigne les coordonnées cartésiennes d'un point du plan, on reconnaît deux équations de droite, une par ligne du système. Par conséquent, toute solution (s_1, s_2) du système correspond aux coordonnées d'un point d'intersection des deux droites. On se ramène donc à un problème *géométrique* très simple dans ce cas particulier. Dans cet exemple, les deux droites se coupent au point de coordonnées $(3,2)$. On a obtenu l'ensemble des solutions $S = \{(3,2)\}$ constitué ici d'un seul élément (on calcule cette solution très simplement en additionnant les deux équations, puis en remplaçant la valeur de x_2 ainsi trouvée).

Il aurait pu tout aussi bien se produire que les deux droites soient parallèles, comme dans l'exemple suivant

$$\left\{ \begin{array}{rrcr} x_1 & -2x_2 & = & -1, \\ -x_1 & +2x_2 & = & 3. \end{array} \right.$$

Dans ce cas, les deux droites ne se coupent pas, donc le système n'a pas de solution. L'ensemble des solutions est l'ensemble vide $S = \emptyset$. Ceci se voit algébriquement en remarquant que le membre de gauche de la première ligne est égal à l'opposé du membre de gauche de la première ligne. Comme $1 \neq 3$, il est impossible de satisfaire en même temps les deux équations linéaires.

Enfin, la troisième et dernière possibilité géométrique est que les deux droites soient confondues.

$$\begin{cases} x_1 & -2x_2 & = & -1, \\ -x_1 & +2x_2 & = & 1. \end{cases}$$

On a alors une infinité de solutions $S = \{\text{coordonnées des points de la droite}\}$.

Ces trois cas de figure obtenus dans le cas de systèmes 2×2 recouvrent en fait la situation générale, comme on le démontrera plus loin. On a en effet l'alternative suivante pour l'ensemble des solutions d'un système linéaire général $m \times n$.

a) Soit il n'y a aucune solution, $S = \emptyset$. Dans ce cas, on dit que le système est *incompatible*.

b) Soit il y a une solution unique, $S = \{(s_1, s_2, \dots, s_n)\}$ l'ensemble des solutions contient un seul n-uplet. Dans ce cas, on dit que le système est *compatible*.

c) Soit il y a une infinité de solutions, et on dit aussi dans ce cas que le système est compatible.

Un cas particulier important est celui des *systèmes homogènes* pour lesquels $b_1 = b_2 = \ldots = b_m = 0$, c'est-à-dire dont le second membre est nul. De tels systèmes sont toujours compatibles car ils admettent toujours la solution $s_1 = s_2 = \ldots = s_n = 0$. Cette solution est appelée *solution triviale*. Géométriquement dans le cas 2×2, un système homogène correspond à deux droites qui passent par l'origine des coordonnées, cette origine $(0,0)$ étant donc toujours solution. Dans le cas des systèmes homogènes, on s'attachera par conséquent à déterminer s'il n'y a que la solution triviale ou s'il y en a d'autres. $\square$

2.2 Notation matricielle

En réfléchissant un petit peu, on se rend compte que dans la donnée d'un système linéaire, seuls comptent les coefficients du système et le second membre. Écrire les équations avec les inconnues permet de visualiser le système, mais n'est pas autrement utile. On introduit donc une façon plus compacte d'écrire un système linéaire : la *notation matricielle.* Il s'agit simplement de ranger les coefficients et le second membre dans des tableaux rectangulaires en suivant l'ordre naturel des lignes et des colonnes.

Plus précisément, on introduit les objets suivants

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

L'objet A s'appelle *la matrice du système linéaire*. Elle a m lignes et n colonnes, c'est une *matrice* $m\times n$ (à coefficients réels). Le coefficient a_{ij} se trouve à l'intersection de la ligne numéro i et de la colonne numéro j. On note aussi de façon générique $A=(a_{ij})_{i=1,\dots,m,j=1,\dots,n}$ ou $A=(a_{ij})$ si la taille de la matrice est sous-entendue.

On introduit aussi

$$\tilde{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} & b_1\\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} & b_2\\ \vdots & \vdots & & \vdots & & \vdots & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} & b_i\\ \vdots & \vdots & & \vdots & & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} & b_m \end{pmatrix}$$

On l'appelle *la matrice augmentée du système.* C'est une matrice $m\times(n+1)$. Elle contient la matrice des coefficients avec une colonne supplémentaire ajoutée à sa droite et contenant le second membre, c'est-à-dire toute l'information nécessaire à déterminer le système.

Exemple 3 Il est très facile de passer d'un système linéaire à sa matrice augmentée et vice-versa : il suffit de lire les coefficients au bon endroit. Considérons l'exemple du système 3×3 suivant

$$\begin{cases} x_1 & -2x_2 & +x_3 & = & 0,\\ & 2x_2 & -8x_3 & = & 8,\\ -4x_1 & +5x_2 & +9x_3 & = & -9. \end{cases}$$

Sa matrice est

$$A=\begin{pmatrix} 1 & -2 & 1\\ 0 & 2 & -8\\ -4 & 5 & 9 \end{pmatrix}$$

et sa matrice augmentée

$$\tilde{A}=\begin{pmatrix} 1 & -2 & 1 & 0\\ 0 & 2 & -8 & 8\\ -4 & 5 & 9 & -9 \end{pmatrix}.$$

□

Définition 10 *On dit que deux matrices $A=(a_{ij})$ et $B=(b_{ij})$ sont égales si elles sont de la même taille et si $a_{ij}=b_{ij}$ pour tout couple d'indices i et j.*

En d'autres termes, deux matrices sont égales si et seulement si elles ont les mêmes coefficients aux mêmes endroits.

2.3 Systèmes échelonnés réduits

Essayons d'imaginer le cas le plus favorable, un système dont la résolution soit parfaitement triviale :

$$\begin{cases} x_1 & & & & = b_1, \\ & x_2 & & & = b_2, \\ & & \ddots & & \vdots \quad \vdots \\ & & & x_n & = b_n. \end{cases}$$

Sa matrice augmentée n'est autre que

$$\tilde{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & b_1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & b_2 \\ 0 & 0 & 1 & \cdots & 0 & 0 & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & b_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 & b_n \end{pmatrix}$$

Cette matrice montre une disposition de coefficients nuls bien particulière. Généralisons tout de suite.

Définition 11 *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.*

Clairement, la matrice $\tilde{A}$ précédente est échelonnée réduite.

Remarque 7 Grâce 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, toujours à cause de 2) et de 1), on voit 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. $\square$

Exemple 4

$$A = \begin{pmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

est échelonnée et

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 & 25 \\ 0 & 1 & -2 & 0 & 16 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

est échelonnée réduite. On reconnaît (à l'œil) les matrices échelonnées à la disposition caractéristique des zéros en escalier descendant du haut à gauche vers le bas à droite. □

Les notions précédentes sont uniquement relatives aux matrices, mais il se trouve que les systèmes linéaires dont la matrice augmentée est échelonnée réduite — appelés systèmes échelonnés réduits pour aller plus vite — sont particulièrement simples à résoudre. Commençons par un exemple.

Exemple 5 Supposons que

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 & 25 \\ 0 & 1 & -2 & 0 & 16 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

soit en fait la matrice augmentée d'un système linéaire. Ce système sera alors 3×4 et s'écrira

$$\begin{cases} x_1 \phantom{{}+x_2} + 2x_3 \phantom{{}+x_4} = 25, \\ \phantom{x_1 +{}} x_2 - 2x_3 \phantom{{}+x_4} = 16, \\ \phantom{x_1 + x_2 - 2x_3 +{}} x_4 = 1. \end{cases}$$

Ce système se résout trivialement en

$$\begin{cases} x_1 = 25 - 2x_3, \\ x_2 = 16 + 2x_3, \\ x_4 = 1. \end{cases}$$

En d'autres termes, pour toute valeur de x_3 réelle, les valeurs de x_1, x_2 et x_4 calculées ci-dessus fournissent une solution du système, et on les a ainsi toutes obtenues. On peut donc décrire entièrement l'ensemble des solutions

$$S = \{(25 - 2x_3, 16 + 2x_3, x_3, 1); x_3 \in \mathbb{R}\}.$$

Il s'agit d'une *représentation paramétrique* de S. On parle encore de *solution générale* du système. □

L'exemple qui précède montre que les inconnues d'un système échelonné réduit ne jouent pas toutes le même rôle. Ceci conduit aux définitions suivantes.

Définition 12 *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 11.*

Ainsi, dans l'exemple 5, on voit trois positions de pivot : $(1,1)$, $(2,2)$ et $(3,4)$. 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 11, 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 rôle important plus loin dans les questions de dimension d'un espace vectoriel.

Les positions de pivot permettent d'introduire une classification des inconnues.

Définition 13 *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, voir le tout premier exemple de cette section. Nous pouvons maintenant résoudre les systèmes échelonnés réduits dans tous les cas.

Théorème 6 *Un système échelonné réduit est compatible si et seulement si sa matrice augmentée ne contient aucune ligne de la forme*

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix} \text{ avec } b \neq 0.$$

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

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix} \text{ avec } b \neq 0.$$

Cette ligne correspond à l'équation linéaire

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

laquelle n'a évidemment aucune solution. Le système est par conséquent incompatible, $S = \emptyset$.

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

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

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

$$x_{i_l} + B_l(x_{\text{libres}}) = b_l,$$

où x_{i_l} est la l-ème variable essentielle (qui n'apparaît que dans cette équation située à la ligne l), $B_l(x_{\text{libres}})$ 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, $B_l(x_{\text{libres}}) = 0$ s'il n'y en a pas, et b_l est la l-ème ligne du second membre. Par conséquent,

$$x_{i_l} = -B_l(x_{\text{libres}}) + b_l,$$

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

On a ainsi établi dans le cas des systèmes échelonnés réduits l'alternative sur l'ensemble des solutions déjà vue géométriquement dans le cas 2×2.

Corollaire 7 *Dans le cas d'un système échelonné réduit $m \times n$ on a l'alternative suivante.*

a) Soit il n'y a aucune solution s'il y a une ligne de la forme

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix} \text{ avec } b \neq 0.$$

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.

2.4 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.

Arrêtons nous quelque peu sur la notion d'algorithme. Il s'agit d'une description précise d'une suite d'opérations à effectuer, dans quel ordre et dans quel cas, qui aboutit au bout d'un nombre fini d'étapes si possible connu à l'avance au résultat voulu. Il y a deux raisons pour introduire un algorithme dans le contexte de la résolution des systèmes linéaires.

La première raison est que l'on peut certes résoudre les systèmes 2×2 ou 3×3 par des manipulations *ad hoc* des équations — résolution par rapport à une variable puis remplacement dans les autres équations, additions ou soustractions d'équations — menées au petit bonheur la chance et qui aboutissent à un résultat après un plus ou moins grand nombre d'opérations. Or l'expérience montre que ces opérations sont le plus souvent inutiles, redondantes, et surtout cause d'erreurs de calculs. Il est bien préférable de se laisser guider par une méthode stricte dont l'application garantit un nombre minimal de calculs (en général).

La seconde raison est que dans les applications pratiques de l'algèbre linéaire, lesquelles sont extrêmement nombreuses et importantes, 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 en tel ou tel langage d'un algorithme.

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

Définition 14 *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-même plus un multiple d'une autre ligne (substitution).

Les opérations ii) et iii) sont à entendre colonne par colonne.

Exemple 6 Considérons la matrice

$$A=\begin{pmatrix} 1 & 2 & 3 & 4\\ 0 & 2 & 4 & 6\\ -1 & 0 & 1 & 0\end{pmatrix}.$$

L'échange des lignes 2 et 3 de A produit la nouvelle matrice

$$A' = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 0 & 1 & 0 \\ 0 & 2 & 4 & 6 \end{pmatrix}.$$

Multiplier la ligne 1 de A par 5 produit la nouvelle matrice

$$A'' = \begin{pmatrix} 5 & 10 & 15 & 20 \\ 0 & 2 & 4 & 6 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

Remplacer une ligne 2 de A par elle-même plus $(-1)\times$ la ligne 1 produit la nouvelle matrice

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

Plus généralement, remplacer la ligne i par elle-même plus $\lambda\times$ la ligne k revient à remplacer dans la colonne j le coefficient a_{ij} par $a_{ij} + \lambda a_{kj}$ pour tous les j de 1 à n.

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. $\square$

Les matrices obtenues après une opération élémentaire ne sont pas égales à la matrice de départ. On introduit donc une nouvelle notion.

Définition 15 *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 15 *Il s'agit d'une relation d'équivalence.*

Démonstration. Cette relation est réflexive. En effet, on a $A \sim A$ puisque A se déduit de A par une suite de zéro opérations élémentaires.

Elle est transitive. En effet, si $A \sim B$ et $B \sim C$, alors on déduit C de A en effectuant d'abord la suite d'opérations élémentaires qui passe de A à B, puis celle qui passe de B à C.

Elle est enfin symétrique. Ce dernier point est un peu plus délicat. Il repose sur le fait que les trois opérations élémentaires sont *inversibles*, c'est-à-dire que l'on peut revenir en arrière par une autre opération élémentaire. Ce fait est évident pour les opérations d'échange et d'homothétie. En effet, il suffit de rééchanger les mêmes lignes dans le cas de l'échange, et de multiplier la ligne par l'inverse de la constante non nulle dans le cas de l'homothétie pour se retrouver dans la configuration de départ. Dans le cas de la substitution, supposons que l'on ait remplacé la ligne i par elle-même plus $\lambda\times$ la ligne k, c'est-à-dire remplacé le coefficient a_{ij} par $a'_{ij} = a_{ij} + \lambda a_{kj}$, $j = 1,\dots,n$. Pour revenir en arrière, il suffit d'effectuer la substitution remplaçant la ligne i par elle-même *moins* $\lambda\times$ la ligne k. En effet, on remplace ainsi a'_{ij} par $a'_{ij} - \lambda a_{kj} = a_{ij}$, $j = 1,\dots,n$. Soient maintenant deux matrices telles que $A \sim B$. On passe de B à A en effectuant les opérations élémentaires inverses de celles qui permettent de passer de A à B dans l'ordre inverse, c'est-à-dire que $B \sim A$. □

La notion d'équivalence de matrices est directement liée à celle d'équivalence des systèmes linéaires de la Définition 9.

Proposition 16 *Si les matrices augmentées de deux systèmes linéaires sont équivalentes, alors les systèmes linéaires sont équivalents.*

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

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = a_{i,n+1}.$$

Multipliant les deux membres par λ, on voit que

$$a'_{i1}s_1 + a'_{i2}s_2 + \cdots + a'_{in}s_n = a'_{i,n+1},$$

et comme les autres équations du système associé à A' sont les mêmes que celles de A, on en déduit que $(s_1, s_2, \dots, s_n) \in S_{A'}$. 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 : $a'_{ij} = a_{ij} + \lambda a_{kj}$ pour un certain i, un certain k et tous $j = 1,\dots,n+1$. Soit $(s_1, s_2, \dots, s_n) \in S_A$. Ce n-uplet vérifie en particulier les équations numéros i et k

$$\begin{aligned} a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n &= a_{i,n+1} \\ a_{k1}s_1 + a_{k2}s_2 + \cdots + a_{kn}s_n &= a_{k,n+1} \end{aligned}$$

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

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n + \lambda(a_{k1}s_1 + a_{k2}s_2 + \cdots + a_{kn}s_n) = a_{i,n+1} + \lambda a_{k,n+1}.$$

On factorise le membre de gauche

$$(a_{i1} + \lambda a_{k1})s_1 + (a_{i2} + \lambda a_{k2})s_2 + \cdots + (a_{in} + \lambda a_{kn})s_n = a_{i,n+1} + \lambda a_{k,n+1},$$

qui n'est autre que

$$a'_{i1}s_1 + a'_{i2}s_2 + \cdots + a'_{in}s_n = a'_{i,n+1}.$$

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

Démonstration. 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).

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.

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_{11} on passe directement à l'étape 2, si c'est un terme a_{i1} 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

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} = A$$

dans le premier cas, ou bien

$$\begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2j} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{i1} & a'_{i2} & \cdots & a'_{ij} & \cdots & a'_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mj} & \cdots & a'_{mn} \end{pmatrix} \sim A$$

avec $a'_{11} \neq 0$ dans le deuxième cas.

Étape 2 : Élimination. On ne touche plus à la ligne 1, et on se sert du pivot pour éliminer tous les termes a'_{i1}, $i \geq 2$. Pour cela, il suffit de remplacer la ligne i par elle-même moins $\frac{a'_{i1}}{a'_{11}} \times$ la ligne 1, ceci pour $i = 2, \ldots, m$.

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

$$\begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1j} & \cdots & a'_{1n} \\ 0 & a''_{22} & \cdots & a''_{2j} & \cdots & a''_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{i2} & \cdots & a''_{ij} & \cdots & a''_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a''_{m2} & \cdots & a''_{mj} & \cdots & a''_{mn} \end{pmatrix} \sim A$$

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

$$\begin{pmatrix} a^1_{11} & a^1_{12} & \cdots & a^1_{1j} & \cdots & a^1_{1n} \\ 0 & a^1_{22} & \cdots & a^1_{2j} & \cdots & a^1_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a^1_{i2} & \cdots & a^1_{ij} & \cdots & a^1_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & a^1_{m2} & \cdots & a^1_{mj} & \cdots & a^1_{mn} \end{pmatrix} \sim A$$

dont la première colonne est bien celle d'une matrice échelonnée. On va donc conserver cette première colonne, ainsi que 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

$$\begin{pmatrix} a^1_{22} & \cdots & a^1_{2j} & \cdots & a^1_{2n} \\ \vdots & & \vdots & & \vdots \\ a^1_{i2} & \cdots & a^1_{ij} & \cdots & a^1_{in} \\ \vdots & & \vdots & & \vdots \\ a^1_{m2} & \cdots & a^1_{mj} & \cdots & a^1_{mn} \end{pmatrix}.$$

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

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1j}^1 & \cdots & a_{1n}^1 \\ 0 & a_{22}^2 & \cdots & a_{2j}^2 & \cdots & a_{2n}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{ij}^2 & \cdots & a_{in}^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{mj}^2 & \cdots & a_{mn}^2 \end{pmatrix} \sim A$$

et ainsi de suite.

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

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.

Voir plus loin un exemple de l'algorithme de Gauss en action.

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.

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

pour un certain $k \leq m$.

Cas 2 : la première colonne différente est une colonne de pivot.

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

De même pour $\tilde{U}'$, on obtient

Cas 1 : la première colonne différente n'est pas une colonne de pivot.

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

Cas 2 : la première colonne différente est une colonne de pivot.

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

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 partis de l'hypothèse $\tilde{U} \neq \tilde{U}'$. Les deux systèmes sont donc compatibles. Or ce sont des systèmes sans variable libre et l'on a donc $S_{\tilde{U}} = \{(u_1, u_2, \ldots, u_k)\} = S_{\tilde{U}'} = \{(u'_1, u'_2, \ldots, u'_k)\}$. On en déduit dans ce cas aussi que $\tilde{U} = \tilde{U}'$, ce qui est également impossible. Contradiction. □

Remarque 8 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. □

Exemple 7 Soit

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

Passage à une forme échelonnée.

Première itération de la boucle, étape 1. Le choix du pivot est tout fait, on garde $a_{11} = 1$.

Première itération de la boucle, étape 2. On remplace la ligne 2 par elle-même moins $0\times$ la ligne 1 (c'est-à-dire qu'on ne fait rien sur cette ligne qui contient déjà un zéro en bonne position) et la ligne 3 par elle-même moins $(-1)\times$ la ligne 1. On obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 2 & 4 & 4 \end{pmatrix}.$$

Deuxième itération de la boucle, étape 1. Le choix du pivot est tout fait, on garde $a^1_{22} = 2$.

Deuxième itération de la boucle, étape 2. On remplace la ligne 3 par elle-même moins $(2/2)\times$ la ligne 1. On obtient

$$A \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

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 $1/2$ et la ligne 3 par $-1/2$ et l'on obtient

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

Étape 2, première itération. On ne touche plus à la ligne 3 et on remplace la ligne 2 par elle-même moins $3\times$ la ligne 3 et la ligne 1 par elle-même moins $4\times$ la ligne 3. On obtient

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

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

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

qui est bien échelonnée réduite.

Remarquons au passage que si A est considérée comme la matrice augmentée d'un système 3×3, alors ce système est incompatible. $\square$

Le théorème 8 permet d'étendre un certain nombre de définitions aux matrices quelconques.

Définition 16 *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 7, on voit trois positions de pivot : $(1,1)$, $(2,2)$ et $(3,4)$ 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 9 *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*

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix} \text{ avec } b \neq 0.$$

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.

De même,

Corollaire 10 *Soit un système linéaire $m \times n$ quelconque, A sa matrice augmentée et U l'unique matrice échelonnée réduite équivalente à A. On a l'alternative suivante.*

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

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix} \text{ avec } b \neq 0.$$

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 9 Nulle part dans les raisonnements qui précèdent on n'a utilisé le fait que les nombres soient des nombres réels. On a seulement utilisé les propriétés de l'addition, de la multiplication et de la division par un nombre non nul, à savoir les axiomes de *corps commutatif*. Par conséquent, tout ce que l'on a dit reste vrai pour des systèmes linéaires $m \times n$ à coefficients dans un corps commutatif quelconque, à résoudre dans ce même corps. Ainsi, on peut résoudre sans difficulté des systèmes à coefficients dans $\mathbb{C}$. $\square$

Remarque 10 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. $\square$

Chapitre 3

Vecteurs dans $\mathbb{R}^m$

Dans ce chapitre, nous allons nous familiariser avec la notion de vecteur du point de vue algébrique. Nous reviendrons le moment venu sur le point de vue géométrique

3.1 Définition et opérations sur les vecteurs

Définition 17 *Soit $m \geq 1$ un entier. Un* vecteur *de $\mathbb{R}^m$ est une matrice à m lignes et 1 colonne. Un* scalaire *est un nombre réel (par opposition à un vecteur).*

On parle aussi de vecteur-colonne. On va également employer une notation provisoire avec des petites flèches, que l'on abandonnera dès que l'on aura acquis les notions de base qui suivent. Ainsi

$$\vec{u} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 0{,}2 \\ \sqrt{5} \end{pmatrix}$$

sont des vecteurs de $\mathbb{R}^2$ et

$$\vec{w} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

est un vecteur de $\mathbb{R}^4$.

Comme un vecteur n'a qu'une seule colonne, il est inutile de conserver la notation à double indice des matrices : un seul indice suffit. La règle d'égalité entre matrices s'applique bien sûr au cas particulier des vecteurs. Ainsi

$$\text{si } \vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \text{ alors } \vec{u} = \vec{v} \text{ si et seulement si } u_1 = v_1 \text{ et } u_2 = v_2.$$

En d'autres termes, les lignes d'un vecteur sont ordonnées dans l'ordre de leur indice. Un vecteur générique de $\mathbb{R}^m$ s'écrit normalement

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}.$$

Cette notation naturelle en colonne a le désavantage de gaspiller l'espace sur une page écrite quand elle utilisée comme ici de façon isolée. C'est pourquoi on utilise occasionnellement une notation en ligne

$$\vec{u} = (u_1, u_2, \dots, u_m)$$

pour désigner le même objet. C'est en fait la notation d'un m-uplet de scalaires et il faut éviter de la confondre avec la notation

$$\begin{pmatrix} u_1 & u_2 & \cdots & u_m \end{pmatrix}$$

qui désigne une matrice à 1 ligne et m colonnes (aussi appelée vecteur-ligne). Ce n'est pas du tout la même chose !

Nous allons maintenant introduire les deux opérations fondamentales sur les vecteurs.

Définition 18 *On définit sur $\mathbb{R}^m$ une opération interne appelée* addition *par*

$$\forall \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \in \mathbb{R}^m, \forall \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} \in \mathbb{R}^m, \quad \vec{u} + \vec{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_m + v_m \end{pmatrix} \in \mathbb{R}^m$$

et une opération externe appelée multiplication par un scalaire *par*

$$\forall \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \in \mathbb{R}^m, \forall \lambda \in \mathbb{R}, \quad \lambda \vec{u} = \begin{pmatrix} \lambda u_1 \\ \lambda u_2 \\ \vdots \\ \lambda u_m \end{pmatrix} \in \mathbb{R}^m.$$

Ces deux opérations sont donc définies ligne par ligne.

Exemple 8 On a

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 1+2 \\ -2+5 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \quad 2\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \times 1 \\ 2 \times (-2) \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \end{pmatrix}, \text{ etc.}$$

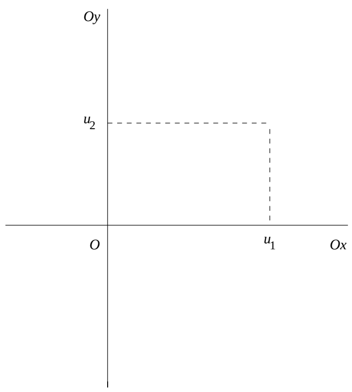
On peut combiner ces opérations. Ainsi, si $\vec{u} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ alors

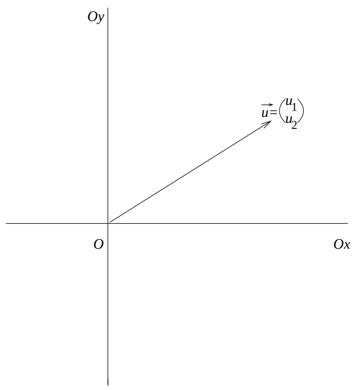
$$2\vec{u}+(-3)\vec{v}=\begin{pmatrix} 2\times 1 \\ 2\times(-2) \end{pmatrix}+\begin{pmatrix} (-3)\times 2 \\ (-3)\times 5 \end{pmatrix}=\begin{pmatrix} 2 \\ -4 \end{pmatrix}+\begin{pmatrix} -6 \\ -15 \end{pmatrix}=\begin{pmatrix} 2-6 \\ -4-15 \end{pmatrix}=\begin{pmatrix} -4 \\ -19 \end{pmatrix},$$

et ainsi de suite. □

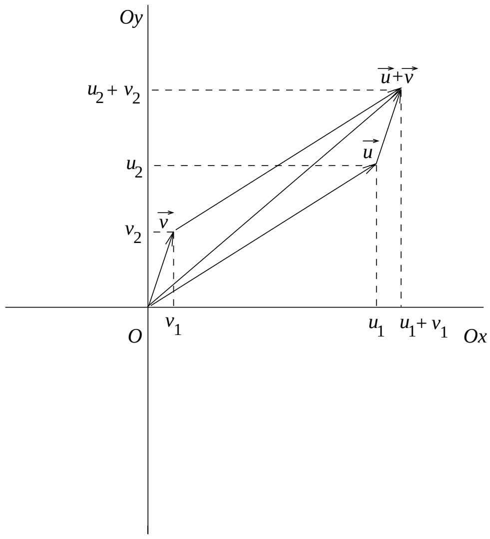
3.2 Interprétation géométrique dans le plan et dans l'espace

On sait que la donnée d'une origine O et de deux axes de coordonnées Ox, Oy munis d'une unité de longueur dans le plan permet de repérer tout point de ce plan par ses deux coordonnées, un couple de nombre réels. Les vecteurs de $\mathbb{R}^2$ étant aussi définis par un couple de nombre réels, on peut se servir de cette correspondance pour représenter géométriquement les vecteurs de $\mathbb{R}^2$. Plus précisément, un vecteur $\vec{u} = (u_1, u_2)$ de $\mathbb{R}^2$ sera représenté par le segment orienté d'origine O et d'extrémité le point de coordonnées (u_1, u_2) dans le repère cartésien choisi. On le dessine traditionnellement à l'aide d'une flèche. De cette façon, les vecteurs — concept algébrique — admettent une représentation géométrique extrêmement utile pour comprendre et avoir l'intuition de comment ils fonctionnent. Par ailleurs, cette correspondance ouvre la voie à un champ d'application de l'algèbre linéaire à la géométrie plane (et dans l'espace), même si nous ne privilégierons pas ce point de vue, néanmoins très important.

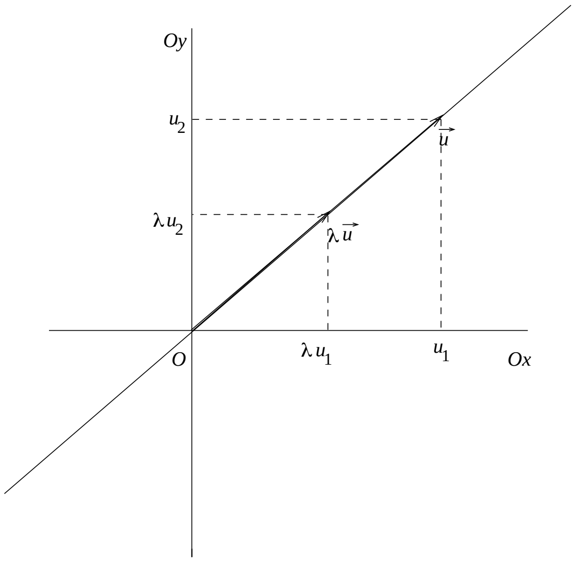




L'addition de deux vecteurs se traduit géométriquement par la *règle du parallélogramme* : la somme de deux vecteurs est représentée par la diagonale du parallélogramme ayant ces deux vecteurs comme côtés.



Par le théorème de Thalès, l'ensemble des multiples scalaires d'un vecteur dont la représentation n'est pas réduite au point O est la droite qui supporte le segment représentant le vecteur. Cette droite passe naturellement par O.



La situation est identique dans l'espace à trois dimensions : celui-ci sert à représenter l'ensemble des vecteurs de $\mathbb{R}^3$, une fois choisi un repère cartésien pour l'espace. Les interprétations géométriques de l'addition et de la multiplication par un scalaire sont identiques au cas plan.

3.3 Structure algébrique de $(\mathbb{R}^m, +, \cdot)$

Une fois muni de ses deux opérations interne et externe, l'ensemble $\mathbb{R}^m$ acquiert une structure algébrique, c'est-à-dire des règles de calcul sur les vecteurs. Un vecteur joue un rôle particulier

Définition 19 *Le vecteur*

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^m$$

est appelé le vecteur nul *de* $\mathbb{R}^m$.

La proposition suivante détaille la structure algébrique de $(\mathbb{R}^m, +, \cdot)$.

Proposition 17 *On a les huit propriétés suivantes. Pour tous* $\vec{u}, \vec{v}, \vec{w}$ *dans* $\mathbb{R}^m$ *et tous* λ, μ *dans* $\mathbb{R}$,

i) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
(commutativité de l'addition).

ii) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
(associativité de l'addition).

iii) $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
($\vec{0}$ est un élément neutre pour l'addition).

iv) $\vec{u} + (-1)\vec{u} = (-1)\vec{u} + \vec{u} = \vec{0}$
(tout élément admet un opposé pour l'addition).

v) $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$
(distributivité de la multiplication par un scalaire par rapport à l'addition vectorielle).

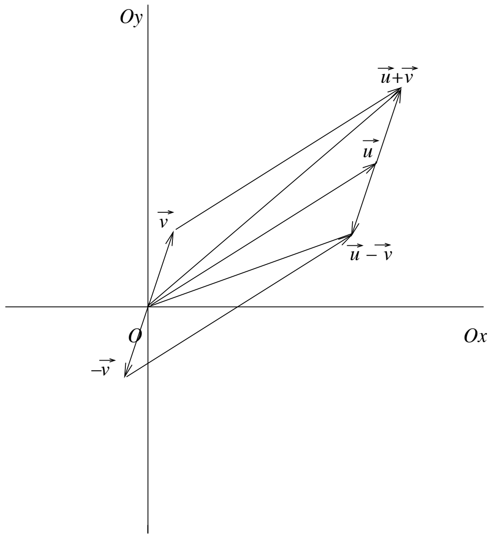
vi) $(\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}$
(distributivité de la multiplication par un scalaire par rapport à l'addition scalaire).

vii) $\lambda(\mu\vec{u}) = (\lambda\mu)\vec{u}$
(« associativité » de la multiplication externe).

viii) $1\vec{u} = \vec{u}$
(1 est un « élément neutre » pour la multiplication externe).

Démonstration. Il s'agit d'une vérification de routine élémentaire, à effectuer quand même une fois dans sa vie. □

Notation On utilisera les notations $-\vec{u} = (-1)\vec{u}$ pour l'opposé d'un vecteur et $\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$ pour la soustraction de deux vecteurs, exactement comme dans le cas scalaire. □



Opposé et soustraction dans $\mathbb{R}^2$.

Remarque 11 Les propriétés ii), iii) et iv) ne sont autres que les axiomes de la structure de groupe. Avec la propriété i), on voit que $(\mathbb{R}^m, +)$ a une structure de groupe commutatif ou abélien. L'ensemble des huit propriétés algébriques ci-dessus confère à $(\mathbb{R}^m, +, \cdot)$ ce que l'on appelle une structure d'*espace vectoriel sur* $\mathbb{R}$. (le point $\cdot$ désigne la multiplication externe). Par ailleurs, on peut maintenant mener toutes sortes de calculs avec l'addition et la multiplication par un scalaire en appliquant les règles usuelles. Ainsi

$$(\lambda + \mu)(\vec{u} + \vec{v}) = (\lambda + \mu)\vec{u} + (\lambda + \mu)\vec{v} = \lambda\vec{u} + \mu\vec{u} + \lambda\vec{v} + \mu\vec{v} = \lambda(\vec{u} + \vec{v}) + \mu(\vec{u} + \vec{v}).$$

Attention quand même à ne pas additionner des scalaires et des vecteurs, et à ne pas multiplier deux vecteurs entre eux. De telles opérations n'ont aucun sens et n'en auront jamais pour la première d'entre elles. □

Proposition 18 *On a, pour tout $\vec{u}$ dans $\mathbb{R}^m$ et tout λ dans $\mathbb{R}$,*

$$0\vec{u} = \vec{0}, \qquad \lambda\vec{0} = \vec{0}.$$

Démonstration. Vérification immédiate à partir des définitions. □

À partir des deux opérations de base appliquées un nombre fini de fois et en utilisant les propriétés i) à viii), on aboutit à une notion fondamentale.

Définition 20 *Soient $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ une famille de n vecteurs de $\mathbb{R}^m$ et $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ une famille de n scalaires. Alors*

$$\vec{y} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_n \vec{v}_n$$

s'appelle une combinaison linéaire *de la famille $\{\vec{v}_j\}_{j=1,\dots,n}$ dont les* coefficients *sont les $\{\lambda_j\}_{j=1,\dots,n}$.*

Les combinaisons linéaires permettent déjà de faire le lien avec les systèmes linéaires. Considérons l'exemple suivant. On se donne trois vecteurs de $\mathbb{R}^3$

$$\vec{a}_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} \quad \text{et} \quad \vec{b} = \begin{pmatrix} 7 \\ 4 \\ -3 \end{pmatrix}$$

et l'on se pose la question de savoir si $\vec{b}$ est combinaison linéaire de la famille $\{\vec{v}_1, \vec{v}_2\}$ ou non. Cela n'a rien d'évident à l'œil nu. Cette question revient à se demander s'il existe deux scalaires x_1 et x_2 tels que $\vec{b} = x_1\vec{a}_1 + x_2\vec{a}_2$. Il s'agit d'une équation vectorielle. Si on l'explicite, il vient

$$x_1 \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ -3 \end{pmatrix},$$

c'est-à-dire

$$\begin{pmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ -3 \end{pmatrix}.$$

D'après la règle d'égalité de deux vecteurs, ceci n'est rien d'autre que le système linéaire 3×2

$$\begin{cases} x_1 & +2x_2 & = & 7, \\ -2x_1 & +5x_2 & = & 4, \\ -5x_1 & +6x_2 & = & -3, \end{cases}$$

et la question est de décider si ce système est compatible ou non. C'est une question à laquelle nous savons répondre grâce à la méthode de Gauss vue au premier chapitre (la réponse est ici oui, avec $x_1 = 3$ et $x_2 = 2$).

Notons que la matrice augmentée de ce système est la matrice

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

Ses colonnes ne sont autres que les vecteurs $\vec{a}_1$, $\vec{a}_2$ et $\vec{b}$ mis côte-à-côte. Ceci amène à introduire la nouvelle notation

$$\begin{pmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{pmatrix} = \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \vec{b} \end{pmatrix}$$

pour une matrice 3×3 vue comme une ligne de 3 vecteurs de $\mathbb{R}^3$. Il s'agit en fait d'une remarque générale.

Proposition 19 *Soient $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \vec{b}$, $n+1$ vecteurs de $\mathbb{R}^m$. Le problème de savoir si $\vec{b}$ est une combinaison linéaire de la famille $\{\vec{v}_j\}_{j=1,\dots,n}$ est équivalent au problème de savoir si l'équation vectorielle*

$$\vec{b} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$$

admet au moins une solution, ce problème étant lui-même équivalent à décider si le système linéaire dont la matrice augmentée $m \times (n+1)$ est $\begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b} \end{pmatrix}$ est compatible ou non.

Démonstration. En effet, si

$$\vec{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{et} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad \text{on a} \quad x_j\vec{a}_j = \begin{pmatrix} x_j a_{1j} \\ x_j a_{2j} \\ \vdots \\ x_j a_{mj} \end{pmatrix},$$

d'où le résultat. □

3.4 Espace engendré par une famille de vecteurs

Définition 21 *Soit $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ une famille de n vecteurs de $\mathbb{R}^m$. L'ensemble des toutes les combinaisons linéaires possibles de la famille est appelé* espace engendré *par cette famille*

$$\text{vect}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} = \{\vec{y} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n \in \mathbb{R}^m, \{\lambda_j\}_{j=1,\dots,n} \in \mathbb{R}^n\}.$$

C'est toujours un sous-ensemble de $\mathbb{R}^m$. Par convention, on posera $\text{vect}(\emptyset) = \{\vec{0}\}$.

Remarque 12 Quelques remarques triviales en vrac.

i) $\vec{0} = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_n \in \text{vect}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ dans tous les cas.

ii) Le système linéaire dont la matrice augmentée est $\begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b} \end{pmatrix}$ est compatible si et seulement si $\vec{b} \in \text{vect}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$.

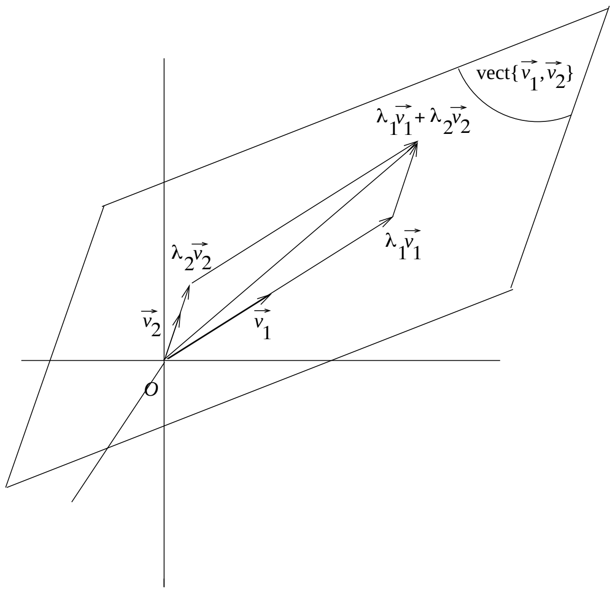
iii) $\forall j$, $\forall \lambda$, $\lambda\vec{v}_j \in \text{vect}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$.

iv) $\text{vect}\{\vec{0}\} = \{\vec{0}\}$. □

Interprétons géométriquement l'espace engendré dans le cas d'un petit nombre de vecteurs de $\mathbb{R}^3$.

Soit $\vec{v} \neq \vec{0}$. Dans ce cas, $\text{vect}\{\vec{v}\} = \{\lambda\vec{v}, \lambda \in \mathbb{R}\}$ est l'ensemble des multiples scalaires de $\vec{v}$ et on a déjà vu que cet ensemble est représenté par la droite qui passe par O et qui s'appuie sur le segment représentant $\vec{v}$.

Soient maintenant $\vec{v}_1 \neq \vec{0}$ et $\vec{v}_2 \neq \vec{0}$ tels que $\vec{v}_2 \notin \text{vect}\{\vec{v}_1\}$, c'est-à-dire n'appartient pas à la droite engendrée par $\vec{v}_1$. Dans ce cas, $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2$ est représenté par la diagonale du parallélogramme construit sur $\lambda_1\vec{v}_1$ et $\lambda_2\vec{v}_2$, et donc balaie le plan passant par O et contenant $\vec{v}_1$ et $\vec{v}_2$ quand λ_1 et λ_2 varient. L'espace engendré est donc représenté par ce plan.



Naturellement, si on a $\vec{v}_2 \in \text{vect}\{\vec{v}_1\}$, alors $\vec{v}_2 = \mu\vec{v}_1$ pour un certain μ et $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 = (\lambda_1 + \lambda_2\mu)\vec{v}_1 \in \text{vect}\{\vec{v}_1\}$. On a donc dans ce cas un espace engendré représenté par une droite.

Nous arrivons à une des définitions les plus importantes.

Définition 22 *On dit qu'une famille* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *de* $\mathbb{R}^m$ *est* génératrice *(ou engendre* $\mathbb{R}^m$*) si*

$$\text{vect}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \mathbb{R}^m.$$

En d'autres termes, si une famille est génératrice, alors *tout* vecteur de $\mathbb{R}^m$ peut s'écrire comme combinaison linéaire des vecteurs de cette famille.

3.5 Produit matrice-vecteur

Il s'agit d'un outil de calcul fondamental.

Définition 23 *Soit A une matrice $m \times n$, de colonnes $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{R}^m$, et soit $\vec{x} \in \mathbb{R}^n$. Le* produit *de A par $\vec{x}$ est le vecteur*

$$A\vec{x} = \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

Le produit $A\vec{x}$ n'est donc que la combinaison linéaire des vecteurs-colonne de A dont les coefficients sont les lignes de $\vec{x}$. C'est par définition un vecteur de $\mathbb{R}^m$ et il n'est visiblement défini que si $\vec{x}$ *a le même nombre de lignes que le nombre de colonnes de A.*

Exemple 9 Voici un exemple pour $m = 2$, $n = 3$,

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 2 \\ -5 \end{pmatrix} + 7\begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \in \mathbb{R}^2.$$

Par contre, le produit

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

n'est pas défini. □

Dans la pratique, on procède au calcul comme suit, ligne par ligne. On commence par poser l'opération

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

Puis on effectue le calcul de la première ligne en groupant les termes qui vont être multipliés ensemble. Ainsi on groupe

$$\begin{pmatrix} \boxed{1} & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} \boxed{4} \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} ? \\ ? \end{pmatrix},$$

puis

$$\begin{pmatrix} 1 & \boxed{2} & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ \boxed{3} \\ 7 \end{pmatrix} = \begin{pmatrix} ? \\ ? \end{pmatrix},$$

et enfin

$$\begin{pmatrix} 1 & 2 & \boxed{-1} \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ \boxed{7} \end{pmatrix} = \begin{pmatrix} ? \\ ? \end{pmatrix}.$$

On obtient la première ligne en sommant les produits obtenus

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \times 4 + 2 \times 3 + (-1) \times 7 \\ ? \end{pmatrix} = \begin{pmatrix} 3 \\ ? \end{pmatrix}.$$

On recommence le même procédé pour les lignes suivantes.

$$\begin{pmatrix} 1 & 2 & -1 \\ \boxed{0} & -5 & 3 \end{pmatrix} \begin{pmatrix} \boxed{4} \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ ? \end{pmatrix},$$

puis

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & \boxed{-5} & 3 \end{pmatrix} \begin{pmatrix} 4 \\ \boxed{3} \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ ? \end{pmatrix},$$

et enfin

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & \boxed{3} \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ \boxed{7} \end{pmatrix} = \begin{pmatrix} 3 \\ ? \end{pmatrix},$$

d'où

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \times 4 + (-5) \times 3 + 3 \times 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

Avec un peu d'habitude, cette technique devient très rapide (on n'écrit plus toutes les étapes, bien sûr). □

Une matrice joue un rôle particulier vis-à-vis du produit matrice-vecteur.

Proposition 20 *La matrice $n \times n$*

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

est telle que $I_n\vec{x} = \vec{x}$ pour tout $\vec{x} \in \mathbb{R}^n$. On l'appelle la matrice identité *$n \times n$.*

Démonstration. Les vecteurs colonne de I_n sont les vecteurs

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

en fait $\vec{e}_j$ ne contient que des zéros sauf sur la ligne j qui contient 1. Il est donc assez clair que

$$x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_j\vec{e}_j + \cdots + x_n\vec{e}_n = \vec{x},$$

pour tout $\vec{x} \in \mathbb{R}^n$. □

Intéressons-nous maintenant aux propriétés algébriques du produit matrice-vecteur.

Proposition 21 *Soit A une matrice $m \times n$, $\vec{u}$ et $\vec{v}$ deux vecteurs de $\mathbb{R}^n$ et λ un scalaire. On a*

i) $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$.

ii) $A(\lambda\vec{u}) = \lambda(A\vec{u})$.

Les propriétés i) et ii) portent le nom de linéarité.

Démonstration. Écrivons A comme une ligne de vecteurs, et $\vec{u}$ et $\vec{v}$ comme des colonnes de scalaires :

$$A = \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{pmatrix}, \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Comme

$$\vec{u} + \vec{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix},$$

on en déduit que

$$\begin{aligned} A(\vec{u}+\vec{v}) &= (u_1+v_1)\vec{a}_1 + (u_2+v_2)\vec{a}_2 + \cdots + (u_n+v_n)\vec{a}_n \\ &= u_1\vec{a}_1 + u_2\vec{a}_2 + \cdots + u_n\vec{a}_n + v_1\vec{a}_1 + v_2\vec{a}_2 + \cdots + v_n\vec{a}_n \\ &= A\vec{u} + A\vec{v}. \end{aligned}$$

La démonstration du ii) est en tout point analogue. □

Il faut remarquer que le signe $+$ dans le membre de gauche de i) désigne l'addition dans $\mathbb{R}^n$, alors que le signe $+$ dans le membre de droite désigne l'addition dans $\mathbb{R}^m$.

Corollaire 11 *On a*

$$A(\lambda_1\vec{u}_1+\lambda_2\vec{u}_2+\cdots+\lambda_k\vec{u}_k)=\lambda_1A\vec{u}_1+\lambda_2A\vec{u}_2+\cdots+\lambda_kA\vec{u}_k.$$

Démonstration. Immédiate à partir de la proposition précédente. □

Le produit matrice-vecteur transforme donc toute combinaison linéaire dans $\mathbb{R}^n$ en la combinaison linéaire dans $\mathbb{R}^m$ avec les mêmes coefficients.

Mentionnons pour clore cette section la formule générale donnant la i-ème ligne d'un produit matrice-vecteur,

$$\begin{aligned}(Ax)_i &= a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\\ &= \textstyle\sum_{j=1}^n a_{ij}x_j.\end{aligned}$$

Remarquer sur cette dernière formule concise la sommation par rapport à l'indice répété j.

3.6 Application aux systèmes linéaires

On va réinterpréter les systèmes linéaires à l'aide de produits matrice-vecteur. Commençons par un exemple. Soit le système linéaire 2×3

$$\begin{cases} x_1 & +2x_2 & -x_3 & = & 3,\\ & -5x_2 & +x_3 & = & 6.\end{cases}$$

Le système est clairement équivalent à l'équation vectorielle

$$\begin{pmatrix} x_1+2x_2-x_3\\ -5x_2+x_3\end{pmatrix}=\begin{pmatrix}3\\6\end{pmatrix}.$$

Cette équation se réécrit encore

$$x_1\begin{pmatrix}1\\0\end{pmatrix}+x_2\begin{pmatrix}2\\-5\end{pmatrix}+x_3\begin{pmatrix}-1\\1\end{pmatrix}=\begin{pmatrix}3\\6\end{pmatrix},$$

forme sous laquelle on reconnaît un produit matrice-vecteur

$$\begin{pmatrix}1 & 2 & -1\\ 0 & -5 & 3\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}3\\6\end{pmatrix}.$$

Posant

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{pmatrix}, \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \vec{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix},$$

on a réécrit le système linéaire sous la forme concise suivante

$$A\vec{x} = \vec{b}.$$

Bien sûr, A n'est autre que la matrice du système et $\vec{b}$ le second membre (écrit sous forme vectorielle).

Proposition 22 *Tout système linéaire de m équations à n inconnues dont la matrice est A et le second membre $\vec{b} \in \mathbb{R}^m$ s'écrit de façon équivalente*

$$A\vec{x} = \vec{b} \quad ou \quad x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}.$$

La démonstration est identique à ce que l'on a fait dans l'exemple, c'est un exercice simple. On en déduit le théorème suivant.

Théorème 12 *Soit A une matrice $m \times n$. Les trois assertions suivantes sont équivalentes.*
i) Pour tout $\vec{b} \in \mathbb{R}^m$, le système $A\vec{x} = \vec{b}$ est compatible.
ii) Les vecteurs-colonne de A forment une famille génératrice de $\mathbb{R}^m$.
iii) La matrice A admet une position de pivot par ligne.

Démonstration. Dire que $A\vec{x} = \vec{b}$ est compatible pour tout $\vec{b} \in \mathbb{R}^m$ est dire que tout vecteur de $\mathbb{R}^m$ s'écrit comme combinaison linéaire des vecteurs-colonne de A, les coefficients étant les solutions du système linéaire. En d'autres termes i) $\Leftrightarrow$ ii).

L'assertion iii) signifie que l'unique matrice échelonnée réduite U équivalente à A ne contient aucune ligne nulle (puisque chaque ligne contient un pivot). Pour tout second membre, la matrice échelonnée réduite équivalente à la matrice augmentée ne peut donc pas contenir de ligne de la forme $\begin{pmatrix} 0 & 0 & \cdots & 0 & b \end{pmatrix}$ avec $b \neq 0$, et le système est donc compatible.

Réciproquement, si A admet une ligne sans pivot, alors la dernière ligne de U est nulle. Il suffit donc de mettre un 1 au second membre dans cette ligne (et 0 ailleurs) pour construire un second membre $\vec{b}$ pour lequel le système n'est pas compatible. La construction explicite du second membre demande de reproduire dans l'ordre inverse la suite des opérations élémentaires inverses de celles qui passent de A à U. On vient de montrer que i) $\Leftrightarrow$ iii). On en déduit immédiatement que ii) $\Leftrightarrow$ iii). $\square$

Corollaire 13 *Une famille $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ de $\mathbb{R}^m$ est génératrice si et seulement si la matrice $\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}$ admet une position de pivot par ligne.*

3.7 Indépendance linéaire

Cette définition est aussi importante que la définition 22 de famille génératrice.

Définition 24 *Une famille* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *de* $\mathbb{R}^m$ *est dite* linéairement indépendante *ou* libre *si toute combinaison linéaire nulle*

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n = \vec{0}$$

est telle que tous ses coefficients sont nuls $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. *Dans le cas contraire, c'est-à-dire s'il existe une combinaison linéaire nulle à coefficients non tous nuls, on dit que la famille est* linéairement dépendante *ou* liée. *Une telle combinaison linéaire s'appelle alors une* relation de dépendance linéaire *entre les* $\vec{v}_j$.

Par convention, on posera que l'ensemble vide est une famille libre.

Théorème 14 *Soit* $A = \begin{pmatrix}\vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n\end{pmatrix}$. *Les trois assertions suivantes sont équivalentes.*

i) La famille $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *est libre.*
ii) Le système homogène $A\vec{x} = \vec{0}$ *n'admet que la solution triviale.*
iii) La matrice A admet une position de pivot par colonne.

Démonstration. Le seul point non trivial consiste à remarquer que le système homogène admet une solution unique si et seulement si il n'y a pas de variable libre, c'est-à-dire si et seulement si la matrice admet une position de pivot dans chacune de ses colonnes. □

Exemple 10 Considérons la famille

$$\left\{\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix}\right\}.$$

On souhaite déterminer si elle est libre ou liée. Pour cela, il suffit d'effectuer la réduction de Gauss sur la matrice associée.

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

Il y a une variable libre correspondant à la troisième colonne sans pivot, donc la famille est liée. on obtient toutes les relations de dépendance linéaire en résolvant le système homogène, ce qui est immédiat à partir de la forme échelonnée réduite ci-dessus, $x_1 = 2x_3$, $x_2 = -x_3$. On a donc

$$2x_3 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - x_3 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \vec{0}$$

pour tout $x_3 \in \mathbb{R}$ et il n'y a pas d'autre relation de dépendance linéaire. □

Considérons le cas particulier des familles de un ou deux vecteurs.

Proposition 23 *i) La famille $\{\vec{v}\}$ est linéairement indépendante si $\vec{v} \neq \vec{0}$ et linéairement dépendante si $\vec{v} = \vec{0}$.*

ii) La famille $\{\vec{v}_1, \vec{v}_2\}$ est linéairement indépendante si et seulement si $\vec{v}_1$ n'est pas un multiple de $\vec{v}_2$ et $\vec{v}_2$ n'est pas un multiple de $\vec{v}_1$.

Démonstration. Le point i) est trivial. Pour le point ii), supposons d'abord la famille liée. Il existe donc λ_1, λ_2 non tous les deux nuls tels que $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 = \vec{0}$. Si c'est λ_1 qui n'est pas nul, on peut diviser par λ_1, ce qui donne $\vec{v}_1 = -\frac{\lambda_2}{\lambda_1}\vec{v}_2$ et $\vec{v}_1$ est un multiple de $\vec{v}_2$. Si c'est λ_2 qui n'est pas nul, alors de même $\vec{v}_2$ est un multiple de $\vec{v}_1$. On vient de montrer que si la famille est liée, alors $\vec{v}_1$ est un multiple de $\vec{v}_2$ *ou* $\vec{v}_2$ est un multiple de $\vec{v}_1$, ce qui est la négation logique de l'assertion « $\vec{v}_1$ n'est pas un multiple de $\vec{v}_2$ *et* $\vec{v}_2$ n'est pas un multiple de $\vec{v}_1$ ».

Réciproquement, si $\vec{v}_1$ est un multiple de $\vec{v}_2$, alors il existe un scalaire μ tel que $\vec{v}_1 = \mu\vec{v}_2$, soit $1\vec{v}_1 + (-\mu)\vec{v}_2 = \vec{0}$ ce qui est une relation de dépendance linéaire entre $\vec{v}_1$ et $\vec{v}_2$ puisque $1 \neq 0$. De même, $\vec{v}_2$ est un multiple de $\vec{v}_1$, alors la famille est liée, d'où la réciproque. □

Généralisons tout de suite le point ii) à une famille d'un nombre quelconque de vecteurs.

Théorème 15 *Une famille $\mathcal{F} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ de $n \geq 2$ vecteurs de $\mathbb{R}^m$ est linéairement dépendante si et seulement si au moins un des vecteurs de $\mathcal{F}$ est combinaison linéaire des autres vecteurs de $\mathcal{F}$.*

Démonstration. C'est essentiellement la même démonstration que ci-dessus. Supposons d'abord $\mathcal{F}$ liée. Il existe donc une relation de dépendance linéaire

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n = \vec{0},$$

avec $\lambda_k \neq 0$ pour au moins un indice k. Passons tous les autres termes à droite du signe égal. Il vient

$$\lambda_k\vec{v}_k = -\lambda_1\vec{v}_1 - \lambda_2\vec{v}_2 - \cdots - \lambda_n\vec{v}_n,$$

où $\vec{v}_k$ ne figure pas au second membre. Comme $\lambda_k \neq 0$, on peut diviser cette égalité par λ_k et l'on obtient

$$\vec{v}_k = -\frac{\lambda_1}{\lambda_k}\vec{v}_1 - \frac{\lambda_2}{\lambda_k}\vec{v}_2 - \cdots - \frac{\lambda_n}{\lambda_k}\vec{v}_n,$$

c'est-à-dire que $\vec{v}_k$ est combinaison linéaire des autres vecteurs de $\mathcal{F}$, ce qui peut encore s'écrire $\vec{v}_k \in \text{vect}\{\mathcal{F} \setminus \{v_k\}\}$ (la notation ensembliste $A \setminus B$ désigne l'ensemble des éléments de A qui n'appartiennent pas à B, c'est-à-dire A dont on a ôté B. C'est la différence ensembliste).

Réciproquement, supposons que pour un certain k, on ait $\vec{v}_k \in \text{vect}\{\mathcal{F} \setminus \{v_k\}\}$. Ceci signifie que l'on peut écrire

$$\vec{v}_k = \mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \cdots + \mu_n\vec{v}_n,$$

où $\vec{v}_k$ ne figure pas au second membre. Passant $\vec{v}_k$ au second membre, il vient

$$\vec{0} = \mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \cdots - \vec{v}_k + \cdots + \mu_n\vec{v}_n,$$

ce qui est une relation de dépendance linéaire pour $\mathcal{F}$ puisque $-1 \neq 0$. □

On déduit du théorème 14 le résultat crucial, quoique trivial, qui suit.

Théorème 16 *Soit $n > m$. Alors toute famille de n vecteurs de $\mathbb{R}^m$ est automatiquement liée.*

Démonstration. La matrice formée à partir de la famille a m lignes, donc au plus m positions de pivot. Or elle a $n > m$ colonnes, donc au moins une colonne n'a pas de pivot. □

Remarque 13 Dans ce cas, inutile donc de se lancer dans des calculs de méthode de Gauss. ATTENTION : par contre si $n \leq m$, on ne peut rien dire *a priori*. La famille peut être libre ou liée, seul le calcul peut le dire. □

La contraposée du théorème précédent est non moins importante, c'est pourquoi on l'explicite.

Théorème 17 *Toute famille libre de vecteurs de $\mathbb{R}^m$ a au plus m éléments.*

Mentionnons deux autres propriétés triviales de moindre importance.

Proposition 24 *Si une famille contient le vecteur nul, alors elle est liée. Si elle contient deux vecteurs égaux, alors elle est liée.*

3.8 Bases de $\mathbb{R}^m$

Encore une définition fondamentale.

Définition 25 *Une* base *de $\mathbb{R}^m$ est une famille à la fois libre et génératrice.*

Exemple 11 Il existe des bases (ce n'est pas évident *a priori*). Ainsi, la famille des vecteurs-colonne de la matrice identité I_m

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \vec{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

forme manifestement une base de $\mathbb{R}^m$ (un pivot par ligne et par colonne). On l'appelle *la base canonique* de $\mathbb{R}^m$, car c'est la plus simple de toutes les bases de $\mathbb{R}^m$. Mais il y en a une infinité d'autres ! □

Les bases possèdent plusieurs propriétés absolument remarquables. En particulier

Théorème 18 *i) Toute base de $\mathbb{R}^m$ admet exactement m éléments.*
ii) Toute famille libre de m éléments est une base.
iii) Toute famille génératrice de m éléments est une base.

Démonstration. i) Soit $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ une base de $\mathbb{R}^m$ à n éléments. Alors, la matrice $m \times n$ $\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}$ admet une position de pivot par ligne et une position de pivot par colonne. Il y a donc autant de lignes et de colonnes que de positions de pivot, d'où le même nombre de lignes et de colonnes, à savoir $n = m$.

ii) Soit $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ une famille libre à m éléments. Alors, la matrice $m \times m$ $\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}$ admet une position de pivot par colonne, c'est-à-dire m positions de pivot. Comme elle a m lignes, il y a une position de pivot par ligne. La famille est donc génératrice.

iii) Soit $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ une famille génératrice à m éléments. Alors, la matrice $m \times m$ $\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}$ admet une position de pivot par ligne, c'est-à-dire m positions de pivot. Comme elle a m colonnes, il y a une position de pivot par colonne. La famille est donc libre. □

Comme le nombre d'éléments d'une base ne dépend pas du choix de la base, c'est qu'il s'agit d'une quantité caractéristique de l'espace tout entier. On pose donc la définition suivante.

Définition 26 *La* dimension *de $\mathbb{R}^m$ est le nombre d'éléments de toute base de $\mathbb{R}^m$,*

$$\dim \mathbb{R}^m = m.$$

Remarque 14 Cette notion algébrique de dimension coïncide avec la notion géométrique de dimension pour $m = 1, 2$ et 3. En effet, $\mathbb{R}$ est représenté géométriquement par la droite, c'est-à-dire un objet monodimensionnel, $\mathbb{R}^2$ par le plan, c'est-à-dire un objet bidimensionnel et $\mathbb{R}^3$ par l'espace qui est tridimensionnel. La géométrie élémentaire s'arrête là, mais pas l'algèbre linéaire. On considère donc sans difficulté aucune des espaces de dimension 4, 5, 10, 100 ou 1000. Naturellement, c'est par le biais de l'algèbre linéaire que l'on pourra envisager d'aborder la géométrie de ces espaces de grande dimension. □

Une autre propriété remarquable des bases est de permettre de décomposer les vecteurs.

Théorème 19 *Soit $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ une base de $\mathbb{R}^m$. Pour tout $\vec{x}$ de $\mathbb{R}^m$, il existe un unique m-uplet de scalaires $(\lambda_1, \lambda_2, \dots, \lambda_m)$ tel que*

$$\vec{x} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_m\vec{v}_m.$$

Démonstration. L'existence du m-uplet découle directement du caractère générateur de $\mathcal{B}$. Pour l'unicité, on va utiliser le caractère libre de $\mathcal{B}$. Supposons donc que l'on ait deux m-uplets $(\lambda_1, \lambda_2, \dots, \lambda_m)$ et $(\lambda'_1, \lambda'_2, \dots, \lambda'_m)$ tels que

$$\begin{array}{rcl} \vec{x} & = & \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_m\vec{v}_m \\ \vec{x} & = & \lambda'_1\vec{v}_1 + \lambda'_2\vec{v}_2 + \cdots + \lambda'_m\vec{v}_m. \end{array}$$

Soustrayant membre à membre, on obtient

$$\vec{0} = (\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_m\vec{v}_m) - (\lambda'_1\vec{v}_1 + \lambda'_2\vec{v}_2 + \cdots + \lambda'_m\vec{v}_m),$$

soit

$$\vec{0} = (\lambda_1 - \lambda'_1)\vec{v}_1 + (\lambda_2 - \lambda'_2)\vec{v}_2 + \cdots + (\lambda_m - \lambda'_m)\vec{v}_m.$$

Comme la famille est libre, c'est que tous les coefficients de cette combinaison linéaire sont nuls, c'est-à-dire

$$\lambda_1 = \lambda'_1, \lambda_2 = \lambda'_2, \dots, \lambda_m = \lambda'_m,$$

d'où l'unicité. □

On est alors fondé à poser la définition suivante.

Définition 27 *Le m-uplet* $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ *s'appelle les* composantes de $\vec{x}$ dans la base $\mathcal{B}$. *On les range naturellement en colonne et l'on note*

$$(\vec{x})_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}.$$

Remarque 15 Attention : en général $(\vec{x})_{\mathcal{B}} \neq \vec{x}$, sauf si $\mathcal{B}$ est la base canonique. Les composantes de $\vec{x}$ dans la base $\mathcal{B}$ nous donnent une nouvelle représentation du vecteur $\vec{x}$. □

Théorème 20 *Les composantes dans la base* $\mathcal{B}$ *ont des propriétés de linéarité*

$$(\vec{x}+\vec{y})_{\mathcal{B}} = (\vec{x})_{\mathcal{B}} + (\vec{y})_{\mathcal{B}}, \quad (\lambda\vec{x})_{\mathcal{B}} = \lambda(\vec{x})_{\mathcal{B}}.$$

Démonstration. Il suffit de l'écrire. □

Comment calcule-t-on les composantes d'un vecteur $\vec{x}$ dans une base $\mathcal{B}$? C'est très simple : on commence par écrire la définition

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_m\vec{v}_m = \vec{x},$$

puis on reconnaît au premier membre un produit matrice-vecteur

$$\mathcal{P}_{\mathcal{B}}(\vec{x})_{\mathcal{B}} = \vec{x},$$

où

$$\mathcal{P}_{\mathcal{B}} = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \end{pmatrix}$$

s'appelle la matrice de passage. Pour calculer $(\vec{x})_{\mathcal{B}}$, il suffit donc de résoudre ce système linéaire $m \times m$ dont le second membre est $\vec{x}$.

Chapitre 4

Sous-espaces vectoriels de $\mathbb{R}^m$

À partir de maintenant, nous abandonnons l'écriture des vecteurs avec des flèches. Le contexte doit être suffisant pour savoir si l'on parle d'un vecteur ou d'un scalaire.

Certains sous-ensembles de $\mathbb{R}^m$ partagent les mêmes propriétés algébriques que $\mathbb{R}^m$ tout entier. Ce sont les sous-espaces vectoriels.

4.1 Sous-espaces vectoriels

Définition 28 *Soit F un sous-ensemble de $\mathbb{R}^m$. On dit que F est un* sous-espace vectoriel de $\mathbb{R}^m$ *si et seulement si*

i) $0 \in F$.

ii) pour tous (u,v) dans F, $u+v \in F$ (F est stable pour l'addition).

iii) pour tout u dans F et tout λ dans $\mathbb{R}$, $\lambda u \in F$ (F est stable pour la multiplication par un scalaire).

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 16 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 iii). Un sev F muni des deux opérations satisfait donc les huit propriétés algébriques de base. C'est donc également un espace vectoriel sur $\mathbb{R}$. $\square$

Exemple 12 1) Pour tout m, $\{0\}$ et $\mathbb{R}^m$ sont des sev de $\mathbb{R}^m$.

2) L'ensemble $F = \{x \in \mathbb{R}^2; x_2 = 0\} = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, x_1 \in \mathbb{R} \right\}$ est un sev de $\mathbb{R}^2$. $\square$

Voyons comment on peut fabriquer des sev.

Proposition 25 *Soit* $\{v_1, v_2, \ldots, v_n\}$ *une famille de vecteurs de* $\mathbb{R}^m$. *Alors,* $F = \mathrm{vect}\{v_1, v_2, \ldots, v_n\}$ *est un sev de* $\mathbb{R}^m$.

Démonstration. Il s'agit de vérifier les trois points de la définition de sev.

i) $0 = 0v_1 + 0v_2 + \cdots + 0v_n \in F$.

ii) Soient x et y deux éléments de F. On a donc

$$\begin{aligned} x &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, \\ y &= \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n. \end{aligned}$$

Additionnant membre à membre, il vient

$$x + y = (\lambda_1 + \mu_1)v_1 + (\lambda_2 + \mu_2)v_2 + \cdots + (\lambda_n + \mu_n)v_n \in F.$$

iii) De même, pour tout $x \in F$ et tout $\lambda \in \mathbb{R}$,

$$\lambda x = (\lambda\lambda_1)v_1 + (\lambda\lambda_2)v_2 + \cdots + (\lambda\lambda_n)v_n \in F$$

d'où le résultat. □

À partir de maintenant, on dira donc *espace vectoriel engendré* par une famille de vecteurs à la place de « espace engendré ».

Certaines opérations ensemblistes conservent les sev.

Proposition 26 *Soient* F_1 *et* F_2 *deux sev de* $\mathbb{R}^m$. *Alors,* $F = F_1 \cap F_2$ *est un sev de* $\mathbb{R}^m$.

Démonstration. Il s'agit de vérifier les trois points de la définition de sev.

i) $0 \in F_1$ et $0 \in F_2$, donc $0 \in F_1 \cap F_2$.

ii) Soient x et y deux éléments de F. Comme x et y appartiennent à F_1, on a $x + y \in F_1$. Comme x et y appartiennent à F_2, on a $x + y \in F_2$. Finalement, $x + y \in F_1 \cap F_2$

iii) Identique. □

Remarque 17 La démonstration montre clairement qu'une intersection *quelconque* de sev est un sev. *Attention :* par contre, en général, $F_1 \cup F_2$ *n'est pas* un sev, sauf si $F_1 \subset F_2$ ou $F_2 \subset F_1$. □

L'opération ensembliste qui remplace la réunion pour les sev est un peu plus sophistiquée.

Définition 29 *Soient $F_1, F_2, \dots, F_k$ une famille de sev de $\mathbb{R}^m$. L'ensemble*

$$F = \{x \in \mathbb{R}^m; \exists x_i \in F_i, i = 1, \dots, k, \text{ tels que } x = x_1 + x_2 + \dots + x_k\}$$

c'est-à-dire l'ensemble des vecteurs de $\mathbb{R}^m$ qui peuvent s'écrire comme une somme de k vecteurs, chacun d'entre eux appartenant à un F_i, est un sev de $\mathbb{R}^m$ appelé somme *des $(F_i)_{i=1,\dots,k}$ et noté $F = F_1 + F_2 + \dots + F_k = \sum_{i=1}^k F_i$.*

Démonstration. Il s'agit de vérifier les trois points de la définition de sev.

i) $0 \in F_i$ pour tout i, donc $0 = 0 + 0 + \dots + 0 \in F$.

ii) Soient x et y deux éléments de F. Il existe donc x_i, y_i dans F_i pour tout i tels que

$$\begin{aligned} x &= x_1 + x_2 + \dots + x_k, \\ y &= y_1 + y_2 + \dots + y_k. \end{aligned}$$

Additionnant membre à membre, il vient

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

Comme F_i est un sev, on a $x_i + y_i \in F_i$ pour tout i par stabilité pour l'addition, d'où finalement $x + y \in F$.

iii) Identique. □

Remarque 18 On peut montrer que $\sum_{i=1}^k F_i$ est le plus petit sev de $\mathbb{R}^m$ qui contient $\cup_{i=1}^k F_i$. □

Un cas particulier important de somme de sev.

Définition 30 *Soit $F = \sum_{i=1}^k F_i$. Si pour tout x dans F la décomposition en somme d'éléments de F_i est unique, alors on dit que la somme est* directe *et l'on note dans ce cas $F = F_1 \oplus F_2 \oplus \dots \oplus F_k = \bigoplus_{i=1}^k F_i$.*

On a une caractérisation des sommes directes.

Proposition 27 *L'espace $F = \sum_{i=1}^k F_i$ est une somme directe si et seulement si on a l'équivalence*

$$x_i \in F_i, x_1 + x_2 + \dots + x_k = 0 \iff \forall i = 1, \dots k, x_i = 0.$$

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

$$x_i \in F_i, x_1 + x_2 + \dots + x_k = 0 \iff \forall i = 1, \dots k, x_i = 0.$$

Soit $y \in F$ et $y_i, y'_i \in F_i$ deux décompositions de y. On a donc

$$\begin{aligned} y &= y_1 + y_2 + \cdots + y_k, \\ y &= y'_1 + y'_2 + \cdots + y'_k. \end{aligned}$$

Soustrayant membre à membre, il vient

$$0 = (y_1 - y'_1) + (y_2 - y'_2) + \cdots + (y_k - y'_k),$$

avec $y_i - y'_i \in F_i$ puisque F_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 28 *Soient F_1 et F_2 deux sev de $\mathbb{R}^m$. On a $F = F_1 \oplus F_2$ si et seulement si $F = F_1 + F_2$ et $F_1 \cap F_2 = \{0\}$.*

Démonstration.

Condition nécessaire : si $F = F_1 \oplus F_2$ alors en particulier $F = F_1 + F_2$. De plus, prenant $x \in F_1 \cap F_2$, on voit que $0 = x + (-x)$ avec $x \in F_1$ et $-x \in F_2$. Par conséquent, $x = 0$.

Condition suffisante : si $F = F_1 + F_2$ et $F_1 \cap F_2 = \{0\}$, prenons $x_1 \in F_1$ et $x_2 \in F_2$ tels que $x_1 + x_2 = 0$. On en déduit que $x_2 = -x_1 \in F_1$, d'où $x_2 \in F_1 \cap F_2$, d'où $x_2 = x_1 = 0$. □

Remarque 19 Attention, cette proposition est fausse à partir de trois sev. On peut parfaitement avoir $F_1 \cap F_2 \cap F_3 = \{0\}$ sans que la somme $F_1 + F_2 + F_3$ ne soit directe. □

Définition 31 *Si $E = F_1 \oplus F_2$, on dit que F_1 et F_2 sont des* sev supplémentaires.

4.2 Bases, composantes, dimension

La notion de base s'étend immédiatement aux sev de $\mathbb{R}^m$.

Définition 32 *Soit F un sev de $\mathbb{R}^m$. Une famille de vecteurs $\{v_1, \ldots, v_k\}$ de F est une base de F si elle est à la fois* libre *et* génératrice *de F, c'est-à-dire que $F = \text{vect}\{v_1, \ldots, v_k\}$.*

Il y a deux grands théorèmes importants pour les bases.

Théorème 21 (de la base extraite) *Soit F un sev de $\mathbb{R}^m$ engendré par une famille finie de vecteurs S. Alors il existe un sous-ensemble de S qui est une base de F.*

Démonstration. Soit $S = \{v_1, v_2, \ldots, v_k\}$ 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 $\operatorname{Card} S = 1$ alors c'est que $v_1 = 0$ et donc $F = \{0\}$. Dans ce cas, $S' = \emptyset$ est la base cherchée. Si $\operatorname{Card} S \geq 2$ alors il existe un indice p tel que v_p soit combinaison linéaire des autres vecteurs de S, d'après le Théorème 15. 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

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

(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

$$x = \lambda_1 v_1 + \cdots + \lambda_{p-1} v_{p-1} + \lambda_p v_p + \lambda_{p+1} v_{p+1} + \cdots + \lambda_k v_k.$$

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

$$\begin{aligned} x &= (\lambda_1 + \lambda_p \mu_1) v_1 + \cdots + (\lambda_{p-1} + \lambda_p \mu_{p-1}) v_{p-1} \\ &\qquad + (\lambda_{p+1} + \lambda_p \mu_{p+1}) v_{p+1} + \cdots + (\lambda_k + \lambda_p \mu_k) v_k \end{aligned}$$

d'où S' génératrice. De plus, $\operatorname{Card} S' = \operatorname{Card} S - 1$.

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 22 (de la base incomplète) *Soit F un sev de $\mathbb{R}^m$ et S une famille libre de vecteurs de F. Alors il existe une base $\mathcal{B}$ de F qui contient S.*

Démonstration. Soit S la famille libre en question. Par le théorème 17, on a $\operatorname{Card} S \leq m$, en particulier c'est une famille finie $\{v_1, \ldots, v_k\}$ avec $k \leq m$ (le cas de l'ensemble vide est trivial). Soit cette famille engendre F et on a terminé, soit elle n'engendre pas F. Dans ce cas, il existe un vecteur $v \in F$ n'appartenant pas à $\operatorname{vect} S$. On pose alors $S' = S \cup \{v\}$. C'est encore une famille libre, en effet soient des scalaires λ_i et μ tels que

$$0 = \lambda_1 v_1 + \cdots + \lambda_k v_k + \mu v.$$

Si $\mu \neq 0$, on peut diviser par μ et l'on obtient en passant les autres termes au membre de gauche

$$v = -\frac{\lambda_1}{\mu} v_1 - \cdots - \frac{\lambda_k}{\mu} v_k,$$

ce qui contredit le fait que $v \notin \operatorname{vect} S$. Donc $\mu = 0$. Mais alors,

$$0 = \lambda_1 v_1 + \cdots + \lambda_k v_k,$$

et comme S est libre, on en déduit que $\lambda_1 = \cdots = \lambda_k = 0$. Donc S' est libre.

On recommence. Comme $\operatorname{Card} S' = \operatorname{Card} S + 1 \leq m$, le processus doit s'arrêter après au plus m étapes, et il s'arrête nécessairement sur une famille à la fois libre et génératrice de F. $\square$

Remarque 20 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 $\mathcal{B}$ et que $\operatorname{Card} \mathcal{B} \leq m$. $\square$

Le théorème suivant est totalement identique à sa version du chapitre précédent dans $\mathbb{R}^m$.

Théorème 23 *Soit F un sev de $\mathbb{R}^m$ et $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ une base de F. Pour tout x de F, il existe un unique k-uplet de scalaires $(\lambda_1, \lambda_2, \ldots, \lambda_k)$ tel que*

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k.$$

Ce k-uplet s'appelle les composantes de x dans la base $\mathcal{B}$, *notées*

$$(x)_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} \in \mathbb{R}^k.$$

De plus, on a les propriétés de linéarité, pour tous x, y dans E et tout λ dans $\mathbb{R}$,

$$(x+y)_{\mathcal{B}} = (x)_{\mathcal{B}} + (y)_{\mathcal{B}}, \quad (\lambda x)_{\mathcal{B}} = \lambda (x)_{\mathcal{B}}.$$

Remarque 21 Le Théorème 23 permet de ramener l'étude d'un sev quelconque de $\mathbb{R}^m$ à celle de l'espace $\mathbb{R}^k$ pour un certain $k \leq m$. Il n'y a plus aucune raison de confondre $x \in \mathbb{R}^m$ et $(x)_{\mathcal{B}} \in \mathbb{R}^k$ puisque ces deux vecteurs n'habitent pas dans le même espace ! $\square$

Voici un corollaire immédiat, mais crucial, du Théorème 23.

Corollaire 24 *Soit F un sev de $\mathbb{R}^m$ et $\mathcal{B}$ une base de F de cardinal k. Alors on a*

i) $(\lambda_1 x_1 + \cdots + \lambda_\ell x_\ell)_{\mathcal{B}} = \lambda_1 (x_1)_{\mathcal{B}} + \cdots + \lambda_\ell (x_\ell)_{\mathcal{B}}$.

ii) Une famille $\{x_1, \ldots, x_\ell\}$ de E est libre si et seulement si la famille des composantes $\{(x_1)_{\mathcal{B}}, \ldots, (x_\ell)_{\mathcal{B}}\}$ est libre dans $\mathbb{R}^k$.

iii) Les familles $\{x_1, \ldots, x_\ell\}$ de E et $\{(x_1)_{\mathcal{B}}, \ldots, (x_\ell)_{\mathcal{B}}\}$ de $\mathbb{R}^k$ ont exactement les mêmes relations de dépendance linéaire.

Démonstration. i) n'est qu'une application répétée du Théorème 23. Montrons iii). Soit k le cardinal de $\mathcal{B}$. On a bien évidemment les composantes du vecteur nul de F dans la base $\mathcal{B}$

$$(0)_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^k,$$

c'est-à-dire le vecteur nul de $\mathbb{R}^k$. Donc si la famille $\{x_1, \dots, x_\ell\}$ admet une relation de dépendance linéaire, la famille $\{(x_1)_{\mathcal{B}}, \dots, (x_\ell)_{\mathcal{B}}\}$ admet la même relation de dépendance linéaire d'après i). Réciproquement, si la famille $\{(x_1)_{\mathcal{B}}, \dots, (x_\ell)_{\mathcal{B}}\}$ admet une relation de dépendance linéaire

$$\lambda_1(x_1)_{\mathcal{B}} + \cdots + \lambda_\ell(x_\ell)_{\mathcal{B}} = 0 \in \mathbb{R}^k$$

avec les λ_i non tous nuls, on en déduit par i) que

$$(\lambda_1 x_1 + \cdots + \lambda_\ell x_\ell)_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Par définition de ce que sont les composantes d'un vecteur dans une base, on en déduit que $\lambda_1 x_1 + \cdots + \lambda_\ell x_\ell = 0 \in \mathbb{R}^m$.

Enfin ii) découle de iii) (absence de relation de dépendance linéaire). □

Remarque 22 Dans le raisonnement précédent, attention à bien comprendre la signification du vecteur nul 0 qui, suivant le contexte, est le vecteur nul de $\mathbb{R}^m$ ou bien celui de $\mathbb{R}^k$. □

Un corollaire également immédiat de tout ce qui précède.

Corollaire 25 *Soit F un sev de $\mathbb{R}^m$ et $\mathcal{B}$ une base de F de cardinal k, $\{u_1, u_2, \dots, u_n\}$ une famille de vecteurs de F. On introduit la matrice $k \times n$*

$$A = \big((u_1)_{\mathcal{B}} \ \ (u_2)_{\mathcal{B}} \ \ \cdots \ \ (u_n)_{\mathcal{B}}\big),$$

à coefficients dans $\mathbb{R}$. Alors on a

i) La famille $\{u_1, u_2, \dots, u_n\}$ est génératrice de F si et seulement si le système linéaire $Ax = b$ est compatible pour tout second membre b dans $\mathbb{R}^k$ si et seulement si la matrice A admet une position de pivot par ligne.

ii) La famille $\{u_1, u_2, \dots, u_n\}$ est libre si et seulement si le système linéaire homogène $Ax = 0$ n'admet que la solution triviale si et seulement si la matrice A admet une position de pivot par colonne.

Démonstration. Après application du Corollaire 24, identique au cas de $\mathbb{R}^m$. □

Le Corollaire 25 permet donc de déterminer si une famille de vecteurs d'un sev quelconque F est libre ou génératrice par un simple calcul d'algorithme de Gauss.

Théorème 26 *Soit F un sev de $\mathbb{R}^m$ et $\mathcal{B}$ une base de F de cardinal k. Toute famille de n vecteurs de E avec $n > k$ est liée.*

Démonstration. Soit $n > k$ et $S = \{u_1, u_2, \dots, u_n\}$ une famille de vecteurs de F. Considérons la matrice A du Corollaire 25. Cette matrice a n colonnes et k lignes, donc au moins une colonne n'a pas de position de pivot. □

On en déduit un théorème fondamental.

Théorème 27 (de la dimension) *Toutes les bases d'un F un sev de $\mathbb{R}^m$ ont le même nombre d'éléments.*

Démonstration. Soit $\mathcal{B}$ une base de F avec $\operatorname{Card}\mathcal{B} = k$ et $\mathcal{B}'$ une autre base de F avec $\operatorname{Card}\mathcal{B}' = k'$. Comme la deuxième base est une famille libre, on déduit du Théorème 26 que $k' \leq k$. Inversant les rôles de $\mathcal{B}$ et $\mathcal{B}'$, il vient aussi $k \leq k'$. □

On peut donc maintenant poser raisonnablement la définition suivante.

Définition 33 *La dimension d'un sous-espace vectoriel F de $\mathbb{R}^m$ est le nombre d'éléments de chacune de ses bases,*

$$\dim F = \operatorname{Card}\mathcal{B}, \textit{ pour toute base } \mathcal{B} \textit{ de } F.$$

Exemple 13 Quelques exemples simples. Tout d'abord, $\dim\{0\} = 0 = \operatorname{Card}\emptyset$. Aussi $\dim\mathbb{R}^m = m$ à l'aide de la base canonique, comme on l'a déjà vu. □

Théorème 28 *Si F est un sev de $\mathbb{R}^m$ de dimension k, alors toute famille libre de F a* au plus *k éléments et une famille libre de k éléments est une base de F. De même, toute famille génératrice de F a* au moins *k éléments et toute famille génératrice de k éléments est une base.*

Démonstration. Traitons le cas des familles libres. Si on a une famille libre S de F, alors d'après le théorème de la base incomplète, on peut lui ajouter des vecteurs pour construire une base. Cette base ayant k éléments, on a donc $\operatorname{Card} S \leq k$. Si $\operatorname{Card} S = k$, alors on ne peut pas ajouter de vecteurs, c'est donc que S était déjà la base complétée.

Le cas des familles génératrices est tout à fait analogue. Soit S une famille génératrice de F. D'après le théorème de base extraite, on peut lui enlever des vecteurs pour construire une base. Cette base ayant k éléments, on a donc $\operatorname{Card} S \geq k$. Si $\operatorname{Card} S = k$, alors on ne peut pas enlever de vecteurs, c'est donc que S était déjà la base extraite. □

Remarque 23 Ce résultat est important car il montre que, quand on connaît déjà la valeur de la dimension d'un sev et *seulement dans ce cas*, pour montrer qu'une famille donnée est une base, il suffit de compter ses éléments, s'assurer qu'il est égal à la dimension, puis vérifier soit que la famille est libre, soit qu'elle est génératrice. Évidemment, il est plus astucieux de vérifier la propriété la plus simple, qui est souvent le caractère libre, mais pas toujours ! □

Continuons dans les résultats importants.

Théorème 29 *Soit F_1 et F_2 deux sev de $\mathbb{R}^m$ tels que $F_1 \subset F_2$. Alors $\dim F_1 \leq \dim F_2 \leq m$.*

Démonstration. D'après la remarque 20, F_1 admet une base, qui est donc aussi une famille libre de F_2. Donc son cardinal est inférieur à $\dim F_2$. Comme $F_2 \subset \mathbb{R}^m$, la même remarque implique que $\dim F_2 \leq m$ □

Corollaire 30 *Si de plus $\dim F_1 = \dim F_2$, alors $F_1 = F_2$.*

Démonstration. Soit $\mathcal{B}_1$ une base de F_1. On a donc $F_1 = \operatorname{vect} \mathcal{B}_1$. Mais c'est aussi une famille libre de F_2 à $\dim F_2$ éléments. Donc, c'est aussi une base de F_2, ce qui implique que $\operatorname{vect} \mathcal{B}_1 = F_2$. □

Ceci s'applique en particulier au cas $F_2 = \mathbb{R}^m$.

Exemple 14 Soit $u \neq 0$ un vecteur non nul de $\mathbb{R}^m$. Alors $F = \operatorname{vect}\{u\} = \{v = \lambda u, \lambda \in \mathbb{R}\} = \mathbb{R}u$ est un sev de $\mathbb{R}^m$ de dimension 1. On dit que c'est une *droite vectorielle*.

Soit $\{u, v\}$ une famille libre de $\mathbb{R}^m$. Alors $F = \operatorname{vect}\{u, v\}$ est un sev de $\mathbb{R}^m$ de dimension 2. On dit que c'est un *plan vectoriel*.

Soit F un sev de $\mathbb{R}^m$ tel que $\dim F = m - 1$. On dit que F est un *hyperplan*. Dans $\mathbb{R}^2$, un hyperplan n'est autre qu'une droite vectorielle, dans $\mathbb{R}^3$, un hyperplan n'est autre qu'un plan vectoriel. □

En général, un sev admet une infinité de bases. La question se pose donc de savoir comment les composantes d'un vecteur dans une base sont reliées aux composantes de ce même vecteur dans une autre base. C'est le problème du *changement de base*. Ce problème se résout simplement.

Considérons donc un sev F de $\mathbb{R}^m$ de dimension k, et deux bases de F, $\mathcal{B} = \{u_1, u_2, \dots, u_k\}$ et $\mathcal{B}' = \{v_1, v_2, \dots, v_k\}$.

Définition 34 *On appelle* matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ *la matrice*

$$\mathcal{P}_{\mathcal{B} \to \mathcal{B}'} = \big((v_1)_{\mathcal{B}} \;\; (v_2)_{\mathcal{B}} \;\; \cdots \;\; (v_k)_{\mathcal{B}} \big).$$

C'est une matrice $k \times k$.

Théorème 31 *Pour tout $x \in F$, on a*

$$(x)_{\mathcal{B}} = \mathcal{P}_{\mathcal{B}\to\mathcal{B}'}(x)_{\mathcal{B}'}.$$

Démonstration. C'est immédiat. En effet

$$(x)_{\mathcal{B}'} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} \Longleftrightarrow x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k.$$

Prenant les composantes dans $\mathcal{B}$ des deux membres, il vient

$$(x)_{\mathcal{B}} = \lambda_1 (v_1)_{\mathcal{B}} + \lambda_2 (v_2)_{\mathcal{B}} + \cdots + \lambda_k (v_k)_{\mathcal{B}} = \mathcal{P}_{\mathcal{B}\to\mathcal{B}'}(x)_{\mathcal{B}'},$$

par définition du produit matrice-vecteur. □

En d'autres termes, si on connaît les composantes des vecteurs de E dans la base $\mathcal{B}$, on commence par construire la matrice de passage, puis on calcule les composantes d'un vecteur x dans la base $\mathcal{B}'$ en résolvant le système linéaire ci-dessus.

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

Théorème 32 *Soit $\mathcal{B}$ une base de $\mathbb{R}^m$. On se donne une partition $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$ de $\mathcal{B}$, c'est-à-dire un découpage de $\mathcal{B}$ en parties disjointes ($\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ pour $i \neq j$ et $\mathcal{B} = \cup_{i=1}^k \mathcal{B}_i$). Alors on a*

$$\mathbb{R}^m = \operatorname{vect} \mathcal{B}_1 \oplus \operatorname{vect} \mathcal{B}_2 \oplus \cdots \oplus \operatorname{vect} \mathcal{B}_k.$$

Réciproquement, si $\mathbb{R}^m = F_1 \oplus F_2 \oplus \cdots \oplus F_k$ et $\mathcal{B}_i$ est une base de F_i, alors $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_k$ est une base de $\mathbb{R}^m$.

Démonstration. On commence par la première partie du théorème. Comme tout vecteur x de $\mathbb{R}^m$ est combinaison linéaire des vecteurs de $\mathcal{B}$, famille génératrice, en regroupant ensemble les termes correspondant à chaque $\mathcal{B}_i$, on écrit ainsi x comme une somme de vecteurs appartenant à $\operatorname{vect} \mathcal{B}_i$, $i = 1, \ldots, k$, soit $\mathbb{R}^m = \operatorname{vect} \mathcal{B}_1 + \operatorname{vect} \mathcal{B}_2 + \cdots + \operatorname{vect} \mathcal{B}_k$. Il reste à montrer que la somme est directe. Pour cela on décompose le vecteur nul

$$0 = x_1 + x_2 + \cdots + x_k \text{ avec } x_i \in \operatorname{vect} \mathcal{B}_i.$$

Chaque x_i est combinaison linéaire des vecteurs de $\mathcal{B}_i$ avec certains coefficients, donc la somme ci-dessus est en fait une combinaison linéaire des vecteurs de $\mathcal{B}$. Comme chaque vecteur de $\mathcal{B}_i$ n'apparaît qu'à un seul endroit dans cette somme, puisque $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ 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 $\mathcal{B}$ 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 $\mathbb{R}^m = F_1 + F_2 + \cdots + F_k$ est clairement dire que $\mathcal{B}$ 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 $\mathcal{B}$ est libre. $\square$

Corollaire 33 *Si $\mathbb{R}^m = F_1 \oplus F_2 \oplus \cdots \oplus F_k$, alors*

$$m = \dim F_1 + \dim F_2 + \cdots + \dim F_k.$$

Démonstration. On compte les vecteurs de base grâce au théorème précédent. $\square$

Un cas particulier important est que quand F_1 et F_2 sont deux sev supplémentaires de $\mathbb{R}^m$, alors

$$m = \dim F_1 + \dim F_2.$$

Corollaire 34 *Si F_1 est un sous-espace vectoriel de $\mathbb{R}^m$, alors F_1 admet au moins un supplémentaire.*

Démonstration. Soit $\mathcal{B}_1$ une base de F_1. C'est une famille libre de $\mathbb{R}^m$, donc d'après le théorème de la base incomplète, on peut lui ajouter un ensemble fini de vecteurs $\mathcal{B}_2$ de telle sorte que $\mathcal{B}_1 \cup \mathcal{B}_2$ soit une base de $\mathbb{R}^m$. Le supplémentaire cherché n'est autre que $F_2 = \operatorname{vect} \mathcal{B}_2$. $\square$

Proposition 29 *On a la formule plus générale*

$$\dim(F_1 + F_2) = \dim F_1 + \dim F_2 - \dim(F_1 \cap F_2).$$

Démonstration. Nous admettrons ce résultat qui utilise le même type d'arguments que précédemment, (commencer avec une base de $F_1 \cap F_2$). $\square$

Corollaire 35 *L'intersection de deux hyperplans de $\mathbb{R}^m$ est soit un hyperplan, soit un sev de dimension $m-2$. 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. Soient F_1 et F_2 deux hyperplans de $\mathbb{R}^m$, c'est-à-dire deux sev de dimension $m-1$. De la formule précédente, on déduit que

$$\dim(F_1 \cap F_2) = \dim F_1 + \dim F_2 - \dim(F_1 + F_2) = 2m - 2 - \dim(F_1 + F_2).$$

Comme $F_1 \subset F_1 + F_2 \subset \mathbb{R}^m$, on a $m-1 = \dim F_1 \leq \dim(F_1+F_2) \leq \dim \mathbb{R}^m = m$. On en déduit un encadrement de la dimension de $F_1 \cap F_2$:

$$m - 2 \leq \dim(F_1 \cap F_2) \leq m - 1.$$

Il y a donc deux possibilités, soit $\dim(F_1 \cap F_2) = m-1$, ce qui n'est possible que si $F_1 \cap F_2 = F_1 = F_2$ (les deux hyperplans sont confondus), soit $\dim(F_1 \cap F_2) = m-2$.

Pour $m = 2$, on a donc soit $\dim(F_1 \cap F_2) = 1$ c'est une droite, soit $\dim(F_1 \cap F_2) = 0$, les deux droites sont supplémentaires.

Pour $m = 3$, on a donc soit $\dim(F_1 \cap F_2) = 2$ c'est un plan, soit $\dim(F_1 \cap F_2) = 1$ c'est une droite. □

Remarque 24 Tout ce que l'on a raconté dans ces deux derniers chapitres reste vrai mot pour mot si l'on remplace systématiquement $\mathbb{R}$ par $\mathbb{C}$ partout où il apparaît. On a alors affaire à des espaces vectoriels complexes, des sev de $\mathbb{C}^m$, etc. Il n'y a *aucune difficulté supplémentaire.* □

Chapitre 5

Applications linéaires

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

5.1 Définition et premiers exemples

Dans ce qui suit, et sauf mention expresse du contraire, E_1 désigne un sev de $\mathbb{R}^{m_1}$, E_2 un sev de $\mathbb{R}^{m_2}$, et ainsi de suite.

Définition 35 *On dit qu'une application $f\colon E_1 \to E_2$ est une* application linéaire *si et seulement si*

i) $\forall x,y \in E_1,\ f(x+y) = f(x) + f(y)$.

ii) $\forall x \in E_1, \forall \lambda \in \mathbb{R},\ f(\lambda x) = \lambda f(x)$.

Il faut noter que les opérations qui apparaissent au membre de gauche de ces égalités sont des opérations dans E_1, alors que celles qui apparaissent au membre de droite sont dans E_2, même si on ne les distingue pas dans la notation.

On vérifie facilement que la Définition 35 est équivalente aux deux caractérisations suivantes.

i)$'$ $\forall x,y \in E_1, \forall \lambda, \mu \in \mathbb{R},\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

ou

i)$''$ $\forall x,y \in E_1, \forall \lambda \in \mathbb{R},\ f(\lambda x + y) = \lambda f(x) + f(y)$.

L'équivalence de la Définition 35 avec i)$'$ est claire. Pour i)$''$, si f est linéaire, alors elle vérifie bien sûr i)$''$ par application successive de i) et ii). Soit f vérifiant i)$''$. Notons d'abord, en prenant $\lambda = 1$ et $y = 0$, que $f(x) = f(x+0) = f(x) + f(0)$. On en déduit que $f(0) = 0$. Prenant ensuite $\lambda = 1$, on obtient i), puis $y = 0$ et on obtient ii). Par conséquent, f est linéaire.

Remarquons au passage que toute application linéaire satisfait $f(0) = 0$.

Exemple 15 1) L'application nulle $E_1 \to E_2$, $x \mapsto 0$, est linéaire.

2) Si $\dim E_1 = k$ et $\mathcal{B}$ est une base de E_1, alors l'application $E \to \mathbb{R}^k$, $x \mapsto (x)_{\mathcal{B}}$ est linéaire.

3) Si A est une matrice $m \times n$, alors l'application $\mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$ est linéaire.

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 $f^{-1}\colon Y \to X$ telle que $\forall x \in X$, $f^{-1}(f(x)) = x$ et $\forall y \in Y$, $f(f^{-1}(y)) = y$.

Si f est une application linéaire bijective de E_1 dans E_2, on dit que c'est un *isomorphisme* entre E_1 et E_2. Dans ce cas, l'application réciproque f^{-1} est également linéaire de E_2 dans E_1, et c'est donc aussi un isomorphisme.

Démonstration. Soit $f\colon E_1 \to E_2$ un isomorphisme. L'application réciproque est déterminée par
$$\forall y \in E_2, \quad f^{-1}(y) = x \Longleftrightarrow y = f(x).$$
Utilisons la caractérisation i)$''$. On s'intéresse donc à $f^{-1}(\lambda y_1 + y_2)$. Ce vecteur de E_1 est déterminé par
$$f^{-1}(\lambda y_1 + y_2) = x \Longleftrightarrow \lambda y_1 + y_2 = f(x).$$
Par ailleurs, on a aussi
$$f^{-1}(y_1) = x_1 \Leftrightarrow y_1 = f(x_1) \quad \text{et} \quad f^{-1}(y_2) = x_2 \Leftrightarrow y_2 = f(x_2).$$
En combinant ces deux dernières relations, il vient
$$\lambda y_1 + y_2 = \lambda f(x_1) + f(x_2) = f(\lambda x_1 + x_2)$$
par linéarité de f. En comparant avec la caractérisation ci-dessus, on obtient
$$f^{-1}(\lambda y_1 + y_2) = \lambda x_1 + x_2 = \lambda f^{-1}(y_1) + f^{-1}(y_2),$$
et f^{-1} est bien linéaire. □

Cette propriété permet donc dire que s'il existe un isomorphisme de E_1 dans E_2, alors il en existe un entre E_2 et E_1, et l'on est alors fondé à déclarer que E_1 et E_2 sont *isomorphes*.

Si maintenant f est une application linéaire de E dans E, on dit que c'est un *endomorphisme*.

Exemple 16 L'exemple 2) plus haut est un exemple d'isomorphisme. Il montre que tout sev de dimension k est isomorphe à $\mathbb{R}^k$. □

Définition 36 *L'ensemble de toutes les applications linéaires de E_1 dans E_2 est noté $\mathcal{L}(E_1,E_2)$. Si $E_1 = E_2 = E$, on le note plus simplement $\mathcal{L}(E)$.*

La linéarité est préservée par composition des applications.

Proposition 30 *Si $f \in \mathcal{L}(E_1,E_2)$ et $g \in \mathcal{L}(E_2,E_3)$, alors $g \circ f \in \mathcal{L}(E_1,E_3)$.*

Démonstration. Utilisons la caractérisation i)$''$. On a, pour tous $x,y \in E_1$ et tout $\lambda \in \mathbb{R}$,

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

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

5.2 Image, noyau, rang

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

Définition 37 *Soit $f \in \mathcal{L}(E_1,E_2)$. On appelle* noyau *de f l'ensemble*

$$\ker f = \{x \in E_1; f(x) = 0\} = f^{-1}(\{0\}).$$

On appelle image *de f l'ensemble*

$$\operatorname{im} f = \{y \in E_2; \exists x \in E_1, y = f(x)\} = f(E_1).$$

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 31 *Le noyau de f est un sev de E_1 et l'image de f est un sev de E_2.*

Démonstration. On va montrer que ces ensembles contiennent le vecteur nul de leur espace respectif et sont stables par addition et multiplication par un scalaire.

Comme $f(0_{E_1}) = 0_{E_2}$ (on distingue exceptionnellement les vecteurs nuls), on voit que $0_{E_1} \in \ker f$ et $0_{E_2} \in \operatorname{im} f$.

Soient $x_1, x_2 \in \ker f$ et $\lambda \in \mathbb{R}$. On a donc $f(x_1) = f(x_2) = 0$. Par conséquent, $f(x_1 + x_2) = f(x_1) + f(x_2) = 0$ et $f(\lambda x_1) = \lambda f(x_1) = 0$, d'où $x_1 + x_2 \in \ker f$ et $\lambda x_1 \in \ker f$.

Soient $y_1, y_2 \in \operatorname{im} f$ et $\lambda \in \mathbb{R}$. Il existe donc $x_1, x_2 \in E_1$ tels que $y_1 = f(x_1)$ et $y_2 = f(x_2)$. Par conséquent, $y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2)$ et $\lambda y_1 = \lambda f(x_1) = f(\lambda x_1)$, d'où $y_1 + y_2 \in \operatorname{im} f$ et $\lambda y_1 \in \operatorname{im} f$. □

Le noyau et l'image d'une application linéaire sont liés de façon cruciale aux propriétés ensemblistes de cette application.

Théorème 36 *Soit $f \in \mathcal{L}(E_1, E_2)$. On a*

i) L'application f est injective si et seulement si $\ker f = \{0\}$.

ii) L'application f est surjective si et seulement si $\operatorname{im} f = E_2$.

Démonstration. ii) est trivial, c'est une question de vocabulaire.

Soit f une application linéaire injective. Comme $f(0) = 0$, c'est donc que $0 \in E_1$ est l'unique vecteur dont l'image est $0 \in E_2$, d'où $\ker f = \{0\}$. Réciproquement, supposons que $\ker f = \{0\}$ et donnons nous deux vecteurs x_1 et x_2 de E_1 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 $x_1 - x_2 \in \ker f$, donc $x_1 = x_2$ et f est injective. □

Remarque 25 Pour démontrer qu'une application linéaire est injective, on utilisera *TOUJOURS* la caractérisation par le noyau i) et non pas la définition générale de l'injectivité. □

Exemple 17 1) Pour l'application nulle, notée 0, on a $\ker 0 = E_1$ et $\operatorname{im} 0 = \{0\}$.

2) Si A est une matrice $m \times n$, on notera par abus de langage $\ker A$ et $\operatorname{im} A$ le noyau et l'image de l'application linéaire $\mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$. On a alors

$$\ker A = \{x \in \mathbb{R}^n \text{ solutions du système homogène } Ax = 0\},$$

et

$$\operatorname{im} A = \{b \in \mathbb{R}^m \text{ tels que le système } Ax = b \text{ est compatible}\},$$

ce qui montre que ces deux notions sont également intimement liées aux questions de systèmes linéaires. On parle dans ce cas de noyau et image d'une matrice. □

Une définition importante.

Définition 38 *Soit* $f \in \mathcal{L}(E_1, E_2)$*. On définit son* rang *par*

$$\operatorname{rg} f = \dim \operatorname{im} f.$$

Proposition 32 *Si* $\{v_1, v_2, \ldots, v_{m_1}\}$ *est une base de* E_1*, alors*

$$\operatorname{im} f = \operatorname{vect}\{f(v_1), f(v_2), \ldots, f(v_{m_1})\} \text{ et } \operatorname{rg} f \leq \dim E_1 = m_1.$$

Démonstration. Comme $\{v_1, v_2, \ldots, v_{m_1}\}$ est une base de E_1, pour tout $y \in \operatorname{im} f$, il existe $x = \lambda_1 v_1 + \cdots + \lambda_{m_1} v_{m_1} \in E_1$ tel que $y = f(x)$. Par linéarité de f, il vient

$$y = f(\lambda_1 v_1 + \cdots + \lambda_{m_1} v_{m_1}) = \lambda_1 f(v_1) + \cdots + \lambda_{m_1} f(v_{m_1}) \in \operatorname{vect}\{f(v_1), \ldots, f(v_{m_1})\}.$$

Donc $\operatorname{im} f \subset \operatorname{vect}\{f(v_1), \ldots, f(v_{m_1})\}$.

Réciproquement, si $y \in \operatorname{vect}\{f(v_1), \ldots, f(v_{m_1})\}$, alors il existe des scalaires λ_j tels que $y = \lambda_1 f(v_1) + \cdots + \lambda_{m_1} f(v_{m_1})$. Posant $x = \lambda_1 v_1 + \cdots + \lambda_{m_1} v_{m_1} \in E_1$, on voit que $y = f(x)$, donc $y \in \operatorname{im} f$. Finalement, $\operatorname{im} f = \operatorname{vect}\{f(v_1), \ldots, f(v_{m_1})\}$.

La famille de m_1 vecteurs $\{f(v_1), \ldots, f(v_{m_1})\}$ est une famille génératrice de $\operatorname{im} f$. Par conséquent, par le Théorème 28, $\dim \operatorname{im} f \leq m_1 = \dim E_1$. □

Proposition 33 *Soit* $f \in \mathcal{L}(E_1, E_2)$*, alors*

i) $\operatorname{rg} f \leq \min(\dim E_1, \dim E_2)$.

ii) L'application f *est surjective si et seulement si* $\operatorname{rg} f = \dim E_2$.

iii) L'application f *est injective si et seulement si* $\operatorname{rg} f = \dim E_1$.

Démonstration. i) On a déjà vu que $\operatorname{rg} f \leq \dim E_1$. Par ailleurs, comme $\operatorname{im} f$ est un sev de E_2, on évidemment aussi $\operatorname{rg} f \leq \dim E_2$.

ii) f est surjective si et seulement si $\operatorname{im} f = E_2$ si et seulement si $\dim \operatorname{im} f = \dim E_2$ (*cf.* Corollaire 30).

iii) Si $\operatorname{rg} f = \dim E_1$, alors $\{f(v_1), \ldots, f(v_{m_1})\}$ est une famille génératrice de $\operatorname{im} f$ dont le nombre d'élément est égal à la dimension de $\operatorname{im} f$. C'en est donc une base et en particulier une famille libre. Or, si $x \in \ker f$, alors on a $x = \sum_{i=1}^{m_1} \lambda_i v_i$ et

$$0 = f(x) = f(\lambda_1 v_1 + \cdots + \lambda_{m_1} v_{m_1}) = \lambda_1 f(v_1) + \cdots + \lambda_{m_1} f(v_{m_1}).$$

On en déduit que $\lambda_1 = \lambda_2 = \cdots = \lambda_{m_1} = 0$, d'où $x = 0$. Par conséquent, $\ker f = \{0\}$ et f est injective.

Réciproquement, si f est injective, alors $\ker f = \{0\}$, donc en remontant le raisonnement précédent, on en déduit que la famille $\{f(v_1), \ldots, f(v_{m_1})\}$ est libre, donc c'est une base de $\operatorname{im} f$. Par conséquent, $\operatorname{rg} f = m_1 = \dim E_1$. □

On en déduit une classification *complète* des sev modulo un isomorphisme.

Corollaire 37 *Deux sev E_1 et E_2 sont isomorphes si et seulement si* $\dim E_1 = \dim E_2$.

Démonstration. On a déjà vu que si $\dim E_1 = \dim E_2 = k$ alors E_1 et E_2 sont isomorphes à $\mathbb{R}^k$. Par transitivité et symétrie, ils sont donc isomorphes.

Réciproquement, si f est un isomorphisme entre E_1 et E_2, alors elle est injective, donc $\operatorname{rg} f = \dim E_1$, et elle est surjective, donc $\operatorname{rg} f = \dim E_2$. □

Remarque 26 1) Il s'agit d'une nouvelle méthode pour calculer la dimension d'un sev : il suffit de montrer qu'il est isomorphe à un autre sev de dimension déjà connue. □

Corollaire 38 *Si f est un isomorphisme de E_1 dans E_2 alors l'image de toute base (resp. famille libre, famille génératrice) de E_1 est une base (resp. famille libre, famille génératrice) de E_2.*

Démonstration. Évident. □

Remarque 27 Il s'agit d'une nouvelle méthode pour construire une base d'un sev : il suffit de prendre l'image directe ou réciproque par un isomorphisme d'une base déjà connue d'un autre sev. □

5.3 Matrice d'une application linéaire dans des bases

Dans cette section, on se donne deux sev E_1 et E_2 avec $\dim E_1 = n$ et $\dim E_2 = m$. Le théorème suivant relie les notions de matrice et d'application linéaire.

Théorème 39 *Soit $\mathcal{B}_1$ une base de E_1 et $\mathcal{B}_2$ une base de E_2. Pour toute application linéaire f de E_1 dans E_2, il existe une unique matrice $m \times n$ A telle que*

$$\forall x \in E_1, \quad (f(x))_{\mathcal{B}_2} = A(x)_{\mathcal{B}_1}.$$

La matrice A est la matrice de f dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$.

Démonstration. Notons $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$ et $\mathcal{B}_2 = \{w_1, w_2, \dots, w_m\}$, avec $v_j \in E_1$ et $w_i \in E_2$. Pour tout $x \in E_1$, par définition des composantes dans la base $\mathcal{B}_1$, on a

$$(x)_{\mathcal{B}_1} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \Longleftrightarrow x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Appliquant f aux deux membres de cette égalité et utilisant la linéarité de f, on obtient

$$f(x) = \lambda_1 f(v_1) + \lambda_2 f(v_2) + \cdots + \lambda_n f(v_n).$$

On prend maintenant les composantes des deux membres dans la base $\mathcal{B}_2$, ce qui est aussi une opération linéaire, d'où

$$(f(x))_{\mathcal{B}_2} = \lambda_1 (f(v_1))_{\mathcal{B}_2} + \lambda_2 (f(v_2))_{\mathcal{B}_2} + \cdots + \lambda_n (f(v_n))_{\mathcal{B}_2}.$$

On reconnaît enfin au second membre le produit matrice-vecteur $A(x)_{\mathcal{B}_1}$ où

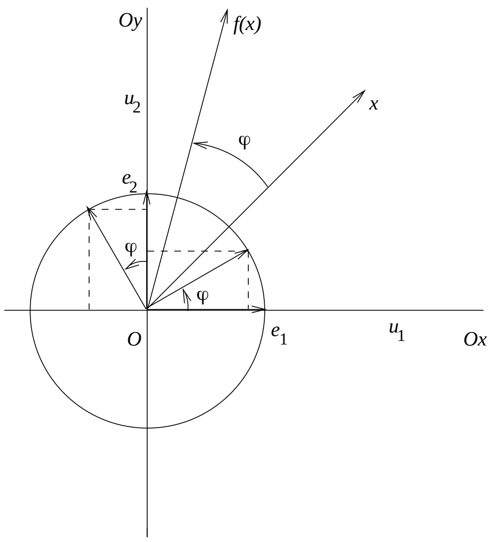
$$A = \big((f(v_1))_{\mathcal{B}_2} \;\; (f(v_2))_{\mathcal{B}_2} \;\; \cdots \;\; (f(v_n))_{\mathcal{B}_2} \big)$$

est la matrice de f dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$. □

Cette matrice est bien une matrice $m \times n$: elle a n colonnes, autant que de vecteurs de $\mathcal{B}_1$, et chacun de ces vecteurs-colonnes — les composantes de $f(v_j)$ dans la base $\mathcal{B}_2$ — a m lignes, autant que de vecteurs de $\mathcal{B}_2$.

Le Théorème 39 permet donc de ramener l'étude des applications linéaires à celles des produits matrice-vecteur et des matrices elles-mêmes, une fois que l'on a choisi une base dans chacun des deux espaces.

Exemple 18 1) Revenons à l'interprétation géométrique de $E_1 = E_2 = \mathbb{R}^2$ comme plan cartésien rapporté à un repère orthonormé pour l'occasion. On montre facilement, à l'aide de la règle du parallélogramme et du théorème de Thalès, que la rotation f_φ du plan de centre O et d'angle $\varphi \in [0, 2\pi]$ est en fait une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$.



Par définition des fonctions trigonométriques, on voit aisément que

$$f_\varphi(e_1) = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \text{ et } f_\varphi(e_2) = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}.$$

On en déduit la matrice de la rotation d'angle φ

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

qui permet de calculer l'image de n'importe quel vecteur par cette rotation. $\square$

Exemple 19 Une application linéaire ℓ de $\mathbb{R}^m$ dans $\mathbb{R}$ s'appelle une *forme linéaire*. La matrice d'une forme linéaire est donc une matrice $1 \times m$, c'est une matrice ligne.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \end{pmatrix}$$

et

$$\ell(x) = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m.$$

Le noyau de ℓ est l'ensemble des solutions de l'équation linéaire homogène $\ell(x) = 0$. $\square$

Les propriétés d'une application linéaire se lisent aisément sur sa matrice dans des bases données.

Proposition 34 *Soient E_1 et E_2 deux sev, $\mathcal{B}_1$ une base de E_1 et $\mathcal{B}_2$ une base de E_2, f une application linéaire de E_1 dans E_2 et A sa matrice dans ces bases. Alors*

i) L'application f est injective si et seulement si les colonnes de A sont linéairement indépendantes si et seulement si A a une position de pivot par colonne.

ii) L'application f est surjective si et seulement si les colonnes de A engendrent $\mathbb{R}^m$ si et seulement si A a une position de pivot par ligne.

Démonstration. i) L'application f est injective si et seulement $\ker f = \{0\}$. On a $\ker f = \{x \in E_1; f(x) = 0\} = \{x \in E_1; A(x)_{\mathcal{B}_1} = 0\}$. Les composantes dans $\mathcal{B}_1$ des vecteurs de $\ker f$ sont donc exactement les solutions du système linéaire homogène $A(x)_{\mathcal{B}_1} = 0$. Or on sait bien que ce système n'admet que la solution triviale si et seulement les colonnes de A sont linéairement indépendantes, c'est-à-dire si et seulement A a une position de pivot par colonne.

ii) L'application f est surjective si et seulement $\operatorname{im} f = E_2$. On a $\operatorname{im} f = \{y \in E_2; \exists x \in E_1, y = f(x)\} = \{y \in E_2; \exists x \in E_1, (y)_{\mathcal{B}_2} = A(x)_{\mathcal{B}_1}\}$. Les composantes dans $\mathcal{B}_2$ des vecteurs de $\operatorname{im} f$ sont donc exactement les seconds membres pour lesquels le système linéaire est compatible. Or on sait bien que c'est le cas pour tout second membre si et seulement les colonnes de A engendrent $\mathbb{R}^m$, c'est-à-dire si et seulement A a une position de pivot par ligne. $\square$

Ainsi, une forme linéaire non nulle est automatiquement surjective.

Les positions de pivot de la matrice d'une application linéaire nous donnent aussi la dimension de son noyau.

Proposition 35 *Mêmes hypothèses, soit k le nombre de variables libres associées à la matrice A. On a* $\dim\ker f = k$.

Démonstration. On vient de voir que $x \in \ker f$ si et seulement si $(x)_{\mathcal{B}_1}$ est solution du système linéaire homogène $A(x)_{\mathcal{B}_1} = 0$. Soit U la matrice échelonnée réduite équivalente à A. Les systèmes homogènes associés étant équivalents, on a aussi que $x \in \ker f$ si et seulement si $U(x)_{\mathcal{B}_1} = 0$.

S'il n'y a pas de variable libre, $k = 0$, alors on sait que l'application est injective, donc $\dim\ker f = 0$.

S'il y a $k \geq 1$ variables libres, alors on sait que l'application n'est pas injective et son noyau est isomorphe à l'ensemble des solutions du système homogène via l'isomorphisme $x \mapsto (x)_{\mathcal{B}_1}$. Pour résoudre ce système homogène échelonné réduit, on exprime les variables essentielles en fonction des variables libres. Plus précisément, soient $\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_k}$ les k variables libres, et $\lambda_{j'_1}, \lambda_{j'_2}, \ldots, \lambda_{j'_{n-k}}$ les $n-k$ variables essentielles restantes. On a

$$\lambda_{j'_l} = -\lambda_{j_1} u_{l\,j_1} - \lambda_{j_2} u_{l\,j_2} - \cdots - \lambda_{j_k} u_{l\,j_k},$$

où u_{ij} est le coefficient générique de la matrice U. Par conséquent, pour tout $x \in \ker f$, on a

$$(x)_{\mathcal{B}_1} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_{j_1} w_1 + \lambda_{j_2} w_2 + \cdots + \lambda_{j_k} w_k,$$

où les vecteurs $w_1, w_2, \ldots, w_k \in \mathbb{R}^n$ ont pour lignes

$$(w_p)_i = \begin{cases} -u_{l\,j_p} & \text{si } i = j'_l \text{ correspond à une variable essentielle,} \\ 1 & \text{si } i = j_p \text{ correspond à la } p\text{-ème variable libre,} \\ 0 & \text{si } i = j_{p'} \text{ correspond à toute autre variable libre.} \end{cases}$$

(Cette construction est plus facile à comprendre sur un exemple, voir ci-après). C'est une famille libre en raison de la disposition des 1 et des 0 correspondant aux lignes associées aux variables libres. Par ailleurs, comme $x \mapsto (x)_{\mathcal{B}_1}$ est un isomorphisme de E_1 dans $\mathbb{R}^n$, il existe un unique k-uplet de vecteurs de E_1, z_p, $p = 1, \ldots, k$, tels que $w_p = (z_p)_{\mathcal{B}_1}$, et cette famille est libre. On en déduit que tout $x \in \ker f$ s'écrit $x = \lambda_{j_1} z_1 + \lambda_{j_2} z_2 + \cdots + \lambda_{j_k} z_k$. On a donc affaire à une famille génératrice de $\ker f$, c'en est donc une base. Par conséquent, on a bien $\dim\ker f = k$. □

Remarque 28 Cette résolution de système homogène donne également une construction d'une base de $\ker f$, ce qui détermine complètement ce sous-espace vectoriel. □

Exemple 20 Considérons l'application linéaire f de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par $x \mapsto Ax$ avec
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$
La réduction de Gauss de cette matrice est immédiate et conduit à la matrice échelonnée réduite
$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$
Il y a deux variables essentielles λ_1 et λ_2 et une variable libre λ_3. Par conséquent, $\ker f$ est une droite vectorielle de $\mathbb{R}^3$. Pour en donner une base, on résout le système homogène échelonné réduit. Il vient
$$\lambda_1 = 0, \quad \lambda_2 = -\lambda_3.$$
Pour tout $x \in \ker f$, on peut donc écrire
$$(x)_{\mathcal{B}_1} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\lambda_3 \\ \lambda_3 \end{pmatrix} = \lambda_3 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \lambda_3 w_1.$$
Remarquer le coefficient 1 à la troisième ligne correspondant à la variable libre. Les coefficients 0 et -1 sont des coefficients de la forme $-u_{l j_p}$ dans le raisonnement précédent. Une base du noyau est donc donnée par le vecteur $w_1 \in \mathbb{R}^3$. □

Pour ce qui concerne l'image, on a la proposition suivante.

Proposition 36 *Les images par f des vecteurs de $\mathcal{B}_1$ correspondant aux colonnes de pivot de A forment une base de $\operatorname{im} f$ et le rang de f est égal au nombre de variables essentielles de A.*

Démonstration. Si $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$, on a déjà vu que la famille $\{f(v_1), f(v_2), \dots, f(v_n)\}$ engendre $\operatorname{im} f$. Par le théorème de la base extraite, on peut en supprimer des vecteurs pour fabriquer une base.

Comme l'application $x \mapsto (x)_{\mathcal{B}_1}$ est un isomorphisme de E_1 dans $\mathbb{R}^n$, elle conserve les relations de dépendances linéaires. Par ailleurs, quand on a deux matrices équivalentes au sens de l'algorithme de Gauss, on rappelle que leurs colonnes ont également exactement les mêmes relations de dépendance linéaire.

Or on voit clairement sur la forme échelonnée réduite que les colonnes de pivot sont linéairement indépendantes, donc les $f(v_j)$ correspondants forment aussi une famille libre.

De plus, on voit aussi sur la forme échelonnée réduite que les autres colonnes sont combinaison linéaire des colonnes de pivot. On peut donc les supprimer sans changer l'espace vectoriel engendré, et en faire de même parmi les $f(v_j)$. Les images des vecteurs de $\mathcal{B}_1$ correspondant aux colonnes de pivot de A forment encore une famille génératrice de $\operatorname{im} f$, donc une base. Leur nombre, soit la dimension de $\operatorname{im} f$, est égal au nombre de variables essentielles. □

Un corollaire très important de ces remarques est le fameux *théorème du rang.*

Théorème 40 *Si $f \in \mathcal{L}(E_1, E_2)$, alors*

$$\dim \ker f + \operatorname{rg} f = \dim E_1.$$

Démonstration. Comme les variables libres et essentielles partitionnent l'ensemble des colonnes, s'il y a k variables essentielles, il y a forcément $n-k$ variables libres, d'où le résultat. □

Remarque 29 Dans le théorème du rang, il faut bien se rappeler que c'est la dimension de l'espace de DÉPART qui compte. Celle de l'espace d'arrivée E_2 n'a aucune importance. □

Corollaire 41 *Le noyau d'une forme linéaire non nulle est un hyperplan de $\mathbb{R}^m$ et tout hyperplan est le noyau d'une forme linéaire.*

Démonstration. En effet, une forme linéaire ℓ non nulle est surjective, donc de rang 1. Par conséquent, $\dim \ker \ell = m-1$.

Réciproquement, si H est un hyperplan de $\mathbb{R}^m$, on prend une base $\{v_1, \dots, v_{m-1}\}$ de H que l'on complète en une base de $\mathbb{R}^m$ en ajoutant un vecteur v_m, par le théorème de la base incomplète. L'application qui à x associe sa m-ème composante dans cette base est une forme linéaire dont le noyau est H. □

Exemple 21 En dimension $m = 2$, on représente donc toute droite vectorielle à l'aide d'une équation linéaire

$$a_{11}x_1 + a_{12}x_2 = 0.$$

En dimension $m = 3$, on représente donc tout plan vectoriel à l'aide d'une équation linéaire

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0.$$

Comme l'intersection de deux plans vectoriels non confondus est une droite, on représente une droite vectorielle dans $\mathbb{R}^3$ par deux équations linéaires

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0. \end{cases}$$

Aucune de ces représentations n'est unique. Si par ailleurs, on résout les systèmes linéaires homogènes, on obtient des bases de ces droites et plans. $\square$

5.4 Le groupe linéaire

Terminons ce chapitre par quelques remarques spécifiques au cas des endomorphismes.

Définition 39 *Soit E un sev de $\mathbb{R}^m$. L'ensemble des endomorphismes bijectifs de E est appelé* groupe linéaire de E *et est noté $GL(E)$.*

Proposition 37 *$(GL(E), \circ)$ est un groupe (non commutatif).*

Démonstration. En effet, la composition des applications est associative, son élément neutre — l'application identité $x \mapsto x$ — est un isomorphisme de E sur E, et si $f \in GL(E)$, on a déjà vu que son application réciproque f^{-1} est aussi un isomorphisme. Or c'est l'inverse pour la composition. $\square$

Quand on considère un endomorphisme d'un sev, on peut lui associer une matrice pour tout choix d'un couple de bases, comme dans le cas général. Comme l'espace de départ et l'espace d'arrivée coïncident, il n'y a souvent aucune raison particulière de choisir deux bases différentes, et on parle alors de matrice dans une base.

Théorème 42 *Soit E un sev de dimension k et $u \in \mathcal{L}(E)$. Alors les propriétés suivantes sont équivalentes :*

i) $u \in GL(E)$.

ii) $\operatorname{rg} u = k$.

iii) u est injective.

iv) u est surjective.

v) La matrice A de u dans une base admet un pivot par ligne.

vi) La matrice A de u dans une base admet un pivot par colonne.

vii) $A \sim I_k$.

Démonstration. La plupart des équivalences sont évidentes après tout ce que l'on a vu précédemment. Attardons nous sur ii), iii) et iv) et leur équivalence avec i).

ii) Si $\operatorname{rg} u = k = \dim E$, alors u est surjective. De plus, par le théorème du rang, $\dim \ker u = k - k = 0$, donc $\ker u = \{0\}$, donc u est injective. Par conséquent, on a bien i).

iii) Si u est surjective, alors $\operatorname{rg} u = k$, c'est-à-dire ii).

iv) Si u est injective, alors $\dim \ker = 0$, donc par le théorème du rang, $0 + \operatorname{rg} f = k$, c'est-à-dire ii). □

Remarque 30 Pour montrer qu'un endomorphisme d'un sev est un isomorphisme, il suffit de montrer qu'il est soit injectif, soit surjectif, mais il est inutile de montrer les deux. Plus généralement, comme deux sev de même dimension sont isomorphes, ceci reste vrai pour une application linéaire entre deux tels sev.

ATTENTION, ceci est FAUX si l'on considère une application linéaire entre deux sev de dimensions différentes. Il n'existe bien sûr aucun isomorphisme entre ces deux sev dans ce cas. □

Remarque 31 Encore une fois, on peut remplacer dans tout ce chapitre $\mathbb{R}$ par $\mathbb{C}$ sans autre changement. □

Chapitre 6

Algèbre matricielle

En plus d'être des tableaux de nombres susceptibles d'être manipulés par des algorithmes pour la résolution des systèmes linéaires et des outils de calcul pour les applications linéaires, les matrices sont également munies de diverses structures algébriques.

6.1 Opérations linéaires sur les matrices

Tout d'abord quelques points de notation et de vocabulaire. L'ensemble des matrices $m \times n$ à coefficients dans $\mathbb{R}$ (ou $\mathbb{C}$) sera noté $\mathcal{M}_{mn}(\mathbb{R})$ (ou $\mathcal{M}_{mn}(\mathbb{C})$). Dans le cas où $m = n$, on dit que l'on a affaire à des *matrices carrées* et l'on note simplement $\mathcal{M}_m(\mathbb{R})$.

Dans une matrice carrée $A = (a_{ij})$ de taille $m \times m$, les coefficients de la forme a_{ii} pour $i = 1, 2, \dots, m$ forment ce que l'on appelle la *diagonale*, ou diagonale principale, de la matrice. Une matrice carrée A telle que $a_{ij} = 0$ dès que $i \neq j$, c'est-à-dire hors de la diagonale, est appelée *matrice diagonale*.

Une matrice carrée A telle que $a_{ij} = 0$ dès que $i > j$, c'est-à-dire en dessous de la diagonale, est appelée *matrice triangulaire supérieure*. Les matrices carrées échelonnées sont triangulaires supérieures. Une matrice carrée A telle que $a_{ij} = 0$ dès que $i < j$, c'est-à-dire au dessus de la diagonale, est appelée *matrice triangulaire inférieure*.

$$
\underset{\text{diagonale}}{\begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & \boxed{a_{22}} & a_{23} & \cdots & a_{2m} \\ a_{31} & a_{32} & \boxed{a_{33}} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & \boxed{a_{mm}} \end{pmatrix}}
\qquad
\underset{\text{matrice diagonale}}{\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{mm} \end{pmatrix}}
$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ 0 & a_{22} & a_{23} & \cdots & a_{2m} \\ 0 & 0 & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{mm} \end{pmatrix} \qquad \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mm} \end{pmatrix}$$

matrice triangulaire supérieure   matrice triangulaire inférieure

Définition 40 *On définit sur $\mathcal{M}_{mn}(\mathbb{R})$ une loi de composition interne appelée addition par :*

Pour tous $A = (a_{ij})$, $B = (b_{ij})$ dans $\mathcal{M}_{mn}(\mathbb{R})$, $A+B$ est la matrice de $\mathcal{M}_{mn}(\mathbb{R})$ dont les coefficients sont donnés par $(A+B)_{ij} = a_{ij} + b_{ij}$, pour $i = 1,\dots,m$ et $j = 1,\dots,n$,

et une loi de composition externe appelée multiplication par un scalaire par :

Pour tout $A = (a_{ij})$ dans $\mathcal{M}_{mn}(\mathbb{R})$ et λ dans $\mathbb{R}$, λA est la matrice de $\mathcal{M}_{mn}(\mathbb{R})$ dont les coefficients sont donnés par $(\lambda A)_{ij} = \lambda a_{ij}$, pour $i = 1,\dots,m$ et $j = 1,\dots,n$.

Remarque 32 Il s'agit de la généralisation immédiate à toutes les matrices des opérations déjà définies sur les matrices à une colonne. En fait, si l'on écrit les matrices sous forme de ligne de colonnes, $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}$ avec $a_j, b_j \in \mathbb{R}^m$, on a

$$A+B = \begin{pmatrix} a_1+b_1 & a_2+b_2 & \cdots & a_n+b_n \end{pmatrix} \text{ et } \lambda A = \begin{pmatrix} \lambda a_1 & \lambda a_2 & \cdots & \lambda a_n \end{pmatrix}.$$

On en déduit immédiatement par la définition du produit matrice-vecteur que $(A+B)x = Ax+Bx$ et $(\lambda A)x = \lambda(Ax)$, pour tout $x \in \mathbb{R}^n$. □

En particulier, on montre très facilement le résultat suivant.

Proposition 38 *L'ensemble $\mathcal{M}_{mn}(\mathbb{R})$ muni de ces deux opérations est un espace vectoriel sur $\mathbb{R}$.*

L'élément neutre pour l'addition est la matrice $m \times n$ nulle, notée 0, dont tous les coefficients sont nuls.

Introduisons mn matrices particulières $E_{kl} \in \mathcal{M}_{mn}(\mathbb{R})$ définies par

$$(E_{kl})_{ij} = \begin{cases} 1 & \text{si } i = k \text{ et } j = l, \\ 0 & \text{sinon.} \end{cases}$$

Proposition 39 *La famille $\{E_{kl}, k = 1,\dots,m, l = 1,\dots,n\}$ forme une base de $\mathcal{M}_{mn}(\mathbb{R})$, appelée base canonique de $\mathcal{M}_{mn}(\mathbb{R})$. En particulier,* $\dim \mathcal{M}_{mn}(\mathbb{R}) = mn$.

Définition 41 *Si $A \in \mathcal{M}_{mn}(\mathbb{R})$, on définit sa matrice* transposée $A^T \in \mathcal{M}_{nm}(\mathbb{R})$ *par $(A^T)_{ij} = a_{ji}$ pour $i = 1,\dots,n$ et $j = 1,\dots,m$.*

La notation traditionnelle française pour la transposée d'une matrice A est ${}^tA$. Malheureusement, cette notation est nettement plus pénible à taper en TEX que la notation anglo-saxonne A^T (`$^t\!A$` au lieu de `$A^T$`), que l'on préférera donc par pure paresse dans la suite de ces notes.

Exemple 22 La transposition opère une symétrie par rapport à la diagonale. Ainsi

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \Longrightarrow A^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Elle échange le rôle des lignes et des colonnes. □

Proposition 40 *L'application de transposition $A \mapsto A^T$ est un isomorphisme de $\mathcal{M}_{mn}(\mathbb{R})$ dans $\mathcal{M}_{nm}(\mathbb{R})$, et l'on a $(A^T)^T = A$ pour tout A.*

6.2 Produit matriciel

La grande nouveauté est que l'on peut multiplier certaines matrices entre elles.

Définition 42 *Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$ et $B \in \mathcal{M}_{np}(\mathbb{R})$. Notons $b_1, b_2, \dots, b_p \in \mathbb{R}^n$ les p colonnes de B. On définit le* produit *AB comme étant la matrice $m \times p$ dont les colonnes sont données par*

$$AB = \begin{pmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{pmatrix}.$$

Remarque 33 Comme A est une matrice $m \times n$ et les b_i sont des vecteurs de $\mathbb{R}^n$, chaque produit matrice-vecteur Ab_i est bien défini et fournit un vecteur de $\mathbb{R}^m$. Le résultat du produit est donc bien une matrice à m lignes et p colonnes.

Retenons donc que le produit *AB n'est défini que quand le nombre de colonnes de A est égal au nombre de lignes de B, sinon il n'est pas défini.* En particulier, on ne peut toujours pas multiplier un vecteur-colonne par un autre vecteur-colonne (sauf s'ils n'ont qu'une ligne). Quand le produit AB est défini, son nombre de lignes est égal au nombre de lignes du premier facteur A et son nombre de colonnes est égal à celui du second facteur B.

Remarquons enfin que si $B \in \mathcal{M}_{n1}(\mathbb{R}) = \mathbb{R}^n$, alors AB n'est autre que le produit matrice-vecteur habituel.

Exemple 23 Si $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, alors le produit AB est bien défini, c'est une matrice 2×2, $AB = \begin{pmatrix} Ab_1 & Ab_2 \end{pmatrix}$, avec

$$Ab_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \text{ et } Ab_2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$

d'où

$$AB = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$$

dans ce cas.

Par contre, si $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, alors le produit AB n'est pas défini, car le nombre de colonnes de A (=3) n'est pas égal au nombre de lignes de B (=4). Il est impossible d'effectuer les produits matrice-vecteur de la définition de AB. □

Proposition 41 *Soient A et B comme dans la Définition* 42*. Alors, pour tout $x \in \mathbb{R}^p$, on a*

$$(AB)x = A(Bx).$$

Démonstration. Notons d'abord que tous les produits matriciels et matrice-vecteur de cette formule sont bien définis.

Si $x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix}$, alors on a par définition du produit matriciel et du produit matrice-vecteur

$$(AB)x = \lambda_1(Ab_1) + \lambda_2(Ab_2) + \cdots + \lambda_p(Ab_p).$$

D'un autre côté, toujours par définition du produit matrice-vecteur,

$$Bx = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_p b_p,$$

d'où par linéarité du produit matrice-vecteur

$$A(Bx) = A(\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_p b_p) = \lambda_1(Ab_1) + \lambda_2(Ab_2) + \cdots + \lambda_p(Ab_p).$$

On en déduit le résultat. □

Le produit matriciel possède les propriétés algébriques suivantes.

Théorème 43 *On a (en supposant les produits matriciels définis)*

i) $(A+A')B = AB + A'B$ *et* $A(B+B') = AB + AB'$.

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

iii) $(AB)C = A(BC)$ *(associativité).*

Démonstration. i) Comme $(A+A')b_j = Ab_j + A'b_j$, la première relation est vraie. De même, les colonnes de $B+B'$ sont les $b_j + b'_j$, et l'on a $A(b_j + b'_j) = Ab_j + Ab'_j$, d'où la deuxième relation, *cf.* Remarque 32.

ii) On utilise encore la Remarque 32.

iii) Assurons-nous d'abord que tous les produits sont bien définis. Pour définir AB, on doit avoir $A \in \mathcal{M}_{mn}(\mathbb{R})$ et $B \in \mathcal{M}_{np}(\mathbb{R})$. Dans ce cas, $AB \in \mathcal{M}_{mp}(\mathbb{R})$. Pour définir BC, on doit avoir $B \in \mathcal{M}_{np}(\mathbb{R})$ et $C \in \mathcal{M}_{pr}(\mathbb{R})$. Dans ce cas, $BC \in \mathcal{M}_{nr}(\mathbb{R})$. Par conséquent, le produit $(AB)C$ est bien défini (nb de colonnes de AB = nb de lignes de C) et $(AB)C \in \mathcal{M}_{mr}(\mathbb{R})$. De même, le produit $A(BC)$ est bien défini et appartient à $\mathcal{M}_{mr}(\mathbb{R})$. En particulier, $(AB)C$ et $A(BC)$ sont deux matrices de même taille, ce qui est la moindre des choses si elles sont censées être égales.

Montrons maintenant qu'elles sont bien égales. Pour cela, on écrit C comme une ligne de vecteurs

$$C = \begin{pmatrix} c_1 & c_2 & \cdots & c_r \end{pmatrix}, c_k \in \mathbb{R}^p.$$

Il vient

$$\begin{aligned}(AB)C &= \begin{pmatrix} (AB)c_1 & (AB)c_2 & \cdots & (AB)c_r \end{pmatrix} \\ &= \begin{pmatrix} A(Bc_1) & A(Bc_2) & \cdots & A(Bc_r) \end{pmatrix}\end{aligned}$$

par la Proposition 41

$$\begin{aligned} &= A\begin{pmatrix} Bc_1 & Bc_2 & \cdots & Bc_r \end{pmatrix} \\ &= A\left(B\begin{pmatrix} c_1 & c_2 & \cdots & c_r \end{pmatrix}\right) = A(BC).\end{aligned}$$

□

Remarque 34 Comme le produit matriciel est associatif, il est inutile d'utiliser des parenthèses du moment que tous les produits intermédiaires soient bien définis. □

Les matrices identités sont des éléments neutres à gauche et à droite pour le produit matriciel.

Proposition 42 *Pour tout* $A \in \mathcal{M}_{mn}(\mathbb{R})$*, on a* $I_m A = A I_n = A$.

Démonstration. Évident car $I_m a_i = a_i$ et $Ae_i = a_i$ pour $i = 1, \ldots, n$. □

On a une formule générale pour l'élément générique du produit de deux matrices.

Proposition 43 *Soit $A = (a_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$ et $B = (b_{jk}) \in \mathcal{M}_{np}(\mathbb{R})$. Alors on a*

$$(AB)_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}$$

pour $i = 1, \ldots, m, k = 1, \ldots, p$.

Démonstration. Il suffit de l'écrire : $(AB)_{ik}$ est la ligne i du vecteur Ab_k. □

Cette formule est assez mnémotechnique, surtout sous sa forme condensée avec le signe somme. Noter la position de l'indice de sommation j, qui est l'indice répété dans cette formule.

Proposition 44 *Si le produit AB est défini, alors le produit $B^T A^T$ est aussi défini et l'on a $(AB)^T = B^T A^T$.*

Démonstration. Soit $A = (a_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$ et $B = (b_{jk}) \in \mathcal{M}_{np}(\mathbb{R})\mathcal{M}_{mp}(\mathbb{R})$, d'où $AB \in \mathcal{M}_{mp}(\mathbb{R})$. On voit donc que $B^T \in \mathcal{M}_{pn}(\mathbb{R})$ et $A^T \in \mathcal{M}_{nm}(\mathbb{R})\mathcal{M}_{np}(\mathbb{R})$. Par conséquent, $B^T A^T$ est bien défini et de la même taille que $(AB)^T$.

Utilisons la formule générale ci-dessus.

$$(B^T A^T)_{ik} = \sum_{j=1}^{n} (B^T)_{ij}(A^T)_{jk} = \sum_{j=1}^{n} b_{ji}a_{kj} = \sum_{j=1}^{n} a_{kj}b_{ji} = (AB)_{ki} = ((AB)^T)_{ik}$$

d'où le résultat. □

Remarque 35 ATTENTION ! Le produit matriciel *n'est pas commutatif*. En effet, il peut se faire que AB soit défini mais pas BA, ou que AB et BA soient tous deux définis mais pas de la même taille. Mais même dans le cas où AB et BA sont définis et de la même taille, on a en général $AB \neq BA$. Considérons l'exemple suivant.

$$\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 3 \\ -2 & -6 \end{pmatrix} \text{ mais } \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 29 & -2 \end{pmatrix}.$$

C'est là la situation générale. L'ordre des facteurs dans un produit matriciel ne doit donc *jamais* être modifié, sous peine de fausser le résultat (sauf si l'on sait que l'on est dans un cas particulier où deux matrices *commutent*, c'est-à-dire sont telles que $AB = BA$. Mais c'est rare...).

En fait, le produit matriciel est le premier exemple que l'on rencontre de produit non commutatif. □

Le calcul pratique d'un produit matriciel se fait en remarquant qu'il ne s'agit que d'une suite de produits matrice-vecteur. Il est commode quand on débute de disposer les calculs de la façon suivante.

$$
\begin{array}{ccc}
 & \begin{pmatrix} & & & \times & & \\ & & & \times & & \\ & & & \times & & \\ & & & \times & & \end{pmatrix} & \leftarrow B \\
A \rightarrow \begin{pmatrix} & & & \\ & & & \\ \times & \times & \times & \times \\ & & & \end{pmatrix} & \begin{pmatrix} & & & | & & \\ & & & | & & \\ - & - & - & | & - & - \\ & & & | & & \end{pmatrix} & \leftarrow AB
\end{array}
$$

On a fait apparaître une ligne générique de A et une colonne générique de B avec les coefficients qui doivent multipliés les uns avec les autres (représentés par des $\times$ dans l'ordre de gauche à droite dans A et de haut en bas dans B) puis additionnés pour donner le coefficient de AB situé à l'intersection de cette ligne et de cette colonne.

Avec un peu plus de pratique, on pose directement l'opération en ligne comme dans l'exemple ci-dessus.

Remarque 36 *Deux erreurs grossières à éviter.* Les règles du calcul des produits de matrices diffèrent de celles des produits dans un corps par d'autres aspects.

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

ii) Si $AB = 0$, on ne peut pas en déduire que soit $A = 0$ soit $B = 0$. C'est faux en général.

Si on a un exemple de ii), on a aussi un exemple de i) puisque $0 = A \times 0$. Il suffit de prendre

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \text{ Alors, } AB = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

mais ni A ni B n'est nulle. $\square$

6.3 Cas où $m = n$, matrices inversibles

Théorème 44 $(\mathcal{M}_n(\mathbb{R}), +, \times)$ *est un anneau unitaire, non commutatif, non intègre.*

Démonstration. En effet, dans le cas $m = n$, la multiplication des matrices est une opération interne dans $\mathcal{M}_n(\mathbb{R})$. Les axiomes d'anneau sont satisfaits. La matrice

I_n est élément neutre pour la multiplication, donc c'est un anneau unitaire. La multiplication n'est pas commutative. Il y a des diviseurs de zéro, (*cf.* l'exemple ci-dessus), donc cet anneau n'est pas intègre. □

Les opérations usuelles d'un anneau sont également définies ici.

Définition 43 *Pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on définit les puissances successives de A par $A^0 = I_n$ et $A^{p+1} = AA^p = A^pA$ pour tout $p \in \mathbb{N}$.*

Pour $p \geq 1$, A^p est le produit de A par elle-même p fois. On définit alors pour tout $P \in \mathbb{R}[X]$ polynôme à une indéterminée sur $\mathbb{R}$, $P(X) = a_0 + a_1X + \cdots a_pX^p$, la matrice $P(A) = a_0I_n + a_1A + \cdots a_pA^p \in \mathcal{M}_n(\mathbb{R})$.

Remarque 37 Comme la multiplication n'est pas commutative, les identités binomiales usuelles sont fausses. En particulier, $(A+B)^2 \neq A^2 + 2AB + B^2$, mais bien $(A+B)^2 = A^2 + AB + BA + B^2$. □

Définition 44 *On dit que $A \in \mathcal{M}_n(\mathbb{R})$ est* inversible *si et seulement si il existe une autre matrice $A' \in \mathcal{M}_n(\mathbb{R})$ telle que $AA' = A'A = I_n$. Dans le cas contraire, on dit qu'elle est* singulière. *L'ensemble $GL_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}), A \text{ inversible}\}$ est un groupe pour la multiplication.*

Ce sont les notions usuelles dans un anneau unitaire. On note bien sûr $A' = A^{-1}$, et plus généralement $A^{-p} = (A^{-1})^p$ pour tout $p \in \mathbb{N}$ quand A est inversible.

Théorème 45 *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Le système linéaire $n \times n$, $Ax = b$ admet une solution et une seule pour tout b si et seulement si A est inversible. Dans ce cas, on a $x = A^{-1}b$.*

Démonstration. Si A est inversible, alors $Ax = b$ implique en multipliant par A^{-1} que $A^{-1}Ax = A^{-1}b$, c'est-à-dire $x = I_nx = A^{-1}b$, c'est-à-dire l'unicité de la solution. Vérifions l'existence : on a bien $A(A^{-1}b) = (AA^{-1})b = I_nb = b$, d'où l'existence.

Réciproquement, supposons que le système $Ax = b$ admette une solution et une seule pour tout $b \in \mathbb{R}^n$. Ceci signifie que l'application linéaire $f : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Ax$, est un isomorphisme. L'application inverse f^{-1} admet une matrice A' dans la base canonique et l'on a $f^{-1}(x) = A'x$ (x est égal dans ce cas à la colonne des ses composantes dans la base canonique). Par conséquent, comme $f \circ f^{-1} = id$, il vient $x = f \circ f^{-1}(x) = A(A'x) = (AA')x$ pour tout $x \in \mathbb{R}^n$. Prenant $x = e_i$ les vecteurs de la bases canonique, on en déduit que $AA' = I_n$. De même, le fait que $f^{-1} \circ f = id$ implique que $A'A = I_n$ et A est inversible. □

Remarque 38 Résolvons le système $Av_i = e_i$, c'est-à-dire $v_i = A^{-1}e_i$. On en déduit que $A^{-1} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$. Donc, le calcul de A^{-1} se ramène à la résolution de n systèmes linéaires $n \times n$, que l'on peut effectuer par la méthode de Gauss de la façon suivante : on pose $\tilde{A} = \begin{pmatrix} A & I_n \end{pmatrix}$ matrice $n \times 2n$. Alors la matrice échelonnée réduite équivalente à $\tilde{A}$ n'est autre que $\tilde{U} = \begin{pmatrix} I_n & A^{-1} \end{pmatrix}$. □

Théorème 46 *Si $A, B \in GL_n(\mathbb{R})$, alors $(AB)^{-1} = B^{-1}A^{-1}$ et $A^T \in GL_n(\mathbb{R})$, avec $(A^T)^{-1} = (A^{-1})^T$.*

Démonstration. La formule pour l'inverse du produit a lieu dans n'importe quel groupe. Pour la transposée, on remarque que comme $I_n = AA^{-1}$ et que $I_n^T = I_n$, on a

$$I_n = (AA^{-1})^T = (A^{-1})^T A^T.$$

De même, le fait que $I_n = A^{-1}A$ implique que $I_n = A^T(A^{-1})^T$, donc A^T est inversible et on a son inverse, la transposée de A^{-1}. □

Une dernière propriété évidente découlant du Théorème 45 en identifiant une matrice A à l'application linéaire $x \mapsto Ax$.

Proposition 45 *On a $A \in GL_n(\mathbb{R})$ si et seulement si $\operatorname{rg} A = n$, si et seulement si $\ker A = \{0\}$.*

6.4 Le lien avec les applications linéaires

On vérifie facilement que l'ensemble des applications linéaires de E_1 dans E_2 peut être muni d'une addition

$$\forall f, g \in \mathcal{L}(E_1, E_2), \quad (f+g)(x) = f(x) + g(x)$$

et d'une multiplication par un scalaire

$$\forall f \in \mathcal{L}(E_1, E_2), \forall \lambda \in \mathbb{R}, \quad (\lambda f)(x) = \lambda f(x),$$

qui en font un *espace vectoriel sur* $\mathbb{R}$.

Soient donc deux espaces vectoriels E_1 et E_2 de dimension finie et $\mathcal{B}_i$ une base de E_i.

Théorème 47 *Soient $f, g \in \mathcal{L}(E_1, E_2)$ et $\lambda \in \mathbb{R}$, A la matrice de f dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$, et B celle de g. Alors la matrice de $f+g$ dans ces mêmes bases est $A+B$ et celle de λf est λA.*

Démonstration. Immédiat en utilisant les définitions et la linéarité des composantes dans une base. □

Ajoutons maintenant un troisième sev E_3 muni d'une base $\mathcal{B}_3$.

Théorème 48 *Soient $f \in \mathcal{L}(E_1, E_2)$ de matrice A dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$, et $g \in \mathcal{L}(E_2, E_3)$ de matrice B dans les bases $\mathcal{B}_2$ et $\mathcal{B}_3$. Alors, la matrice de $g \circ f$ dans les bases $\mathcal{B}_1$ et $\mathcal{B}_3$ est BA.*

Démonstration. Soit C la matrice de $g \circ f$. Elle est déterminée par le fait que

$$\forall x \in E_1, \quad (g \circ f(x))_{\mathcal{B}_3} = C(x)_{\mathcal{B}_1}.$$

Par ailleurs, on a

$$(f(x))_{\mathcal{B}_2} = A(x)_{\mathcal{B}_1} \text{ et } (g(y))_{\mathcal{B}_3} = B(y)_{\mathcal{B}_2}.$$

Donc

$$(g(f(x)))_{\mathcal{B}_3} = B(f(x))_{\mathcal{B}_2} = B(A(x)_{\mathcal{B}_1}) = (BA)(x)_{\mathcal{B}_1}$$

d'où l'identification de $C = BA$. □

Donc la multiplication des matrices n'est que la traduction dans des bases de la composition des applications linéaires. On comprend mieux pourquoi elle n'est pas toujours définie, pourquoi elle est associative et pourquoi elle n'est pas commutative.

Dans le cas d'un seul espace $E_1 = E_2 = E$, on en déduit le corollaire suivant.

Proposition 46 *On a $f \in GL(E)$ si et seulement si sa matrice A dans une base appartient à $GL_n(\mathbb{R})$. Dans ce cas, la matrice de f^{-1} dans la même base est A^{-1}.*

Démonstration. On a que $f \in GL(E)$ si et seulement si il existe $f^{-1} \in \mathcal{L}(E)$ tel que $f \circ f^{-1} = f^{-1} \circ f = id$. Soit $\mathcal{B}$ 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. □

Le rang d'une matrice $m \times n$ est le rang de l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ qui lui est associée. Ce rang est aussi celui de n'importe quelle application linéaire représentée par cette même matrice dans un choix de bases. On en déduit que $\mathrm{rg}(BA) \leq \min(\mathrm{rg}\, B, \mathrm{rg}\, A)$.

6.5 Changement de bases

On a déjà vu le changement de bases pour les composantes dans un même sev de dimension n

$$(x)_{\mathcal{B}} = \mathcal{P}_{\mathcal{B} \to \mathcal{B}'}(x)_{\mathcal{B}'}.$$

Proposition 47 *On a $\mathcal{P}_{\mathcal{B}\to\mathcal{B}'} \in GL_n(\mathbb{R})$ et $(\mathcal{P}_{\mathcal{B}\to\mathcal{B}'})^{-1} = \mathcal{P}_{\mathcal{B}'\to\mathcal{B}}$.*

Démonstration. Évident d'après la relation précédente appliquée également au passage de $\mathcal{B}'$ à $\mathcal{B}$. □

Considérons maintenant le même problème pour une application linéaire et sa matrice dans plusieurs choix de bases.

Théorème 49 *Soit E_1 un sev muni de deux bases $\mathcal{B}_1$ et $\mathcal{B}_1'$, E_2 de même muni de deux bases $\mathcal{B}_2$ et $\mathcal{B}_2'$, f une application linéaire de E_1 dans E_2, A sa matrice dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ et A' sa matrice dans les bases $\mathcal{B}_1'$ et $\mathcal{B}_2'$. Alors on a la* formule de changement de bases

$$A' = \mathcal{P}_{\mathcal{B}_2'\to\mathcal{B}_2} A \mathcal{P}_{\mathcal{B}_1\to\mathcal{B}_1'}.$$

Démonstration. On applique les divers résultats connus :

$$(x)_{\mathcal{B}_1} = \mathcal{P}_{\mathcal{B}_1\to\mathcal{B}_1'}(x)_{\mathcal{B}_1'}, \quad (y)_{\mathcal{B}_2} = \mathcal{P}_{\mathcal{B}_2\to\mathcal{B}_2'}(y)_{\mathcal{B}_2'},$$

pour les changements de base et

$$(f(x))_{\mathcal{B}_2} = A(x)_{\mathcal{B}_1}, \quad (f(x))_{\mathcal{B}_2'} = A'(x)_{\mathcal{B}_1'},$$

pour les matrices d'applications linéaires. Par conséquent,

$$\mathcal{P}_{\mathcal{B}_2\to\mathcal{B}_2'}(f(x))_{\mathcal{B}_2'} = A\mathcal{P}_{\mathcal{B}_1\to\mathcal{B}_1'}(x)_{\mathcal{B}_1'}.$$

Multipliant cette égalité à gauche par $\mathcal{P}_{\mathcal{B}_2'\to\mathcal{B}_2} = (\mathcal{P}_{\mathcal{B}_2\to\mathcal{B}_2'})^{-1}$, on obtient

$$(f(x))_{\mathcal{B}_2'} = \mathcal{P}_{\mathcal{B}_2'\to\mathcal{B}_2} A \mathcal{P}_{\mathcal{B}_1\to\mathcal{B}_1'}(x)_{\mathcal{B}_1'} = A'(x)_{\mathcal{B}_1'},$$

pour tout $x \in E_1$, d'où le résultat voulu. □

Dans le cas particulier où $E_1 = E_2$, $\mathcal{B}_1 = \mathcal{B}_2$ et $\mathcal{B}_1' = \mathcal{B}_2'$, alors si l'on pose $P = \mathcal{P}_{\mathcal{B}_1\to\mathcal{B}_1'}$ alors la formule de changement de bases prend la forme plus compacte

$$A' = P^{-1}AP.$$

Définition 45 *On dit que deux matrices A et B de $\mathcal{M}_n(k)$ sont* semblables *s'il existe une matrice $P \in GL_n(\mathbb{R})$ telle que $B = P^{-1}AP$.*

En d'autres termes, deux matrices carrées sont semblables si et seulement si elles représentent la même application linéaire dans deux bases différentes. Il est facile de montrer qu'il s'agit *d'une relation d'équivalence.*

6.6 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 46 *On appelle* matrice élémentaire *toute matrice qui résulte de l'application d'une opération élémentaire sur les lignes à la matrice identité I_m.*

Exemple 24 Dans le cas 3×3,

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

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 48 *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. Écrivons la matrice A comme une ligne de n vecteurs de $\mathbb{R}^m$, soit $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$. Par définition du produit matriciel, il vient donc $EA = \begin{pmatrix} Ea_1 & Ea_2 & \cdots & Ea_n \end{pmatrix}$. 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 $\{e_1, e_2, \ldots, e_m\}$ 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. $\square$

Théorème 50 *Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$ et $U \in \mathcal{M}_{mn}(\mathbb{R})$ l'unique matrice échelonnée réduite qui lui est équivalente. Alors il existe une unique matrice $M \in GL_m(\mathbb{R})$ telle que*

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

Démonstration. 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

1ère étape : $A_1 = E_1A$.

2ème étape : $A_2 = E_2A_1 = E_2(E_1A) = (E_2E_1)A$.

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

$p$ème étape : $U = A_p = E_pA_{p-1} = (E_pE_{p-1}\cdots E_2E_1)A$.

On pose donc $M = E_pE_{p-1}\cdots E_2E_1$. Comme chacune des opérations élémentaires est inversible, chaque matrice élémentaire E_k appartient à $GL_m(\mathbb{R})$, d'où $M \in GL_m(\mathbb{R})$. $\square$

Remarque 39 Si la matrice M n'est pas très facile à calculer, on peut montrer que la matrice $M^{-1} = E_1^{-1}E_2^{-1}\cdots E_{p-1}^{-1}E_p^{-1}$ 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. $\square$

Remarque 40 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 $\tilde{A} = \begin{pmatrix} A & I_m \end{pmatrix}$. En effet, $\tilde{U} = M\tilde{A} = \begin{pmatrix} MA & MI_m \end{pmatrix} = \begin{pmatrix} I_m & A^{-1} \end{pmatrix}$. $\square$

Chapitre 7

Déterminants

Ce chapitre ne concerne que *les matrices carrées*. On travaillera systématiquement dans $\mathcal{M}_n(\mathbb{R})$.

7.1 Définition et premières propriétés

Le déterminant est une application de $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$ qui a de nombreuses propriétés importantes, mais dont l'existence est un peu délicate. Nous allons donc admettre le théorème suivant.

Théorème 51 *Il existe une unique application de $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$, appelée* déterminant, *telle que*

i) Le déterminant est linéaire par rapport à chaque vecteur-colonne, les autres étant fixés.

ii) Si une matrice A a deux vecteurs colonnes égaux, alors son déterminant est nul.

iii) Le déterminant de la matrice identité I_n vaut 1.

On a plusieurs notations pour les déterminants :

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

On utilisera aussi la notation associée aux lignes de vecteurs $a_i \in \mathbb{R}^n$

$$\det A = \begin{vmatrix} a_1 & a_2 & \cdots & a_n \end{vmatrix}$$

Une application de $\mathcal{M}_n(\mathbb{R})$ dans $\mathbb{R}$ qui satisfait la propriété i) est appelée *forme multilinéaire.* Si elle satisfait ii), on dit qu'elle est *alternée.* Le déterminant est donc la seule forme multilinéaire alternée qui vaut 1 sur la matrice I_n. Les autres formes multilinéaires alternées sont les multiples scalaires du déterminant. On verra plus loin comment on peut calculer effectivement les déterminants.

Donnons maintenant quelques propriétés importantes du déterminant.

Proposition 49 *Soit A une matrice $n \times n$ et A' la matrice obtenue en échangeant deux colonnes distinctes de A. Alors on a* $\det A' = -\det A$.

Démonstration. Soit $A = \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix}$. On va échanger les colonnes i et j, ce qui donne la matrice $A' = \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix}$, où le vecteur a_j se retrouve en colonne i et le vecteur a_i en colonne j (on pris ici $i < j$, sans perte de généralité).

Introduisons alors une troisième matrice

$$\tilde{A} = \begin{pmatrix} a_1 & \cdots & a_i + a_j & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix}.$$

Cette matrice a deux colonnes distinctes égales, donc d'après ii)

$$\det \tilde{A} = 0.$$

D'un autre côté, nous pouvons développer ce déterminant en utilisant la propriété i) de multilinéarité, c'est-à-dire linéarité par rapport à chaque colonne. Ceci donne

$$\begin{aligned}
\det \tilde{A} &= \det\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix} + \det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix} \\
&= \det\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix} + \det\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_i & \cdots & a_n \end{pmatrix} \\
&\quad + \det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_j & \cdots & a_n \end{pmatrix} + \det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix} \\
&= \det A + 0 + 0 + \det A',
\end{aligned}$$

encore grâce à i) pour les deux déterminants du milieu. ☐

Proposition 50 *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. Soit $A = \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix}$ et donnons nous des scalaires λ_j, $j = 1, \ldots, n$, $j \neq i$. On pose

$$A' = \begin{pmatrix} a_1 & \cdots & a_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda_j a_j & \cdots & a_n \end{pmatrix}.$$

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

$$\det A' = \det A + \sum_{\substack{j=1\\ j\neq i}}^{n} \lambda_j \det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_n \end{pmatrix}.$$

Or chacun des déterminants apparaissant sous le signe de sommation est nul, puisqu'il concerne une matrice dont les colonnes i et j sont égales. □

Corollaire 52 *Si une colonne de A est combinaison linéaire des autres colonnes alors* $\det A = 0$.

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

7.2 Déterminants et matrices inversibles

Un des usages des déterminants est de caractériser les matrices inversibles.

Proposition 51 *Si A est une matrice triangulaire supérieure ou inférieure, alors on a*

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

Autrement dit, pour une matrice triangulaire, et seulement pour une telle matrice, le déterminant est égal au produit des termes diagonaux.

Démonstration. On traite le cas des matrices triangulaires supérieures, le cas des matrices triangulaires inférieures est identique. Soit donc

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Par linéarité par rapport à la première colonne, on a

$$\det A = a_{11} \begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

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

$$\det A = a_{11} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

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

$$\det A = a_{11}a_{22} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix},$$

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

$$\det A = a_{11}a_{22}a_{33}\cdots a_{nn} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = a_{11}a_{22}a_{33}\cdots a_{nn} \det I_n,$$

d'où le résultat par iii). □

Remarque 41 La façon de procéder doit rappeler l'algorithme de Gauss. C'est en fait le même argument mais avec des substitutions de colonne.

Notons que le résultat s'applique en particulier aux *matrices diagonales*, lesquelles sont à la fois triangulaires supérieures et triangulaires inférieures. □

Corollaire 53 *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. 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. □

Nous arrivons au théorème fondamental de cette section.

Théorème 54 *Une matrice A est inversible si et seulement si* $\det A \neq 0$.

Démonstration. 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. D'après le Corollaire 52, on en déduit que $\det A = 0$.

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 la matrice I_n. Il existe donc des matrices élémentaires E_j telles que

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

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 élémentaire 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

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

puis par récurrence que

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

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. □

Remarque 42 On obtient de la sorte que

$$\det A = \frac{1}{\lambda_p \cdots \lambda_2 \lambda_1} = \frac{1}{\det E_p \cdots \det E_2 \det E_1},$$

ce qui donne une méthode efficace pour calculer le déterminant d'une matrice en général, puisqu'elle ne nécessite pas beaucoup plus d'opérations que l'algorithme de Gauss (on montre que ce nombre d'opérations est de l'ordre de $2n^3/3$). Dans le cas d'une petite matrice 25×25, on calcule ainsi le déterminant en environ 10 000 opérations. C'est un peu trop pour un être humain, mais très peu pour un ordinateur. $\square$

On en déduit aussi une moitié du premier théorème admis.

Corollaire 55 *Le déterminant est unique.*

Démonstration. Soient D et D' deux formes multilinéaires alternées qui valent 1 sur I_n. Si A est inversible, on applique le raisonnement précédent à D et D', lequel utilise seulement les propriétés i), ii) et iii), et l'on trouve

$$D(A) = D(A') = \frac{1}{\lambda_p \cdots \lambda_2 \lambda_1}.$$

En effet, les coefficients λ_i viennent de l'application de l'algorithme de Gauss et sont les mêmes pour D et D'. Si A n'est pas inversible, ses colonnes sont liées, donc $0 = D(A) = D(A')$ également dans ce cas. $\square$

On tire de la démonstration précédente une autre propriété importante du déterminant, la multiplicativité.

Théorème 56 *On a*

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

Démonstration. 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

$$B(E_1^T E_2^T \cdots E_p^T) = I_n \text{ et } \det B = \frac{1}{\lambda_p \cdots \lambda_2 \lambda_1}.$$

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

$$\det(C(E_1^T E_2^T \cdots E_p^T)) = \lambda_p \cdots \lambda_2 \lambda_1 \det C.$$

Or

$$C(E_1^T E_2^T \cdots E_p^T) = A(B(E_1^T E_2^T \cdots E_p^T)) = A,$$

d'où le résultat dans ce cas.

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

Corollaire 57 *Si $A \in GL_n(\mathbb{R})$ alors* $\det(A^{-1}) = \frac{1}{\det A}$.

Démonstration. En effet, $AA^{-1} = I_n$, donc $\det A \det(A^{-1}) = \det I_n = 1$. □

Corollaire 58 *On a* $\det(A^T) = \det A$.

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{R})$. Par l'algorithme de Gauss, on a une factorisation $A = M^{-1}U$ avec U échelonnée réduite, donc en particulier triangulaire supérieure et $M^{-1} = E_1^{-1}E_2^{-1}\cdots E_{p-1}^{-1}E_p^{-1}$. Par conséquent, en transposant on a aussi $A^T = U^T(M^{-1})^T$ avec $(M^{-1})^T = (E_p^{-1})^T(E_{p-1}^{-1})^T\cdots(E_2^{-1})^T(E_1^{-1})^T$.

Utilisant la multiplicativité du déterminant, on en déduit

$$\begin{cases} \det A &= \dfrac{\det U}{\det E_1 \det E_2 \cdots \det E_p}, \\ \det A^T &= \dfrac{\det U^T}{\det E_1^T \det E_2^T \cdots \det E_p^T}. \end{cases}$$

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. □

Remarque 43 Tout ce que l'on a dit des déterminants à propos des colonnes est donc vrai pour les lignes. Ainsi, le déterminant est multilinéaire par rapport aux lignes, si une matrice a deux lignes égales, son déterminant est nul, on ne modifie pas un déterminant en ajoutant à une ligne une combinaison linéaire des autres lignes, etc. □

Corollaire 59 *Deux matrices semblables ont même déterminant.*

Démonstration. Soit $A' = P^{-1}AP$ avec $P \in GL_n(\mathbb{R})$. Par multiplicativité du déterminant, on en déduit que

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

puisque $\det P^{-1} = 1/\det P$. □

Terminons cette section par deux avertissements sans frais. D'une part

$$\boxed{\det(A+B) \neq \det A + \det B.}$$

L'exemple suivant

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

le montre amplement. Le déterminant *n'est pas* linéaire. En fait, il n'y a pas de formule simple pour exprimer le déterminant d'une somme de matrices (il y a des formules relativement compliquées). D'autre part

$$\boxed{\det(\lambda A) \neq \lambda \det A.}$$

Ici il y a une formule simple. En effet

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

puisque λI_n est diagonale et son déterminant est le produit de ses termes diagonaux, soit ici λ^n.

7.3 Cofacteurs, développements du déterminant

Définition 47 *Soit A une matrice $n \times n$ et A_{ij} la matrice $(n-1)\times(n-1)$ obtenue en effaçant la ligne i et la colonne j de A. On appelle* mineur de A relatif à a_{ij} *le déterminant $\Delta_{ij} = \det A_{ij}$. On appelle* cofacteur de A relatif à a_{ij} *le nombre $C_{ij} = (-1)^{i+j}\Delta_{ij}$.*

Le théorème suivant est en fait un résultat d'existence du déterminant, c'est pourquoi nous l'admettrons également.

Théorème 60 (Développement par rapport à la première ligne) *On a la formule suivante*

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\Delta_{1j} = \sum_{j=1}^{n} a_{1j}C_{1j}.$$

Remarque 44 Le développement par rapport à la première ligne permet de ramener le calcul d'un déterminant $n \times n$ à celui de n déterminants $(n-1)\times(n-1)$. Par récurrence descendante, on se ramène ainsi au calcul de $n!$ déterminants 1×1. Il faut remarquer que le nombre $n!$ croît extrêmement vite avec n. Ainsi, pour une matrice 25×25, on a $25! \approx 1{,}5 \times 10^{25}$. Un ordinateur téraflops, c'est-à-dire capable d'effectuer 10^{12} opérations en virgule flottante par seconde aurait besoin d'au moins 500 000 ans de fonctionnement ininterrompu pour effectuer ce calcul. Ce chiffre est à comparer avec le nombre d'opérations de la méthode de Gauss, dans ce cas de l'ordre de 10^5, que le même ordinateur effectuera en 0,1 millionnièmes de secondes. À titre d'information, l'ordinateur le plus puissant du monde en 2003 (le Earth Simulator) atteint une puissance maximale de 35,86 téraflops (http ://www.top500.org/list/2003/11/). $\square$

Exemple 25 On déduit du développement par rapport à la première ligne des expressions explicites pour les déterminants 2×2 et 3×3.

Il faut d'abord remarquer qu'un déterminant 1×1 est de la forme $\det\big(a\big) = a$. C'est en effet visiblement la seule forme multilinéaire alternée qui vaut 1 sur la matrice $\big(1\big)$.

Considérons maintenant un déterminant 2×2.

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

En effet, $A_{11} = \big(a_{22}\big)$ et $A_{12} = \big(a_{21}\big)$, d'où $C_{11} = a_{22}$ et $C_{12} = -a_{21}$. Cette formule de développement de déterminant est la seule formule explicite à connaître par cœur.

Le cas des déterminants 3×3 est déjà beaucoup plus compliqué.

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

Cette expression, qu'il est inutile de chercher à retenir, contient $6 = 3!$ produits de trois coefficients de A (un par colonne) affectés des signes $+$ ou $-$ suivant la nature d'une certaine permutation associée au produit en question. Pour un déterminant 4×4, on aurait $24 = 4!$ produits de quatre coefficients de A, et pour un déterminant $n\times n$, on aurait $n!$ produits de n coefficients de A affectés de signes $+$ ou $-$.

Mentionnons la *règle de Sarrus*, une astuce mnémotechnique qui permet de retrouver les déterminants 3×3 (et seulement ceux-là, cette règle ne se généralise pas à d'autres dimensions). On écrit la matrice en tableau et on lui ajoute en bas ses deux premières lignes. On obtient ainsi un tableau 5×3

$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}$$

Les produits de trois termes affectés du signe $+$ apparaissent dans les trois diagonales « descendantes » du haut à gauche vers le bas à droite, tandis que les produits de trois termes affectés du signe $-$ apparaissent dans les trois diagonales « montantes » du bas à gauche vers le haut à droite.

En résumé, le développement par rapport à la première ligne n'est utile pour calculer explicitement un déterminant que si la matrice dont on part a des propriétés particulières, par exemple beaucoup de zéros, ou s'y ramène par des opérations qui ne modifient pas le déterminant. $\square$

Proposition 52 *On a aussi*

$$\begin{array}{lll} \forall i, & \det A = & \sum_{j=1}^n (-1)^{i+j} a_{ij} \Delta_{ij} = \sum_{j=1}^n a_{ij} C_{ij} \\ & & \quad \textit{(développement par rapport à la ligne i),} \\ \forall j, & \det A = & \sum_{i=1}^n (-1)^{i+j} a_{ij} \Delta_{ij} = \sum_{i=1}^n a_{ij} C_{ij} \\ & & \quad \textit{(développement par rapport à la colonne j).} \end{array}$$

Démonstration. Comme $\det A = \det A^T$, on en déduit le développement par rapport à la première colonne $j = 1$. Par échange de deux colonnes, qui multiplie le déterminant par -1, on en déduit le développement par rapport à n'importe quelle colonne. Par une nouvelle transposition, on en déduit le développement par rapport à n'importe quelle ligne. $\square$

Définition 48 *On introduit la* matrice des cofacteurs de A, $\mathrm{cof} A = (c_{ij})$ *avec* $c_{ij} = (-1)^{i+j} \det A_{ij}$.

C'est aussi une matrice $n \times n$. Noter la disposition en échiquier des signes $(-1)^{i+j}$, commençant par un $+$ au coin supérieur gauche

$$\begin{pmatrix} + & - & + & \cdots & (-1)^{1+n} \\ - & + & - & \cdots & (-1)^{2+n} \\ + & - & + & \cdots & (-1)^{3+n} \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}.$$

Théorème 61 *On a pour toute matrice A*

$$A(\operatorname{cof} A)^T = (\operatorname{cof} A)^T A = (\det A) I_n.$$

En particulier, si A est inversible, alors

$$A^{-1} = \frac{1}{\det A}(\operatorname{cof} A)^T.$$

Démonstration. Posons $B = A(\operatorname{cof} A)^T$. Par la formule générale du produit matriciel,

$$(B)_{ij} = \sum_{k=1}^{n} a_{ik}(\operatorname{cof} A)^T_{kj} = \sum_{k=1}^{n} a_{ik} c_{jk}.$$

Si $i = j$, on reconnaît le développement du déterminant de A par rapport à la ligne i, donc

$$\forall i, \quad (B)_{ii} = \det A.$$

Si $i \neq j$, on reconnaît le développement par rapport à la ligne j du déterminant de la matrice $\tilde{A}$ dans laquelle la ligne j a été remplacée par la ligne i (ce qui ne change pas les cofacteurs considérés). Donc

$$\forall i \neq j, \quad (B)_{ij} = \det \tilde{A} = 0.$$

On procède de même pour $(\operatorname{cof} A)^T A$ avec les développements par rapport aux colonnes.

Dans le cas où A est inversible, alors $\det A \neq 0$ et il suffit de diviser par $\det A$ pour obtenir la formule pour A^{-1}. □

Remarque 45 Sauf pour $n = 2$, ou pour des matrices très particulières, *ce n'est pas la bonne façon de calculer explicitement l'inverse d'une matrice.* Si on a vraiment besoin de l'inverse d'une matrice, alors on a tout intérêt à utiliser la méthode de Gauss dès que $n \geq 3$. Par contre, la formule précédente est intéressante pour la théorie. □

Théorème 62 (Formules de Cramer) *Soit A une matrice inversible. L'unique solution du système linéaire $Ax = b$ est donnée par*

$$x_i = \frac{\det(A_i(b))}{\det A}, \text{ où } A_i(b) = \begin{pmatrix} a_1 & a_2 & \cdots & b & \cdots & a_n \end{pmatrix}$$

est la matrice A dans laquelle on a remplacé la colonne a_i par b.

Démonstration. Écrivons

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_i & \cdots & a_n \end{pmatrix} \text{ et } I_n = \begin{pmatrix} e_1 & e_2 & \cdots & e_i & \cdots & e_n \end{pmatrix}$$

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

$$I_{n,i}(x) = \begin{pmatrix} e_1 & e_2 & \cdots & x & \cdots & e_n \end{pmatrix}$$

dans laquelle on a remplacé e_i par x. Il vient

$$AI_{n,i}(x) = \begin{pmatrix} Ae_1 & Ae_2 & \cdots & Ax & \cdots & Ae_n \end{pmatrix} = A_i(b).$$

Par conséquent,

$$\det A \det I_{n,i}(x) = \det A_i(b),$$

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

Remarque 46 Sauf pour $n = 2$, à l'extrême rigueur $n = 3$, 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). Par contre, les formules de Cramer sont intéressantes pour la théorie. □

Encore une fois, le fait que l'on ait travaillé sur des matrices à coefficients réels n'a joué aucun rôle. Tout marche pareil dans $\mathcal{M}_n(\mathbb{C})$ avec un déterminant à valeurs complexes.

7.4 Interprétation géométrique des déterminants

On a une interprétation géométrique de $\mathbb{R}^n$ pour $n = 1, 2, 3$. On va voir qu'en dimension 2, les déterminants sont liés aux questions de surface et en dimension 3 aux questions de volume.

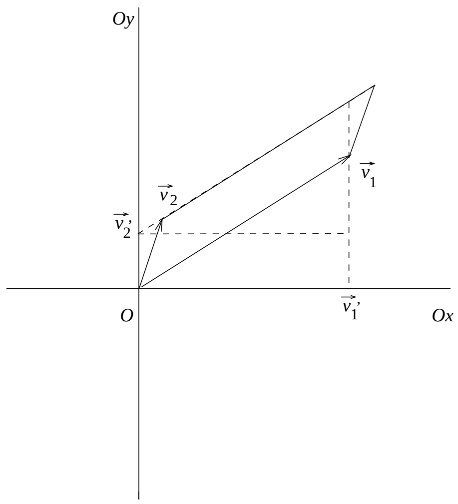
En dimension 2, deux vecteurs v_1, v_2 déterminent un parallélogramme, alors qu'en dimension 3, trois vecteurs v_1, v_2, v_3 déterminent un parallélépipède.

On prendra comme unité de surface la surface du carré unité dont les côtés 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 53 *La surface du parallélogramme est donnée par* $|\det\begin{pmatrix}v_1 & v_2\end{pmatrix}|$. *Le volume du parallélépipède est donné par* $|\det\begin{pmatrix}v_1 & v_2 & v_3\end{pmatrix}|$.

Démonstration. Traitons le cas $n = 2$. Le résultat est vrai si $\begin{pmatrix}v_1 & v_2\end{pmatrix} = \begin{pmatrix}a & 0\\ 0 & d\end{pmatrix}$. En effet, dans ce cas on a affaire à un rectangle de côtés $|a|$ et $|d|$, donc de surface $|ad|$, alors que le déterminant de la matrice vaut ad.

Supposons que $\{v_1, v_2\}$ est une famille libre. Notons $\begin{pmatrix}v_1 & v_2\end{pmatrix} = \begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$. Si $a_{11} \neq 0$, alors $v_2' = v_2 - \frac{a_{12}}{a_{11}}v_1$ 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_{22}' non nulle. On pose alors $v_1' = v_1 - \frac{a_{21}}{a_{22}'}v'2$, 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 $a_{11} = 0$, alors $a_{12} \neq 0$ 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 parallélogramme est réduit à un segment et est donc de surface nulle.

Le cas tridimensionnel se traite de façon analogue. □

Chapitre 8

Réduction des matrices

La réduction des matrices constitue le premier pas de ce que l'on appelle la théorie spectrale, vaste sujet. Ses applications pratiques sont nombreuses : modélisation des vibrations, dynamique des populations, analyse de données en composantes principales en statistique, mécanique quantique, économie mathématique, etc.

8.1 Valeurs propres, vecteurs propres

Pour la première fois, il va être utile de distinguer entre $\mathbb{R}$ et $\mathbb{C}$. On adoptera donc la notation générique $K = \mathbb{R}$ ou $\mathbb{C}$, et l'on précisera si besoin est.

Définition 49 *Soit $A \in \mathcal{M}_n(K)$. On dit que $x \in K^n$ est un* vecteur propre *de A si et seulement si*

i) $x \neq 0$.

ii) $\exists \lambda \in K, Ax = \lambda x$.

Si x est un vecteur propre de A, alors le scalaire λ de ii) est la valeur propre *de A associée à x.*

Remarque 47 La condition i) est essentielle. En effet, si on ne l'impose pas, alors la condition ii) ne dit rien car pour tout λ, on a $0 = A \times 0 = \lambda 0$.

Un vecteur propre de A est donc un vecteur non nul dont l'image par A en est un multiple scalaire (on dit encore, lui est colinéaire). □

Remarque 48 Notons que $\lambda = 0$ est valeur propre de A si et seulement si $\ker A \neq \{0\}$, c'est-à-dire si et seulement si $x \mapsto Ax$ n'est pas injective. □

Proposition 54 *Un scalaire $\lambda \in K$ est valeur propre de $A \in \mathcal{M}_n(K)$ si et seulement si* $\ker(A - \lambda I_n) \neq \{0\}$.

Démonstration. En effet ; λ est valeur propre si et seulement si il existe un $x \neq 0$ tel que $Ax = \lambda x$, c'est-à-dire $Ax - \lambda x = 0$, ou encore $(A - \lambda I_n)x = 0$. $\square$

Définition 50 *L'ensemble de tous les vecteurs propres associés à une valeur propre λ, complété par le vecteur nul, $E_\lambda(A) = \ker(A - \lambda I_n)$ est appelé* sous-espace vectoriel propre *associé à λ. Sa dimension* $\dim E_\lambda(A)$ *est appelé* multiplicité géométrique de la valeur propre λ.

Définition 51 *L'ensemble de toutes les valeurs propres de f est appelé spectre de A et est noté $\sigma(A) \subset K$.*

Proposition 55 *Deux matrices semblables ont même spectre.*

Démonstration. En effet, si $A' = P^{-1}AP$ et si x est un vecteur propre de A, alors $P^{-1}x$ est un vecteur propre de A', puisque $A'P^{-1}x = P^{-1}A(PP^{-1})x = P^{-1}Ax = \lambda P^{-1}x$. $\square$

Elles n'ont par contre pas les mêmes espaces propres, puisque ceux-ci se correspondent par changement de base.

Remarque 49 i) Si on connaît une valeur propre λ d'une matrice A, alors la détermination du sous-espace propre associé dans K^n se ramène en principe à la résolution du système linéaire homogène $(A - \lambda I_n)x = 0$, dont on sait qu'il n'a pas que la solution triviale. En pratique, ce n'est pas si simple. En effet, en général, les valeurs propres ne seront connues qu'approximativement avec une certaine précision. Donc les nombres $\tilde{\lambda}$ dont on dispose en réalité ne sont pas exactement des valeurs propres, et le système linéaire homogène associé $(A - \tilde{\lambda} I_n)x = 0$ n'a que la solution triviale. Ce n'est donc pas de cette façon que l'on calcule les sous-espaces propres dans la pratique.

ii) Ce que l'on ne sait pas encore faire, même en principe, c'est trouver les valeurs propres. $\square$

Proposition 56 *On a $\lambda \in \sigma(A)$ si et seulement si* $\det(A - \lambda I_n) = 0$.

Démonstration. C'est évident d'après ce qui précède, puisque il faut et il suffit que la matrice $A - \lambda I_n$ soit singulière. $\square$

Définition 52 *Le polynôme $P_A(X) = \det(A - XI_n)$ s'appelle le* polynôme caractéristique de A.

C'est bien un polynôme à une indéterminée car

$$\det(A - XI_n) = \begin{vmatrix} a_{11} - X & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - X & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - X \end{vmatrix},$$

et l'on voit en développant par rapport à la première colonne que $d^\circ P_A(X) = n$ et que le terme de plus haut degré de $P_A(X)$ est $(-1)^n X^n$.

Remarque 50 i) Le terme de plus bas degré de $P_A(X)$ est $P_A(0) = \det A$.

ii) Si deux matrices A et B sont semblables, elles ont même polynôme caractéristique. En effet, si $B = P^{-1}AP$, alors $B - XI_n = P^{-1}(A - XI_n)P$ et on utilise la multiplicativité du déterminant (cet argument triche un peu, puisque ce n'est pas à proprement parler un déterminant de matrice au sens précédent car l'indéterminée X apparaît dans les expressions. Néanmoins, on peut rendre le raisonnement rigoureux et le résultat est correct). □

Corollaire 63 *On a $\lambda \in \sigma(A)$ si et seulement si λ 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 un tel polynôme a au plus n racines, on peut tout de suite en déduire le corollaire suivant.

Corollaire 64 *Toute matrice $n \times n$ admet au plus n valeurs propres.*

On voit donc que l'existence de valeurs propres dépend *fondamentalement* de la nature du corps de base K, contrairement à tout ce que l'on a pu voir jusqu'à maintenant.

Corollaire 65 *Si $K = \mathbb{C}$, toute matrice admet au moins une valeur propre.*

Démonstration. C'est le théorème de d'Alembert qui affirme que tout polynôme à coefficients complexes admet au moins une racine dans $\mathbb{C}$ (on dit que $\mathbb{C}$ est *algébriquement clos*). □

Exemple 26 Il existe des matrices à coefficients réels sans valeur propre réelle. Par exemple, les matrices de rotation d'angle $\varphi \neq k\pi$, $k \in \mathbb{Z}$, n'ont pas de vecteur propre dans $\mathbb{R}^2$, ce qui est géométriquement évident, puisque l'image d'un vecteur par une telle rotation n'est certainement pas colinéaire au vecteur en question. En termes de polynôme caractéristique, il vient

$$P_A(X) = \begin{vmatrix} \cos\varphi - X & -\sin\varphi \\ \sin\varphi & \cos\varphi - X \end{vmatrix} = X^2 - 2\cos\varphi X + 1,$$

qui n'a pas de racine réelle pour ces valeurs de φ.

Par contre, la même matrice de rotation peut être interprétée comme celle d'une application linéaire de $\mathbb{C}^2$ dans $\mathbb{C}^2$, où elle admet des vecteurs et valeurs propres (complexes, bien sûr). □

Corollaire 66 *On a* $\sigma(A^T) = \sigma(A)$.

Démonstration. En effet, $P_{A^T}(X) = \det(A^T - XI_n) = \det((A - XI_n)^T) = \det(A - XI_n) = P_A(X)$, puisque $I_n^T = I_n$. □

8.2 Trigonalisation, diagonalisation

Proposition 57 *Si A est triangulaire, alors elle a n valeurs propres qui sont ses éléments diagonaux.*

Démonstration. D'après le corollaire précédent, il suffit de traiter le cas des matrices triangulaires supérieures. Dans ce cas

$$P_A(X) = \begin{vmatrix} a_{11} - X & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - X & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - X \end{vmatrix} = (a_{11} - X)(a_{22} - X)\cdots(a_{nn} - X),$$

comme on le voit en développant par rapport à la première colonne. Le polynôme P_A est un produit de facteurs du premier degré (on dit qu'il est *scindé*) et ses n racines sont visibles, ce sont les a_{ii}, $i = 1, \ldots, n$. □

Définition 53 *i) On dit qu'une matrice A est* trigonalisable *si elle est semblable à une matrice triangulaire.*

ii) On dit qu'une matrice A est diagonalisable *si elle est semblable à une matrice diagonale.*

En d'autres termes, A est trigonalisable (resp. diagonalisable) s'il existe une matrice inversible P telle que $P^{-1}AP$ est triangulaire (resp. diagonale). Dans ce cas, on dit que la matrice P *trigonalise* (resp. *diagonalise*) la matrice A.

Proposition 58 *Une matrice A est diagonalisable si et seulement si il existe une base de K^n constituée de vecteurs propres de A.*

Démonstration. Supposons qu'il existe $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ base de K^n telle que $Au_i = \lambda_i u_i$, alors la matrice A' de l'application linéaire $x \mapsto Ax$ dans la base $\mathcal{B}$ est diagonale, ses éléments diagonaux étant les valeurs propres λ_i. Si P désigne la matrice de passage de la base canonique dans $\mathcal{B}$, on a bien que $A' = P^{-1}AP$, donc A et A' sont semblables.

Réciproquement, supposons A semblable à une matrice diagonale A', c'est-à-dire qu'il existe P inversible telle que $A' = P^{-1}AP$. Interprétons P comme la matrice de passage de la base canonique à une autre base $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ (les composantes de u_i dans la base canonique sont simplement les colonnes de P). Alors on voit que $(Au_i)_{\mathcal{B}} = A'e_i = \lambda_i e_i$ où e_i est le $i$ème vecteur de la base canonique de K^n (en effet, $(u_i)_{\mathcal{B}} = e_i$). On en déduit que $(Au_i)_{\mathcal{B}} = \lambda_i (u_i)_{\mathcal{B}}$, donc que $Au_i = \lambda_i u_i$ et les u_i forment une base de vecteurs propres de A. $\square$

Remarque 51 Par définition d'un espace propre, la restriction de l'application linéaire $x \mapsto Ax$ à $E_{\lambda_i}(A)$ est l'homothétie de rapport λ_i. $\square$

On a besoin de critères de trigonalisabilité et diagonalisabilité.

Théorème 67 *Une matrice A est trigonalisable si et seulement si son polynôme caractéristique P_A est scindé sur K.*

Démonstration. On rappelle qu'un polynôme est dit être scindé s'il est factorisable sur K en un produit de termes du premier degré.

Supposons d'abord A trigonalisable. Elle est donc semblable à une matrice triangulaire (mettons supérieure),

$$P^{-1}AP = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{pmatrix}.$$

On en déduit immédiatement que

$$P_A(X) = (b_{11} - X)(b_{22} - X) \cdots (b_{nn} - X),$$

et le polynôme P_A est scindé.

On pourra admettre la réciproque. En voici une démonstration. Supposons P_A scindé. On raisonne par récurrence sur la dimension de l'espace.

Pour $n = 1$, on a $A = \begin{pmatrix} a_{11} \end{pmatrix}$ donc $P_A(X) = a_{11} - X$ est scindé et A est bien triangulaire supérieure.

Supposons la propriété vraie en dimension $n-1$. Soit A une matrice $n \times n$ telle que $P_A(X) = \prod_{i=1}^{n}(\lambda_i - X)$ avec $\lambda_i \in K$. Comme λ_1 est une valeur propre de A, il existe un vecteur propre $v_1 \in K^n$, $v_1 \neq 0$, associé. C'est donc une famille libre, que l'on peut compléter à l'aide de vecteurs $v_2, \dots, v_n$ en une base $\mathcal{B}$ de K^n. La matrice de l'endomorphisme associé à A dans cette base est semblable à A et de la forme

$$A' = \left(\begin{array}{c|ccc} \lambda_1 & b_2 & \cdots & b_n \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right),$$

où $\tilde{A} \in \mathcal{M}_{n-1}(K)$ (on utilise la notation dite par blocs). Il vient donc

$$P_A(X) = \left|\begin{array}{c|ccc} \lambda_1 - X & b_2 & \cdots & b_n \\ \hline 0 & & & \\ \vdots & & \tilde{A} - XI_{n-1} & \\ 0 & & & \end{array}\right| = (\lambda_1 - X) P_{\tilde{A}}(X)$$

en développant par rapport à la première colonne. Comme P_A est scindé, on en déduit que $P_{\tilde{A}}(X) = \prod_{i=2}^{n}(\lambda_i - X)$ est également scindé. L'hypothèse de récurrence s'applique : il existe $\tilde{P} \in GL_{n-1}(K)$ telle que $\tilde{P}^{-1}\tilde{A}\tilde{P} = \tilde{T}$ avec $\tilde{T}$ matrice triangulaire supérieure $(n-1) \times (n-1)$.

On pose alors

$$P = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{array}\right),$$

d'où $\det P = \det \tilde{P} \neq 0$, ce qui montre que P est inversible. De plus on vérifie aisément que

$$P^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{P}^{-1} & \\ 0 & & & \end{array}\right),$$

puis que

$$P^{-1}AP = \left(\begin{array}{c|ccc} \lambda_1 & c_1 & \cdots & c_n \\ \hline 0 & & & \\ \vdots & & \tilde{T} & \\ 0 & & & \end{array}\right),$$

qui est triangulaire supérieure. □

Corollaire 68 *Si $K = \mathbb{C}$, toute matrice est trigonalisable.*

Revenons à l'étude générale des sous-espaces propres.

Théorème 69 *Les sous-espaces propres d'une matrice A associés à des valeurs propres distinctes sont en somme directe.*

Démonstration. Soient $\lambda_1, \ldots, \lambda_p \in K$ les valeurs propres distinctes de A et soient $E_{\lambda_1}, \ldots, E_{\lambda_p}$ les sous-espaces propres associés dans K^n. 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 $x_i \in E_{\lambda_i}$, $i = 1, \ldots, k$ tels que $x_1 + x_2 + \cdots + x_k = 0$.

Multiplions cette relation par λ_k. Il vient

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

Multiplions également cette égalité par la matrice A. Il vient

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

Soustrayons les deux égalités obtenues membre à membre. On obtient

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

Posant $y_i = (\lambda_k - \lambda_i)x_i \in E_{\lambda_i}$, $i = 1, \ldots, k-1$, 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 $i = 1, \ldots, k-1$. Or on a pris des valeurs propres distinctes, donc $\lambda_k - \lambda_i \neq 0$. Par conséquent, $x_i = 0$ pour $i = 1, \ldots, k-1$. Reportant ceci dans la première relation, on en déduit finalement que $x_k = 0$, et les sous-espaces propres sont bien en somme directe. □

Corollaire 70 *Une matrice A est diagonalisable si et seulement si K^n est somme directe de sous-espaces propres.*

Démonstration. En effet, la réunion de bases des sous-espaces propres forme alors une base de K^n constituée de vecteurs propres de A. □

Corollaire 71 *Si une matrice $n \times n$ a exactement n valeurs propres distinctes, alors elle est diagonalisable.*

Démonstration. En effet, dans ce cas la matrice A a n sous-espaces propres $E_{\lambda_i}(A)$ en somme directe. Pour chaque espace propre, on a $\dim E_{\lambda_i}(A) \geq 1$. Donc

$$n \leq \sum_{i=1}^{n} \dim E_{\lambda_i}(A) = \dim\Big(\bigoplus_{i=1}^{n} E_{\lambda_i}(A)\Big) \leq \dim K^n = n,$$

puisque $\bigoplus_{i=1}^{n} E_{\lambda_i}(A)$ est un sev de K^n. On en déduit que

$$\dim\Big(\bigoplus_{i=1}^{n} E_{\lambda_i}(A)\Big) = \dim K^n,$$

et donc que

$$\bigoplus_{i=1}^{n} E_{\lambda_i}(A) = K^n,$$

d'où la diagonalisabilité de A. □

Ce corollaire n'est qu'une *condition suffisante.* De nombreuses matrices avec des valeurs propres multiples sont aussi diagonalisables.

Remarque 52 Si une matrice est diagonalisable, la décomposition de l'espace K^n en somme directe de sous-espaces propres permet de mieux comprendre l'action de cette matrice. En effet, on a déjà noté que sa restriction à un sous-espace propre est l'application linéaire la plus simple qui soit, une homothétie. □

Remarque 53 ATTENTION, même sur $\mathbb{C}$ il existe des matrices qui ne sont pas diagonalisables. Par exemple

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

est triangulaire supérieure. Elle admet donc une seule valeur propre $\lambda = 0$. C'est une racine double, puisque $P_A(X) = X^2$, mais sa multiplicité géométrique est visiblement 1 (c'est la dimension de son noyau). Donc $\mathbb{C}^2$ ne peut pas être somme directe des espaces propres, et A n'est pas diagonalisable

On peut aussi le voir directement. Si A est diagonalisable, alors

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

contradiction. Donc A n'est pas diagonalisable. □

À quoi peut bien servir la diagonalisation des matrices ? À bien des choses. Donnons en un exemple.

Proposition 59 *Soit A une matrice diagonalisable. Pour tout $k \in \mathbb{N}$, les valeurs propres de A^k sont les λ_i^k. De plus, si P est une matrice qui diagonalise A, alors*

$$A^k = P \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix} P^{-1}.$$

Plus généralement, si Q est un polynôme à une indéterminée, on a

$$Q(A) = P \begin{pmatrix} Q(\lambda_1) & 0 & \cdots & 0 \\ 0 & Q(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q(\lambda_n) \end{pmatrix} P^{-1}.$$

Démonstration. La formule pour un polynôme Q se déduit immédiatement de celles pour les monômes A^k. Si P diagonalise A, ceci veut dire que $P^{-1}Ap$ est diagonale, ou encore

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} P^{-1} = PDP^{-1}.$$

Par conséquent,

$$A^k = (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) = PD(P^{-1}P)D(P^{-1}P)\cdots(P^{-1}P)DP^{-1} = PD^kP^{-1}.$$

Or, il est évident que

$$D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix},$$

d'où le résultat. □

Donnons pour finir un exemple d'application concrète en dynamique des populations. Supposons que l'on observe une population animale divisée en deux classes d'individus, les individus immatures et les individus adultes. On recense la population à intervalles réguliers, disons tous les ans. Son état à l'année n est donc représenté par un vecteur de $\mathbb{R}^2$

$$x(n) = \begin{pmatrix} \text{nombre d'individus immatures} \\ \text{nombre d'individus adultes} \end{pmatrix} = \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix}.$$

On observe que, d'une année sur l'autre, une proportion $0 \leq p_1 \leq 1$ d'individus immatures meurt, le reste devient adulte. Pour les individus adultes, une proportion $0 \leq p_2 \leq 1$ meurt tandis qu'une proportion $0 \leq q \leq 1$ des adultes donne naissance à un individu immature. Par conséquent, à l'année $n+1$ on recensera $(1-p_1)x_1(n)+(1-p_2)x_2(n)$ individus adultes et $qx_2(n)$ individus immatures. On obtient donc une évolution (on dit une dynamique) de la forme

$$x(n+1) = \begin{pmatrix} qx_2(n) \\ (1-p_1)x_1(n)+(1-p_2)x_2(n) \end{pmatrix} = \begin{pmatrix} 0 & q \\ 1-p_1 & 1-p_2 \end{pmatrix} \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} = Ax(n).$$

Par récurrence, on en déduit immédiatement que

$$x(n) = A^n x(0)$$

où le vecteur $x(0)$ représente la population initiale. La survie ou l'extinction de la population est donc gouvernée par le comportement des puissances successives de la matrice A et par la population initiale. On est donc amené à étudier les valeurs propres de A en fonction des paramètres p_1, p_2 et q. En effet, si celle-ci se diagonalise en D, on aura

$$x(n) = PD^nP^{-1}x(0).$$

Étudions un cas extrême. Supposons les taux de mortalité p_1 et p_2 nuls. On obtient

$$A = \begin{pmatrix} 0 & q \\ 1 & 1 \end{pmatrix} \text{ d'où } P_A(X) = -X(1-X)-q = X^2-X-q.$$

Il y a deux racines réelles $\lambda_1 = \frac{1-\sqrt{1+4q}}{2}$ et $\lambda_2 = \frac{1+\sqrt{1+4q}}{2}$. Si le taux de naissance q est strictement positif, on aura $\lambda_2 > 1$ et on peut voir que la population se comportera en général comme λ_2^n, c'est-à-dire qu'elle explose exponentiellement avec le temps (ce qui n'est pas surprenant avec ces hypothèses).

Dans le cas général, on trouve toujours deux valeurs propres réelles distinctes $\lambda = \frac{1-p_2 \pm \sqrt{(1-p_2)^2+4q(1-p_1)}}{2}$. Par exemple, pour $p_1 = 0{,}1$ et $p_2 = 0{,}2$, on obtient pour diverses valeurs de q

$$\text{si } q = 2/9, \text{ alors } \lambda_1 = -0{,}2, \lambda_2 = 1$$

et la population tend vers un état d'équilibre quand le temps tend vers l'infini. On le voit plus facilement sur le vecteur $y(n) = P^{-1}x(n)$. En effet, ce vecteur suit la dynamique $y(n) = D^n y(0)$ avec $D^n = \begin{pmatrix} (-0{,}2)^n & 0 \\ 0 & 1 \end{pmatrix}$. Donc il est clair que $y(n) = \begin{pmatrix} (-0{,}2)^n y_1(0) \\ y_2(0) \end{pmatrix} \to \begin{pmatrix} 0 \\ y_2(0) \end{pmatrix}$ quand $n \to +\infty$, soit $x(n) \to P \begin{pmatrix} 0 \\ y_2(0) \end{pmatrix}$.

Prenons un taux de naissance légèrement supérieur :

$$\text{si } q = 0{,}3, \text{ alors } \lambda_1 = -0{,}255\ldots, \lambda_2 = 1{,}055\ldots$$

et la population explose exponentiellement car $\lambda_1^n \to 0$ et $\lambda_2^n \to +\infty$. Avec un taux de naissance légèrement inférieur,

$$\text{si } q = 0{,}2, \text{ alors } \lambda_1 = -0{,}024\ldots, \lambda_2 = 0{,}824\ldots$$

et la population s'éteint exponentiellement car $\lambda_1^n \to 0$ et $\lambda_2^n \to 0$.

Naturellement, il s'agit d'une modélisation très grossière de l'évolution d'une population, mais elle est déjà riche d'enseignements.

Chapitre 9

Produits scalaires dans $\mathbb{R}^m$, produit vectoriel dans $\mathbb{R}^3$

9.1 Produits scalaires dans $\mathbb{R}^m$

Nous allons introduire des produits de vecteurs, mais dont les valeurs sont scalaires, d'où leur nom.

Définition 54 *On appelle* forme bilinéaire *sur* $\mathbb{R}^m$ *toute application* B *de* $\mathbb{R}^m \times \mathbb{R}^m$ *à valeurs dans* $\mathbb{R}$ *telle que pour tous* $x, x', y, y' \in \mathbb{R}^m$ *et tout* $\lambda \in \mathbb{R}$,

$$B(\lambda x + x', y) = \lambda B(x,y) + B(x',y), \quad \textit{et} \quad B(x, \lambda y + y') = \lambda B(x,y) + B(x,y').$$

Une forme bilinéaire est donc linéaire par rapport à chacun de ses arguments, l'autre étant fixé. On en déduit

$$B(\lambda x + \lambda' x', \mu y + \mu' y') = \lambda\mu B(x,y) + \lambda\mu' B(x,y') + \lambda'\mu B(x',y) + \lambda'\mu' B(x',y').$$

Attention, une forme bilinéaire n'est pas linéaire par rapport au couple (x,y) !

Définition 55 *On dit qu'une forme bilinéaire* B *est un* produit scalaire *sur* $\mathbb{R}^m$ *si et seulement si*

i) B *est* symétrique

$$\forall x, y \in \mathbb{R}^m, \quad B(x,y) = B(y,x),$$

ii) B *est* définie

$$\forall x \in \mathbb{R}^m, \quad B(x,x) = 0 \Longleftrightarrow x = 0,$$

iii) et B *est* positive

$$\forall x \in \mathbb{R}^m, \quad B(x,x) \geq 0.$$

En abrégé, on dit aussi une forme bilinéaire sdp. Les conditions ii) et iii) impliquent que si $x \neq 0$, alors $B(x,x) > 0$.

Notation On utilise traditionnellement plusieurs notations pour les produits scalaires. Ainsi, en mathématiques, on trouve $x \cdot y$, (x,y) ou encore $(x|y)$. Les physiciens utilisent volontiers la notation $\langle x|y\rangle$, notamment en mécanique quantique.

Exemple 27 Il existe une infinité de produits scalaires différents sur $\mathbb{R}^m$. En particulier
$$(x|y) = x_1y_1 + x_2y_2 + \cdots + x_my_m$$
est un produit scalaire sur $\mathbb{R}^m$, appelé produit scalaire canonique.

Définition 56 *L'espace $\mathbb{R}^m$ muni d'un produit scalaire $(\cdot|\cdot)$ est appelé* espace euclidien*. L'application de $\mathbb{R}^m$ dans $\mathbb{R}_+$ définie par*
$$x \mapsto \|x\| = \sqrt{(x|x)}$$
est appelée norme euclidienne *associée au produit scalaire.*

Il y a donc autant de normes euclidiennes différentes sur $\mathbb{R}^m$ que de produits scalaires différent. La norme euclidienne canonique est associée au produit scalaire canonique :
$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}.$$

La propriété la plus importante des produits scalaires est la suivante.

Théorème 72 (Inégalité de Cauchy-Schwarz) *Soit $(\cdot|\cdot)$ un produit scalaire sur $\mathbb{R}^m$ et $\|\cdot\|$ la norme euclidienne associée. Alors, pour tous $x, y \in \mathbb{R}^m$, on a*
$$|(x|y)| \leq \|x\|\|y\|.$$

Démonstration. Si $x = 0$, l'inégalité est triviale $0 \leq 0$. Supposons donc $x \neq 0$ et considérons la fonction de $\mathbb{R}$ dans $\mathbb{R}$, $f(t) = \|tx + y\|^2$. Par construction, $f(t) \geq 0$ pour tout t. Par bilinéarité et symétrie du produit scalaire, on a aussi
$$f(t) = (tx+y|tx+y) = t^2(x|x) + t(x|y) + t(y|x) + (y|y) = \|x\|^2t^2 + 2(x|y)t + \|y\|^2.$$

C'est un trinôme du second degré de la forme $f(t) = at^2 + 2bt + c$ avec $a > 0$, il atteint donc son minimum en $t = -b/a$ et ce minimum vaut $c - (b^2/a)$. Comme il est positif, on déduit que $b^2 \leq ac$, ce qui n'est autre que l'inégalité de Cauchy-Schwarz (au carré). $\square$

Théorème 73 *Une norme euclidienne satisfait les trois propriétés fondamentales suivantes.*

i) Inégalité triangulaire

$$\forall x, y \in \mathbb{R}^m, \quad \|x+y\| \leq \|x\| + \|y\|.$$

ii) Homogénéité positive

$$\forall x \in \mathbb{R}^m, \forall \lambda \in \mathbb{R} \quad \|\lambda x\| = |\lambda| \|x\|.$$

iii) Caractère défini

$$\|x\| = 0 \Longleftrightarrow x = 0.$$

Démonstration. Le point iii) découle directement du caractère défini du produit scalaire. Pour le point ii), on remarque par bilinéarité que

$$\|\lambda x\|^2 = (\lambda x|\lambda x) = \lambda^2 (x|x) = \lambda^2 \|x\|^2$$

d'où le résultat en prenant les racines carrées des deux membres, puisque $\sqrt{\lambda^2} = |\lambda|$, bien sûr.

L'inégalité triangulaire est une conséquence de l'inégalité de Cauchy-Schwarz. En effet

$$\|x+y\|^2 = (x+y|x+y) = \|x\|^2 + 2(x|y) + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

d'où le résultant en prenant les racines carrées. $\square$

Remarque 54 Il y a égalité dans Cauchy-Schwarz, c'est-à-dire $|(x|y)| = \|x\| \|y\|$, si et seulement si $x = 0$ ou bien si le minimum de $f(t)$ vaut 0. Dans ce cas, il existe donc $t \in \mathbb{R}$ tel que $\|tx + y\| = 0$. Comme une norme euclidienne est définie, point iii), ceci implique que $tx + y = 0$ ou encore $y = tx$. Dans tous les cas, l'égalité n'a lieu dans Cauchy-Schwarz que si x et y sont colinéaires. En d'autres termes, si la famille $\{x, y\}$ est libre, alors $|(x|y)| < \|x\| \|y\|$. $\square$

Remarque 55 On peut récupérer le produit scalaire de deux vecteurs à partir de normes. Par exemple, il est facile de vérifier que

$$(x|y) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2),$$

en développant la première norme au carré par bilinéarité et symétrie. $\square$

9.2 Interprétation géométrique dans $\mathbb{R}^2$ et $\mathbb{R}^3$

Interprétons géométriquement le produit scalaire et la norme canoniques dans $\mathbb{R}^2$. On a dans ce cas

$$(x|y) = x_1y_1 + x_2y_2 \quad \text{et} \quad \|x\| = \sqrt{x_1^2 + x_2^2}.$$

On identifie donc $\mathbb{R}^2$ et le plan rapporté à un repère cartésien Ox_1, Ox_2 dont les axes sont à angle droit. Par le théorème de Pythagore, on voit que la norme de x n'est autre que la longueur du segment OP qui représente le vecteur x. Cette interprétation explique le terme d'inégalité triangulaire appliqué à la première propriété d'une norme : par la règle du parallélogramme, on voit que deux côtés et la diagonale du parallélogramme forment un triangle dont les longueurs sont $\|x\|$ et $\|y\|$ pour les côtés et $\|x+y\|$ pour la diagonale. Or celle-ci est plus courte que la somme des longueurs des côtés.

Pour interpréter le produit scalaire, il est commode d'utiliser l'interprétation du plan comme plan complexe. Tout vecteur x de $\mathbb{R}^2$ est donc identifié au complexe $z = x_1 + ix_2$. On voit d'abord que $\|x\| = |z|$, ce qui est conforme à l'interprétation géométrique de la norme que l'on vient de faire. De plus, on voit que si $z' = y_1 + iy_2$, alors

$$(x|y) = \Re(z\overline{z'}).$$

Supposons x et y non nuls. On passe alors en modules et arguments : $z = |z|e^{i\theta}$ et $z' = |z'|e^{i\theta'}$. Comme $\overline{z'} = |z'|e^{-i\theta'}$, il vient

$$(x|y) = \Re(|z|e^{i\theta}|z'|e^{-i\theta'}) = |z||z'|\Re(e^{i(\theta-\theta')}) = |z||z'|\cos(\theta-\theta').$$

Si l'on note $\psi = \theta - \theta'$ l'angle entre les vecteurs x et y, on a donc obtenu l'interprétation géométrique suivante du produit scalaire

$$(x|y) = \|x\|\,\|y\|\cos\psi.$$

On peut aussi écrire

$$\cos\psi = \frac{(x|y)}{\|x\|\,\|y\|},$$

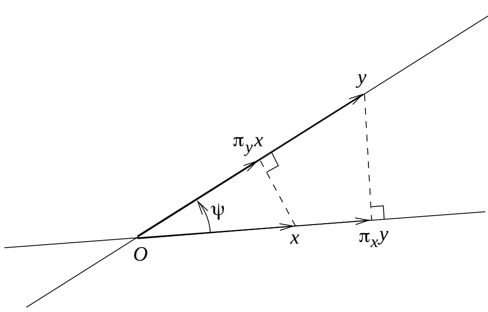
et l'inégalité de Cauchy-Schwarz nous montre que le cosinus d'un angle est toujours compris entre -1 et 1.

Un corollaire est que $(x|y) = 0$ si et seulement si l'angle ψ est un angle droit. Un autre corollaire est que, notant $\pi_y x$ la projection orthogonale de x sur la droite engendrée par y, on a $(x|y) = \pm\|\pi_y x\|\,\|y\|$ avec le signe $+$ si $\pi_y x$ et y pointent dans la même direction (angle aigu) et le signe $-$ s'ils pointent dans des directions opposées (angle obtus). On a aussi bien sûr $(x|y) = \pm\|x\|\|\pi_x y\|$ avec les mêmes signes.

L'interprétation géométrique dans $\mathbb{R}^3$ est identique, quoique un peu plus difficile à écrire (on ne peut pas s'aider du plan complexe...). En particulier, $\|x\|$ est toujours la longueur du segment qui représente le vecteur x dans l'espace, et

$$(x|y) = \|x\|\|y\|\cos\psi,$$

où ψ est l'angle entre les segments représentant x et y dans l'espace, ainsi que l'interprétation en termes de projection orthogonale.



Interprétation géométrique du produit scalaire dans $\mathbb{R}^2$.

9.3 Orthogonalité

Dans cette section, on se place dans $\mathbb{R}^m$, muni de son produit scalaire canonique.

Définition 57 *On dit que deux vecteurs x et y sont* orthogonaux*, et l'on note $x \perp y$ si et seulement si $(x|y) = 0$.*

Cette définition est parfaitement raisonnable, puisqu'on a vu en dimension 2 et 3 que la nullité du produit scalaire correspond à deux vecteurs à angle droit. On a donc étendu cette notion géométrique en dimension quelconque.

Définition 58 *On dit qu'une base de $\mathbb{R}^m$ est une* base orthogonale *si elle est formée de vecteurs orthogonaux deux à deux. On dit qu'elle est* orthonormée *si en outre ces vecteurs sont de norme* 1.

Une base $\{v_1, v_2, \ldots, v_m\}$ est donc orthogonale si $(v_i|v_j) = 0$ pour tous $i \neq j$, et elle est orthonormée si en plus $(v_i|v_i) = 1$ pour tout i.

Exemple 28 La base canonique de $\mathbb{R}^m$ est trivialement orthonormée (pour le produit scalaire canonique, pas forcément pour un autre produit scalaire). $\square$

Il y a une infinité d'autres bases orthonormées. Voici comment en construire.

Proposition 60 (Procédé d'orthonormalisation de Gram-Schmidt) *Soit une base de $\mathbb{R}^m$, $\{v_1, v_2, \dots, v_m\}$. Il existe une base orthonormée $\{u_1, u_2, \dots, u_m\}$ telle que*

$$\mathrm{vect}\{u_1, u_2, \dots, u_k\} = \mathrm{vect}\{v_1, v_2, \dots, v_k\}$$

pour tout $k = 1, \dots, m$.

Démonstration. On procède par récurrence sur k. Pour $k = 1$, on prend $u_1 = \frac{v_1}{\|v_1\|}$. En effet, $v_1 \neq 0$, et par homogénéité positive, $\|u_1\| = 1$. De plus, trivialement, $\mathrm{vect}\{u_1\} = \mathrm{vect}\{v_1\}$.

Faisons l'hypothèse de récurrence que l'on a construit $u_1, u_2, \dots, u_{k-1}$ orthonormés et tels que $\mathrm{vect}\{u_1, u_2, \dots, u_{k-1}\} = \mathrm{vect}\{v_1, v_2, \dots, v_{k-1}\}$. On pose

$$w_k = v_k - \sum_{i=1}^{k-1} (v_k|u_i)u_i.$$

Comme $v_k = w_k + \sum_{i=1}^{k-1} (v_k|u_i)u_i$, on a bien

$$\mathrm{vect}\{u_1, u_2, \dots, u_{k-1}, w_k\} = \mathrm{vect}\{v_1, v_2, \dots, v_k\}$$

grâce à la deuxième partie de l'hypothèse de récurrence. De plus, par linéarité du produit scalaire par rapport au deuxième vecteur, pour tout $1 \leq j \leq k-1$,

$$(w_k|u_j) = (v_k|u_j) - \sum_{i=1}^{k-1} (v_k|u_i)(u_i|u_j).$$

Or $(u_i|u_j) = 0$ pour tout $i \neq j$ et $(u_j|u_j) = 1$. Il vient donc $(w_k|u_j) = (v_k|u_j) - (v_k|u_j) = 0$ pour tout $j \leq k-1$. Pour conclure, on remarque que w_k n'est pas nul, puisque la famille des v_i est libre, et l'on pose $u_k = \frac{w_k}{\|w_k\|}$ ce qui ne modifie pas les espaces vectoriels engendrés. □

Le calcul des produits scalaires et des normes à l'aide des composantes dans une base orthonormée sont particulièrement simples.

Proposition 61 *Soit $\mathcal{B}$ une base orthonormée de $\mathbb{R}^m$,*

$$(x)_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}, \quad (y)_{\mathcal{B}} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix},$$

les composantes de deux vecteurs x et y dans la base $\mathcal{B}$. Alors

$$(x|y) = \sum_{i=1}^{m} \lambda_i \mu_i \quad et \quad \|x\| = \sqrt{\sum_{i=1}^{m} \lambda_i^2}.$$

Démonstration. Soit $\mathcal{B} = \{u_1, u_2, \dots, u_m\}$ la base orthonormée. Par définition de ce que sont les composantes, on a

$$x = \sum_{i=1}^{m} \lambda_i u_i \quad \text{et} \quad y = \sum_{j=1}^{m} \mu_j u_j.$$

Par bilinéarité, il vient

$$(x|y) = \sum_{i,j=1}^{m} \lambda_i \mu_j (u_i|u_j).$$

Or, dans cette somme double, tous les termes $(u_i|u_j)$, $i \neq j$ sont nuls, et les termes $(u_i|u_i)$ qui restent valent 1, d'où la formule du produit scalaire. La norme s'en déduit immédiatement. $\square$

Définition 59 *Soit F un sev de $\mathbb{R}^m$. On appelle* orthogonal de F *l'ensemble*

$$F^\perp = \{y \in \mathbb{R}^m; \forall x \in F, (x|y) = 0\}.$$

Deux cas particuliers importants $\{0\}^\perp = \mathbb{R}^m$ et $(\mathbb{R}^m)^\perp = \{0\}$.

Proposition 62 *L'orthogonal de F est un sev de $\mathbb{R}^m$ et l'on a*

$$F \oplus F^\perp = \mathbb{R}^m.$$

Démonstration. Montrons d'abord que $F^\perp$ est un sev. Il contient bien le vecteur nul. Il est stable pour l'addition car si $y_1, y_2 \in F^\perp$, alors pour tout $x \in F$

$$(x|y_1 + y_2) = (x|y_1) + (x|y_2) = 0 + 0 = 0.$$

On montre de même que $F^\perp$ est stable pour la multiplication par un scalaire.

Montrons maintenant que les deux sev F et $F^\perp$ sont supplémentaires. Tout d'abord, soit $y \in F \cap F^\perp$. On peut donc prendre $x = y \in F$, ce qui donne $0 = (y|y) = \|y\|^2$. Donc $y = 0$, c'est-à-dire $F \cap F^\perp = \{0\}$. Ils sont donc en somme directe.

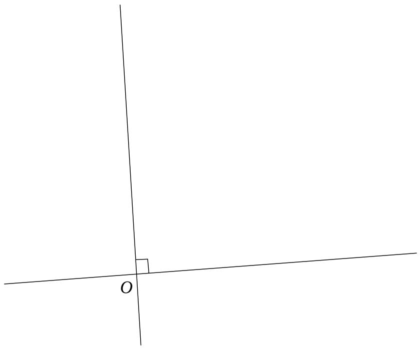
Posons $k = \dim F$ et soit $\{v_1, v_2, \dots, v_k\}$ une base de F. Par le théorème de la base incomplète, on lui ajoute des vecteurs $\{v_{k+1}, \dots v_m\}$ pour en faire une base de $\mathbb{R}^m$. Appliquons le procédé de Gram-Schmidt à cette base. Il fournit une nouvelle base orthonormée $\{u_1, u_2, \dots, u_m\}$ de $\mathbb{R}^m$ telle que $\{u_1, u_2, \dots, u_k\}$ est une base (orthonormée) de F. De plus, par construction, $u_{k+1}, \dots, u_m$ appartiennent tous à $F^\perp$, puisqu'ils sont orthogonaux à une base de F. Pour tout $x \in \mathbb{R}^m$, on peut donc écrire

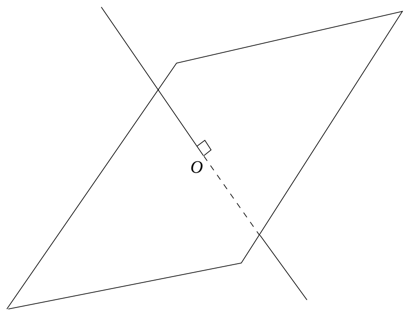
$$x = \lambda_1 u_1 + \cdots + \lambda_k u_k + \lambda_{k+1} u_{k+1} + \cdots \lambda_m u_m = x_F + x_{F^\perp},$$

avec $x_F = \lambda_1 u_1 + \cdots + \lambda_k u_k \in F$ et $x_{F^\perp} = \lambda_{k+1} u_{k+1} + \cdots \lambda_m u_m \in F^\perp$. On a donc $F + F^\perp = \mathbb{R}^m$. $\square$

Remarque 56 On a obtenu aussi que $\{u_{k+1},\dots u_m\}$ est une base orthonormée de $F^\perp$. De plus, évidemment, $\dim F + \dim F^\perp = m$. Enfin, l'application $\Pi_F : \mathbb{R}^m \to \mathbb{R}^m$, $x \mapsto x_F$ est une application linéaire appelée *projection orthogonale* sur F. Son image est F et son noyau est $F^\perp$. C'est une projection car $\Pi_F \circ \Pi_F = \Pi_F$. Idem pour la projection orthogonale sur $F^\perp$. Notons pour finir que $(F^\perp)^\perp = F$. □

Exemple 29 Dans $\mathbb{R}^2$, les sev non triviaux sont des droites. L'orthogonal d'une droite est aussi une droite, c'est la droite perpendiculaire passant par 0. Dans $\mathbb{R}^3$, les sev non triviaux sont les droites et les plans. L'orthogonal d'une droite est le plan qui lui est perpendiculaire passant par 0, alors que l'orthogonal d'un plan est la droite qui lui est perpendiculaire passant par 0.





Droites orthogonales dans $\mathbb{R}^2$, droite et plan orthogonaux dans $\mathbb{R}^3$.

9.4 Matrices orthogonales, rotations, symétries

On commence par une remarque simple.

Proposition 63 *Pour tous $x, y \in \mathbb{R}^m$, on a*

$$(x|y) = x^T y.$$

Cette formule est à comprendre au sens où la *matrice* 1×1 $x^T y$ contient comme unique coefficient le *nombre* $(x|y)$.

Démonstration. Par la définition de la multiplication des matrices, le produit matriciel $x^T y$ est bien défini et est une matrice 1×1. De plus,

$$x^T y = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} x_1 y_1 + x_2 y_2 + \cdots + x_m y_m \end{pmatrix}$$

que l'on identifie au nombre $(x|y)$. □

Définition 60 *On dit qu'une matrice $m \times m$ Q est* orthogonale *si pour tous $x, y \in \mathbb{R}^m$, on a*

$$(Qx|Qy) = (x|y).$$

En d'autres termes, l'application linéaire associée à Q, $x \mapsto Qx$, conserve le produit scalaire.

Proposition 64 *L'application linéaire $\mathbb{R}^m \to \mathbb{R}^m$, $x \mapsto Qx$ avec Q matrice orthogonale conserve les normes (on dit que c'est une* isométrie*) et les angles entre vecteurs.*

Démonstration. En effet, $\|Qx\|^2 = (Qx|Qx) = (x|x) = \|x\|^2$. Par conséquent, si ψ est l'angle entre les vecteurs x et y et ψ_Q l'angle entre Qx et Qy, on a

$$\cos\psi_Q = \frac{(Qx|Qy)}{\|Qx\|\|Qy\|} = \frac{(x|y)}{\|x\|\|y\|} = \cos\psi$$

d'où l'égalité des angles. □

Proposition 65 *Une matrice Q est orthogonale si et seulement si*

$$QQ^T = Q^T Q = I_m.$$

Démonstration. On a, pour tous $x, y \in \mathbb{R}^m$,

$$(Qx|Qy) = (Qx)^T Qy = x^T Q^T Qy = x^T y,$$

ou encore

$$x^T (Q^T Q - I_m)y = 0,$$

ou encore

$$(x|(Q^T Q - I_m)y) = 0$$

pour tout $x \in \mathbb{R}^m$. On en déduit que $(Q^T Q - I_m)y \in (\mathbb{R}^m)^\perp = \{0\}$ pour tout $y \in \mathbb{R}^m$, donc $Q^T Q - I_m = 0$. En particulier Q est inversible, et son inverse est Q^T. □

Corollaire 74 *Une matrice est orthogonale si et seulement si ses vecteurs colonne forment une base orthonormée de $\mathbb{R}^m$.*

Démonstration. En effet, si l'on écrit Q comme une ligne de vecteurs colonne, $Q = \begin{pmatrix} a_1 & a_2 & \cdots & a_m \end{pmatrix}$, alors il est facile de voir que le coefficient d'indices i et j du produit $Q^T Q$ n'est autre que $a_i^T a_j = (a_i|a_j)$. Comme $Q^T Q = I_m$, on en déduit que $(a_i|a_j) = 0$ pour tous $i \neq j$ et $(a_i|a_i) = 1$ pour tout i, et réciproquement. □

Corollaire 75 *Le produit de deux matrices orthogonales est orthogonale, l'inverse d'une matrice orthogonale, qui est égale à sa transposée, est orthogonale.*

Démonstration. Soient Q_1 et Q_2 deux matrices orthogonales. On a

$$(Q_1Q_2)^T Q_1Q_2 = Q_2^T(Q_1^T Q_1)Q_2 = Q_2^T Q_2 = I_m$$

et de même pour l'autre relation.

On a vu que $Q_1^{-1} = Q_1^T$, donc

$$Q_1^{-T} Q_1^{-1} = (Q_1^T)^T Q_1^T = Q_1 Q_1^T = I_m$$

et de même pour l'autre relation. □

Proposition 66 *Le déterminant d'une matrice orthogonale vaut* ± 1.

Démonstration. Prenant le déterminant de l'égalité $Q^T Q = I_m$, il vient

$$\det Q^T \det Q = \det I_m = 1$$

puisque le déterminant est multiplicatif. Mais $\det Q^T = \det Q$, d'où finalement $(\det Q)^2 = 1$. □

Définition 61 *Les matrices orthogonales de déterminant* $+1$ *sont appelées* matrices de rotation, *celles de déterminant* -1 matrices de symétrie orthogonale.

Définition 62 *L'ensemble de toutes les matrices orthogonales*

$$O(m) = \{Q \in \mathcal{M}_m(\mathbb{R}); Q^T Q = QQ^T = I_m\}$$

s'appelle le groupe orthogonal. *L'ensemble des matrices de rotation*

$$SO(m) = \{Q \in \mathcal{M}_m(\mathbb{R}); Q^T Q = QQ^T = I_m, \det Q = 1\}$$

s'appelle le groupe des rotations.

Manifestement, les produits, inverses et transposées de rotations sont des rotations. Par contre, le produit d'un nombre pair de symétries est une rotation, alors que le produit d'un nombre impair de symétries est une symétrie.

Enfin, une somme de matrices orthogonales ou le produit d'une matrice orthogonale par un scalaire *ne sont pas* des matrices orthogonales.

Déterminons $O(2)$ et $SO(2)$. Soit $Q \in O(2)$. Son premier vecteur colonne est de norme 1, il est donc situé sur le cercle unité et est de la forme $\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ pour un certain θ. Le second vecteur colonne est aussi de norme 1 et il est perpendiculaire au premier, donc nécessairement de la forme $\begin{pmatrix}\cos(\theta \pm \frac{\pi}{2})\\ \sin(\theta \pm \frac{\pi}{2})\end{pmatrix}$. Les formules de trigonométrie élémentaires nous donnent que $\cos(\theta + \frac{\pi}{2}) = -\sin\theta$, $\sin(\theta + \frac{\pi}{2}) = \cos\theta$ et $\cos(\theta - \frac{\pi}{2}) = \sin\theta$, $\sin(\theta - \frac{\pi}{2}) = -\cos\theta$. On a donc deux formes possibles

$$Q = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \quad \text{d'où} \quad \det Q = 1,$$

c'est la rotation d'angle θ que nous avons déjà rencontré, ou bien

$$Q = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix} \quad \text{d'où} \quad \det Q = -1,$$

et l'on peut voir que c'est la symétrie orthogonale par rapport à la droite passant par 0 et qui fait un angle de $\frac{\theta}{2}$ avec l'axe des abscisses.

On peut décrire $O(3)$ et $SO(3)$ de façon analogue, mais c'est nettement plus compliqué, donc nous ne le ferons pas ici.

9.5 Produit vectoriel dans $\mathbb{R}^3$

Nous allons enfin introduire un produit de vecteurs qui donne un autre vecteur, mais ce produit n'existe qu'en dimension 3. On travaillera donc dans $\mathbb{R}^3$.

Définition 63 *Soient*

$$x = \begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} \quad \textit{et} \quad y = \begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix}$$

deux vecteurs de $\mathbb{R}^3$. *Le* produit vectoriel *de x et y est le vecteur de* $\mathbb{R}^3$

$$x \wedge y = \begin{pmatrix}x_2y_3 - x_3y_2\\ x_3y_1 - x_1y_3\\ x_1y_2 - x_2y_1\end{pmatrix}.$$

Donnons tout de suite quelques propriétés évidentes du produit vectoriel.

Proposition 67 *Le produit vectoriel est bilinéaire alterné.*

Démonstration. C'est clair puisqu'il est fabriqué à l'aide de déterminants 2×2.
□

En particulier, on a

$$x \wedge y = -y \wedge x \quad \text{et} \quad x \wedge x = 0.$$

Remarque 57 Règle pratique pour calculer un produit vectoriel (mis à part mémoriser la définition) : on forme la matrice

$$A = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix}$$

et l'on remarque que $x \wedge y$ est la première colonne de la matrice des cofacteurs de A, $\operatorname{cof} A$.

Par exemple, si $\{e_1, e_2, e_3\}$ désigne la base canonique de $\mathbb{R}^3$, on a

$$e_1 \wedge e_2 = e_3, \quad e_2 \wedge e_3 = e_1, \quad e_3 \wedge e_1 = e_2.$$

□

Proposition 68 *Pour tous $x, y, z \in \mathbb{R}^3$, on a*

$$(x \wedge y | z) = \det \begin{pmatrix} x & y & z \end{pmatrix}.$$

Démonstration. On a

$$\det \begin{pmatrix} x & y & z \end{pmatrix} = z_1(x_2y_3 - x_3y_2) - z_2(x_1y_3 - x_3y_1) + z_3(x_1y_2 - x_2y_1) = (x \wedge y | z)$$

en développant par rapport à la troisième colonne (voir aussi la remarque ci-dessus). □

Définition 64 *La quantité $(x \wedge y | z)$ s'appelle le* produit mixte *des trois vecteurs.*

Proposition 69 *0n a*

$$(x \wedge y | z) = (y \wedge z | x) = (z \wedge x | y)$$

(invariance du produit mixte par permutation circulaire.

Démonstration. En effet,

$$\det \begin{pmatrix} x & y & z \end{pmatrix} = -\det \begin{pmatrix} y & x & z \end{pmatrix} = \det \begin{pmatrix} y & z & x \end{pmatrix} = (y \wedge z | x)$$

par deux échanges successifs de colonne, et de même pour la deuxième égalité. □

Proposition 70 *Le produit vectoriel de deux vecteurs est orthogonal à chacun de ces vecteurs.*

Démonstration. On a
$$(x \wedge y|x) = \det\begin{pmatrix} x & y & x \end{pmatrix} = 0$$
puisqu'on a deux colonnes égales, et de même $(x \wedge y|y) = 0$. □

Corollaire 76 *Si la famille $\{x,y\}$ est libre, alors $x \wedge y$ appartient à l'orthogonal du plan* $\mathrm{vect}\{x,y\}$. *Si elle est liée, alors $x \wedge y = 0$.*

Démonstration. Le vecteur $x \wedge y$ étant orthogonal à x et à y, il est orthogonal à toute combinaison linéaire de ces deux vecteurs, donc au sev $\mathrm{vect}\{x,y\}$. Si la famille est libre, ce sev est un plan, donc $x \wedge y$ appartient à la droite orthogonale à ce plan. Si elle est liée, on peut supposer sans perte de généralité que $y = tx$ pour un certain $t \in \mathbb{R}$. Par conséquent, $x \wedge y = x \wedge (tx) = tx \wedge x = 0$. □

Proposition 71 *La famille $\{x,y\}$ est liée si et seulement si $x \wedge y = 0$.*

Démonstration. On vient de voir qu'une famille liée est telle que $x \wedge y = 0$. Réciproquement, si $x \wedge y = 0$, on en déduit que pour tout $z \in \mathbb{R}^3$, $\det\begin{pmatrix} x & y & z \end{pmatrix} = 0$, c'est-à-dire que la famille $\{x,y,z\}$ est liée. Or si $\{x,y\}$ était libre, il suffirait de prendre $z \neq 0$ dans la droite orthogonale au plan $\mathrm{vect}\{x,y\}$ pour construire une famille libre. On en déduit que $\{x,y\}$ est liée. □

Corollaire 77 *Si la famille $\{x,y\}$ est libre, alors $x \wedge y$ engendre la droite orthogonale au plan engendré par x et y.*

Définition 65 *On dit que deux bases de $\mathbb{R}^3$ ont la même* orientation *si le déterminant de la la matrice de passage de l'une à l'autre est positif. Elles sont d'orientations opposées si ce déterminant est négatif.*

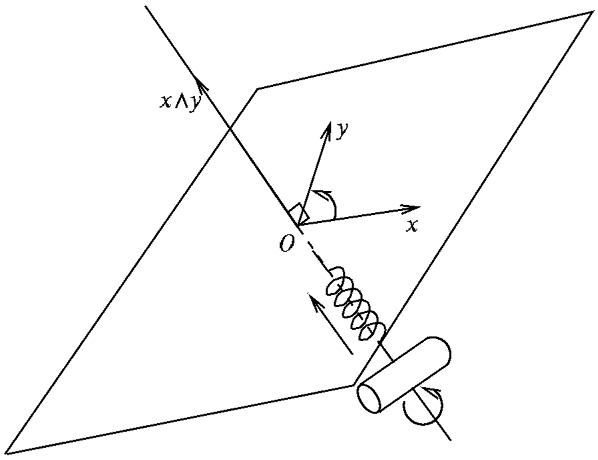
Remarquons qu'une matrice de passage est toujours inversible, donc son déterminant n'est jamais nul, il est soit > 0 soit < 0. On partitionne ainsi l'ensemble des bases de $\mathbb{R}^3$ en deux parties disjointes :

Définition 66 *On dit qu'une base de $\mathbb{R}^3$ est* directe *si elle a la même* orientation *que la base canonique et qu'elle est* indirecte *sinon.*

Proposition 72 *Soit $\{x,y\}$ une famille libre. Alors, la base $\{x,y,x \wedge y\}$ est directe.*

Démonstration. La matrice de passage n'est autre que $P = \begin{pmatrix} x & y & x \wedge y \end{pmatrix}$. On a donc
$$\det P = (x \wedge y|x \wedge y) = \|x \wedge y\|^2 > 0,$$
car la famille est libre. □

Remarque 58 On peut visualiser la proposition précédente à l'aide de la *règle du tire-bouchon*. Si l'on place la poignée d'un tire-bouchon dans le plan $\text{vect}\{x,y\}$ et qu'on la fait tourner dans le sens de x vers y, alors le tire-bouchon progresse dans le sens de $x \wedge y$. Si on n'aime pas les tire-bouchons, on peut faire la même chose avec un tournevis. □



La règle du tire-bouchon.

On peut également procéder avec les doigts de la main droite : placer le majeur, l'index et le pouce à angles droits. Si le majeur représente x et le pouce y, l'index représente $x \wedge y$. □

Pour finir de caractériser géométriquement le produit vectoriel, il nous reste à interpréter sa norme.

Proposition 73 *La norme de $x \wedge y$ est égale à l'aire du parallélogramme engendré par x et y. Si ψ désigne l'angle entre x et y, on a aussi*

$$\|x \wedge y\| = \|x\| \|y\| |\sin\psi|.$$

Démonstration. On sait que $|\det P|$ est le volume du parallélépipède engendré par x, y et $x \wedge y$. Comme $x \wedge y$ est orthogonal aux deux autres vecteurs, ce volume est aussi égal au produit de la longueur de $x \wedge y$, c'est-à-dire $\|x \wedge y\|$ par l'aire de la base du parallélépipède, c'est-à-dire l'aire du parallélogramme engendré par x et y, d'où le résultat. □

Remarque 59 Si $x, y \in \text{vect}\{e_1, e_2\}$, alors

$$x \wedge y = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ x_1 y_2 - x_2 y_1 \end{pmatrix},$$

d'où

$$\|x \wedge y\| = |x_1 y_2 - x_2 y_1|.$$

Or on a déjà vu que la valeur absolue de ce déterminant est l'aire du parallélogramme en question. On reconnaît aussi le produit $\|x\|\|y\||\sin\psi|$ en passant dans le plan complexe comme on l'a fait pour le produit scalaire dans $\mathbb{R}^2$. Attention, le produit vectoriel n'est pas une opération interne à $\text{vect}\{e_1, e_2\}$, puisqu'il est colinéaire à e_3 dans ce cas. $\square$