Université de Bordeaux
Licence de Sciences, Technologies, Santé
Mentions Mathématiques et Informatique
4TPU207U Algèbre Linéaire 1

Algèbre Linéaire 1

Année 2021-2022

Table des matières

Chapitre 1

Systèmes linéaires

Dans ce chapitre et dans toute la suite du cours, la notation $\mathbb{K}$ désigne l'ensemble $\mathbb{R}$ des nombres réels ou l'ensemble $\mathbb{C}$ des nombres complexes.

Nous allons définir les *systèmes linéaires* et apprendre à les résoudre par la *méthode du pivot de Gauss.*

1.1 Introduction et quelques définitions

Définition 1.1.1
Un système linéaire (S) sur $\mathbb{K}$ à n inconnues (appelées aussi variables) $x_1, x_2, \ldots, x_n$ et k équations $(E_1), \ldots, (E_k)$ se présente sous la forme :

$$(S) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n & = b_1 \quad (E_1) \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n & = b_2 \quad (E_2) \\ & \vdots \\ a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,n}x_n & = b_k \quad (E_k) \end{cases}$$

Les coefficients *du système sont les nombres $a_{i,j} \in \mathbb{K}$. Ils sont affectés de deux indices i, j où i est l'indice de l'équation (E_i) et j l'indice de la variable x_j.*

On dit que le système (S) est homogène *si $b_1 = b_2 = \cdots = b_k = 0$*

On note $\mathcal{S}$ l'ensemble de ses solutions, c'est-à-dire

$$\mathcal{S} = \big\{(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n \text{ tels que } (x_1, \ldots, x_n) \text{ vérifie } (E_1), \ldots, (E_k)\big\}.$$

Résoudre le système (S) *signifie déterminer explicitement et exactement l'ensemble $\mathcal{S}$ de ses solution.*

Notre but dans ce chapitre est de décrire un *algorithme* permettant de résoudre un système linéaire en toutes circonstances et de façon efficace. C'est *l'algorithme du pivot de Gauss.* Tout d'abord introduisons un peu de vocabulaire.

Définition 1.1.2
Deux systèmes linéaires sont dits équivalents *s'ils ont le même ensemble de solutions. On note alors*

$$(S) \Longleftrightarrow (S')$$

Exemple et définition 1.1.3
Voici un système linéaire à trois inconnues et trois équations ($n = k = 3$).

$$(S_1) \quad \left\{ \begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ & & -y & + & 2z & = & 1 \\ & & & & 2z & = & 5 \end{array} \right.$$

Un système de ce type est dit triangulaire *pour des raisons évidentes. Il se résout facilement, de bas en haut, en calculant successivement z, puis y, puis x. On trouve une solution unique et on écrit :*

$$\mathcal{S} = \big\{(-31/2, 4, 5/2)\big\}.$$

1.2 Système linéaire échelonné

Les systèmes linéaires *échelonnés* sont un peu plus généraux que les systèmes triangulaires ; ils se résolvent presque aussi facilement.

Définition 1.2.1
Informellement, un système linéaire (S) est dit échelonné *si chacune de ses équations commence avec un décalage à droite par rapport à la précédente, comme dans la représentation qui suit.*

$$\left\{ \begin{array}{ccccccccc} a_{1,j_1}x_{j_1} & +\cdots+ & a_{1,j_2}x_{j_2} & +\cdots+ & a_{1,j_k}x_{j_k} & +\cdots+ a_{1,n}x_n & = & b_1 \\ & & a_{2,j_2}x_{j_2} & +\cdots+ & a_{2,j_k}x_{j_k} & +\cdots+ a_{2,n}x_n & = & b_2 \\ & & & & \vdots & & & \vdots \\ & & & & a_{k,j_k}x_{j_k} & +\cdots+ a_{k,n}x_n & = & b_k \end{array} \right.$$

Plus formellement, notons pour chaque équation (E_i) son premier coefficient non nul *a_{i,j_i}. On dit que le système est échelonné si $1 \leq j_1 < j_2 < \cdots < j_k \leq n$.*

Les termes $a_{1,j_1}x_{j_1}$, $a_{2,j_2}x_{j_2}, \ldots, a_{k,j_k}x_{j_k}$ s'appellent les pivots *du système. Les variables $x_{j_1}, x_{j_2}, \ldots, x_{j_k}$ s'appellent les* variables pivot *et il y en a exactement autant que d'équations (remarquons qu'on a donc $k \leq n$). Les autres variables s'appellent les* variables auxiliaires.

On peut raffiner la notion de système échelonné en introduisant celle de système *échelonné et réduit* :

Définition 1.2.2
Un système linéaire est dit échelonné réduit *s'il est échelonné, si ses pivots sont égaux à* 1*, et si ses variables pivot apparaissent dans une seule équation (autrement dit, si les coefficients* a_{i,j_ℓ} *au-dessus des pivots* a_{ℓ,j_ℓ} *sont nuls).*

Pour résoudre un système linéaire échelonné, on se ramène à un système triangulaire en passant dans toutes les équations les variables auxiliaires à droite du signe "=". Ensuite on exprime les variables pivot en fonction des variables auxiliaires., en procédant successivement de bas en haut. On obtient un ensemble de solutions paramétrées par les variables auxiliaires, c'est-à-dire les variables x_i pour $i \notin \{j_1, \ldots, j_k\}$.

Exemple numérique :

$$(S_2) \quad \left\{ \begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ & & x_2 & + & x_3 & - & 6x_4 & = & 0 \\ & & & & & & 3x_4 & = & 1 \end{array} \right.$$

On a

$$(S_2) \iff \left\{ \begin{array}{ccccccc} x_1 & + & x_2 & + & x_4 & = & -x_3 \\ & & x_2 & - & 6x_4 & = & -x_3 \\ & & & & 3x_4 & = & 1 \end{array} \right.$$

$$\iff \left\{ \begin{array}{ccc} x_1 & = & -x_2 - x_3 - x_4 \\ x_2 & = & -x_3 + 6x_4 \\ x_4 & = & 1/3 \end{array} \right. \iff \left\{ \begin{array}{ccc} x_1 & = & -7/3 \\ x_2 & = & -x_3 + 2 \\ x_4 & = & 1/3 \end{array} \right.$$

À chaque $x_3 \in \mathbb{K}$ correspond une solution du système. On obtient

$$\mathcal{S} = \big\{(-7/3, -x_3 + 2, x_3, 1/3) \mid x_3 \in \mathbb{K}\big\}$$

ou encore

$$\mathcal{S} = \big\{(-7/3, 2, 0, 1/3) + x_3(0, -1, 1, 0) \mid x_3 \in \mathbb{K}\big\}.$$

1.3 L'algorithme du pivot de Gauss

On va décrire maintenant les opérations que l'on peut exécuter sur un système sans changer son ensemble de solutions, c'est-à-dire celles qui le transforment en un système équivalent.

Définition 1.3.1
On appelle opérations élémentaires sur le système linéaire (S) *les opérations suivantes :*

(1) Échange de deux équations.

(2) Remplacement de l'équation (E_i) *par* $\lambda(E_i)$ *pour* $\lambda \in \mathbb{K}$, $\lambda \neq 0$.

(3) Remplacement de l'équation (E_i) *par* $(E_i) + \mu(E_j)$ *pour* $j \neq i$ *et* $\mu \in \mathbb{K}$.

Remarque : Il faut bien sûr exécuter ces opérations sur les deux membres des équations, i.e. à gauche comme à droite du signe "=".

> **Proposition 1.3.2**
> *On obtient un système linéaire équivalent en appliquant à (S) une opération élémentaire, ou une succession de telles opérations.*

Preuve. Il est clair que les opérations (1) et (2) ne changent pas l'ensemble des solutions $\mathcal{S}$. Vérifions-le pour (3) : soit $j \neq i$. Il est clair que si (E_i) et (E_j) sont vérifiées alors $(E_i)+\mu(E_j)$ l'est aussi. Réciproquement, supposons avoir remplacé (E_i) par $(E_i)+\mu(E_j)$. L'équation (E_j) elle n'a pas changé. On remarque que $(E_i) = ((E_i) + \mu(E_j)) - \mu(E_j)$ donc, si $(E_i) + \mu(E_j)$ et (E_j) sont vérifiées alors (E_i) l'est aussi. $\square$

La *méthode du pivot de Gauss* que nous allons étudier maintenant est un algorithme qui, par une succession d'opérations élémentaires, permet de ramener la résolution d'un système linéaire à celle d'un système échelonné.

Description informelle de l'algorithme : On procède variable par variable, en travaillant successivement sur les colonnes du système.

Soit x_{j_1} la première variable qui apparait dans le système. Supposons pour simplifier les notations que $j_1 = 1$. On cherche de haut en bas la première équation qui contienne un terme en x_1, et si ce n'est pas la première, on l'échange avec (E_1). Puis, on divise (E_1) par le coefficient de x_1. Ensuite, on utilise (E_1) pour éliminer les termes en x_1 dans les équations $(E_2), \dots, (E_k)$. Pour cela on remplace (E_i) par $(E_i) - a_{i,1}(E_1)$. On ne modifiera plus (E_1).

Ensuite on recommence le procédé avec les équations $(E_2), \dots (E_k)$: on passe à la première variable présente dans $(E_2), \dots (E_k)$, soit x_{j_2}. On a $j_2 > j_1$. On cherche parmi $(E_2), \dots (E_k)$, de haut en bas, la première équation qui contienne un terme en x_{j_2}, et on l'échange avec (E_2). On divise (E_2) par a_{2,j_2}, puis, comme avec x_1, on élimine la variable x_{j_2} des équations $(E_3), \dots (E_k)$. On procède ainsi jusqu'à la dernière variable.

Au cours de l'algorithme, on a pu transformer une équation en une équation du type "$0 = b$". Si $b \neq 0$, il n'y a pas de solutions, et on peut conclure que $\mathcal{S} = \emptyset$. Si $b = 0$, on enlève cette équation du système.

On a démontré le résultat suivant :

> **Proposition 1.3.3**
> *L'algorithme du pivot de Gauss conduit soit à une équation sans solutions du type "$0 = b$" avec $b \neq 0$ (et dans ce cas on peut conclure que le système initial n'a pas de solutions) soit à un système linéaire équivalent, échelonné, et dont le nombre d'équations est inférieur ou égal à celui du système initial.*

Preuve. En ce qui concerne le nombre d'équations, on remarque que les opérations élémentaires (définition 1.3.1) le laissent inchangé. La seule chose qui peut arriver au cours de l'algorithme du pivot de Gauss, c'est que l'on élimine une équation du type "$0 = 0$". Donc le système linéaire final contient au plus le même nombre d'équation que le système de départ. $\square$

Exemple numérique :

$$(S_3) \quad \left\{ \begin{array}{rcrcrcrcl} & & & & x_3 & + & x_4 & = & 4 \\ x_1 & - & 3x_2 & + & x_3 & - & x_4 & = & 0 \\ -8x_1 & + & 24x_2 & - & 4x_3 & + & 16x_4 & = & 8 \end{array} \right.$$

$$(S_3) \Longleftrightarrow \left\{ \begin{array}{rcrcrcrcll} x_1 & - & 3x_2 & + & x_3 & - & x_4 & = & 0 & \quad (E_1) \leftrightarrow (E_2) \\ & & & & x_3 & + & x_4 & = & 4 & \\ -8x_1 & + & 24x_2 & - & 4x_3 & + & 16x_4 & = & 8 & \end{array} \right.$$

$$\Longleftrightarrow \left\{ \begin{array}{rcrcrcrcll} x_1 & - & 3x_2 & + & x_3 & - & x_4 & = & 0 & \\ & & & & x_3 & + & x_4 & = & 4 & \\ & & & & 4x_3 & + & 8x_4 & = & 8 & \quad (E_3) \leftarrow (E_3) + 8(E_1) \end{array} \right.$$

$$\Longleftrightarrow \left\{ \begin{array}{rcrcrcrcrl} x_1 & - & 3x_2 & + & x_3 & - & x_4 & = & 0 & \\ & & & & x_3 & + & x_4 & = & 4 & \\ & & & & & & 4x_4 & = & -8 & \quad (E_3) \leftarrow (E_3) - 4(E_2) \end{array} \right.$$

On a obtenu un système linéaire échelonné, que l'on résout suivant la méthode du paragraphe 1.2 :

$$(S_3) \Longleftrightarrow \left\{ \begin{array}{lcl} x_1 & = & 3x_2 - x_3 + x_4 = 3x_2 - 8 \\ x_3 & = & 6 \\ x_4 & = & -2 \end{array} \right.$$

On obtient

$$\begin{aligned} \mathcal{S} &= \big\{(3x_2 - 8, x_2, 6, -2) \mid x_2 \in \mathbb{K}\big\} \\ &= \big\{(-8, 0, 6, -2) + x_2(3, 1, 0, 0) \mid x_2 \in \mathbb{K}\big\} \end{aligned}$$

Remarque 1.3.4

L'algorithme du pivot de Gauss peut non seulement échelonner un système linéaire mais aussi le réduire. Prenons un exemple numérique : le système suivant (exemple de la section 1.2) est échelonné mais pas réduit.

$$(S_2) \quad \left\{ \begin{array}{rcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ & & x_2 & + & x_3 & - & 6x_4 & = & 0 \\ & & & & & & 3x_4 & = & 1 \end{array} \right.$$

On peut le transformer en un système réduit par les opérations élémentaires suivantes :

- $(E_3) \leftarrow 1/3 * (E_3)$ *qui met le troisième pivot à 1*
- $(E_1) \leftarrow (E_1) - (E_2)$ *qui élimine la variable* x_2 *de* (E_1)

- *$(E_1) \leftarrow (E_1) - (E_3)$ et $(E_2) \leftarrow (E_2) - (E_3)$ qui élimine la variable x_4 de (E_1) et de (E_2).*

On obtient

$$(S_2) \Longleftrightarrow \left\{ \begin{array}{rcrcrcrcr} x_1 & & & & & & & = & -7/3 \\ & & x_2 & + & x_3 & & & = & 2 \\ & & & & & & x_4 & = & 1/3 \end{array} \right.$$

Pour résoudre (S_2), il ne reste plus qu'à faire passer x_3 (la variable auxiliaire) à droite du signe $=$.

1.4 Conclusions

En plus de l'aspect calculatoire, on retiendra de ce chapitre le résultat théorique suivant, qui nous sera utile par la suite :

Proposition 1.4.1
Pour qu'un système d'équations linéaires à n inconnues et k équations possède une solution unique, il est nécessaire *d'avoir $k \geq n$ i.e. au moins autant d'équations que d'inconnues.*

Preuve. Soit (S) un système linéaire, à n inconnues et k équations. On suppose que (S) possède une solution unique, et on veut montrer que $k \geq n$. D'après la proposition 1.3.3, le système (S) est donc équivalent à un système (S') échelonné, à n inconnues et $k' \leq k$ équations. Supposons d'abord $k' < n$. Comme il y a autant de variables pivot que d'équations, il y a alors dans (S') des variables qui ne sont pas des variables pivot, c'est-à-dire des variables auxiliaires, qui vont paramétrer les solutions. Il y aurait donc forcément une infinité de solutions ce qui est contraire à notre hypothèse. Nous pouvons conclure que $k' = n$ et donc $n \leq k$. □

Remarque : La condition $k \geq n$ (au moins autant d'équations que d'inconnues) est nécessaire mais bien évidemment pas suffisante pour avoir unicité de la solution, comme le montrent les exemples stupides suivants :

$$\left\{ \begin{array}{rl} x+y & = 0 \\ -x-y & = 0 \end{array} \right. \qquad \text{(infinité de solutions)}$$

$$\left\{ \begin{array}{rl} x+y & = 0 \\ x+y & = 1 \end{array} \right. \qquad \text{(aucune solution)}$$

Remarque 1.4.2
Nous avons vu que dans le cas général, l'ensemble des solutions d'un système linéaire est paramétré par les variables auxiliaires d'un système échelonné équivalent. Par définition, le nombre de ces variables auxiliaires est égal à la différence (nombre de variables) - (nombre de variables pivots). Nous verrons au chapitre 5 comment interpréter ce nombre en termes de dimension (théorème 5.6.3 et remarque 5.6.4)

Chapitre 2

L'espace vectoriel $\mathbb{K}^n$

On étudie au lycée les vecteurs du plan et de l'espace. Les premiers sont repérés par deux coordonnées $(x, y) \in \mathbb{R}^2$ tandis que les second sont repérés par trois coordonnées $(x, y, z) \in \mathbb{R}^3$. Nous allons maintenant généraliser les propriétés des vecteurs du plan et de l'espace aux éléments de $\mathbb{R}^n$ et de $\mathbb{C}^n$ lorsque n est en entier quelconque, en se détachant de notre vision intuitive de la géométrie. Dans ce chapitre, un vecteur est simplement un élément de $\mathbb{R}^n$ ou de $\mathbb{C}^n$. La notation $\mathbb{K}$ désigne indifféremment $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Définition de l'espace vectoriel $\mathbb{K}^n$

Notation

Un élément de $\mathbb{K}^n$ sera noté $\vec{v} = (v_1, v_2, \ldots, v_n)$. On dit que $\vec{v}$ est un vecteur *et que les $v_i \in \mathbb{K}$ pour $i = 1, \ldots, n$, sont* ses coordonnées.

Définition et notation 2.1.1

On définit deux opérations dans $\mathbb{K}^n$:

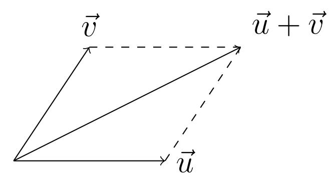
- *L'*addition *de deux vecteurs $\vec{u} = (u_1, u_2, \ldots, u_n)$ et $\vec{v} = (v_1, v_2, \ldots, v_n)$ est le vecteur noté $\vec{u} + \vec{v}$ et obtenu en additionnant terme à terme les coordonnées de chacun des vecteurs :*

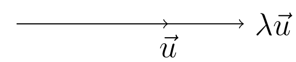
$$\vec{u} + \vec{v} = (u_1 + v_1, \ldots, u_n + v_n).$$

- *La multiplication d'un vecteur $\vec{u} = (u_1, u_2, \ldots, u_n)$ par un scalaire $\lambda \in \mathbb{K}$ est le vecteur obtenu en multipliant par λ les coordonnées de $\vec{u}$:*

$$\lambda\vec{u} = (\lambda u_1, \ldots, \lambda u_n).$$

On appelle vecteur nul *et on note $\vec{0}$ le vecteur $\vec{0} = (0, 0, \ldots, 0)$. On appelle* opposé *de $\vec{v}$ le vecteur $-\vec{v} = (-v_1, -v_2, \ldots, -v_n)$.*

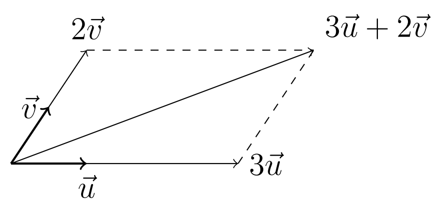




Proposition 2.1.2
Les opérations d'addition et de multiplication par un scalaire dans $\mathbb{K}^n$ ont les propriétés suivantes : pour tout $\vec{u}, \vec{v}, \vec{w} \in \mathbb{K}^n$ et tout $\lambda, \mu \in \mathbb{K}$,

1. $\vec{u}+\vec{v} = \vec{v}+\vec{u}$ *(commutativité de l'addition)*
2. $(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w})$ *(associativité de l'addition)*
3. $\vec{u}+\vec{0} = \vec{u}$
4. $\vec{u}+(-\vec{u}) = \vec{0}$
5. $1\,\vec{u} = \vec{u}$, $0\,\vec{u} = \vec{0}$.
6. $\lambda(\mu\vec{u}) = (\lambda\mu)\vec{u}$
7. $\lambda(\vec{u}+\vec{v}) = \lambda\vec{u}+\lambda\vec{v}$
8. $(\lambda+\mu)\vec{u} = \lambda\vec{u}+\mu\vec{u}$

En combinant addition et multiplication scalaire, on obtient des *combinaisons linéaires* de vecteurs. Par exemple, $3\vec{u}+2\vec{v}$ est une combinaison linéaire des vecteurs $\vec{u}$ et $\vec{v}$.



Définition 2.1.3
On dit que $\vec{u} \in \mathbb{K}^n$ est une combinaison linéaire *des vecteurs $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ s'il existe des scalaires $\lambda_1, \dots, \lambda_k$ tels que*

$$\vec{u} = \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k.$$

Les nombres $\lambda_1, \dots, \lambda_k$ s'appellent les coefficients de la combinaison linéaire.

Exercice
Dans $\mathbb{R}^3$, est-ce que $(1,1,0)$ est combinaison linéaire de $(1,2,0)$ et de $(0,1,0)$?

Pour répondre à cette question, on cherche s'il existe λ_1 et λ_2 tels que $(1,1,0) =$

$\lambda_1(1,2,0)+\lambda_2(0,1,0)$. *On a*

$$\begin{aligned}(1,1,0)=\lambda_1(1,2,0)+\lambda_2(0,1,0) &\Longleftrightarrow (1,1,0)=(\lambda_1, 2\lambda_1+\lambda_2, 0)\\ &\Longleftrightarrow \left\{\begin{array}{rcl}1 & = & \lambda_1\\ 1 & = & 2\lambda_1+\lambda_2\end{array}\right.\\ &\Longleftrightarrow \left\{\begin{array}{rcl}\lambda_1 & = & 1\\ \lambda_2 & = & -1\end{array}\right.\end{aligned}$$

On a trouvé la relation $(1,1,0)=(1,2,0)-(0,1,0)$. *On remarque que, pour calculer les coefficients* λ_1,λ_2, *on a été amenés à résoudre un système linéaire.*

2.2 Sous-espaces vectoriels de $\mathbb{K}^n$

Définition 2.2.1
Soit $F \subset \mathbb{K}^n$. *On dit que* F *est un* sous-espace vectoriel de $\mathbb{K}^n$ *si* F *vérifie les propriétés suivantes :*

1. *F est non vide*
2. *Pour tout* $\vec{u}\in F$, $\vec{v}\in F$, $\vec{u}+\vec{v}\in F$.
3. *Pour tout* $\vec{u}\in F$, $\lambda\in\mathbb{K}$, $\lambda\vec{u}\in F$.

Remarque 2.2.2
En combinant les propriétés 2. et 3., on voit qu'un sous-espace vectoriel est stable par combinaison linéaire, *c'est-à-dire que toute combinaison linéaire de vecteurs d'un sous-espace vectoriel appartient à ce sous-espace vectoriel.*

Exemple 2.2.3
$\{\vec{0}\}$ *et* $\mathbb{K}^n$ *sont des sous-espaces vectoriels de* $\mathbb{K}^n$.

Remarque 2.2.4
Il est important de remarquer qu'un sous-espace vectoriel contient toujours $\vec{0}$. *En effet, si* F *est un sous-espace vectoriel de* $\mathbb{K}^n$ *alors il contient au moins un élément* $\vec{u}$ *; mais alors il contient aussi* $0\,\vec{u}=\vec{0}$. *Par conséquent, si* F *est un ensemble candidat à être un sous-espace vectoriel, pour démontrer qu'il est non vide le mieux est de montrer qu'il contient* $\vec{0}$. *(En fait dans la définition de sous-espace vectoriel on pourrait remplacer la condition "*F *non vide" par la condition "*F *contient* $\vec{0}$*").*

Il y a deux exemples très importants de sous-espaces vectoriels de $\mathbb{K}^n$, que l'on va voir maintenant.

2.3 Premier exemple

Le premier exemple de sous-espace vectoriel de $\mathbb{K}^n$ est l'ensemble des solutions d'un système d'équations linéaires homogène.

Proposition 2.3.1
L'ensemble des solutions d'un système d'équations linéaires homogène à k équations et n inconnues est un sous-espace vectoriel de $\mathbb{K}^n$.

Preuve. Soit $\mathcal{S}$ l'ensemble des solutions du système linéaire homogène (S) :

$$(S) \quad \left\{ \begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n & = & 0 \\ & \vdots & \\ a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,n}x_n & = & 0 \end{array} \right.$$

On remarque d'abord que $\vec{0} \in \mathcal{S}$, car les équations sont bien vérifiées lorsque $x_1 = \cdots = x_n = 0$. Donc cet ensemble n'est pas vide. Soit $\vec{u} \in \mathcal{S}$ et $\vec{v} \in \mathcal{S}$, on doit vérifier que $\vec{u} + \vec{v} \in \mathcal{S}$. Notons $\vec{w} = \vec{u} + \vec{v}$. On a, pour tout $i = 1, \ldots, k$:

$$\begin{aligned} a_{i,1}w_1 + a_{i,2}w_2 + \cdots + a_{i,n}w_n &= a_{i,1}(u_1 + v_1) + a_{i,2}(u_2 + v_2) + \cdots + a_{i,n}(u_n + v_n) \\ &= a_{i,1}u_1 + a_{i,1}v_1 + a_{i,2}u_2 + a_{i,2}v_2 + \cdots + a_{i,n}u_n + a_{i,n}v_n \\ &= (a_{i,1}u_1 + a_{i,2}u_2 + \cdots + a_{i,n}u_n) + (a_{i,1}v_1 + a_{i,2}v_2 + \cdots + a_{i,n}v_n) \\ &= \underbrace{(a_{i,1}u_1 + a_{i,2}u_2 + \cdots + a_{i,n}u_n)}_{=0} + \underbrace{(a_{i,1}v_1 + a_{i,2}v_2 + \cdots + a_{i,n}v_n)}_{=0} = 0. \end{aligned}$$

On doit vérifier également que $\lambda\vec{u} \in \mathcal{S}$:

$$\begin{aligned} a_{i,1}(\lambda\vec{u})_1 + a_{i,2}(\lambda\vec{u})_2 + \cdots + a_{i,n}(\lambda\vec{u})_n &= a_{i,1}(\lambda u_1) + a_{i,2}(\lambda u_2) + \cdots + a_{i,n}(\lambda u_n) \\ &= \lambda(a_{i,1}u_1 + a_{i,2}u_2 + \cdots + a_{i,n}u_n) \\ &= \lambda \underbrace{(a_{i,1}u_1 + a_{i,2}u_2 + \cdots + a_{i,n}u_n)}_{=0} = 0. \end{aligned}$$

□

Exercice
Montrez que, si le système linéaire n'est pas homogène, l'ensemble de ses solutions n'est pas un sous-espace vectoriel.

2.4 Deuxième exemple

Le deuxième exemple de sous-espace vectoriel est construit à partir d'un ensemble fini de vecteurs $\vec{v}_1, \ldots, \vec{v}_k$ en prenant toutes leurs combinaisons linéaires.

Définition et proposition 2.4.1
Soit $\vec{v}_1, \ldots, \vec{v}_k$ des vecteurs de $\mathbb{K}^n$. On note $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$ l'ensemble des combinaisons linéaires de $\vec{v}_1, \ldots, \vec{v}_k$:

$$\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k) = \{\vec{u} \in \mathbb{K}^n \mid \exists(\lambda_1, \ldots, \lambda_k) \in \mathbb{K}^n \mid \vec{u} = \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k\}$$

Alors $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$ est un sous-espace vectoriel de $\mathbb{K}^n$.

Preuve. Posons $F = \mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k)$. On remarque d'abord que $\vec{0} = 0\,\vec{v}_1 + \cdots + 0\,\vec{v}_k \in F$ donc F est non vide. Soit $\vec{u} \in F$ et $\vec{w} \in F$. Il existe $(\lambda_1, \dots, \lambda_k) \in \mathbb{K}^n$ tels que $\vec{u} = \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k$ et il existe $(\mu_1, \dots, \mu_k) \in \mathbb{K}^n$ tels que $\vec{w} = \mu_1\vec{v}_1 + \cdots + \mu_k\vec{v}_k$. Alors,

$$\begin{aligned} \vec{u} + \vec{w} &= (\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k) + (\mu_1\vec{v}_1 + \cdots + \mu_k\vec{v}_k) \\ &= (\lambda_1 + \mu_1)\vec{v}_1 + \cdots + (\lambda_k + \mu_k)\vec{v}_k \in F. \end{aligned}$$

On a aussi, pour tout $\lambda \in \mathbb{K}$: $\lambda\vec{u} = \lambda(\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k) = (\lambda\lambda_1)\vec{v}_1 + \cdots + (\lambda\lambda_k)\vec{v}_k \in F$. Donc F est bien un sous-espace vectoriel de $\mathbb{K}^n$. $\square$

2.5 Somme et intersection de sous-espaces vectoriels

Dans ce paragraphe nous allons décrire deux façons de construire un nouveau sous-espace vectoriel à partir de deux sous-espaces vectoriels.

Proposition 2.5.1
Si F et G sont deux sous-espaces vectoriels de $\mathbb{K}^n$, alors leur intersection $F \cap G$ est aussi un sous-espace vectoriel de $\mathbb{K}^n$.

Preuve. Posons $V = F \cap G$. On a vu que $\vec{0} \in F$ et $\vec{0} \in G$ donc on a $\vec{0} \in V$ et celui-ci est non vide. Soit $\vec{u}$ et $\vec{v}$ deux éléments de V. En particulier, $\vec{u}$ et $\vec{v}$ appartiennent à F, donc, puisque F est un sous-espace vectoriel de $\mathbb{K}^n$, $\vec{u} + \vec{v} \in F$. Par le même raisonnement, on montre que $\vec{u} + \vec{v} \in G$. On a donc que $\vec{u} + \vec{v} \in F \cap G = V$. De la même façon, on montre que, si $\lambda \in \mathbb{K}$ et $\vec{u} \in V$, alors $\lambda\vec{u} \in V$. $\square$

Remarque 2.5.2
Attention, la réunion de deux sous-espaces vectoriels n'est pas en général un sous-espace vectoriel. La bonne notion est celle de somme de deux sous-espaces vectoriels.

Définition et proposition 2.5.3
Si F et G sont deux sous-espaces vectoriels de $\mathbb{K}^n$, on note $F + G$ et on appelle somme de F et G *l'ensemble*

$$F + G = \{\vec{u} + \vec{v} \text{ tels que } \vec{u} \in F \text{ et } \vec{v} \in G\}.$$

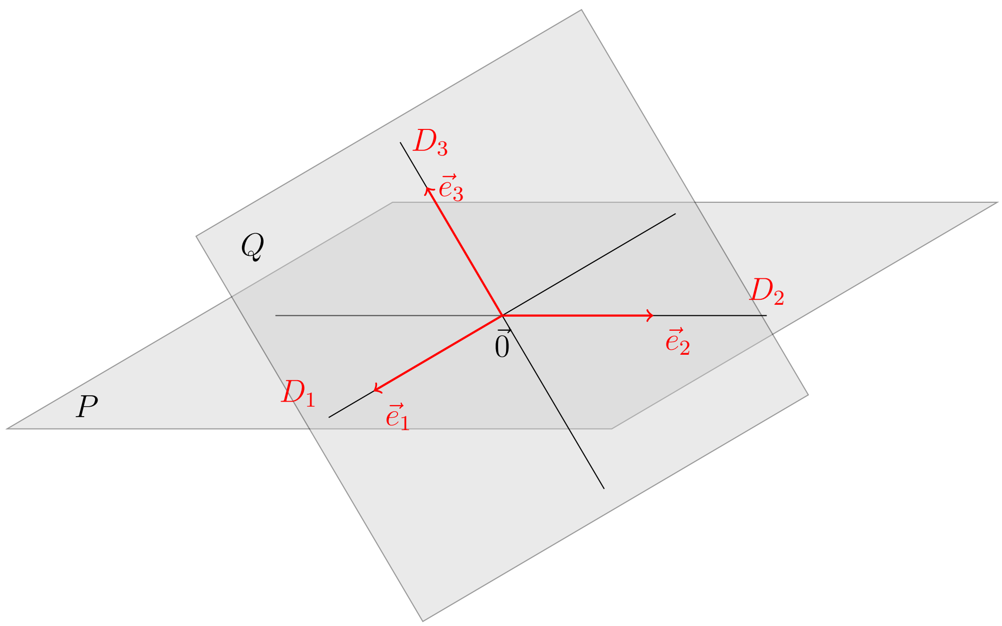
Alors $F + G$ est un sous-espace vectoriel de $\mathbb{K}^n$.

Preuve. Tout d'abord, $F + G$ contient $\vec{0}$ car $\vec{0} = \vec{0} + \vec{0}$ donc il est non vide. Ensuite, on vérifie qu'il est stable par addition : en effet, si $(\vec{u}, \vec{u}') \in F^2$ et $(\vec{v}, \vec{v}') \in G^2$, $(\vec{u} + \vec{v}) + (\vec{u}' + \vec{v}') = (\vec{u} + \vec{u}') + (\vec{v} + \vec{v}') \in F + G$. L'ensemble $F + G$ est aussi stable par multiplication par les scalaires car $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v} \in F + G$. $\square$

Remarque 2.5.4
Remarquons que $F+G$ contient F et G. En effet, si $\vec{u} \in F$, on peut écrire $\vec{u} = \vec{u}+\vec{0}$ donc $\vec{u} \in F+G$. On peut démontrer que $F+G$ est le plus petit sous-espace vectoriel de $\mathbb{K}^n$ contenant F et G.

2.6 Illustration géométrique

Le dessin suivant représente des sous-espaces vectoriels de $\mathbb{R}^3$:



On a

$$D_1 = \mathrm{Vect}(\vec{e}_1) \qquad D_2 = \mathrm{Vect}(\vec{e}_2) \qquad D_3 = \mathrm{Vect}(\vec{e}_3)$$

qui sont des droites. On a aussi représenté deux plans :

$$P = \mathrm{Vect}(\vec{e}_1, \vec{e}_2) \qquad Q = \mathrm{Vect}(\vec{e}_1, \vec{e}_3)$$

Un bon exercice consiste à identifier toutes les intersections et sommes de ces sous-espaces. Par exemple :

$$D_1 \cap D_2 = \{\vec{0}\} \qquad P \cap D_3 = \{\vec{0}\} \qquad P \cap D_1 = D_1 \qquad P \cap Q = D_1$$

et

$$D_1 + D_2 = P \qquad P + D_3 = \mathbb{R}^3 \qquad P + D_1 = P \qquad P + Q = \mathbb{R}^3.$$

Chapitre 3

Espaces vectoriels abstraits

Dans ce chapitre nous généralisons la notion d'espace vectoriel dans un cadre abstrait. L'intérêt de cette démarche abstraite est de constituer une "boite à outil" de résultats qui pourront s'appliquer dans une grande variété de situations, et à des objets mathématiques à priori éloignés de la notions naive de vecteurs (par exemple : des fonctions, des suites, des polynômes, des matrices, etc..). Comme d'habitude la notation $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

3.1 Définitions

Définition 3.1.1
Soit E un ensemble non vide, muni de deux opérations : une opération (ou loi) interne appelée addition *et notée $+$:*

$$\begin{aligned} E \times E &\to E \\ (\vec{u}, \vec{v}) &\mapsto \vec{u} + \vec{v} \end{aligned}$$

et une opération (ou loi) externe appelée multiplication *et notée $\cdot$:*

$$\begin{aligned} \mathbb{K} \times E &\to E \\ (\lambda, \vec{u}) &\mapsto \lambda \cdot \vec{u} \end{aligned}$$

On dit que $(E, +, \cdot)$ est un espace vectoriel sur $\mathbb{K}$ *ou un* $\mathbb{K}$-espace vectoriel *si les propriétés suivantes sont vérifiées :*

- *L'addition sur E est :*
  (1) commutative : $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ pour tout $(\vec{u}, \vec{v}) \in E^2$
  (2) associative : $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ pour tout $(\vec{u}, \vec{v}, \vec{w}) \in E^3$
  (3) possède un 'zéro' (appelé vecteur nul) : il existe $\vec{0} \in E$ tel que $\vec{u} + \vec{0} = \vec{u}$, pour tout $\vec{u} \in E$
  (4) tout élément possède un opposé : pour tout $\vec{u} \in E$, il existe un élément noté $-\vec{u}$ tel que $\vec{u} + (-\vec{u}) = \vec{0}$.

- *La multiplication externe vérifie :*

  (5) $1 \cdot \vec{u} = \vec{u}$ *pour tout* $\vec{u} \in E$

  (6) $\lambda \cdot (\mu \cdot \vec{u}) = (\lambda\mu) \cdot \vec{u}$ *pour tout* $(\lambda, \mu) \in \mathbb{K}^2$, $\vec{u} \in E$.
- *La multiplication est* distributive sur l'addition *:*

  (7) $\lambda \cdot (\vec{u} + \vec{v}) = \lambda \cdot \vec{u} + \lambda \cdot \vec{v}$ *pour tout* $\lambda \in \mathbb{K}$, *pour tout* $(\vec{u}, \vec{v}) \in E^2$

  (8) $(\lambda + \mu) \cdot \vec{u} = \lambda \cdot \vec{u} + \mu \cdot \vec{u}$ *pour tout* $(\lambda, \mu) \in \mathbb{K}^2$, *pour tout* $\vec{u} \in E$.

On appelle les éléments d'un espace vectoriel des *vecteurs* et les éléments de $\mathbb{K}$ des *scalaires*. L'élément $\vec{0}$ est appelé le *vecteur nul*.

Les opérations d'espaces vectoriels vérifient quelques propriétés supplémentaires qui se déduisent de celles listées dans la définition (lesquelles sont parfois appelées les *axiomes d'espaces vectoriels*) :

Proposition 3.1.2
Si $(E, +, \cdot)$ *est un* $\mathbb{K}$*-espace vectoriel, alors :*

1. *Son vecteur nul est unique*
2. *Tout vecteur de E possède un unique opposé*
3. $\lambda \cdot \vec{0} = \vec{0}$ *pour tout* $\lambda \in \mathbb{K}$
4. $0 \cdot \vec{u} = \vec{0}$ *pour tout* $\vec{u} \in E$
5. *Si* $\lambda \cdot \vec{u} = \vec{0}$ *alors* $\lambda = 0$ *ou* $\vec{u} = \vec{0}$.
6. $-\vec{u} = (-1) \cdot \vec{u}$ *pour tout* $\vec{u} \in E$

Preuve.
1. Si $\vec{0}'$ est un autre vecteur nul, alors on a en appliquant la propriété (3) d'espace vectoriel, $\vec{0} + \vec{0}' = \vec{0}'$ mais aussi $\vec{0} + \vec{0}' = \vec{0}$ donc $\vec{0} = \vec{0}'$.

2. Si $\vec{u}$ possède deux opposés notés $\vec{v}$ et $\vec{w}$, alors $\vec{v} = \vec{v} + \vec{0} = \vec{v} + (\vec{u} + \vec{w}) = (\vec{v} + \vec{u}) + \vec{w} = \vec{0} + \vec{w} = \vec{w}$ (en appliquant successivement (3) (4) (2) (4) (3)).

3. On a $\lambda \cdot \vec{0} = \lambda \cdot (\vec{0} + \vec{0}) = \lambda \cdot \vec{0} + \lambda \cdot \vec{0}$ en appliquant successivement les propriétés d'espaces vectoriels (3) et (7), donc, en ajoutant l'opposé de $\lambda \cdot \vec{0}$, on trouve $\vec{0} = \lambda \cdot \vec{0}$.

4. On écrit $0 \cdot \vec{u} = (0 + 0) \cdot \vec{u} = 0 \cdot \vec{u} + 0 \cdot \vec{u}$ en appliquant (8) et on conclut de même que $0 \cdot \vec{u} = \vec{0}$.

5. Supposons que $\lambda \cdot \vec{u} = \vec{0}$ avec $\lambda \neq 0$, et montrons qu'alors $\vec{u} = \vec{0}$. On a $\vec{u} = 1 \cdot \vec{u} = (\lambda^{-1}\lambda) \cdot \vec{u} = \lambda^{-1} \cdot (\lambda \cdot \vec{u}) = \lambda^{-1} \cdot \vec{0} = \vec{0}$ en appliquant successivement (5), (6), et 4. que l'on vient de montrer.

6. On remarque que $\vec{u} + (-1) \cdot \vec{u} = 1 \cdot \vec{u} + (-1) \cdot \vec{u} = (1 + (-1)) \cdot \vec{u} = 0 \cdot \vec{u} = \vec{0}$ et on conclut grâce à l'unicité de l'opposé de $\vec{u}$ qu'on a démontré en 2.

□

Notation
Dorénavant, on va simplifier les notations en posant $\vec{u} + (-\vec{v}) = u - \vec{v}$ *et* $\lambda \cdot \vec{u} = \lambda\vec{u}$.

3.2 Exemples

On va maintenant décrire quelques exemples d'espaces vectoriels classiques :

Exemple 1 : $\mathbb{K}^n$ muni des opérations d'addition et de multiplication scalaire définies au chapitre précédent est bien sûr un $\mathbb{K}$-espace vectoriel au sens de notre définition abstraite.

Exemple 2 : $\{\vec{0}\}$ où $\vec{0}$ est un 'symbole'. Les lois internes et externes sont définies par $\vec{0}+\vec{0}=\vec{0}$ et $\lambda \cdot \vec{0}=\vec{0}$ et les propriétés (1) à (8) sont bien vérifiées. On dit que $\{\vec{0}\}$ est *l'espace vectoriel nul.*

Exemple 3 : On va maintenant décrire une famille vraiment nouvelle d'espaces vectoriels. Soit I un ensemble, et soit $\mathcal{F}(I,\mathbb{K})$ l'ensemble des applications $f : I \to \mathbb{K}$. Cet ensemble est aussi souvent noté $\mathbb{K}^I$. On définit sur $\mathcal{F}(I,\mathbb{K})$ une loi d'addition : si $f \in \mathcal{F}(I,\mathbb{K})$ et $g \in \mathcal{F}(I,\mathbb{K})$, la somme $f+g$ est l'application de I dans $\mathbb{K}$ définie par $(f+g)(x)=f(x)+g(x)$. Pour définir la loi externe, on doit définir une application $\lambda \cdot f$ de I dans $\mathbb{K}$. Pour cela on pose $(\lambda \cdot f)(x)=\lambda f(x)$ pour tout $x \in I$. On a donc bien défini nos deux opérations

$$\begin{array}{ccc} \mathcal{F}(I,\mathbb{K}) \times \mathcal{F}(I,\mathbb{K}) & \to & \mathcal{F}(I,\mathbb{K}) \\ (f,g) & \mapsto & f+g \end{array} \qquad \begin{array}{ccc} \mathbb{K} \times \mathcal{F}(I,\mathbb{K}) & \to & \mathcal{F}(I,\mathbb{K}) \\ (\lambda,f) & \mapsto & \lambda \cdot f \end{array}$$

On retrouve dans cette famille d'espaces des objets mathématiques familiers : les n-uplets, les suites, les applications numériques :

- Si $I=\{1,2,\ldots,n\}$, alors $\mathcal{F}(I,\mathbb{K})$ s'identifie à K^n.
- Si $I=\mathbb{N}$, alors $\mathcal{F}(\mathbb{N},\mathbb{K})$ est l'ensemble des suites réelles ou complexes.
- Si $I=[a,b]$, alors $\mathcal{F}([a,b],\mathbb{R})$ est l'ensemble des applications à valeurs réelles définies sur l'intervalle $[a,b]$.

Proposition 3.2.1
L'ensemble $\mathcal{F}(I,\mathbb{K})$ des applications définies sur un ensemble I et à valeurs dans $\mathbb{K}$, muni des opérations d'addition et de multiplication scalaire définies ci-dessus est un $\mathbb{K}$-espace vectoriel. Son vecteur nul est la fonction nulle prenant constamment la valeur 0.

Preuve. Les axiomes (1) à (8) sont bien vérifiés, essentiellement parce que les opérations ont lieu dans $\mathbb{K}$. □

3.3 Sous-espaces vectoriels

Définition 3.3.1
Soit $(E,+,\cdot)$ un espace vectoriel sur $\mathbb{K}$. Un sous-espace vectoriel de E est un sous-ensemble F de E vérifiant :

1. $F \neq \emptyset$
2. *Pour tout* $(\vec{u},\vec{v}) \in F^2$, $\vec{u}+\vec{v} \in F$
3. *Pour tout* $\lambda \in \mathbb{K}$, $\vec{u} \in F$, $\lambda\vec{u} \in F$.

Remarque 3.3.2
On peut condenser les propriétés 2. et 3. ci-dessus en une seule, qui s'énonce : pour tout $(\vec{u}, \vec{v}) \in F^2$*, pour tout* $(\lambda, \mu) \in \mathbb{K}^2$*,* $\lambda\vec{u} + \mu\vec{v} \in F$*.*

Théorème 3.3.3
Si F *est un sous-espace vectoriel de* $(E, +, \cdot)$ *alors* $(F, +, \cdot)$ *est lui-même un espace vectoriel sur* $\mathbb{K}$*.*

Preuve. Par définition, dire que F est un sous-espace vectoriel de E c'est dire que les opérations d'addition et de multiplication scalaire de E se restreignent en des opérations de F. Les axiomes d'espace vectoriel sont automatiquement vérifiés pour F par restriction, à ceci près qu'il faut vérifier que $\vec{0} \in F$ et que l'opposé d'un vecteur de F appartient à F. Montrons donc ces deux propriétés :

Puisque F est supposé non vide, soit $\vec{u} \in F$. Alors $0 \cdot \vec{u} \in F$ par la propriété 3.; comme E est un espace vectoriel, $0 \cdot u = \vec{0}$, donc $\vec{0} \in F$.

Soit $\vec{u} \in F$; $\vec{u}$ possède un opposé $-\vec{u}$ dans E. On a vu que $-\vec{u} = (-1) \cdot \vec{u}$ donc par la propriété 3., $-\vec{u} \in F$. □

Remarque 3.3.4
Le théorème 3.3.3 est très important dans la pratique. En effet, pour montrer qu'un certain ensemble est un espace vectoriel, plutôt que de vérifier tous les axiomes de la définition 3.1.1, ce qui est vite fastidieux, il est préférable de montrer que c'est un sous-espace vectoriel d'un espace vectoriel déjà connu (bien sûr, quand c'est possible). Voici quelques exemples, qui pour certains seront détaillés en TD :

- *L'ensemble des suites de support fini (où le support d'une suite* $(x_n)_{n\geq 0}$ *est l'ensemble des entiers* k *tels que* $x_k \neq 0$*) est un sous-espace vectoriel de* $\mathcal{F}(\mathbb{N}, \mathbb{K})$*.*
- *L'ensemble des suites de limite* 0 *est un sous-espace vectoriel de* $\mathcal{F}(\mathbb{N}, \mathbb{K})$*.*
- *L'ensemble des fonctions continues (respectivement dérivables) sur l'intervalle* $[a, b]$ *est un sous-espace vectoriel de* $\mathcal{F}([a, b], \mathbb{R})$*.*

3.4 Combinaisons linéaires

Soit E un $\mathbb{K}$-espace vectoriel. Si $\lambda_1, \ldots, \lambda_k$ appartiennent à $\mathbb{K}$, et si $\vec{v}_1, \ldots, \vec{v}_k$ appartiennent à E, alors

$$\vec{u} = \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k$$

appartient à E. On dit que $\vec{u}$ est une *combinaison linéaire* des vecteurs $\vec{v}_1, \ldots, \vec{v}_k$.

Proposition 3.4.1
Si F *est un sous-espace vectoriel de* E *alors toute combinaison linéaire d'éléments de* F *est dans* F *(on dit qu'un sous-espace vectoriel est* stable par combinaisons linéaires*).*

Réciproquement, un sous-ensemble non vide de E qui est stable par combinaisons linéaires est un sous-espace vectoriel de E.

Preuve. Soit $\vec{v}_1, \dots, \vec{v}_k$ des vecteurs de F et soit $\lambda_1, \dots, \lambda_k$ des éléments de $\mathbb{K}$. Par la propriété 3. de sous-espace vectoriel, $\lambda_i \vec{v}_i \in F$ pour tout $i = 1, 2, \dots, k$. Par la propriété 2., $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \in F$. Donc aussi $(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2) + \lambda_3 \vec{v}_3 = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3$, et ainsi de suite jusqu'à $\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k$.

La réciproque est évidente puisque $\vec{u} + \vec{v}$ et $\lambda \vec{u}$ sont des combinaisons linéaires d'éléments de F particulières. □

Comme dans $\mathbb{K}^n$, on obtient des sous-espaces vectoriels en prenant l'ensemble des combinaisons linéaires d'une famille finie de vecteurs.

Définition et proposition 3.4.2
Soit $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ des vecteurs de E. On note $\mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k)$ *l'ensemble des combinaisons linéaires de $\vec{v}_1, \dots, \vec{v}_k$:*

$$\mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k) = \{\vec{u} \in E \mid \exists (\lambda_1, \dots, \lambda_k) \in \mathbb{K}^k \mid \vec{u} = \lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k\}$$

Alors $\mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k)$ *est un sous-espace vectoriel de E appelé* le sous-espace vectoriel engendré par $\vec{v}_1, \dots, \vec{v}_k$.

Plus généralement, on peut définir le sous-espace vectoriel engendré par une partie quelconque (non nécessairement finie) de E :

Définition et proposition 3.4.3
Soit A une partie quelconque de E. On note $\mathrm{Vect}(A)$ *et on appelle* sous-espace vectoriel engendré par *A l'ensemble des combinaisons linéaires d'éléments de A. C'est le plus petit sous-espace vectoriel de E contenant A.*

3.5 Somme et intersection de sous-espaces vectoriels

On a les mêmes notions de somme et d'intersection de deux sous-espaces vectoriels que dans le cas de $\mathbb{K}^n$ vu au chapitre précédent. On se contente de répéter les définitions et résultats sans répéter les démonstrations qui sont identiques.

Proposition 3.5.1
Si F et G sont deux sous-espaces vectoriels de E, alors leur intersection $F \cap G$ est aussi un sous-espace vectoriel de E.

Définition et proposition 3.5.2
Si F et G sont deux sous-espaces vectoriels de E, on note $F + G$ et on appelle somme de F et G *l'ensemble*

$$F + G = \{\vec{u} + \vec{v} \text{ tels que } \vec{u} \in F \text{ et } \vec{v} \in G\}.$$

Alors $F + G$ est un sous-espace vectoriel de E.

Exercice
Montrez que, pour toute partie A de E, $\mathrm{Vect}(A)$ *est l'intersection de tous les sous-espaces vectoriels de E contenant A.*

Exercice
Montrez que, si F et G sont deux sous-espaces vectoriels de E, $F+G$ est le plus petit sous-espace vectoriel de E contenant $F \cup G$.

3.6 Produit direct d'espaces vectoriels

Dans ce paragraphe, on suppose que U et V sont deux espaces vectoriels sur $\mathbb{K}$ quelconques. On va construire un nouvel espace vectoriel noté $U \times V$.

Définition et proposition 3.6.1
Soit $(U,+,\cdot)$ et $(V,+,\cdot)$ deux espaces vectoriels sur $\mathbb{K}$. Le produit cartésien $U \times V$ muni des opérations :

- *Pour tout $(\vec{u},\vec{v}) \in U \times V$, $(\vec{u'},\vec{v'}) \in U \times V$, $(\vec{u},\vec{v}) + (\vec{u'},\vec{v'}) = (\vec{u}+\vec{u'},\vec{v}+\vec{v'})$*
- *Pour tout $\lambda \in \mathbb{K}$, $(\vec{u},\vec{v}) \in U \times V$, $\lambda \cdot (\vec{u},\vec{v}) = (\lambda \cdot \vec{u}, \lambda \cdot \vec{v})$*

est un espace vectoriel sur $\mathbb{K}$. On l'appelle le produit direct de U et V.

Preuve. Il faut vérifier les huit propriétés d'espace vectoriel. Le vecteur nul de $U \times V$ est la paire $(\vec{0}_U, \vec{0}_V)$; l'opposé de $(\vec{u},\vec{v})$ est $-(\vec{u},\vec{v}) = (-\vec{u},-\vec{v})$. La vérification ne pose pas de difficultés, elle est laissée au lecteur. □

Exemple 3.6.2
Si $U = \mathbb{K}^k$ et $V = \mathbb{K}^n$ alors $U \times V = \mathbb{K}^k \times \mathbb{K}^n$ s'identifie à $\mathbb{K}^{k+n}$ (on dira plus tard qu'ils sont isomorphes*).*

Chapitre 4

Familles génératrices, libres, bases

Dans tout le chapitre, E désigne un $\mathbb{K}$-espace vectoriel. On garde en tête le cas très important de $E = \mathbb{K}^n$, sur lequel on va se baser pour les exemples illustratifs.

On a vu comme exemples d'espaces vectoriels les espaces de la forme $E = \text{Vect}(\vec{e}_1, \ldots, \vec{e}_k)$. On dit dans ce cas que $(\vec{e}_1, \ldots, \vec{e}_k)$ *engendre* ou *est une famille génératrice* de E. L'avantage, c'est qu'on peut exprimer tous les vecteurs de E avec seulement k d'entre eux. On va voir dans ce chapitre que, d'une part, il y a des parties génératrices plus intéressantes que d'autres parce que minimales (ce sont les bases de E) et, d'autres part, que tous les espaces vectoriels de la forme $E = \text{Vect}(\vec{e}_1, \ldots, \vec{e}_k)$ possèdent des bases.

4.1 Familles génératrices

Définition 4.1.1
Soit $\vec{e}_1, \ldots, \vec{e}_k$ des vecteurs de E. On dit que $(\vec{e}_1, \ldots, \vec{e}_k)$ est une famille génératrice de E *si tout vecteur de E est combinaison linéaire de $\vec{e}_1, \ldots, \vec{e}_k$, ou, de façon équivalente, si on a*

$$E = \text{Vect}(\vec{e}_1, \ldots, \vec{e}_k).$$

Par convention, la famille vide $\emptyset$ est génératrice de l'espace nul $\{\vec{0}\}$.

Exemple 4.1.2
Dans $\mathbb{K}^n$, on défini les vecteurs :

$$\begin{aligned}
\vec{\epsilon}_1 &= (1, 0, 0, \ldots, 0) \\
\vec{\epsilon}_2 &= (0, 1, 0, \ldots, 0) \\
&\vdots \\
\vec{\epsilon}_n &= (0, 0, 0, \ldots, 1)
\end{aligned}$$

Alors la famille $(\vec{\epsilon}_1, \ldots, \vec{\epsilon}_n)$ est génératrice de $\mathbb{K}^n$.

En effet, soit $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{K}^n$. Il est clair que $\vec{v} = v_1\vec{\epsilon}_1 + \cdots + v_n\vec{\epsilon}_n$.

Exercice
Soit F et G deux sous-espaces vectoriels de E tels que $F = \mathrm{Vect}(\vec{f}_1, \dots, \vec{f}_k)$ et $G = \mathrm{Vect}(\vec{g}_1, \dots, \vec{g}_\ell)$. Montrez que $F \subset G$ si et seulement si pour tout $1 \leq i \leq k$, $\vec{f}_i$ est combinaison linéaire de $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_\ell$.

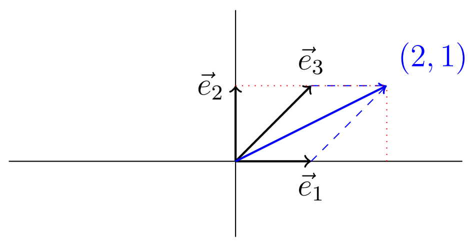
4.2 Familles libres et liées, bases

S'il est avantageux de pouvoir exprimer des vecteurs comme combinaison linéaire d'une famille fixée $(\vec{e}_1, \dots, \vec{e}_k)$, on remarque qu'une telle écriture n'est pas forcément unique. Par exemple, si on prend dans $\mathbb{R}^2$ les vecteurs $\vec{e}_1 = (1,0)$, $\vec{e}_2 = (0,1)$, $\vec{e}_3 = (1,1)$, on a

$$(2,1) = 2(1,0) + (0,1) = (1,0) + (1,1)$$

soit

$$(2,1) = 2\vec{e}_1 + \vec{e}_2 = \vec{e}_1 + \vec{e}_3.$$



En calculant la différence des deux expressions du vecteur $(2,1)$, on obtient la relation

$$\vec{e}_1 + \vec{e}_2 - \vec{e}_3 = \vec{0}.$$

Quand une telle relation existe, on dit que les vecteurs sont *liés*. Comme on le voit sur cet exemple, c'est l'existence d'une telle relation de liaison qui fait obstacle à l'unicité des coefficients dans l'expression d'un vecteur comme combinaison linéaire de $\vec{e}_1, \vec{e}_2, \vec{e}_3$.

Définition 4.2.1
La famille $(\vec{e}_1, \dots, \vec{e}_k)$ de vecteurs de E est une famille liée *ou* linéairement dépendante *s'il existe des scalaires $\lambda_1, \dots, \lambda_k$* non tous nuls *tels que*

$$\lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k = \vec{0}.$$

La famille $(\vec{e}_1, \dots, \vec{e}_k)$ de vecteurs de E est une famille libre *ou* linéairement indépendante *s'ils ne sont pas liés, c'est-à-dire si l'implication suivante est vraie :*

$$\lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k = \vec{0} \implies \lambda_1 = \cdots = \lambda_k = 0$$

Par convention, la famille vide $\emptyset$ est libre.

Remarque 4.2.2
Une famille de vecteurs qui contient le vecteur nul, c'est-à-dire une famille de la forme $(\vec{0}, \vec{e}_2, \dots, \vec{e}_k)$*, est toujours liée. En effet, on a la relation :*

$$1.\vec{0} + 0.\vec{e}_2 + \cdots + 0.\vec{e}_k = \vec{0}$$

et les coefficients $1, 0, \dots, 0$ *sont bien non tous nuls.*

De façon plus générale, une famille de vecteurs qui contient une sous-famille liée est liée.

Exemple 4.2.3
La famille $(\vec{\epsilon}_1, \dots, \vec{\epsilon}_n)$ *de* $\mathbb{K}^n$ *est libre. En effet,*

$$\sum_{i=1}^{n} \lambda_i \vec{\epsilon}_i = (\lambda_1, \dots, \lambda_n)$$

donc si ce vecteur est égal au vecteur nul, c'est que tous les coefficients λ_i *sont nuls.*

Exemple 4.2.4
Une famille à un élément $(\vec{u})$ *est libre si et seulement si* $\vec{u} \neq \vec{0}$*. Pour une famille à deux éléments, il est facile de voir que* $(\vec{u}, \vec{v})$ *est libre si et seulement si* $\vec{u}$ *et* $\vec{v}$ *sont non colinéaires i.e. non nuls et non proportionnels.*

Proposition 4.2.5
Les vecteurs $(\vec{e}_1, \dots, \vec{e}_k)$ *sont liés si et seulement si l'un de ces vecteurs est combinaison linéaire des autres.*

Preuve. Supposons les vecteurs liés ; alors il existe $\lambda_1, \dots, \lambda_k$ *non tous nuls* tels que

$$\lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k = \vec{0}.$$

Puisque les scalaires $\lambda_1, \dots, \lambda_k$ ne sont pas tous nuls, il existe un indice i tel que $\lambda_i \neq 0$. Alors on a $\lambda_i \vec{e}_i = -\sum_{j \neq i} \lambda_j \vec{e}_j$ et donc $\vec{e}_i = -\sum_{j \neq i} (\lambda_j / \lambda_i) \vec{e}_j$.

Réciproquement, supposons que l'un des vecteurs, par exemple $\vec{e}_1$, soit combinaison linéaire des autres. Alors, il existe des coefficients $\lambda_2, \dots, \lambda_k$ tels que $\vec{e}_1 = \lambda_2 \vec{e}_2 + \cdots + \lambda_k \vec{e}_k$. On a alors la relation de dépendance linéaire

$$\vec{e}_1 - \lambda_2 \vec{e}_2 - \cdots - \lambda_k \vec{e}_k = \vec{0}$$

et les coefficients sont bien non tous nuls puisque le premier vaut 1. □

Théorème 4.2.6
Soit $(\vec{e}_1, \dots, \vec{e}_k)$ *une famille de vecteurs. Les deux propositions suivantes sont équivalentes :*

(i) La famille $(\vec{e}_1, \dots, \vec{e}_k)$ *est libre.*

(ii) Pour tout $\vec{v} \in \mathrm{Vect}(\vec{e}_1, \dots, \vec{e}_k)$*, il existe un* unique *k-uplet* $(\lambda_1, \dots, \lambda_k)$ *tel que* $\vec{v} = \lambda_1 \vec{e}_1 + \lambda_2 \vec{e}_2 + \cdots + \lambda_k \vec{e}_k$*.*

Preuve. Supposons d'abord (i). Soit $\vec{v} \in \text{Vect}(\vec{e}_1, \dots, \vec{e}_k)$.

Si $\vec{v} = \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \cdots + \lambda_k\vec{e}_k = \mu_1\vec{e}_1 + \mu_2\vec{e}_2 + \cdots + \mu_k\vec{e}_k$, la différence de ces deux expressions conduit à $\vec{0} = (\lambda_1 - \mu_1)\vec{e}_1 + \cdots + (\lambda_k - \mu_k)\vec{e}_k$. Puisque $(\vec{e}_1, \dots, \vec{e}_k)$ est libre, on peut en déduire $\lambda_1 - \mu_1 = \cdots = \lambda_k - \mu_k = 0$ soit $\lambda_1 = \mu_1, \dots, \lambda_k = \mu_k$.

Réciproquement, supposons (ii) vérifiée. Alors, si les vecteurs $(\vec{e}_1, \dots, \vec{e}_k)$ étaient liés, le vecteur $\vec{0}$ aurait deux expressions comme combinaison linéaire des $\vec{e}_i$: en effet, il existerait des scalaires *non tous nuls* $\lambda_1, \dots, \lambda_k$ tels que

$$\vec{0} = 0\,\vec{e}_1 + \cdots + 0\,\vec{e}_k = \lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k$$

ce qui contredirait (ii). □

Exercice
Soit $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (1,1,0)$ *et* $\vec{e}_3 = (-1,0,1)$. *Sont-ils libres ?*

Solution. Pour répondre à cette question, on considère une combinaison linéaire nulle de ces vecteurs, soit $\lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \lambda_3\vec{e}_3 = \vec{0}$.

$$\begin{aligned}
\lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \lambda_3\vec{e}_3 = \vec{0} &\iff \lambda_1(1,0,0) + \lambda_2(1,1,0) + \lambda_3(-1,0,1) = (0,0,0)\\
&\iff (\lambda_1 + \lambda_2 - \lambda_3, \lambda_2, \lambda_3) = (0,0,0)\\
&\iff \left\{\begin{array}{ccccccc} \lambda_1 & + & \lambda_2 & - & \lambda_3 & = & 0\\ & & \lambda_2 & & & = & 0\\ & & & & \lambda_3 & = & 0 \end{array}\right.\\
&\iff \left\{\begin{array}{ccc} \lambda_1 & = & 0\\ \lambda_2 & = & 0\\ \lambda_3 & = & 0 \end{array}\right.
\end{aligned}$$

Nous avons donc montré que la famille $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est libre. Remarquons que pour cela nous avons dû résoudre un système linéaire d'inconnues $\lambda_1, \lambda_2, \lambda_3$.

En général pour des vecteurs de $\mathbb{K}^n$, si on note, pour $1 \leq j \leq k$, $\vec{e}_j = (e_{1,j}, e_{2,j}, \dots, e_{n,j})$, on a l'équivalence :

$$\lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \cdots + \lambda_k\vec{e}_k = \vec{0} \iff \left\{\begin{array}{ccc} \lambda_1 e_{1,1} + \lambda_2 e_{1,2} + \cdots + \lambda_k e_{1,k} & = & 0\\ \lambda_1 e_{2,1} + \lambda_2 e_{2,2} + \cdots + \lambda_k e_{2,k} & = & 0\\ & \vdots & \\ \lambda_1 e_{n,1} + \lambda_2 e_{n,2} + \cdots + \lambda_k e_{n,k} & = & 0 \end{array}\right.$$

Ainsi, les vecteurs $(\vec{e}_1, \dots, \vec{e}_k)$ sont libres si et seulement si ce système linéaire homogène, d'inconnues $\lambda_1, \dots, \lambda_k$, admet pour unique solution $(0,0,\dots,0)$. On remarque que cela implique $k \leq n$ d'après la Proposition 1.4.1.

Définition 4.2.7
Une base *de E est une famille de vecteurs de E qui est à la fois libre et génératrice de E.*

Exemple 4.2.8 (La base canonique de $\mathbb{K}^n$)
On a vu que la famille $(\vec{\epsilon}_1, \ldots, \vec{\epsilon}_n)$ *où* $\vec{\epsilon}_1 = (1, 0, \ldots, 0), \ldots, \vec{\epsilon}_n = (0, 0, \ldots, 1)$ *est une famille libre et génératrice de* $\mathbb{K}^n$*. C'est donc une base, qu'on appelle* la base canonique de $\mathbb{K}^n$.

Théorème 4.2.9
Une famille $(\vec{e}_1, \ldots, \vec{e}_k)$ *de vecteurs de* E *est une base de* E *si et seulement, pour tout* $\vec{v} \in E$*, il existe un unique k-uplet* $(\lambda_1, \ldots, \lambda_k) \in \mathbb{K}^k$ *tel que* $\vec{v} = \lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k$.

Preuve. L'existence des coefficients pour tout vecteur $\vec{v}$ de E est équivalente au fait que la famille est génératrice de E. Le théorème 4.2.6 montre que l'unicité de ces coefficients est équivalente au fait que la famille est libre. $\square$

Remarque 4.2.10
Si $E = \{\vec{0}\}$*, l'ensemble vide* $\emptyset$ *est une base de* E*, d'après nos conventions.*

Définition et notation 4.2.11
Si $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_k)$ *est une base de* E *et si* $\vec{v} = \lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k$*, on dit que* $\lambda_1, \ldots, \lambda_k$ *sont* les coordonnées de $\vec{v}$ dans la base $\mathcal{B}$. *On note en général les coordonnées de* $\vec{v}$ *sous la forme d'un vecteur colonne :*

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix}$$

Un espace vectoriel qui possède une base finie, c'est-à-dire une base au sens de notre définition, est dit *de dimension finie*. Nous allons voir au paragraphe suivant que les espaces vectoriels engendrés par une famille finie de vecteurs sont de dimension finie.

4.3 Dimension des espaces vectoriels de type fini

Dans cette section on fait l'hypothèse qu'il existe des vecteurs $\vec{g}_1, \ldots, \vec{g}_m$ dans E tels que $E = \mathrm{Vect}(\vec{g}_1, \ldots, \vec{g}_m)$. On dit dans ce cas que E est *finiment engendré* ou encore que E est *de type fini*. Notons que c'est bien le cas pour $\mathbb{K}^n$ qui est engendré par $\vec{\epsilon}_1 = (1, 0, \ldots, 0)$, $\vec{\epsilon}_2 = (0, 1, 0, \ldots, 0), \ldots \vec{\epsilon}_n = (0, 0, \ldots, 1)$.

Nous allons très vite démontrer qu'un espace vectoriel finiment engendré possède des bases.

Théorème 4.3.1
Soit $(\vec{e}_1, \ldots, \vec{e}_k)$ *une famille de vecteurs de* E*. Si* $(\vec{e}_1, \ldots, \vec{e}_k)$ *est une famille génératrice de* E*, et si* $\vec{e}_k$ *est combinaison linéaire de* $\vec{e}_1, \ldots, \vec{e}_{k-1}$*, alors* $(\vec{e}_1, \ldots, \vec{e}_{k-1})$ *est une famille génératrice de* E*.*

Preuve. Par hypothèse, il existe des coefficicients $\mu_1, \ldots, \mu_{k-1}$ tels que $\vec{e}_k = \sum_{i=1}^{k-1} \mu_i \vec{e}_i$. Soit $\vec{v} \in E$; comme $(\vec{e}_1, \ldots, \vec{e}_k)$ est génératrice de E, il existe $(\lambda_1, \ldots, \lambda_k) \in \mathbb{K}^n$ tels que $\vec{v} = \lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k$. On obtient en remplaçant $\vec{e}_k$:

$$\begin{aligned} v &= \lambda_1 \vec{e}_1 + \cdots + \lambda_{k-1} \vec{e}_{k-1} + \lambda_k (\mu_1 \vec{e}_1 + \cdots + \mu_{k-1} \vec{e}_{k-1}) \\ &= (\lambda_1 + \lambda_k \mu_1) \vec{e}_1 + \cdots + (\lambda_{k-1} + \lambda_k \mu_{k-1}) \vec{e}_{k-1} \end{aligned}$$

On voit alors que $\vec{v}$ est combinaison linéaire de $\vec{e}_1, \ldots \vec{e}_{k-1}$. Donc on a montré que $(\vec{e}_1, \ldots \vec{e}_{k-1})$ est une famille génératrice de E. $\square$

Théorème 4.3.2 (Théorème de la base extraite)
Soit $(\vec{e}_1, \ldots, \vec{e}_k)$ une famille de vecteurs de E. Si $(\vec{e}_1, \ldots, \vec{e}_k)$ est une famille génératrice de E, alors il existe un sous-ensemble de cette famille qui est une base de E.

Preuve. Si la famille $(\vec{e}_1, \ldots, \vec{e}_k)$ est aussi libre, on a fini. Sinon, c'est qu'elle est liée, mais alors l'un de ses vecteurs est combinaison linéaire des autres. Sans perte de généralité, supposons que c'est le dernier, soit $\vec{e}_k$. D'après le Théorème 4.3.1, on peut enlever ce vecteur en conservant une famille génératrice de E. On itère ce procédé jusqu'à obtenir une famille libre et génératrice, donc une base. Notons qu'on ne peut pas itérer indéfiniment puisque à chaque étape la famille de vecteurs perd un élément. $\square$

Corollaire 4.3.3
E possède des bases.

Preuve. En effet, d'après le Théorème 4.3.2, on peut extraire de la famille génératrice $(\vec{g}_1, \ldots, \vec{g}_m)$ une base de E. $\square$

Théorème 4.3.4
Si $(\vec{e}_1, \ldots, \vec{e}_k)$ est une famille libre de E à k éléments et si $(\vec{h}_1, \ldots, \vec{h}_s)$ est une famille génératrice de E à s éléments alors $k \leq s$.

Preuve. Par le Théorème 4.3.2, on peut extraire de $(\vec{h}_1, \ldots, \vec{h}_s)$ une base de E ; notons-la $(\vec{f}_1, \ldots \vec{f}_n)$. Notons que nécessairement $n \leq s$.

Pour tout $j = 1, \ldots, k$, puisque $\vec{e}_j$ appartient à E, $\vec{e}_j$ est combinaison linéaire des vecteurs $\vec{f}_1, \ldots, \vec{f}_n$. Il existe donc des coefficients $a_{j,i}$ tels que

$$\vec{e}_j = a_{1,j} \vec{f}_1 + a_{2,j} \vec{f}_2 + \cdots + a_{n,j} \vec{f}_n.$$

On peut alors transformer une combinaison linéaire $\lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k$ des vecteurs $\vec{e}_1, \ldots, \vec{e}_k$ en une combinaison linéaire des vecteurs $\vec{f}_1, \ldots, \vec{f}_n$:

$$\begin{aligned} \lambda_1 \vec{e}_1 + \cdots + \lambda_k \vec{e}_k &= \lambda_1 (a_{1,1} \vec{f}_1 + \cdots + a_{n,1} \vec{f}_n) + \cdots + \lambda_k (a_{1,k} \vec{f}_1 + \cdots + a_{n,k} \vec{f}_n) \\ &= (\lambda_1 a_{1,1} + \cdots + \lambda_k a_{1,k}) \vec{f}_1 + \cdots + (\lambda_1 a_{n,1} + \cdots + \lambda_k a_{n,k}) \vec{f}_n \end{aligned}$$

Comme $(\vec{f_1}, \dots, \vec{f_n})$ est une base de E, on a

$$\lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k = \vec{0} \Longleftrightarrow \left\{ \begin{array}{rcl} \lambda_1 a_{1,1} + \cdots + \lambda_k a_{1,k} & = & 0 \\ \lambda_1 a_{2,1} + \cdots + \lambda_k a_{2,k} & = & 0 \\ & \vdots & \\ \lambda_1 a_{n,1} + \cdots + \lambda_k a_{n,k} & = & 0 \end{array} \right.$$

Dire que $(\vec{e}_1, \dots, \vec{e}_k)$ est libre, c'est dire que ce système linéaire homogène a pour unique solution $(\lambda_1, \dots, \lambda_k) = (0, \dots, 0)$. Ce système a k inconnues et n équations. On fait maintenant appel à un résultat du chapitre 1 : d'après la Proposition 1.4.1, un système linéaire ayant une solution unique a au moins autant d'équations que d'inconnues ; on a donc $n \geq k$, donc a fortiori $s \geq k$. □

Théorème 4.3.5
Soit $(\vec{e}_1, \dots, \vec{e}_k)$ une famille de vecteurs de E. Si $(\vec{e}_1, \dots, \vec{e}_k)$ est libre, et si $\vec{e}_{k+1} \notin \mathrm{Vect}(\vec{e}_1, \dots, \vec{e}_k)$, alors $(\vec{e}_1, \dots, \vec{e}_k, \vec{e}_{k+1})$ est libre.

Preuve. Soit $(\lambda_1, \dots, \lambda_{k+1}) \in \mathbb{K}^{k+1}$ tels que $\lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k + \lambda_{k+1}\vec{e}_{k+1} = \vec{0}$.

Si $\lambda_{k+1} = 0$: alors on a $\lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k = \vec{0}$. Mais $(\vec{e}_1, \dots, \vec{e}_k)$ est libre donc on peut conclure que $\lambda_1 = \cdots = \lambda_k = 0$.

Si $\lambda_{k+1} \neq 0$: alors on a $\vec{e}_{k+1} = -(\lambda_1/\lambda_{k+1})\vec{e}_1 - \cdots - (\lambda_k/\lambda_{k+1})\vec{e}_k$ ce qui contredit l'hypothèse $\vec{e}_{k+1} \notin \mathrm{Vect}(\vec{e}_1, \dots, \vec{e}_k)$.

Donc on peut conclure que la famille $(\vec{e}_1, \dots, \vec{e}_{k+1})$ est libre. □

Théorème 4.3.6 (Théorème de la base incomplète)
Soit $(\vec{e}_1, \dots, \vec{e}_k)$ une famille de vecteurs de E. Si $(\vec{e}_1, \dots, \vec{e}_k)$ est libre, alors il existe un ensemble (éventuellement vide) $(\vec{e}_{k+1}, \dots, \vec{e}_n)$ de vecteurs de E tels que $(\vec{e}_1, \dots, \vec{e}_k, \vec{e}_{k+1}, \dots, \vec{e}_n)$ soit une base de E.

Preuve. Si $(\vec{e}_1, \dots, \vec{e}_k)$ est aussi génératrice de E alors on a fini. Sinon, il existe un vecteur $\vec{e}_{k+1}$ qui est dans E mais pas dans $\mathrm{Vect}(\vec{e}_1, \dots, \vec{e}_k)$. Alors, d'après le Théorème 4.3.5, la famille $(\vec{e}_1, \dots, \vec{e}_k, \vec{e}_{k+1})$ est aussi libre.

On itère cette procédure, jusqu'à obtenir une famille génératrice de E. Cela arrive forcément au bout d'un nombre fini d'étapes parce qu'à chaque étape on augmente le cardinal de la famille, et on sait par le Théorème 4.3.4 qu'une famille libre est de cardinal au plus égal à m. □

Théorème 4.3.7
Tout sous-espace vectoriel de E possède des bases. En particulier, les sous-espaces vectoriels de $\mathbb{K}^n$ possèdent tous des bases.

Preuve. C'est la même preuve que celle du théorème de la base incomplète. Soit F un sous-espace vectoriel de E. On part de $\emptyset \subset F$ et on lui rajoute des vecteurs de F un par un, en les choisissant en dehors de l'espace vectoriel engendré par les précédents, et ce jusqu'à ce que la famille ainsi créée engendre F. Par construction, d'après le théorème

4.3.5, la famille de vecteurs est libre. D'après le théorème 4.3.4, son cardinal est au plus égal à m. Donc la procédure s'arrête au bout d'un nombre fini d'étapes, et donc F possède bien une base. $\square$

Remarque 4.3.8
Il est important de remarquer que les démonstrations des théorèmes de la base extraite et de la base incomplète fournissent des méthodes constructives pour construire des bases d'un espace vectoriel de type fini.

Théorème 4.3.9
Toutes les bases de E ont le même cardinal. Le cardinal d'une base de E s'appelle la dimension de E *et est notée* $\dim(E)$.

Preuve. Soit $(\vec{e}_1, \ldots, \vec{e}_k)$ et $(\vec{f}_1, \ldots, \vec{f}_\ell)$ deux bases de E. Puisque $(\vec{e}_1, \ldots, \vec{e}_k)$ est une famille libre et que $(\vec{f}_1, \ldots, \vec{f}_\ell)$ est une famille génératrice de E, d'après le Théorème 4.3.4, on a $k \leq \ell$. Réciproquement, puisque $(\vec{f}_1, \ldots, \vec{f}_\ell)$ est une famille libre et que $(\vec{e}_1, \ldots, \vec{e}_k)$ est une famille génératrice de E, on a aussi $\ell \leq k$. Donc $k = \ell$. $\square$

Exemple 4.3.10
On a $\dim(\mathbb{K}^n) = n$ *puisque* $(\vec{\epsilon}_1, \ldots, \vec{\epsilon}_n)$ *est une base de* $\mathbb{K}^n$*, et on a* $\dim(\{\vec{0}\}) = 0$ *puisque* $\emptyset$ *est une base de* $\{\vec{0}\}$.

Un espace de dimension 1 s'appelle une droite, et un espace de dimension 2 s'appelle un plan.

Corollaire 4.3.11
Soit $n = \dim(E)$.

1. *Une famille libre de E a au plus n éléments, et, si elle a n éléments, c'est une base de E.*
2. *Une partie génératrice de E a au moins n éléments, et, si elle a n éléments, c'est une base de E.*
3. *Si $F \subset E$ est un sous-espace vectoriel de E, alors $\dim(F) \leq n$, et si $\dim(F) = n$ alors $F = E$.*

Preuve. Conséquences immédiates des théorèmes 4.3.2, 4.3.6 et 4.3.9. $\square$

On termine ce paragraphe par la formule qui lie les dimensions de la somme $F + G$ et de l'intersection $F \cap G$ de deux sous-espaces vectoriels. C'est une conséquence du théorème de la base incomplète (théorème 4.3.6).

Théorème 4.3.12
Soit F et G deux sous-espaces vectoriels de E.

$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Preuve. Soit $d = \dim(F)$, $k = \dim(G)$, $\ell = \dim(F \cap G)$. Soit $(\vec{v}_1, \dots \vec{v}_\ell)$ une base de $F \cap G$. Par le théorème de la base incomplète (Théorème 4.3.6) il existe $(\vec{e}_{\ell+1}, \dots, \vec{e}_d)$ tels que $(\vec{v}_1, \dots, \vec{v}_\ell, \vec{e}_{\ell+1}, \dots, \vec{e}_d)$ soit une base de F et il existe $(\vec{f}_{\ell+1}, \dots, \vec{f}_k)$ tels que $(\vec{v}_1, \dots, \vec{v}_\ell, \vec{f}_{\ell+1}, \dots, \vec{f}_k)$ soit une base de G. On montre que la famille

$$\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_\ell, \vec{e}_{\ell+1}, \dots, \vec{e}_d, \vec{f}_{\ell+1}, \dots, \vec{f}_k)$$

forme une base de $F + G$:

Montrons que c'est une famille libre, et pour cela supposons que

$$\sum_{i=1}^{\ell} \lambda_i \vec{v}_i + \sum_{j=\ell+1}^{d} \mu_j \vec{e}_j + \sum_{s=\ell+1}^{k} \nu_s \vec{f}_s = \vec{0}.$$

On en déduit que $\sum_{j=\ell+1}^{d} \mu_j \vec{e}_j = -\sum_{i=1}^{\ell} \lambda_i \vec{v}_i + \sum_{s=\ell+1}^{k} \nu_s \vec{f}_s \in F \cap G$ et donc que $\mu_i = 0$ pour tout $i = \ell+1, \dots, d$. On obtient de même que $\nu_j = 0$ pour tout $j = \ell+1, \dots, k$. Il reste $\sum_{i=1}^{\ell} \lambda_i \vec{v}_i = \vec{0}$ qui conduit à $\lambda_i = 0$ pour tout $i = 1, \dots, \ell$.

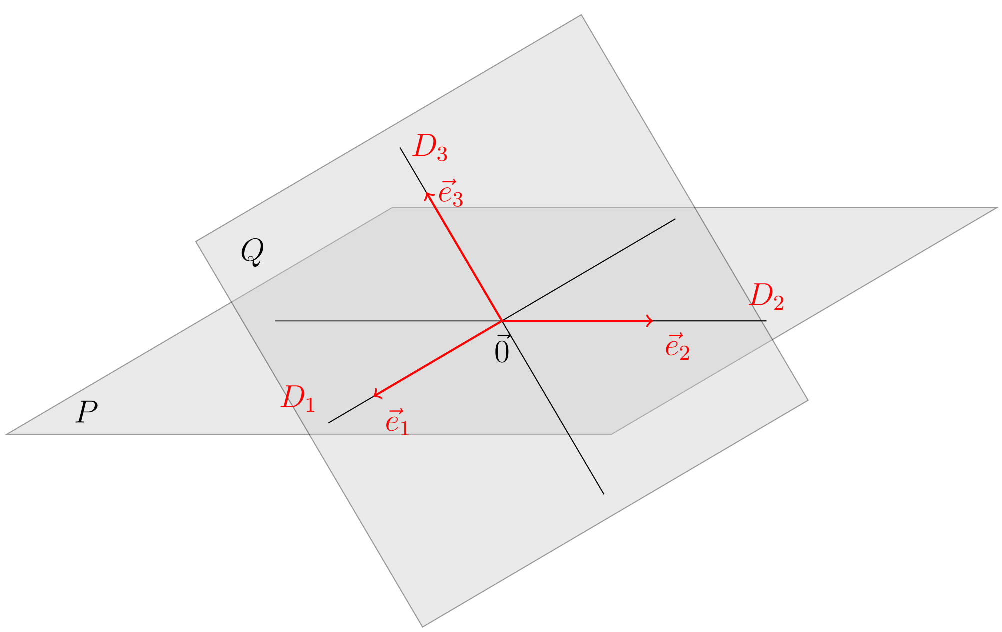
Montrons que $\mathcal{B}$ est une famille génératrice de $F + G$: en effet tout vecteur $\vec{w}$ de $F + G$ se décompose en $\vec{w} = \vec{u} + \vec{v}$ avec $\vec{u} \in F$ et $\vec{v} \in G$. Il reste à écrire ces vecteurs sur les bases respectives de ces espaces et à sommer leurs expressions pour constater que $\vec{w}$ est bien combinaison linéaire de la famille $\mathcal{B}$.

On obtient alors, en comptant les éléments de cette base $\mathcal{B}$ de $F + G$:

$$\dim(F + G) = \ell + (d - \ell) + (k - \ell) = d + k - \ell = \dim(F) + \dim(G) - \dim(F \cap G).$$

□

Illustrons la formule de Grassmann par un dessin :



On a représenté un espace E de dimension 3, et une base $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ de cet espace. On y a matérialisé trois droites : D_1, D_2, D_3 et deux plans : P, Q. On a $P \cap Q = D_1$ et $P + Q = E$. Noter qu'on a bien

$$\dim(P+Q) = \dim(P) + \dim(Q) - \dim(P \cap Q)$$

puisque $3 = 2 + 2 - 1$.

4.4 Le rang d'une famille de vecteurs

Dans cette section, on s'intéresse aux opérations que l'on peut faire subir à une famille de vecteurs sans changer le sous-espace que cette famille engendre. On s'intéresse à cette question dans le but de transformer le plus efficacement possible une famille génératrice d'un sous-espace en une base de ce même sous-espace. Nous reviendrons sur ce point de vue algorithmique dans le chapitre prochain sur les matrices.

On prend les notations suivantes : F est un sous-espace vectoriel de E engendré par la famille $(\vec{v}_1, \ldots, \vec{v}_k)$ de vecteurs de E. Autrement dit, $F = \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$.

Définition 4.4.1
On appelle opérations élémentaires *sur la famille $(\vec{v}_1, \ldots, \vec{v}_k)$ les transformations suivantes :*

1. *Échange de deux vecteurs*
2. *Remplacement d'un vecteur $\vec{v}_i$ par $\lambda\vec{v}_i$, pour $\lambda \in \mathbb{K}$, $\lambda \neq 0$.*
3. *Remplacement d'un vecteur $\vec{v}_i$ par $\vec{v}_i + \mu\vec{v}_j$ pour $j \neq i$ et $\mu \in \mathbb{K}$.*
4. *Suppression d'un vecteur $\vec{v}_i$ lorsque $\vec{v}_i = \vec{0}$*

Proposition 4.4.2
On ne change pas le sous-espace vectoriel engendré par une famille de vecteurs si on effectue une opération élémentaire (ou une succession de telles opérations) sur cette famille.

Preuve. Il est bien clair qu'un échange entre deux vecteurs ne va pas changer l'ensemble $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$ des combinaisons linéaires de $\vec{v}_1, \ldots, \vec{v}_k$.

Soit $\lambda \neq 0$; comme $\vec{v}_i = \lambda^{-1}(\lambda\vec{v}_i)$, on a $\vec{v}_i \in \mathrm{Vect}(\vec{v}_1, \ldots, \lambda\vec{v}_i, \ldots, \vec{v}_k)$; bien sûr, $\lambda\vec{v}_i \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$, donc on a bien $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k) = \mathrm{Vect}(\vec{v}_1, \ldots, \lambda\vec{v}_i, \ldots, \vec{v}_k)$.

Soit $\mu \in \mathbb{K}$ et $i \neq j$; soit $v_i' = \vec{v}_i + \mu\vec{v}_j$. Il est clair que $v_i' \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$; on a aussi $\vec{v}_i = v_i' - \mu\vec{v}_j$ donc $\vec{v}_i \in \mathrm{Vect}(\vec{v}_1, \ldots, v_i', \ldots, \vec{v}_k)$. On peut donc conclure que $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k) = \mathrm{Vect}(\vec{v}_1, \ldots, v_i', \ldots, \vec{v}_k)$. Enfin, si $\vec{0}$ apparait dans la famille de vecteurs, il est clair qu'on peut l'enlever sans changer l'espace engendré F. □

Ces opérations élémentaires sont analogues à celles que nous avons définies sur les systèmes linéaires dans le chapitre 1. Lorsque $E = \mathbb{K}^n$, nous verrons au chapitre suivant qu'il est commode d'organiser les vecteurs dans un tableau rectangulaire (une matrice), et que l'algorithme du pivot de Gauss permet d'échelonner ce tableau et d'en déduire une base de F.

Remarque 4.4.3
On a démontré dans le théorème 4.3.1 que, si l'un des vecteurs est combinaison linéaire des autres, on peut l'enlever de la famille sans changer F. Cette opération est en fait une combinaison des opérations élémentaires définies en 4.4.1. En effet, si $\vec{v}_k = \sum_{i=1}^{k-1} \lambda_i \vec{v}_i$, on peut remplacer $\vec{v}_k$ par $\vec{v}_k - \sum_{i=1}^{k-1} \lambda_i \vec{v}_i$ par $k-1$ opérations successives du type 3. puis supprimer ce vecteur qui est égal à $\vec{0}$ par 4.

On introduit maintenant la notion de rang d'une famille de vecteurs.

Définition 4.4.4
Le rang *d'une famille de vecteurs $(\vec{v}_1, \dots, \vec{v}_k)$ de E est la dimension du sous-espace vectoriel engendré par ses vecteurs. On le note* $\mathrm{rang}(\vec{v}_1, \dots, \vec{v}_k)$. *On a donc*

$$\mathrm{rang}(\vec{v}_1, \dots, \vec{v}_k) = \dim(\mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k)).$$

Proposition 4.4.5
On a les propriétés suivantes :

1. $\mathrm{rang}(\vec{v}_1, \dots, \vec{v}_k) \leq k$.
2. $\mathrm{rang}(\vec{v}_1, \dots, \vec{v}_k) = k$ *si et seulement si la famille* $(\vec{v}_1, \dots, \vec{v}_k)$ *est libre.*
3. $\mathrm{rang}(\vec{v}_1, \dots, \vec{v}_k) = k$ *si et seulement si* $(\vec{v}_1, \dots, \vec{v}_k)$ *est une base de F.*
4. *Une opération élémentaire sur la famille* $(\vec{v}_1, \dots, \vec{v}_k)$ *ne change pas son rang.*

Preuve. On note toujours $F = \mathrm{Vect}(\vec{v}_1, \dots, \vec{v}_k)$.

Les propriétés 1., 2., 3. sont vérifiées car la famille $(\vec{v}_1, \dots .\vec{v}_k)$ engendre F, en appliquant la propriété 2. du corollaire 4.3.11.

On a vu que les opérations élémentaires ne changent pas l'espace F, donc à fortiori elles ne changent pas sa dimension. □

4.5 Espaces vectoriels de dimension infinie

Dans ce paragraphe nous voulons insister sur le fait que tous les espaces vectoriels ne sont pas de dimension finie. L'exemple typique d'un espace vectoriel de dimension infinie est l'espace $\mathcal{F}(\mathbb{N}, \mathbb{K})$ des suites à valeurs dans $\mathbb{K}$. Notons $\mathcal{F}_k$ le sous-ensemble des suites nulles à partir du rang k :

$$\mathcal{F}_k = \{u = (u_n)_{n \geq 0} \in \mathcal{F}(\mathbb{N}, \mathbb{K}) \mid u_n = 0 \text{ pour tout } n \geq k\}.$$

Exercice : Montrez que $\mathcal{F}_k$ est un sous-espace vectoriel de $\mathcal{F}(\mathbb{N}, \mathbb{K})$.

Notons $\delta^0, \delta^1, \dots, \delta^i, \dots$ les suites suivantes :

$$
\begin{array}{rcccccc}
\delta^0 = & 1, & 0, & 0, & \dots & 0, & \dots \\
\delta^1 = & 0, & 1, & 0, & \dots & 0, & \dots \\
\vdots & & & & & & \\
\delta^i = & 0, & 0, & 0, & \dots & 1, & \dots \\
\vdots & & & & & &
\end{array}
$$

Plus précisément la suite δ^i est la suite dont tous les termes sont nuls sauf le terme d'indice i qui vaut 1. On peut aussi l'écrire ainsi :

$$
\delta^i = (\delta^i_n)_{n\geq 0} \quad \text{avec} \quad \delta^i_n = \begin{cases} 1 & \text{si } n = i \\ 0 & \text{si } n \neq i \end{cases}
$$

Il est clair que $\delta^0, \dots, \delta^{k-1}$ appartiennent à $\mathcal{F}_k$, et forment une base de $\mathcal{F}_k$. En effet, une suite $u \in \mathcal{F}_k$ s'écrit de façon unique $u = u_0\delta^0 + \cdots + u_{k-1}\delta^{k-1}$. Donc $\dim(\mathcal{F}_k) = k$.

L'espace vectoriel $\mathcal{F}(\mathbb{N}, \mathbb{K})$ contient donc des sous-espaces vectoriels de dimension arbitrairement grande, donc il ne peut pas lui-même être de dimension finie (en vertu de la propriété 3. du corollaire 4.3.11 qui dit que la dimension d'un espace vectoriel est toujours supérieure ou égale à celles de ses sous-espaces vectoriels).

Chapitre 5

Matrices

Une matrice est simplement un tableau rectangulaire de nombres. On verra que c'est un outil extrêmement pratique pour représenter certains objets de l'algèbre linéaire comme les familles de vecteurs, les systèmes linéaires ou les applications linéaires (dont on parlera au chapitre suivant), et pour effectuer des calculs sur ces objets.

5.1 Définitions

Définition 5.1.1
Une matrice A à ℓ lignes et n colonnes est un tableau rectangulaire d'éléments de $\mathbb{K}$ notée généralement

$$A = (a_{i,j})_{\substack{1 \leq i \leq \ell \\ 1 \leq j \leq n}} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{\ell,1} & a_{\ell,2} & \dots & a_{\ell,n} \end{pmatrix}$$

On dit que A est de taille (ℓ, n) *(ou $\ell \times n$) et que les $a_{i,j}$ sont ses* coefficients*.*

L'ensemble des matrices à ℓ lignes et n colonnes, à coefficients dans $\mathbb{K}$, est noté $\mathcal{M}_{\ell,n}(\mathbb{K})$. Si $\ell = n$, on note $\mathcal{M}_{\ell,n}(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$.

Il faut bien assimiler le repérage des coefficients d'une matrice : ceux-ci sont affectés d'un double indice i, j où i est l'indice de ligne, j est l'indice de colonne, et où les lignes et les colonnes sont numérotées à partir du coin en haut à gauche.

La terminologie relative aux matrices est assez étendue, on regroupe ici les définitions les plus utiles :

Définition et notation 5.1.2

- *La* matrice nulle *notée* $\mathbf{0}$*, est une matrice dont tous les coefficients sont égaux à* 0.
- *Une* matrice ligne *est une matrice de taille $(1, n)$ de la forme :*

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{pmatrix}$$

Une matrice colonne *est une matrice de taille $(\ell,1)$, de la forme :*

$$\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{\ell,1} \end{pmatrix}$$

On identifiera le plus souvent les matrices lignes, les matrices colonnes, et les vecteurs de $\mathbb{K}^n$.

- *Si A, est une matrice de taille (ℓ,n), on note $L_1,\dots,L_\ell$ ses lignes et $C_1,\dots,C_n$ ses colonnes.*

$$A = \begin{pmatrix} \dots & L_1 & \dots \\ & & \\ \dots & L_\ell & \dots \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ C_1 & & C_n \\ \vdots & & \vdots \end{pmatrix}$$

Avec les notations usuelles on a donc $L_i = (a_{i,1},\dots,a_{i,n}) \in \mathbb{K}^n$ et $C_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{\ell,j} \end{pmatrix} \in \mathbb{K}^\ell$.

- *La* matrice transposée *de A est la matrice notée A^T, à n lignes et ℓ colonnes, obtenue en échangeant dans A les lignes avec les colonnes :*

$$A^T = (a_{j,i})_{\substack{1\le i\le n \\ 1\le j\le \ell}}.$$

- *Une* matrice carrée *est une matrice ayant le même nombre de lignes que de colonnes. La* diagonale *d'une matrice carrée de taille n est le n-uplet des coefficients diagonaux $(a_{1,1}, a_{2,2}\dots, a_{n,n})$.*
- *Parmi les matrices carrées, on distingue les* matrices diagonales*, dont les coefficients en dehors de la diagonale sont nuls,*

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

ainsi que les matrices triangulaires supérieures *(ayant des zéros sous la diagonale), et* triangulaires inférieures *(ayant des zéros au-dessus de la diagonale).*

- *La* matrice identité *est la matrice diagonale, dont les coefficients diagonaux sont tous égaux à 1. On la note I ou I_n si elle est de taille n.*

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

5.2 Opérations sur les matrices

On définit d'abord l'addition de deux matrices et la multiplication d'une matrice par un scalaire, qui sont analogues aux opérations déjà vues sur les vecteurs.

Définition 5.2.1
On considère deux opérations dans $\mathcal{M}_{\ell,n}(\mathbb{K})$:

- *L'addition de deux matrices $A = (a_{i,j})_{\substack{1\le i\le \ell\\1\le j\le n}}$ et $B = (b_{i,j})_{\substack{1\le i\le \ell\\1\le j\le n}}$ est la matrice $A+B$ définie par :*
$$A + B = \big(a_{i,j} + b_{i,j}\big)_{\substack{1\le i\le \ell\\1\le j\le n}}$$
- *La multiplication d'une matrice $A = (a_{i,j})_{\substack{1\le i\le \ell\\1\le j\le n}}$ par un scalaire $\lambda \in \mathbb{K}$ est la matrice λA définie par :*
$$\lambda A = (\lambda a_{i,j})_{\substack{1\le i\le \ell\\1\le j\le n}}$$

Proposition 5.2.2
Les opérations d'addition et de multiplication par un scalaire dans $\mathcal{M}_{\ell,n}(\mathbb{K})$ ont les propriétés suivantes : pour tout $A, B, C \in \mathcal{M}_{\ell,n}(\mathbb{K})$ et tout $\lambda, \mu \in \mathbb{K}$,

1. $A + B = B + A$ *(commutativité de l'addition)*
2. $(A + B) + C = A + (B + C)$ *(associativité de l'addition)*
3. $A + \mathbf{0} = A$
4. $A + (-A) = \mathbf{0}$
5. $1\,A = A$, $0\,A = \mathbf{0}$.
6. $\lambda(\mu A) = (\lambda\mu)A$
7. $(\lambda + \mu)A = \lambda A + \mu A$
8. $\lambda(A + B) = \lambda A + \lambda B$

Autrement dit, $\mathcal{M}_{\ell,n}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel. Sa dimension est ℓn.

Preuve. Les propriétés 1. à 8. sont élémentaires et leur vérification est laissée au lecteur.

Définissons la matrice élémentaire $E_{i,j}$: c'est la matrice dont tous les coefficients sont nuls sauf le coefficient (i, j) qui est égal à 1. Il est clair que

$$A = \sum_{\substack{1\le i\le \ell\\1\le j\le n}} a_{i,j} E_{i,j}$$

et que cette expression de A comme combinaison linéaire des matrices $E_{i,j}$ est unique. Par conséquent, l'ensemble des matrices élémentaires $E_{i,j}$ est une base de $\mathcal{M}_{\ell,n}(\mathbb{K})$ (appelée la base canonique) et en particulier $\dim(\mathcal{M}_{\ell,n}(\mathbb{K})) = \ell n$. $\square$

Maintenant, nous allons définir une opération de *multiplication* entre matrices, qui est plus subtile. Nous allons procéder par étapes, en commençant par le produit d'une ligne par une colonne. Attention, il est impératif qu'ils aient le même nombre de coefficients.

Définition 5.2.3
Le produit d'une ligne par une colonne est défini lorsque qu'ils sont de même longueur*. Le résultat est un scalaire donné par la formule :*

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Exemple numérique : Calculons, le produit de $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ par $\begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}$.

$$\begin{pmatrix} 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix} = 1 \cdot 5 + 2 \cdot 2 + 0 \cdot 3 = 9.$$

Nous passons maintenant au produit de deux matrices rectangulaires. *Attention, il est nécessaire que le nombre de colonnes de la première soit égal au nombre de lignes de la deuxième.*

Définition 5.2.4
Soit A et B deux matrices. Le produit de A par B dans cet ordre est défini si le nombre de colonnes de A est égal au nombre de lignes de B. Soit donc A de taille (ℓ, n) et B de taille (n, p). Alors le produit $C = AB$ est la matrice de taille (ℓ, p) dont le coefficient $c_{i,j}$ est le produit de la i-ème ligne de A par la j-ème colonne de B :

$$c_{i,j} = \begin{pmatrix} a_{i,1} & a_{i,2} & \cdots & a_{i,n} \end{pmatrix} \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{n,j} \end{pmatrix} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \cdots + a_{i,n} b_{n,j}$$

c'est-à-dire

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \tag{5.1}$$

Exemple numérique : Calculons le produit suivant :

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

Notons que la première étant de taille $(2,3)$ et la deuxième de taille $(3,2)$, le résultat est une matrice de taille $(2,2)$.

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

$$\begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & -1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 9 & 7 \\ & \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & -1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 9 & 7 \\ 23 & \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & -1 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 9 & 7 \\ 23 & 9 \end{pmatrix}.$$

L'un des intérêts des matrices est de permettre la manipulation efficace des combinaisons linéaires de vecteurs. Pour cela, il est utile d'interpréter celles-ci en termes de produit de matrices, c'est l'objet de la proposition suivante.

Proposition 5.2.5
Le produit d'une matrice A par une matrice ligne à gauche est égal à la combinaison linéaire des lignes de A affectées des coefficients de la matrice ligne :

$$(x_1, x_2, \dots, x_\ell) \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{\ell,1} & a_{\ell,2} & \dots & a_{\ell,n} \end{pmatrix} = x_1 L_1 + x_2 L_2 + \dots + x_\ell L_\ell.$$

Le produit d'une matrice A par une matrice colonne à droite est égal à la combinaison linéaire des colonnes de A affectées des coefficients de la matrice colonne :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{\ell,1} & a_{\ell,2} & \dots & a_{\ell,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 C_1 + x_2 C_2 + \dots + x_n C_n.$$

Preuve. Démontrons la première propriété. Notons d'abord que la ligne étant de taille $(1, \ell)$ et la matrice de taille (ℓ, n), le produit est de taille $(1, n)$ comme les lignes de A. C'est donc un vecteur $(y_1, \ldots, y_n)$. Par définition, $y_k = \sum_{i=1}^{\ell} x_i a_{i,k}$, et c'est bien le k-ième coefficient de $\sum_{i=1}^{\ell} x_i L_i$.

La propriété sur les colonnes se démontre de la même façon. □

Exemple numérique : Illustrons cela sur un exemple. Soit la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Le produit de A à gauche par la matrice ligne $\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$ est donc la somme des deux premières lignes de A :

$$\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 5 & 7 & 9 \end{pmatrix}$$

Le produit de A à gauche par une matrice B s'interprète comme la concaténation des combinaisons linéaires des lignes de A affectées des coefficients des lignes de B. Ainsi

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} L_1 + L_2 \\ L_2 - L_3 \end{pmatrix} = \begin{pmatrix} 5 & 7 & 9 \\ -3 & -3 & -3 \end{pmatrix}$$

Remarque 5.2.6
La proposition 5.2.5 permet aussi de réinterpréter les systèmes linéaires en termes matriciels. En effet, le système linéaire

$$\left\{ \begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n & = & b_2 \\ & \vdots & \\ a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,n}x_n & = & b_k \end{array} \right.$$

peut s'écrire sous la forme de l'équation matricielle :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{k,1} & a_{\ell,2} & \ldots & a_{k,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

La proposition suivante résume les propriétés du produit de matrices.

Proposition 5.2.7
La multiplication des matrices vérifie les propriétés suivantes, valables pour toutes matrices A, B, C sous réserve de la compatibilité de leurs tailles respectives avec les opérations mises en oeuvre :

1. $\mathbf{0}A = A\mathbf{0} = \mathbf{0}$ *et* $I_\ell A = AI_n = A$.
2. $(AB)C = A(BC)$ *(associativité de la multiplication).*
3. $A(B+C) = AB+AC$ *et* $(A+B)C = AC+BC$ *(distributivité de l'addition sur la multiplication).*
4. $(AB)^T = B^T A^T$.

Preuve. On démontre l'associativité, les autres démonstrations sont laissées au lecteur. Notons les tailles respectives de A, B et C par (ℓ, n), (n, p) et (p, q). Posons $D = BC$, $E = AB$, $M = A(BC) = AD$ et $N = (AB)C = EC$. Si l'on note respectivement d_{ij}, e_{ij}, m_{ij}, et n_{ij} le coefficients de ces matrices, on obtient, en appliquant la formule (5.1) :

$$m_{ij} = \sum_{k=1}^{n} a_{i,k} d_{k,j} = \sum_{k=1}^{n} a_{i,k} \left(\sum_{s=1}^{p} b_{k,s} c_{s,j} \right). \tag{5.2}$$

On peut alors "inverser les deux sommations" dans la formule précédente pour obtenir :

$$m_{ij} = \sum_{s=1}^{p} \left(\sum_{k=1}^{n} a_{i,k} b_{k,s} \right) c_{s,j} = \sum_{s=1}^{p} e_{i,s} c_{s,j} = n_{ij}. \tag{5.3}$$

□

Remarque 5.2.8
Attention, la multiplication des matrices n'est pas commutative, c'est-à-dire que l'on n'a pas, en général, pour des matrices carrées, $AB = BA$. Par exemple,

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

alors que

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

Notation
Si A est une matrice carrée de taille (n, n), i.e. si $A \in \mathcal{M}_n(\mathbb{K})$, on peut multiplier A par elle-même autant de fois que l'on veut. On note A^k le produit de A par elle-même itéré k fois, avec la convention $A^0 = I_n$.

5.3 Matrice échelonnée, pivot de Gauss

Soit $(\vec{v}_1, \ldots, \vec{v}_k)$ une famille de vecteurs de $\mathbb{K}^n$. Pour étudier l'espace vectoriel engendré par ces vecteurs, et notamment pour en calculer une base et sa dimension (i.e. le rang de la famille de vecteur introduit dans le paragraphe 4.4), il est pratique d'organiser ces vecteurs dans une matrice de taille (k, n). Nous allons voir un algorithme (le pivot de Gauss, analogue à celui sur les systèmes linéaires) qui permet de transformer cette matrice de façon à faire apparaître une base de $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_k)$.

Nous commençons par la notion de *matrice échelonnée.*

Définition 5.3.1
Informellement, une matrice est dite ligne échelonnée *si chacune de ses lignes a son premier coefficicient non nul décalé à droite par rapport à celui de la ligne précédente, comme dans la représentation suivante, où les $\bullet$ représentent des coefficients non nuls et les $*$ représentent des coefficients quelconques :*

$$\begin{pmatrix} \bullet & * & * & * & * & * & * & * & * \\ 0 & 0 & \bullet & * & * & * & * & * & * \\ 0 & 0 & 0 & \bullet & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Plus formellement, soit A une matrice de taille (ℓ, n). Pour tout $i = 1, \ldots, \ell$, on note j_i la position du premier coefficient non nul de la ligne L_i (s'il y en a un ; si la ligne ne contient que des 0 on note $j_i = \infty$). On dit que A est ligne échelonnée *si on a : $j_1 < j_2 < \cdots < j_\ell$. Les coefficients $a_{i,j_i} \neq 0$, pour $j_i < \infty$, s'appellent* les pivots *de A (matérialisés par des $\bullet$ dans le schéma ci-dessus).*

On dit que A est ligne échelonnée réduite *si A est ligne échelonnée, et si, de plus, les pivots sont égaux à 1 et si les coefficients situés au-dessus d'eux sont égaux à 0 (i.e pour tout $i = 1, \ldots, \ell$, $a_{i,j_i} = 1$, et $a_{s,j_i} = 0$ pour $1 \leq s < j_i$).*

On dit qu'une matrice est colonne échelonnée*, respectivement* colonne échelonnée réduite *si sa transposée est ligne échelonnée, respectivement ligne échelonnée réduite.*

Exemple numérique : Les matrices suivantes sont ligne échelonnées, et la dernière est de plus réduite.

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

Les matrices suivantes sont colonne échelonnées, et la dernière est de plus réduite.

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

L'intéret des matrices échelonnées réside dans le fait que leurs lignes (respectivement leur colonnes) non nulles sont automatiquement linéairement indépendantes, comme l'exprime le théorème suivant :

Théorème 5.3.2
Si A est une matrice ligne échelonnée de taille (ℓ, n), ses lignes non nulles sont linéairement indépendantes. En conséquence, elles forment une base du sous-espace vectoriel de $\mathbb{K}^n$ engendré par les lignes de A. Le rang des lignes de A est donc égal au nombre de ses pivots.

Si A est colonne échelonnée, ses colonnes non nulles sont linéairement indépendantes. En conséquence, elles forment une base du sous-espace vectoriel de $\mathbb{K}^\ell$ engendré par les colonnes de A. Le rang des colonnes de A est donc égal au nombre de ses pivots.

Preuve. On démontre le résultat pour les colonnes. Supposons donc que la matrice soit colonne échelonnée, et soit r l'indice de la dernière colonne non nulle. Soit $(\lambda_1, \ldots, \lambda_r) \in \mathbb{K}^r$ tels que $\lambda_1 C_1 + \cdots + \lambda_r C_r = \mathbf{0}$. De cette équation vectorielle, on retient les r équations linéaires correspondant aux pivots, c'est-à-dire aux coordonnées d'indice $j_1, j_2, \ldots, j_r$. Il est clair que le système correspondant est échelonné, homogène, et qu'il a r équations et r inconnues. Il a donc une solution unique $(0, 0, \ldots, 0)$.

Illustrons cela avec un schéma qui sera plus parlant : supposons que la matrice A soit de la forme

$$\begin{pmatrix} \bullet & 0 & 0 \\ * & 0 & 0 \\ * & \bullet & 0 \\ * & * & \bullet \end{pmatrix}$$

où, comme dans la définition 5.3.1, $\bullet$ représente un coefficient non nul (donc un pivot de A) et $*$ représente un coefficient quelconque. On veut donc montrer que ces trois colonnes sont linéairement indépendantes. La condition $\lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3 = \mathbf{0}$ se traduit matriciellement par

$$\begin{pmatrix} \bullet & 0 & 0 \\ * & 0 & 0 \\ * & \bullet & 0 \\ * & * & \bullet \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

On reconnait le système linéaire homogène de matrice A. On s'intéresse aux équations qui contiennent des pivots, c'est-à-dire aux équations 1, 3, 4, qui sont de la forme

$$\left\{\begin{array}{rcrcrcl} \bullet\,\lambda_1 & & & & & = & 0 \\ *\,\lambda_1 & + & \bullet\,\lambda_2 & & & = & 0 \\ *\,\lambda_1 & + & *\,\lambda_2 & + & \bullet\,\lambda_3 & = & 0 \end{array}\right.$$

On peut noter que l'équation 2 ne nous apporterait pas d'information supplémentaire, c'est pour cela qu'on la laisse tomber. Il est clair que la seule solution de ce système triangulaire est $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$.

Retournons au cas général. Nous avons donc démontré la première affirmation, c'est-à-dire que les colonnes non nulles de la matrice A sont linéairement indépendantes.

Les colonnes d'indice plus grand que r étant nulles, il est clair que $\mathrm{Vect}(C_1, \ldots, C_n) = \mathrm{Vect}(C_1, \ldots, C_r)$ et donc $(C_1, \ldots, C_r)$ est bien une base de l'espace vectoriel engendré par les colonnes de A. Le rang des colonnes de A est la dimension de cet espace ; c'est donc r, et comme il y a un pivot par colonne non nulle, c'est bien le nombre de pivots de A. $\square$

L'algorithme du pivot de Gauss

Cet algorithme peut s'effectuer sur les lignes ou sur les colonnes d'une matrice. Il permet d'échelonner les lignes (ou les colonnes) sans changer l'espace vectoriel engendré par ces lignes (ou par ces colonnes). Il effectue une succession d'opérations élémentaires sur les lignes (ou sur les colonnes) de la matrice. Nous avons déjà introduit ces opérations dans le paragraphe 4.4 du chapitre 4. Cet algorithme est identique à celui déjà vu sur les systèmes linéaires, nous nous contenterons de le décrire sur un exemple.

On retiendra :

Théorème 5.3.3
Soit A une matrice de taille (ℓ, n). Les opérations élémentaires sur les lignes $L_1, \ldots, L_\ell$ de A sont les suivantes :

1. *Échange de deux lignes (noté $L_i \leftrightarrow L_j$).*
2. *Multiplication d'une ligne par un scalaire non nul (noté $L_i \leftarrow \lambda L_i,\ \lambda \neq 0$).*
3. *Ajout à une ligne d'un multiple d'une autre ligne (noté $L_i \leftarrow L_i + \mu L_j$).*

L'algorithme du pivot de Gauss permet, par une succession de ces opérations, de mettre la matrice A sous forme ligne échelonnée de façon efficace et donc en particulier d'obtenir une base du sous-espace vectoriel de $\mathbb{K}^n$ engendré par ses lignes, ainsi que le rang de ces lignes.

Les opérations élémentaires sur les colonnes $C_1, \ldots, C_n$ de A sont les suivantes :

1. *Échange de deux colonnes (noté $C_i \leftrightarrow C_j$).*
2. *Multiplication d'une colonne par un scalaire non nul (noté $C_i \leftarrow \lambda C_i,\ \lambda \neq 0$).*
3. *Ajout à une colonne d'un multiple d'une autre colonne (noté $C_i \leftarrow C_i + \mu C_j$).*

L'algorithme du pivot de Gauss permet, par une succession de ces opérations, de mettre la matrice A sous forme colonne échelonnée de façon efficace et donc en particulier d'obtenir une base du sous-espace vectoriel de $\mathbb{K}^\ell$ engendré par ses colonnes, ainsi que le rang de ces colonnes.

Remarque 5.3.4
On peut donner un sens plus précis au terme "efficace" que nous utilisons pour qualifier l'algorithme du pivot de Gauss, pour cela il faudrait introduire la notion de complexité d'un algorithme *qui évalue le nombre d'opérations effectuées au cours de l'algorithme en fonction de la taille des données d'entrée. Ces notions très intéressantes font partie de ce que l'on appelle l'*algorithmique *et sortent du cadre de ce cours.*

Exemple numérique : On cherche une base du sous-espace de $\mathbb{R}^3$ engendré par les vecteurs $(0,1,-2)$, $(0,3,6)$, $(1,1,-1)$, $(1,-1,4)$. On place ces vecteurs dans les colonnes d'une matrice :

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

On échelonne les colonnes de cette matrice par le pivot de Gauss. On commence par positionner le premier pivot en échangeant les colonnes 1 et 3 ($C_1 \leftrightarrow C_3$).

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

Ensuite on utilise ce pivot pour annuler les coefficients sur sa droite. Pour cela il suffit de remplacer C_4 par $C_4 - C_1$ ($C_4 \leftarrow C_4 - C_1$)

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

Le deuxième pivot est en place, c'est 3. Il est plus facile de travailler avec des pivots égaux à 1, donc on divise C_2 par 3 ($C_2 \leftarrow C_2/3$).

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

Ensuite, pour annuler les coefficients à sa droite on exécute : $C_3 \leftarrow C_3 - C_2$ et $C_4 \leftarrow C_4 + 2C_2$ et on obtient :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 2 & -4 & 9 \end{pmatrix}$$

Il reste à diviser C_3 par -4, puis à remplacer C_4 par $C_4 - 9C_3$ pour obtenir une matrice colonne échelonnée :

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

Cette matrice a trois colonnes non nulles, qui forment une base du sous-espace engendré par les colonnes initiales. La famille de vecteurs $((0,1,-2),(0,3,6),(1,1,-1),(1,-1,4))$ est donc de rang 3.

5.4 Le rang d'une matrice

À une matrice $A \in \mathcal{M}_{\ell,n}(\mathbb{K})$, on peut associer deux familles de vecteurs : ses lignes, qui sont des vecteurs de $\mathbb{K}^n$, et ses colonnes, qui sont des vecteurs de $\mathbb{K}^\ell$. Ces deux familles n'ont à priori rien en commun, puisque, en général, elles ne vivent même pas dans le même espace. En effet, rappelons que les lignes de A sont de longueur n (elles appartiennent à $\mathbb{K}^n$) tandis que les colonnes son de longueur ℓ (elles appartiennent à $\mathbb{K}^\ell$). Pourtant, on va voir un résultat remarquable qui dit qu'elles engendrent des sous-espaces vectoriels *de même dimension*.

Théorème 5.4.1
Soit A une matrice de taille (ℓ, n). Le sous-espaces vectoriel de $\mathbb{K}^n$ engendré par les lignes de A et le sous-espace vectoriel de $\mathbb{K}^\ell$ engendré par les colonnes de A sont de même dimension. Cette dimension commune est appelée le rang de la matrice A *et est noté* $\mathrm{rang}(A)$. *En particulier,* $\mathrm{rang}(A) \leq \min(\ell, n)$.

Preuve. Soit A une matrice de taille (ℓ, n), soit$(L_1, \ldots, L_\ell)$ ses lignes, et $(C_1, \ldots, C_n)$ ses colonnes. Notons $E = \mathrm{Vect}(L_1, \ldots, L_\ell) \subset \mathbb{K}^n$, et $F = \mathrm{Vect}(C_1, \ldots, C_n) \subset \mathbb{K}^\ell$. Notre but est de montrer que $\dim(E) = \dim(F)$.

Soit $d = \dim(E)$ et soit $(\vec{e}_1, \ldots, \vec{e}_d)$ une base de E. Chacun des $\vec{e}_i$ appartient à $\mathbb{K}^n$; notons $\vec{e}_i = (e_{i,1}, \ldots, e_{i,n})$. Soit B la matrice dont les lignes sont les vecteurs $\vec{e}_i$. La matrice B est donc de taille (d, n) et son coefficient d'indice (i, j) est $e_{i,j}$.

Chacun des L_i est combinaison linéaire des $\vec{e}_j$. Notons $L_i = \sum_{j=1}^d \lambda_{i,j} \vec{e}_j$. On a donc (voir la proposition 5.2.5)

$$A = \Lambda B \quad \text{avec} \quad \Lambda = (\lambda_{i,j}) \in \mathcal{M}_{\ell,d}(\mathbb{K}).$$

Maintenant, on regarde l'équation matricielle $A = \Lambda B$ du point de vue des colonnes. Cette équation signifie que les colonnes de A sont des combinaisons linéaires des colonnes de Λ affectées des coefficients de la matrice B (toujours par la proposition 5.2.5). En particulier, les colonnes de A appartiennent à un sous-espace vectoriel engendré par d vecteurs (les d colonnes de Λ) donc elles sont de rang au plus d (par le point 1. de la proposition 4.4.5). Donc, $\dim(F) \leq d = \dim(E)$.

En remplaçant A par A^T, on obtient l'autre inégalité $\dim(E) \leq \dim(F)$.

$\square$

5.5 Matrices carrées inversibles

Dans cette section nous ne considérons que des matrices carrées. Nous avons vu dans les paragraphes précédent que dans l'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées de même taille n, à coefficients dans $\mathbb{K}$, on a des opérations d'addition et de multiplication analogues à celles de $\mathbb{K}$. En effet, la plupart des rêgles de calcul dans $\mathbb{K}$ sont vraies pour les matrices (voir les propositiones 5.2.2 et 5.2.7), ce qui fait qu'on peut calculer avec des matrices carrées presque "comme avec des nombres". Toutefois, nous avons déjà remarqué une différence notoire pour la multiplication, : celle-ci n'est plus commutative dans les matrices, c'est-à-dire que en général $AB \neq BA$.

Du côté des analogies qui marchent bien, le 0 de $\mathbb{K}$ est remplacé par la matrice nulle $\mathbf{0}$ puiqu'elle vérifie $\mathbf{0} + A = A + \mathbf{0} = A$ et, comme pour les éléments de $\mathbb{K}$, tout matrice A a un opposé $-A$ qui vérifie $A - A = \mathbf{0}$. Le 1 de $\mathbb{K}$ est remplacé par la matrice identité I_n puisque $I_nA = AI_n = A$. Par contre, nous allons voir que l'inversibilité des matrices est plus compliquée. Dans $\mathbb{K}$, on sait que tout élément non nul est inversible : pour tout $\lambda \in \mathbb{K}$, $\lambda \neq 0$, il existe un inverse λ^{-1} aussi noté $1/\lambda$, qui vérifie $\lambda\lambda^{-1} = 1$. L'inversion d'un élément de $\mathbb{K}$ est typiquement utilisé dans les équations de la façon suivante : si on a une équation $\lambda x = y$ et si $\lambda \neq 0$, on obtient une équation équivalente $x = \lambda^{-1}y$.

Nous introduisons maintenant la notion de matrice inversible :

Définition 5.5.1
Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite inversible *s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$.*

Il est important de prendre conscience que toutes les matrices non nulles ne sont pas inversibles. Pour cela, considérons la matrice carrée de taille 2 suivante : $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.
Si on multiplie A par une matrice $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ quelconque, on obtient :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix}$$

On constate que le résultat du produit AB est une matrice dont la deuxième ligne est nulle. Quelle que soit la matrice B, le résultat du produit AB ne peut donc jamais être la matrice $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Un raisonnement analogue montre qu'on ne peut pas non plus avoir $BA = I_2$. Donc A n'est pas une matrice inversible.

Dans ce paragraphe, nous allons caractériser les matrices inversibles (par leur rang) et nous allons également voir un algorithme (sans surprise, basé sur le pivot de Gauss) qui permet de décider si une matrice est inversible et de calculer son inverse lorqu'elle l'est.

On commence par un petit lemme technique qui nous sera utile pour montrer l'unicité de l'inverse d'une matrice, et que nous utiliserons à nouveau plus tard.

Lemme 5.5.2
S'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ *telle que* $AB = I_n$*, et une matrice* $C \in \mathcal{M}_n(\mathbb{K})$ *telle que* $CA = I_n$*, alors* $B = C$.

Preuve. On a donc $AB = I_n$ et $CA = I_n$. On peut alors multiplier l'égalité $AB = I_n$ à gauche par C ce qui donne $C(AB) = C$, soit par associativité du produit $(CA)B = C$, et donc $I_nB = C$ ce qui donne $B = C$. $\square$

Corollaire 5.5.3
Si A *est inversible, il existe une* unique *matrice* B *vérifiant* $AB = BA = I_n$*. on l'appelle alors* l'inverse de A *et on la note* A^{-1}.

Preuve. Cela résulte immédiatement du lemme précédent. En effet, si A a deux inverse disons B et C, alors on a $AB = BA = I_n$ et $AC = CA = I_n$ et d'après le lemme on obtient que $B = C$. $\square$

Nous allons dans le prochain théorème donner une caractérisation des matrices inversibles par leur rang.

Théorème 5.5.4
Soit $A \in \mathcal{M}_n(\mathbb{K})$*. Les propriétés suivantes sont équivalentes :*

1. A *est inversible.*
2. *Il existe* $B \in \mathcal{M}_n(\mathbb{K})$ *telle que* $AB = I_n$
3. *Il existe* $C \in \mathcal{M}_n(\mathbb{K})$ *telle que* $CA = I_n$.
4. $\mathrm{rang}(A) = n$.

Remarque 5.5.5
L'équivalence entre 1. et 4. est le point le plus important du théorème : les matrices inversibles sont les matrices dont le rang est maximal (c'est-à-dire égal à la taille de la matrice).

Il faut bien comprendre la différence entre le point 1. et les points 2. et 3. . Dans 2., on fait l'hypothèse à priori plus faible que A *a un inverse* à droite *(et dans 3. que* A *a un inverse* à gauche*). L'équivalence entre 1. et 2. signifie en particulier que si on a* $AB = I_n$*, alors l'égalité* $BA = I_n$ *est automatiquement vérifiée.*

Preuve. On montre les implications $1 \Rightarrow 2 \Rightarrow 4 \Rightarrow 1$. Les implications $1 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$ se démontrent de la même façon (ou par transposition).

L'implication $1 \Rightarrow 2$ est triviale. Montrons que $2 \Rightarrow 4$: Notons que les colonnes de I_n sont les vecteurs $\vec{\epsilon}_1, \ldots, \vec{\epsilon}_n$ de la base canonique de $\mathbb{R}^n$. L'existence d'une matrice B telle que $AB = I_n$ est donc équivalente à l'existence de n vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ (les colonnes de B) tels que pour tout $j = 1, \ldots, n$, $A\vec{e}_j = \vec{\epsilon}_j$. Cela montre que les vecteurs $\vec{\epsilon}_j$ appartiennent au sous-espace engendré par les colonnes de A (d'après la proposition 5.2.5). Comme $(\vec{\epsilon}_1, \ldots, \vec{\epsilon}_n)$ forment une base de $\mathbb{K}^n$ (c'est la base canonique), les colonnes de A engendrent $\mathbb{R}^n$ tout entier et on peut en déduire qu'elles sont de rang n.

Montrons maintenant que $4 \Rightarrow 1$: puisque A est de rang n, l'espace vectoriel engendré par ses colonnes est égal à $\mathbb{K}^n$, donc pour tout $j = 1, \ldots, n$, il existe des vecteurs $\vec{e}_j$ tels que $A\vec{e}_j = \vec{\epsilon}_j$, soit une matrice B telle que $AB = I_n$. Grâce au théorème 5.4.1 on sait que l'espace vectoriel engendré par les lignes de A est aussi égal à $\mathbb{K}^n$. Donc pour tout $i = 1, \ldots, n$, il existe des vecteurs ligne $\vec{v}_i$ tels que $\vec{v}_i A = \vec{\epsilon}_i$ (où $\vec{\epsilon}_i$ est maintenant un vecteur ligne). Si C est la matrice dont les lignes sont $\vec{v}_1, \ldots, \vec{v}_n$, on a donc $CA = I_n$. On a vu dans la démonstration de la proposition 5.5.2 que si $AB = CA = I_n$ alors $B = C$, donc la matrice A est bien inversible.

□

Proposition 5.5.6 (Exemples d'inverses de matrices, à connaitre)

- *Inverse d'une matrice diagonale :*

$$A = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

Le rang de A est le nombre de ses coefficients diagonaux non nuls (en effet, elle est échelonnée..). Donc A est inversible si et seulement si $d_i \neq 0$ pour tout i, et il est facile de vérifier que

$$A^{-1} = \begin{pmatrix} d_1^{-1} & & \\ & \ddots & \\ & & d_n^{-1} \end{pmatrix}$$

- *Inverse d'une matrice $(2, 2)$:*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

A est inversible si et seulement si $ad - bc \neq 0$ (car c'est la condition pour que ses colonnes ne soient pas proportionnelles, c'est-à-dire pour que A soit de rang 2), et il est facile de vérifier que dans ce cas,

$$A^{-1} = \frac{1}{(ad - bc)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Avant de passer à l'algorithme de calcul de l'inverse d'une matrice, nous démontrons quelques propriétés de l'inverse d'une matrice :

Proposition 5.5.7

On note $\mathcal{G}\ell_n(\mathbb{K})$ l'ensemble des matrices carrées (n, n) inversibles. On a les propriétés, si $A \in \mathcal{G}\ell_n(\mathbb{K})$, $B \in \mathcal{G}\ell_n(\mathbb{K})$:

1. $I_n^{-1} = I_n$

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

3. $(AB)^{-1} = B^{-1}A^{-1}$

Preuve. la propriété 1. résulte de l'identité $I_nI_n = I_n$. La propriété 2. résulte de l'identité $AA^{-1} = I_n$. En effet, celle-ci dit que A^{-1} est l'inverse de A, mais elle dit aussi que A est l'inverse de A^{-1} ! Vérifions la troisième propriété : On a

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

Donc AB est inversible d'inverse $B^{-1}A^{-1}$ (et non pas $A^{-1}B^{-1}$!). $\square$

Calcul effectif de l'inverse d'une matrice par l'algorithme du pivot de Gauss : Il suffit d'effectuer sur la concaténation $(A \mid I_n)$ de A et I_n des opérations élémentaires de lignes jusqu'à ce que A soit transformée en I_n. Pour cela on procède comme dans l'algorithme du pivot de Gauss. Alors I_n est transformée en A^{-1} ! Effectuons ce calcul sur un exemple avant de justifier cela.

Exemple numérique : Calculons par cette méthode l'inverse de la matrice

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

$$\begin{array}{cccccccl}
1 & 1 & 1 & 1 & 0 & 0 & \\
1 & -1 & 0 & 0 & 1 & 0 & \\
0 & 1 & -1 & 0 & 0 & 1 & \\
\\
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & -2 & -1 & -1 & 1 & 0 & L_2 \leftarrow L_2 - L_1 \\
0 & 1 & -1 & 0 & 0 & 1 & \\
\\
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & 1 & -1 & 0 & 0 & 1 & \\
0 & -2 & -1 & -1 & 1 & 0 & L_2 \leftrightarrow L_3 \\
\\
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & 1 & -1 & 0 & 0 & 1 & \\
0 & 0 & -3 & -1 & 1 & 2 & L_3 \leftarrow L_3 + 2L_2 \\
\\
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & 1 & -1 & 0 & 0 & 1 & \\
0 & 0 & 1 & 1/3 & -1/3 & -2/3 & L_3 \leftarrow -1/3 * L_3 \\
\\
1 & 1 & 0 & 2/3 & 1/3 & 2/3 & L_1 \leftarrow L_1 - L_3 \\
0 & 1 & 0 & 1/3 & -1/3 & 1/3 & L_2 \leftarrow L_2 + L_3 \\
0 & 0 & 1 & 1/3 & -1/3 & -2/3 &
\end{array}$$

$$\begin{array}{ccc|ccc|l}
1 & 0 & 0 & 1/3 & 2/3 & 1/3 & L_1 \leftarrow L_1 - L_2 \\
0 & 1 & 0 & 1/3 & -1/3 & 1/3 & \\
0 & 0 & 1 & 1/3 & -1/3 & -2/3 &
\end{array}$$

De ce calcul on peut conclure (et vérifier à posteriori) que l'inverse de A est

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -2 \end{pmatrix}$$

Justification de cet algorithme : Expliquons maintenant pourquoi on obtient bien l'inverse de A en sortie. Pour cela, il faut comprendre qu'effectuer une opération élémentaire sur les lignes d'une matrice revient à multiplier cette matrice à gauche par une certaine matrice inversible. Cela découle de la proposition 5.2.5 (voir aussi l'exemple numérique qui suit cette proposition). À titre d'exemples, nous allons interpréter ainsi les quatre premières opérations effectuées dans l'exemple précédent :

Opération sur les lignes	**Multiplication à gauche par**
$L_2 \leftarrow L_2 - L_1$	$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$L_2 \leftrightarrow L_3$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
$L_3 \leftarrow L_3 + 2L_2$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$
$L_3 \leftarrow -1/3 * L_3$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/3 \end{pmatrix}$

Effectuer une succession d'opérations sur les lignes d'une matrice revient alors à multiplier cette matrice à gauche par le produit des matrices correspondant à chacune des opérations élémentaires, produit égal à une certaine matrice P. Revenons à l'algorithme de calcul de l'inverse de A. Puisqu'on commence avec la concaténation $(A \mid I_n)$ et qu'on aboutit à $(I_n|B)$ pour une certaine matrice B, c'est que $PA = I_n$ et que $PI_n = P = B$. On a donc bien $P = A^{-1}$ et A^{-1} se retrouve en fin d'algorithme à la place de I_n.

Remarquons que pour effectuer cet algorithme, il n'est pas nécessaire de montrer au préalable que A est inversible. En effet, si A ne l'est pas, l'algorithme conduit à une matrice dont la moitié gauche est échelonnée mais a certaines de ses lignes nulles, ce qui montre que A est de rang strictement plus petit que n et donc qu'elle n'est pas inversible. Au passage, on récupère le rang de A.

5.6 Noyau et image d'une matrice

Dans cette section, A est à nouveau une matrice quelconque, de taille (ℓ, n). On introduit le noyau et l'image de A, en préparation du chapitre suivant sur les applications linéaires.

Définition 5.6.1
Le noyau de A, *noté* $\mathrm{Ker}(A)$, *est l'ensemble des vecteurs colonne* $X \in \mathbb{K}^n$ *tels que* $AX = \mathbf{0}$ *:*

$$\mathrm{Ker}(A) = \{X \in \mathbb{K}^n \mid AX = \mathbf{0}\}.$$

*L'*image de A, *noté* $\mathrm{Im}(A)$ *est le sous-espace vectoriel de* $\mathbb{K}^\ell$ *engendré par les colonnes* $C_1, \ldots, C_n$ *de* A *:*

$$\mathrm{Im}(A) = \mathrm{Vect}(C_1, \ldots, C_n)$$

Remarque 5.6.2
Le noyau $\mathrm{Ker}(A)$ *est donc tout simplement l'ensemble des solutions d'un système linéaire homogène (voir la remarque 5.2.6). On remarque également que l'algorithme du pivot de Gauss sur les colonnes (théorème 5.3.3) nous permet de calculer une base de* $\mathrm{Im}(A)$.

Une variante de l'algorithme vu à la section précédente pour le calcul de l'inverse d'une matrice se généralise en un algorithme qui calcule *simultanément* le noyau et l'image d'une matrice. On concatène cette fois les matrices A et I_n *en colonnes* sous la forme :

$$\left(\begin{array}{c} A \\ \hline I_n \end{array}\right)$$

puis on effectue des opérations élémentaires *sur les colonnes* jusqu'à ce que les colonnes de A soient échelonnées.

À l'issue du calcul, A est transformée en une matrice colonne échelonnée. Notons r le nombre de ses colonnes non nulles. Cette matrice colonne échelonnée issue de A, a donc r colonnes non nulles suivies de $n-r$ colonnes nulles, et $r = \mathrm{rang}(A)$. Notons B la partie non nulle. On distingue quatre parties dans cette matrice, de la façon suivante :

$$\left(\begin{array}{c|c} B & \mathbf{0} \\ \hline C & D \end{array}\right)$$

où C est de taille (n, r) et D de taille $(n, n-r)$. *Alors, les colonnes de* B *forment une base de* $\mathrm{Im}(A)$ *et les colonnes de* D *forment une base de* $\mathrm{Ker}(A)$.

En effet, on sait que les colonnes de B forment une base de $\mathrm{Im}(A)$ d'après le théorème 5.3.3. Notons P la matrice inversible (n, n) résultant de la succession d'opérations élémentaires effectuées sur les colonnes. On a donc

$$\left(\begin{array}{c} AP \\ \hline P \end{array}\right) = \left(\begin{array}{c|c} B & \mathbf{0} \\ \hline C & D \end{array}\right)$$

On voit facilement que $X \in \mathrm{Ker}(AP)$ si et seulement si $PX \in \mathrm{Ker}(A)$. Parce que AP est échelonnée, il est clair que $\mathrm{Ker}(AP)$ est exactement l'ensemble des vecteurs colonne X dont les r premiers coefficients valent 0. Alors, PX parcourt exactement l'ensemble des combinaisons linéaires des colonnes de D. Notons que P est inversible donc ses colonnes sont linéairement indépendantes.

En particulier, on a démontré que $\dim(\mathrm{Ker}(A)) = n - r$ et ce résultat important porte le nom de *théorème du rang* :

Théorème 5.6.3 (Théorème du rang pour les matrices)
Soit A une matrice de taille (ℓ, n). Alors

$$\dim(\mathrm{Ker}(A)) + \mathrm{rang}(A) = n$$

Remarque 5.6.4
Le théorème du rang a des interprétations intéressantes pour un système linéaire homogène de matrice A, dont l'ensemble des solutions n'est autre que $\mathrm{Ker}(A)$.

1. *Si les équations d'un système linéaire homogène à n inconnues et k équations sont indépendantes, alors* $\mathrm{rang}(A) = k$ *et donc la dimension de l'ensemble de ses solutions est $n - k$.*
2. *Dans le même ordre d'idée, nous avions vu au chapitre* 1 *que l'ensemble des solutions d'un système linéaire échelonné est paramétré par les variables auxiliaires de ce système. Le nombre de ces variables auxiliaires est la dimension de l'ensemble des solutions du système linéaire homogène associé. En effet,* $\dim(\mathrm{Ker}(A) = n - \mathrm{rang}(A)$ *et* $\mathrm{rang}(A)$ *est le nombre de pivots (voir la remarque 1.4.2).*

Une des conséquences intéressantes de ce théorème est la caractérisation des matrices inversibles par leur noyau :

Corollaire 5.6.5
Soit A une matrice carrée de taille n. Alors A est inversible si et seulement si $\mathrm{Ker}(A) = \{\mathbf{0}\}$.

Preuve. En effet, on a vu que A est inversible si et seulement si $\mathrm{rang}(A) = n$. D'après le théorème du rang, $\mathrm{rang}(A) = n$ si et seulement si $\dim(\mathrm{Ker}(A)) = 0$ soit $\mathrm{Ker}(A) = \{\vec{0}\}$.
□

Chapitre 6

Applications linéaires

Les applications linéaires sont les applications entre espaces vectoriels qui sont compatibles avec les opérations d'espaces vectoriel. Nous allons voir qu'en dimension finie, les matrices sont un outil très utile pour manipuler les applications linéaires.

6.1 Définition et premières propriétés

Définition 6.1.1
Soit $(E, +, \cdot)$ *et* $(F, +, \cdot)$ *deux* $\mathbb{K}$*-espaces vectoriels. Une* application linéaire *de* E *dans* F *(on dit aussi un* homomorphisme *ou un* morphisme*) est une application* $f : E \to F$ *vérifiant :*

$$\textit{Pour tout } (\vec{u}, \vec{v}) \in E^2,\ (\lambda, \mu) \in \mathbb{K}^2,\ f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}).$$

Si de plus f *est un bijection, on dit que* f *est un* isomorphisme d'espaces vectoriels.

Remarque 6.1.2
Les espaces de départ et d'arrivée, que nous noterons usuellement E *et* F*, peuvent être égaux ou pas. S'ils sont différents, leurs opérations* $+$ *et* $\cdot$*, leurs vecteurs nuls ne sont pas les mêmes ; toutefois, pour ne pas alourdir les notations, nous les noterons avec les mêmes symboles. S'il y a une ambiguité d'interprétation nous rajouterons l'espace en indice : par exemple nous noterons* $\vec{0}_E$ *le vecteur nul de* E *et* $\vec{0}_F$ *le vecteur nul de* F*.*

Remarque 6.1.3
Une application linéaire envoie toujours le vecteur nul de E *sur le vecteur nul de* F*. En effet,* $f(\vec{0}_E) = f(0 \cdot \vec{0}_E) = 0 \cdot f(\vec{0}_E) = \vec{0}_F$ *donc* $f(\vec{0}_E) = \vec{0}_F$*.*

Remarquons que dans le calcul précédent, on distingue les vecteurs nuls de E *et* F *comme expliqué à la remarque précédente, mais que cette distinction est en fait superflue : en effet, si on écrit simplement* $f(\vec{0}) = \vec{0}$*, comme* f *va de* E *dans* F*, le premier* $\vec{0}$ *est forcément celui de* E *et le deuxième celui de* F*. Désormais nous utiliserons la notation commune* $\vec{0}$ *pour les vecteurs nuls de* E *et* F*, sauf en cas d'ambiguité.*

On commence par décrire une famille importante d'applications linéaires définies à partir de matrices. Nous verrons au paragraphe 4 que toutes les applications linéaires entre espaces vectoriels de dimension fini peuvent s'exprimer ainsi.

Proposition 6.1.4
Soit $E = \mathbb{K}^n$ et $F = \mathbb{K}^\ell$, et soit $A \in \mathcal{M}_{\ell,n}(\mathbb{K})$. On note ici les vecteurs de $\mathbb{K}^n$ et de $\mathbb{K}^\ell$ en colonne. Alors l'application suivante, notée f_A, est une application linéaire :

$$\begin{array}{rccc} f_A : & \mathbb{K}^n & \to & \mathbb{K}^\ell \\ & X & \mapsto & AX \end{array}$$

Preuve. Montrons que f_A est linéaire : en effet, $f_A(\lambda X + \mu X') = A(\lambda X + \mu X') = \lambda AX + \mu AX' = \lambda f_A(X) + \mu f_A(X')$. □

Remarque 6.1.5
La raison pour laquelle on note ici les vecteurs de $\mathbb{K}^n$ en colonne, est que l'on va retrouver les applications $X \to f_A(X)$ au paragraphe 4, où X sera le vecteur des coordonnées dans une base., donc un vecteur colonne suivant notre convention.

6.2 Noyau et image d'une application linéaire

Dans ce paragraphe nous allons introduire et étudier deux sous-espaces vectoriels importants associés à une application linéaire : son noyau et son image.

Définition 6.2.1
Soit $f : E \to F$ une application linéaire. On définit :

$$\mathrm{Ker}(f) = \{\vec{u} \in E \mid f(\vec{u}) = \vec{0}\}$$

et

$$\mathrm{Im}(f) = \{f(\vec{u}) \mid \vec{u} \in E\}.$$

Exemple 6.2.2
Si $f = f_A$ est définie comme dans la proposition 6.1.4, on retrouve les notions de noyau et d'image d'une matrice introduite dans la définition 5.6.1 du chapitre 5. En effet, il est clair que $\mathrm{Ker}(f_A) = \mathrm{Ker}(A)$ et que $\mathrm{Im}(f_A) = \mathrm{Im}(A)$.

Théorème 6.2.3
On a les propriétés suivantes :

1. $\mathrm{Ker}(f)$ *est un sous-espace vectoriel de* E.
2. $\mathrm{Im}(f)$ *est un sous-espace vectoriel de* F.
3. f *est injective si et seulement si* $\mathrm{Ker}(f) = \{\vec{0}\}$.
4. f *est surjective si et seulement si* $\mathrm{Im}(f) = F$.

Preuve. Le vecteur nul appartient à $\mathrm{Ker}(f)$ car $f(\vec{0}) = \vec{0}$. Si $\vec{u}$ et $\vec{v}$ appartiennent à $\mathrm{Ker}(f)$ et $(\lambda, \mu) \in \mathbb{K}^2$, alors $f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}) = \lambda\vec{0} + \mu\vec{0} = \vec{0}$ (noter qu'on utilise ici la linéarité de f) donc $\lambda\vec{u} + \mu\vec{v} \in \mathrm{Ker}(f)$. On a démontré que $\mathrm{Ker}(f)$ est un sous-espace vectoriel de E.

Le vecteur nul appartient à $\mathrm{Im}(f)$ car $\vec{0} = f(\vec{0})$. Si $\vec{x}$ et $\vec{y}$ appartiennent à $\mathrm{Im}(f)$ alors il existe $(\vec{u}, \vec{v}) \in E^2$ tels que $\vec{x} = f(\vec{u})$ et $\vec{y} = f(\vec{v})$. Alors $\lambda\vec{x} + \mu\vec{y} = \lambda f(\vec{u}) + \mu f(\vec{v}) = f(\lambda\vec{u} + \mu\vec{v}) \in \mathrm{Im}(f)$. (à nouveau, la linéarité de f est appliquée) Donc $\mathrm{Im}(f)$ est un sous-espace vectoriel de F.

$\mathrm{Ker}(f)$ est l'image réciproque de $\{\vec{0}\}$ et contient $\vec{0}$, donc la condition $\mathrm{Ker}(f) = \{\vec{0}\}$ est clairement nécessaire pour que f soit injective. Réciproquement, supposons $\mathrm{Ker}(f) = \{\vec{0}\}$ et montrons que f est injective. Si $(\vec{u}, \vec{v}) \in E^2$ sont tels que $f(\vec{u}) = f(\vec{v})$, alors $f(\vec{u}) - f(\vec{v}) = \vec{0}$ donc, par linéarité de f, $f(\vec{u} - \vec{v}) = \vec{0}$ c'est-à-dire $\vec{u} - \vec{v} \in \mathrm{Ker}(f)$. Par hypothèse $\mathrm{Ker}(f) = \{\vec{0}\}$, on peut donc conclure que $\vec{u} - \vec{v} = \vec{0}$ soit $\vec{u} = \vec{v}$.

Le point 4. n'est rien d'autre que la définition d'une application surjective. □

On termine ce paragraphe par un théorème important sur les applications linéaires qui met en relation les dimensions du noyau et de l'image, il s'agit du *théorème du rang*.

Définition 6.2.4
Avec les notations précédente, on note $\mathrm{rang}(f)$ *et on appelle rang de f la dimension de* $\mathrm{Im}(f)$.

Théorème 6.2.5 (Théorème du rang)
Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie, et soit $f : E \to F$ une application linéaire. Alors,

$$\dim(\mathrm{Ker}(f)) + \mathrm{rang}(f) = \dim(E).$$

Preuve. Soit $n = \dim(E)$ et soit $k = \dim(\mathrm{Ker}(f))$. Soit $(\vec{e}_1, \ldots, \vec{e}_k)$ une base de $\mathrm{Ker}(f)$. On complète cette base en une base $(\vec{e}_1, \ldots, \vec{e}_n)$ de E. Alors,$(f(\vec{e}_{k+1}), \ldots, f(\vec{e}_n))$ forme une base de $\mathrm{Im}(f)$. En effet, c'est une famille génératrice de $\mathrm{Im}(f)$ car $f(\vec{e}_1) = \cdots = f(\vec{e}_k) = \vec{0}$. Montrons qu'elle est libre : supposons qu'il existe $\lambda_{k+1}, \ldots, \lambda_n$ tels que $\lambda_{k+1} f(\vec{e}_{k+1}) + \cdots + \lambda_n f(\vec{e}_n) = \vec{0}$. Alors, par la propriété de linéarité de f on obtient $f(\lambda_{k+1}\vec{e}_{k+1} + \cdots + \lambda_n\vec{e}_n) = \vec{0}$, et on peut conclure que $\lambda_{k+1}\vec{e}_{k+1} + \cdots + \lambda_n\vec{e}_n \in \mathrm{Ker}(f)$. Comme $(\vec{e}_1, \ldots, \vec{e}_k)$ est une base de $\mathrm{Ker}(f)$, il existe donc $\lambda_1, \ldots, \lambda_k$ tels que $\lambda_{k+1}\vec{e}_{k+1} + \cdots + \lambda_n\vec{e}_n = \lambda_1\vec{e}_1 + \cdots + \lambda_k\vec{e}_k$; mais alors, comme la famille $(\vec{e}_1, \ldots, \vec{e}_n)$ est libre, on obtient que tous les λ_i, $i = 1, \ldots, n$ sont nuls, donc que $(f(\vec{e}_{k+1}), \ldots, f(\vec{e}_n))$ est libre.

Puisque $(f(\vec{e}_{k+1}), \ldots, f(\vec{e}_n))$ forme une base de $\mathrm{Im}(f)$, la dimension de cet espace est égal au cardinal de cette base, soit $\dim(\mathrm{Im}(f)) = n - k = \dim(E) - \dim(\mathrm{Ker}(f))$. □

Remarque 6.2.6
Dans le cas $f = f_A$, on a $\mathrm{rang}(f) = \mathrm{rang}(A)$ *et on retrouve le théorème du rang pour les matrices démontré au chapitre précédent (théorème 5.6.3).*

6.3 Applications et propriétés

On démontre d'abord un résultat très utile dans la pratique, qui est une conséquence du théorème du rang. Rappelons qu'un isomorphisme est une application linéaire qui est bijective.

Proposition 6.3.1 (Corollaire du théorème du rang)
Soit $f : E \to F$ une application linéaire.

1. *Si f est un isomorphisme, alors* $\dim(E) = \dim(F)$.
2. *Si* $\dim(E) = \dim(F)$, *alors :*

$$f \text{ est un isomorphisme } \iff f \text{ est injective } \iff f \text{ est surjective.}$$

Preuve. Démontrons 1. Si f est un isomorphisme, on a $\mathrm{Ker}(f) = \{\vec{0}\}$ et $\mathrm{Im}(f) = F$ d'après le théorème 6.2.3. Donc $\dim(\mathrm{Ker}(f)) = 0$ et $\mathrm{rang}(f) = \dim(F)$. Par le théorème du rang, on a

$$\dim(\mathrm{Ker}(f)) + \mathrm{rang}(f) = \dim(E)$$

qui devient donc ici $\dim(F) = \dim(E)$.

Démontrons 2. et posons $n = \dim(E) = \dim(F)$. Le théorème du rang dit que $n = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f))$. On voit donc que $\dim(\mathrm{Ker}(f)) = 0$ si et seulement si $\dim(\mathrm{Im}(f)) = n$, c'est-à-dire que $\mathrm{Ker}(f) = \{\vec{0}\}$ si et seulement si $\mathrm{Im}(f) = F$. Donc on a bien l'équivalence entre f injective et f surjective. Comme f est un isomorphisme si et seulement si elle est à la fois injective et surjective, on obtient le résultat annoncé. □

Un isomorphisme $f : E \to F$ est un outil qui permet de "transporter" toute question (concernant la structure d'espace vectoriel) dans E en une question dans F. Par exemple, une famille de E est libre (respectivement est génératrice, est une base) si et seulement si la famille des images par f de ses vecteurs a la même propriété dans F ; l'image par f d'un sous-espace vectoriel de E est un sous-espace vectoriel de F de même dimension, etc.. On laisse la vérification de ces affirmations au lecteur.

En particulier, si E est une espace vectoriel de dimension finie n, et si on fixe une base de E, on obtient naturellement un isomorphisme de E sur $\mathbb{K}^n$ en associant à un vecteur ses coordonnées dans la base choisie. C'est le sujet du théorème suivant :

Théorème 6.3.2
Soit $(E, +, \cdot)$ un $\mathbb{K}$-espace vectoriel de dimension finie notée n, et soit $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ une base de E. Alors, l'application :

$$\begin{aligned} \phi \ : \ \mathbb{K}^n &\to E \\ (x_1, \ldots, x_n) &\mapsto x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \end{aligned}$$

est un isomorphisme d'espaces vectoriels. Sa bijection réciproque est l'application qui à un vecteur de E associe ses coordonnées dans la base $\mathcal{B}$ et est aussi un iso-

morphisme :

$$\psi \ : \ E \to \mathbb{K}^n$$
$$\vec{x} \mapsto (x_1, \ldots, x_n)$$

où $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *sont les coordonnées de* $\vec{x}$ *dans la base* $\mathcal{B}$ *i.e.* $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$.

Preuve. On vérifie facilement que ϕ est une application linéaire. On sait que ϕ est injective si et seulement si $\mathrm{Ker}(\phi) = \{\vec{0}\}$ et on voit facilement que la condition $\mathrm{Ker}(\phi) = \{\vec{0}\}$ est équivalente à la propriété que la famille $\mathcal{B}$ est libre. La surjectivité de ϕ est équivalente à la propriété de $\mathcal{B}$ d'être génératrice de E. Donc ϕ est bien un isomorphisme.

Il est clair que ψ est linéaire et que c'est l'application réciproque de ϕ. □

Remarque 6.3.3

Ce théorème est très important dans la pratique puisqu'il permet de ramener tous les problèmes d'algèbre linéaire dans un espace de dimension finie à un problème dans $\mathbb{K}^n$*, et que dans* $\mathbb{K}^n$ *il existe des algorithmes efficaces pour les résoudre (notamment basés sur le pivot de Gauss comme on l'a vu).*

On termine ce paragraphe par quelques propriétés générales sur les applications linéaires.

Proposition 6.3.4

On a les propriétés suivantes :

1. *Si* f *et* g *sont des applications linéaires de* E *dans* F *et si* $\lambda \in \mathbb{K}$, $\mu \in \mathbb{K}$, *alors* $\lambda f + \mu g$ *est aussi une application linéaire de* E *dans* F*. En particulier, l'ensemble* $\mathcal{L}(E, F)$ *des applications linéaires de* E *dans* F *est un* $\mathbb{K}$*-espace vectoriel.*
2. *Si* $f : E \to F$ *et* $g : F \to G$ *sont des applications linéaires alors* $g \circ f : E \to G$ *est encore une application linéaire.*
3. *Si* $f : E \to F$ *est un isomorphisme d'espaces vectoriels, alors* $f^{-1} : F \to E$ *est aussi un isomorphisme d'espaces vectoriels.*

Preuve. Il suffit de l'écrire. Démontrons en détails le point 3 : soit $(\vec{x}, \vec{y}) \in F^2$ et $(\lambda, \mu) \in \mathbb{K}^2$, on doit montrer que $f^{-1}(\lambda\vec{x} + \mu\vec{y}) = \lambda f^{-1}(\vec{x}) + \mu f^{-1}(\vec{y})$. Par la surjectivité de f, il existe $(\vec{u}, \vec{v}) \in E^2$ tels que $\vec{x} = f(\vec{u})$ et $\vec{y} = f(\vec{v})$. Alors

$$f^{-1}(\lambda\vec{x} + \mu\vec{y}) = f^{-1}(\lambda f(\vec{u}) + \mu f(\vec{v})) = f^{-1}(f(\lambda\vec{u} + \mu v)) = \lambda\vec{u} + \mu\vec{v}.$$

□

Exercice
Si $f = f_A$ comme dans la proposition 6.1.4, montrez que :

1. $\lambda f_A + \mu f_B = f_{\lambda A + \mu B}$
2. $f_A \circ f_B = f_{AB}$
3. f_A *est un isomorphisme si et seulement si* A *est une matrice inversible, et dans ce cas,* $f_A^{-1} = f_{A^{-1}}$.

Exercice
Soit $f : E \to F$ une application linéaire, et soit $(\vec{e}_1, \ldots, \vec{e}_k)$ une famille de vecteurs de E. Démontrez les propriétés suivantes :

1. *Si $(\vec{e}_1, \ldots, \vec{e}_k)$ est libre et si f est injective, alors $(f(\vec{e}_1), \ldots, f(\vec{e}_k))$ est libre.*
2. *Si $(\vec{e}_1, \ldots, \vec{e}_k)$ est génératrice de E et si f est surjective, alors $(f(\vec{e}_1), \ldots, f(\vec{e}_k))$ est génératrice de F.*
3. *Si $(\vec{e}_1, \ldots, \vec{e}_k)$ est une base de E et si f est un isomorphisme, alors $(f(\vec{e}_1), \ldots, f(\vec{e}_k))$ est une base de F.*

6.4 Matrices d'une application linéaire

Dans ce paragraphe nous allons faire le lien entre application linéaire et matrices. On ne considère plus que des espaces vectoriels de dimension finie.

Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives n et ℓ. Soit $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ une base de E et soit $\mathcal{B}' = (\vec{f}_1, \ldots, \vec{f}_\ell)$ une base de F.

Si $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \in E$, alors par linéarité on a

$$f(\vec{x}) = x_1 f(\vec{e}_1) + \cdots + x_n f(\vec{e}_n). \tag{6.1}$$

En particulier, notons que cela signifie que f *est uniquement déterminée par les images $f(\vec{e}_1), \ldots, f(\vec{e}_n)$ de $\vec{e}_1, \ldots, \vec{e}_n$ dans F.*

Transformons maintenant l'équation (6.1) en une équation sur les coordonnées dans la base $\mathcal{B}'$. Prenons les notations suivantes : soit $\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{\ell,j} \end{pmatrix}$ les coordonnées de $f(\vec{e}_j)$ dans la base $\mathcal{B}'$, et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_\ell \end{pmatrix}$ les coordonnées de $f(\vec{x})$ dans la base $\mathcal{B}'$. Alors l'équation (6.1) devient :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_\ell \end{pmatrix} = x_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{\ell,1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{\ell,n} \end{pmatrix}$$

soit en termes matriciels :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_\ell \end{pmatrix} = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{\ell,1} & \ldots & a_{\ell,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{6.2}$$

La matrice qui apparait dans (6.2) s'appelle *la matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$*. Nous résumons dans l'énoncé suivant ce que nous venons de démontrer :

Définition et proposition 6.4.1
La matrice de f dans les bases $\mathcal{B}$, $\mathcal{B}'$, *notée $M_{\mathcal{B},\mathcal{B}'}(f)$, est la matrice de taille (ℓ, n) dont les colonnes sont les coordonnées dans la base $\mathcal{B}'$ des images par f des vecteurs de la base $\mathcal{B}$, soit avec les notations précédentes :*

$$M_{\mathcal{B},\mathcal{B}'}(f) = \begin{array}{c} \\ \vec{f}_1 \\ \vec{f}_2 \\ \vdots \\ \vec{f}_\ell \end{array} \begin{array}{c} \begin{array}{cccc} f(\vec{e}_1) & f(\vec{e}_2) & \dots & f(\vec{e}_n) \end{array} \\ \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{\ell,1} & a_{\ell,2} & \dots & a_{\ell,n} \end{pmatrix} \end{array}.$$

Alors, si on note X le vecteur des coordonnées de $\vec{x} \in E$ dans la base $\mathcal{B}$, et Y le vecteur des coordonnées de $\vec{y} = f(\vec{x})$ dans la base $\mathcal{B}'$, on a

$$Y = M_{\mathcal{B},\mathcal{B}'}(f)X.$$

Si $E = F$ et $\mathcal{B} = \mathcal{B}'$, on note plus simplement $M_{\mathcal{B}}(f)$ ou même $M(f)$ la matrice de f dans la base $\mathcal{B}$.

Remarque 6.4.2
La relation que l'on vient de voir :

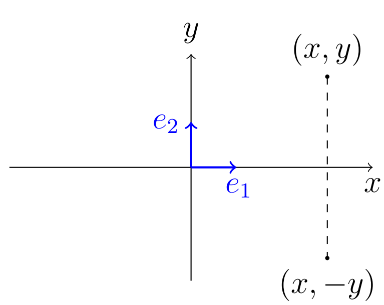
$$Y = M_{\mathcal{B},\mathcal{B}'}(f)X$$

montre que, si on voit f comme une application entre vecteurs de coordonnées dans les bases $\mathcal{B}$ et $\mathcal{B}'$, f s'identifie à l'application f_A pour $A = M_{\mathcal{B},\mathcal{B}'}(f)$.

Exemple numérique : Soit l'exemple très simple suivant : s est l'application linéaire

$$\begin{aligned} s : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (x, -y) \end{aligned}$$

C'est la symétrie par rapport à l'axe des abscisses comme on peut le voir dans le dessin suivant :



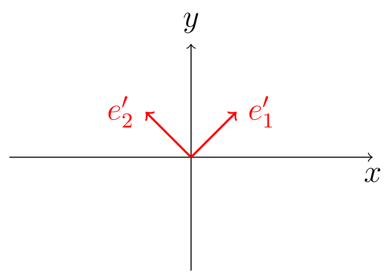
On a $s(e_1) = e_1$ et $s(e_2) = -e_2$ donc

$$M_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Notons que les coordonnées de (x, y) dans la base $\mathcal{B}$ sont $\begin{pmatrix} x \\ y \end{pmatrix}$ et qu'on a bien

$$\begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Il est important de comprendre que la matrice d'un application linéaire n'est pas unique car elle dépend de la base choisie. Par exemple, on peut choisir à la place de $\mathcal{B}$ la base $\mathcal{B}'$ suivante :



On a $s(e'_1) = -e'_2$ et $s(e'_2) = -e'_1$ donc la matrice de s dans $\mathcal{B}'$ est :

$$M_{\mathcal{B}'}(s) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

On voit bien ici que $M_{\mathcal{B}}(s)$ et $M_{\mathcal{B}'}(s)$ sont deux matrices différentes associées à la même application linéaire s, mais exprimées dans des bases différentes. Au prochain paragraphe nous discutons les relations qui existent entre les différentes matrices d'une même application linéaire.

On continue avec des exemples un peu plus généraux :

Exemple 6.4.3

Quelques exemples de matrices d'applications linéaires :

1. $M_{\mathcal{B},\mathcal{B}'}(0) = \mathbf{0}$, *où* 0 *représente l'application nulle* $\vec{x} \to \vec{0}$.
2. *Si* $E = F$ *et* $\mathcal{B} = \mathcal{B}'$, *alors* $M_{\mathcal{B}}(\mathrm{Id}) = I_n$.
3. *Avec les notations de la proposition 6.1.4, dans les bases canoniques de* $\mathbb{K}^n$ *et de* $\mathbb{K}^\ell$, $M(f_A) = A$.

Les opérations sur les applications linéaires se traduisent par des opérations matricielles, en particulier la multiplication des matrices correspond à la composition des applications. C'est le sujet de l'énoncé qui suit :

Théorème 6.4.4
Soit E, F, G trois $\mathbb{K}$-espaces vectoriels de dimensions finies. Soit $\mathcal{B}$, $\mathcal{B}'$, $\mathcal{B}''$ des bases de ces espaces.

1. *Si f et g sont des applications linéaires de E dans F et si $\lambda \in \mathbb{K}$, $\mu \in \mathbb{K}$, alors*
$$M_{\mathcal{B},\mathcal{B}'}(\lambda f + \mu g) = \lambda M_{\mathcal{B},\mathcal{B}'}(f) + \mu M_{\mathcal{B},\mathcal{B}'}(g).$$
2. *Si $f : E \to F$ et $g : F \to G$ sont des applications linéaires, alors*
$$M_{\mathcal{B},\mathcal{B}''}(g \circ f) = M_{\mathcal{B}',\mathcal{B}''}(g) M_{\mathcal{B},\mathcal{B}'}(f)$$
3. *Soit $f : E \to F$ une application linéaire. C'est un isomorphisme si et seulement si la matrice $M_{\mathcal{B},\mathcal{B}'}(f)$ est inversible, et dans ce cas $f^{-1} : F \to E$ est aussi un isomorphisme et*
$$M_{\mathcal{B}',\mathcal{B}}(f^{-1}) = \big(M_{\mathcal{B},\mathcal{B}'}(f)\big)^{-1}.$$

Preuve. Démontrons le point 2. Soit $\vec{x} \in E$, $\vec{y} = f(\vec{x}) \in F$ et $\vec{z} = g(\vec{y}) = (g \circ f)(\vec{x})$. Soit X, Y, Z, les vecteurs colonne des coordonnées respectives de $\vec{x}$, $\vec{y}$, $\vec{z}$ dans les bases $\mathcal{B}$, $\mathcal{B}'$, $\mathcal{B}''$. Alors :
$$Z = M_{\mathcal{B}',\mathcal{B}''}(g)Y = M_{\mathcal{B}',\mathcal{B}''}(g)\big(M_{\mathcal{B},\mathcal{B}'}(f)X\big) = \big(M_{\mathcal{B}',\mathcal{B}''}(g)M_{\mathcal{B},\mathcal{B}'}(f)\big)X$$
donc
$$M_{\mathcal{B},\mathcal{B}''}(g \circ f) = M_{\mathcal{B}',\mathcal{B}''}(g) M_{\mathcal{B},\mathcal{B}'}(f).$$

Le point 3. est essentiellement une conséquence de 2. : en effet, supposons d'abord que f est un isomorphisme. Par la proposition 6.3.1, $\dim(E) = \dim(F)$ et donc les matrices des applications entre E et F sont carrées. On sait qu'il existe une application $f^{-1} : F \to E$ telle que $f \circ f^{-1} = \mathrm{Id}$. Par 2., on en déduit que
$$M_{\mathcal{B},\mathcal{B}'}(f) M_{\mathcal{B}',\mathcal{B}}(f^{-1}) = M_{\mathcal{B}',\mathcal{B}'}(\mathrm{Id}) = I_n$$
et donc que $M_{\mathcal{B},\mathcal{B}'}(f)$ est inversible.

Réciproquement, si la matrice de f est inversible, elle est en particulier carrée et donc cela signifie que $\dim(E) = \dim(F)$. Alors f est bijective si et seulement si elle est injective (Proposition 6.3.1), c'est-à-dire si et seulement si son noyau est réduit au vecteur nul. Par la relation $Y = M_{\mathcal{B},\mathcal{B}'}(f)X$, on voit que cela est vrai car le noyau de $M_{\mathcal{B},\mathcal{B}'}(f)$ est nul (car c'est une matrice inversible).

$\square$

6.5 Changement de base

Dans cette partie, E est un espace vectoriel de dimension finie n sur $\mathbb{K}$. Un vecteur donné de E a plusieurs systèmes de coordonnées, suivant la base choisie. De même, une application linéaire a plusieurs matrices qui la représentent, en fonction des bases choisies. Nous allons dans ce paragraphe expliciter les relations entre ces différentes expressions.

Soit donc $\mathcal{B} = (\vec{e}_1, \ldots, \vec{e}_n)$ et $\mathcal{B}' = (\vec{e}\,'_1, \ldots, \vec{e}\,'_n)$ deux bases de E.

Définition et proposition 6.5.1
La matrice de changement de base de $\mathcal{B}$ à $\mathcal{B}'$, *encore appelée la* matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ *est la matrice carrée de taille n, notée $P_{\mathcal{B},\mathcal{B}'}$ ou plus simplement P, dont les colonnes sont les coordonnées des vecteurs de $\mathcal{B}'$ (la nouvelle base) dans $\mathcal{B}$ (l'ancienne base) :*

$$P = \begin{array}{c} \\ \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{array} \begin{array}{c} \begin{array}{cccc} \vec{e}\,'_1 & \vec{e}\,'_2 & \ldots & \vec{e}\,'_n \end{array} \\ \begin{pmatrix} \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \end{pmatrix} \end{array}.$$

Soit $\vec{x} \in E$, soit X et X' les coordonnées de $\vec{x}$ respectivement dans les bases $\mathcal{B}$ et $\mathcal{B}'$ (ce sont donc des vecteurs colonne de taille n). Alors on a

$$X = PX'.$$

Preuve. Soit $\vec{x} \in E$ et soit $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ les coordonnées de $\vec{x}$ dans la base $\mathcal{B}'$. On a donc :

$$\vec{x} = x'_1\vec{e}\,'_1 + \cdots + x'_n\vec{e}\,'_n.$$

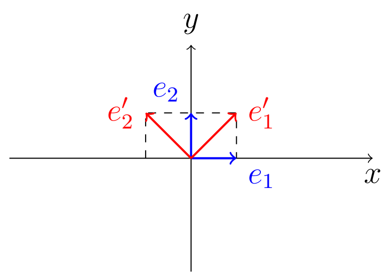
On passe maintenant aux coordonnées dans la base $\mathcal{B}$: cette identité de vecteurs se traduit par l'identité des vecteurs de coordonnées dans la base $\mathcal{B}$. Soit donc $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ les coordonnées de $\vec{x}$ dans la base $\mathcal{B}$, et notons, pour tout $j = 1, \ldots, n$, $\begin{pmatrix} p_{1,j} \\ \vdots \\ p_{n,j} \end{pmatrix}$ les coordonnées de $\vec{e}\,'_j$ dans la base $\mathcal{B}$. D'après la définition de la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$, ce sont les colonnes de P. On a donc :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x'_1 \begin{pmatrix} p_{1,1} \\ \vdots \\ p_{n,1} \end{pmatrix} + \cdots + x'_n \begin{pmatrix} p_{1,n} \\ \vdots \\ p_{n,n} \end{pmatrix}$$
$$= \begin{pmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{pmatrix} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = PX'.$$

□

Remarque 6.5.2
Il est clair que la matrice P de changement de base de $\mathcal{B}$ à $\mathcal{B}'$ est inversible (car elle est de rang n). Son inverse P^{-1} est la matrice de changement de base de $\mathcal{B}'$ à $\mathcal{B}$: en effet on a $X = PX'$ donc $X' = P^{-1}X$.

Exemple numérique : Dans l'exemple des bases $\mathcal{B} = (e_1, e_2)$ et $\mathcal{B}' = (e'_1, e'_2)$ de $\mathbb{R}^2$:



on a $e'_1 = e_1 + e_2$ et $e'_2 = -e_1 + e_2$ donc la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Soit maintenant $f : E \to E$ une application linéaire, on cherche une relation entre les matrices $M_{\mathcal{B}}(f)$ et $M_{\mathcal{B}'}(f)$.

Proposition 6.5.3
Soit P la matrice de changement de base de $\mathcal{B}$ à $\mathcal{B}'$. On a

$$M_{\mathcal{B}'}(f) = P^{-1} M_{\mathcal{B}}(f) P.$$

Preuve. Posons, pour simplifier les notations, $M = M_{\mathcal{B}}(f)$ et $M' = M_{\mathcal{B}'}(f)$. Soit $\vec{x} \in E$, et soit $\vec{y} = f(\vec{x})$. Soit X et X' les coordonnées de $\vec{x}$ respectivement dans les bases $\mathcal{B}$ et $\mathcal{B}'$, et soit Y et Y' les coordonnées de $\vec{y}$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$. On a :

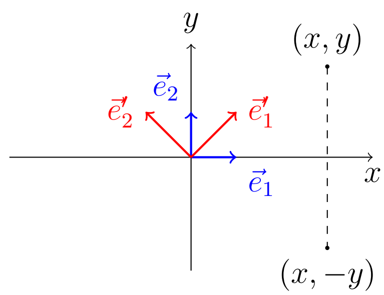
$$X = PX', \qquad Y = PY', \qquad Y = MX, \qquad Y' = M'X'.$$

On en déduit que $PY' = MPX'$, soit que $Y' = P^{-1}MPX'$. On a donc $M'X' = (P^{-1}MP)X'$ pour tout $X' \in \mathbb{K}^n$ ce qui implique

$$M' = P^{-1}MP.$$

□

Exemple numérique : Reprenons l'exemple de la symétrie s et des bases $\mathcal{B}$, $\mathcal{B}'$:



Nous avons déjà vu que

$$M_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad M_{\mathcal{B}'}(s) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

La matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ et son inverse sont :

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

La relation $M_{\mathcal{B}'}(s) = P^{-1}M_{\mathcal{B}}(s)P$ est bien vérifiée :

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

Remarque 6.5.4

On peut donner une formule plus générale pour le changement de bases qui est la suivante : Soit $f : E \to F$, soit $\mathcal{B}_E$ et $\mathcal{B}'_E$ deux bases de E, soit $\mathcal{B}_F$ et $\mathcal{B}'_F$ deux bases de F. Soit P la matrice de passage de $\mathcal{B}_E$ à $\mathcal{B}'_E$, et soit Q la matrice de passage de $\mathcal{B}_F$ à $\mathcal{B}'_F$. Alors

$$M_{\mathcal{B}'_E,\mathcal{B}'_F}(f) = Q^{-1}M_{\mathcal{B}_E,\mathcal{B}_F}(f)P.$$

Nous venons de traiter le cas où $E = F$, $\mathcal{B}_E = \mathcal{B}'_E = \mathcal{B}$, $\mathcal{B}_F = \mathcal{B}'_F = \mathcal{B}'$. La démonstration du cas général est très semblable à celle du cas particulier que nous venons de voir et est laissée au lecteur.

Chapitre 7

Compléments

7.1 Somme directe de sous-espaces vectoriels

Soit $(E, +, \cdot)$ un espace vectoriel et soit U et V deux sous-espaces vectoriels de E. On rappelle la notion de somme $U + V$: c'est le sous-espace vectoriel de E défini par :

$$U + V = \{\vec{u} + \vec{v} \mid \vec{u} \in U, \vec{v} \in V\}.$$

On a vu aussi que l'intersection $U \cap V$ est un sous-espace vectoriel de E et on a démontré dans le cas de la dimension finie la formule (Chapitre 4, Théorème 4.3.12) :

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V).$$

Définition 7.1.1
On dit que E est la somme directe de U et V, *et on note $E = U \oplus V$, si les conditions suivantes sont réalisées :*

1. $E = U + V$

2. $U \cap V = \{\vec{0}\}$

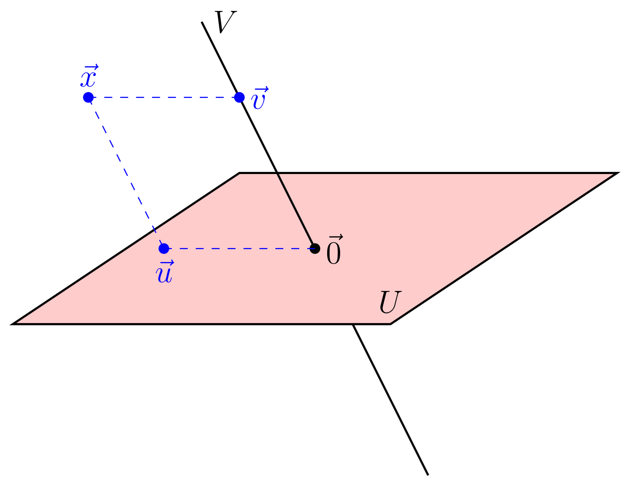
Théorème 7.1.2
Les conditions suivantes sont équivalentes :

(1) $E = U \oplus V$

(2) Tout vecteur $\vec{x} \in E$ s'écrit d'une façon unique *$\vec{x} = \vec{u} + \vec{v}$ avec $\vec{u} \in U$ et $\vec{v} \in V$.*

Preuve. Montrons l'équivalence de (1) et (2). Supposons $E = U \oplus V$, et soit $\vec{x} \in E$. Comme $E = U + V$, on sait qu'il existe $\vec{u} \in U$, $\vec{v} \in V$ tels que $\vec{x} = \vec{u} + \vec{v}$. S'il y avait deux décompositions de $\vec{x}$, on aurait $\vec{x} = \vec{u} + \vec{v} = \vec{u}' + \vec{v}''$ mais alors $\vec{u} - \vec{u}' = \vec{v}' - \vec{v} \in U \cap V$. L'hypothèse $U \cap V = \{\vec{0}\}$ montre que $\vec{u} - \vec{u}' = \vec{v}' - \vec{v} = \vec{0}$ donc $\vec{u} = \vec{u}'$ et $\vec{v} = \vec{v}'$.

Réciproquement, si (2) est vrai, alors on a bien sûr $E = U + V$. Il reste à montrer que $U \cap V = \{\vec{0}\}$. Si $\vec{x} \neq \vec{0}$, $\vec{x} \in U \cap V$, on aurait deux décompositions de $\vec{0}$ en la somme d'un vecteur de U et d'un vecteur de V : $\vec{0} = \vec{0} + \vec{0} = \vec{x} + (-\vec{x})$ ce qui contredirait (2). □



Théorème 7.1.3
Si, en outre, E est de dimension finie, alors on a équivalence de :

(3) $E = U \oplus V$

(4) $E = U + V$ *et* $\dim(E) = \dim(U) + \dim(V)$

(5) $U \cap V = \{\vec{0}\}$ *et* $\dim(E) = \dim(U) + \dim(V)$.

Preuve. Les équivalences de (3), (4), (5), se montrent avec la formule de Grassmann $\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$. □

Exercice : Montrez que, si $E = U \oplus V$, la réunion d'une base de U et d'une base de V est une base de E.

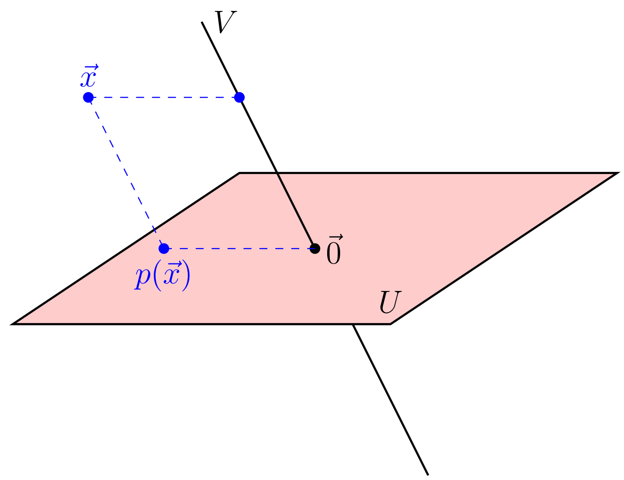
7.2 Projections et symétries

La notion de somme directe de sous-espaces vectoriels nous permet de définir des applications linéaires de grande importance, qui sont les projections et les symétries.

Définition 7.2.1
Soit E un $\mathbb{K}$-espace vectoriel, soit U et V deux sous-espaces vectoriels de E tels que $E = U \oplus V$. Avec les notations du théorème 7.1.2, la projection sur U parallèlement à V *est par définition l'application*

$$p : E \to E$$
$$\vec{x} = \vec{u} + \vec{v} \mapsto p(\vec{x}) = \vec{u}$$

Notons que, parce que $E = U \oplus V$, tout vecteur $\vec{x} \in E$ a une unique décomposition sous la forme $\vec{x} = \vec{u} + \vec{v}$ avec $\vec{u} \in U$ et $\vec{v} \in V$, et donc que l'application p est définie sans ambiguité.



Proposition 7.2.2
Avec les notations de la définition précédente, p est une application linéaire. De plus, on a les propriétés suivantes :

1. $p \circ p = p$
2. $\mathrm{Ker}(p) = V$
3. $\mathrm{Im}(p) = U$.
4. $U = \{\vec{x} \in E \mid p(\vec{x}) = \vec{x}\}$
5. *Si on note p' la projection sur V relativement à U, alors $p + p' = \mathrm{Id}$.*

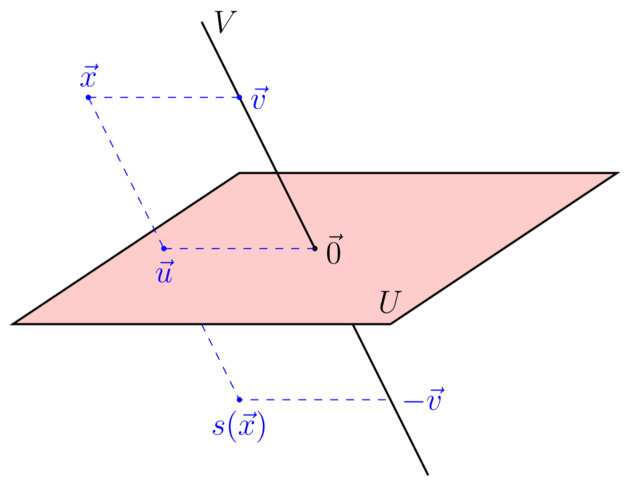
Preuve. Soit $\vec{x} = \vec{u} + \vec{v}$ et $\vec{x}' = \vec{u}' + \vec{v}'$ deux vecteurs de E. Alors $\lambda\vec{x} + \mu\vec{x}' = (\lambda\vec{u} + \mu\vec{u}') + (\lambda\vec{v} + \mu\vec{v}')$ et, comme $\vec{u}'' := \lambda\vec{u} + \mu\vec{u}' \in U$ et $\vec{v}'' := \lambda\vec{v} + \mu\vec{v}' \in V$, par définition de p, on a $p(\lambda\vec{x} + \mu\vec{x}') = \lambda\vec{u} + \mu\vec{u}'$. Donc on a bien $p(\lambda\vec{x} + \mu\vec{x}') = \lambda p(\vec{x}) + \mu p(\vec{x}')$ et p est linéaire.

Les propriétés énoncées se vérifient facilement à partir de la définition de p. Montrons la propriété 1. : pour tout $\vec{x} \in E$, $(p \circ p)(\vec{x}) = p(p(\vec{x})) = p(\vec{u}) = \vec{u} = p(\vec{x})$ donc on a bien $p \circ p = p$. On laisse les autres au lecteur. $\square$

Remarque 7.2.3
On peut montrer que, réciproquement, si f est une application linéaire de E dans E telle que $f \circ f = f$, alors f est une projection (voir l'exercice 68 du fascicule).

Définition 7.2.4
Soit E un $\mathbb{K}$-espace vectoriel, soit U et V deux sous-espaces vectoriels de E tels que $E = U \oplus V$. Avec les notations du théorème 7.1.2, la symétrie par rapport à U parallèlement à V *est par définition l'application*

$$s : E \to E$$
$$\vec{x} = \vec{u} + \vec{v} \mapsto s(\vec{x}) = \vec{u} - \vec{v}$$



Proposition 7.2.5
Avec les notations de la définition précédente, s est une application linéaire. De plus, on a les propriétés suivantes :

1. $s \circ s = \mathrm{Id}$
2. s *est inversible et* $s^{-1} = s$
3. $\mathrm{Im}(s) = E$ *et* $\mathrm{Ker}(s) = \{\vec{0}\}$
4. $U = \{\vec{x} \in E \mid s(\vec{x}) = \vec{x}\} = \mathrm{Ker}(s - \mathrm{Id})$
5. $V = \{\vec{x} \in E \mid s(\vec{x}) = -\vec{x}\} = \mathrm{Ker}(s + \mathrm{Id})$

Preuve. On montre que s est linéaire comme précédemment pour p. Pour tout $\vec{x} \in E$, $(s \circ s)(\vec{x}) = s(s(\vec{x})) = s(\vec{u} - \vec{v}) = \vec{u} - (-\vec{v}) = \vec{u} + \vec{v} = \vec{x}$ donc on a bien $s \circ s = \mathrm{Id}$. Cette identité montre que s est inversible d'inverse elle-même. En particulier, comme tout isomorphisme, $\mathrm{Im}(s) = E$ et $\mathrm{Ker}(s) = \{\vec{0}\}$

On a bien $s(\vec{x}) = \vec{x} \Leftrightarrow \vec{u} - \vec{v} = \vec{u} + \vec{v} \Leftrightarrow \vec{v} = \vec{0} \Leftrightarrow \vec{x} \in U$. De la même façon $s(\vec{x}) = -\vec{x} \Leftrightarrow \vec{x} \in V$. □

Remarque 7.2.6
On peut montrer que, réciproquement, si f est une application linéaire de E dans E telle que $f \circ f = \mathrm{Id}$, alors f est une symétrie (voir l'exercice 69 du fascicule).

7.3 Dualité

Dans cette section, on suppose que E est un espace vectoriel de dimension finie n et on va s'intéresser à l'espace $E^* := \mathcal{L}(E, \mathbb{K})$ des applications linéaires de E dans $\mathbb{K}$. On a déjà vu que c'est un $\mathbb{K}$-espace vectoriel (proposition 6.3.4). Son vecteur nul est l'application nulle de E dans $\mathbb{K}$ qui à tout vecteur $\vec{x}$ de E associe 0. On la notera simplement 0. On appelle cet espace *l'espace dual de E*.

Exemple 7.3.1
Supposons que $E = \mathbb{K}^n$. Soit $a_1, \dots, a_n$ des éléments de $\mathbb{K}$. L'application

$$\begin{aligned} E &\to \mathbb{K} \\ \vec{x} = (x_1, \dots, x_n) &\mapsto a_1 x_1 + \cdots + a_n x_n \end{aligned}$$

est un élément de E^*. En fait, tout élément φ de E^* est de cette forme, il suffit de prendre $a_1 = \varphi(\vec{\epsilon}_1), \dots, a_n = \varphi(\vec{\epsilon}_n)$, où $(\vec{\epsilon}_1, \dots, \vec{\epsilon}_n)$ est la base canonique de $\mathbb{K}^n$, et d'appliquer la propriété de linéarité :

$$\varphi(\vec{x}) = \varphi(x_1\vec{\epsilon}_1 + \cdots + x_n\vec{\epsilon}_n) = x_1\varphi(\vec{\epsilon}_1) + \cdots + x_n\varphi(\vec{\epsilon}_n) = x_1 a_1 + \cdots + x_n a_n.$$

Définition 7.3.2
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soit $(\vec{e}_1, \dots, \vec{e}_n)$ une base de E. On définit $(\vec{e}_1^*, \dots, \vec{e}_n^*)$ par :

$$\begin{aligned} \vec{e}_k^* \ : \ E &\to \mathbb{K} \\ \vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n &\mapsto \vec{e}_k^*(\vec{x}) = x_k \end{aligned}$$

Théorème 7.3.3
La famille $(\vec{e}_1^*, \dots, \vec{e}_n^*)$ est une base de E^*, On l'appelle la base duale *de la base $(\vec{e}_1, \dots, \vec{e}_n)$. En particulier,* $\dim(E^*) = \dim(E)$.

Preuve. On doit démontrer d'abord que $\vec{e}_k^* \in E^*$, c'est-à-dire que $\vec{e}_k^*$ est linéaire. En effet, si $\vec{x} \in E$ et $\vec{y} \in E$, on peut décomposer ces éléments sur la base $(\vec{e}_1, \dots, \vec{e}_n)$: il existe $(x_1, \dots, x_n) \in \mathbb{K}^n$ tels que $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$ et il existe $(y_1, \dots, y_n) \in \mathbb{K}^n$ tels que $\vec{y} = y_1\vec{e}_1 + \cdots + y_n\vec{e}_n$. Alors, pour tout $(\lambda, \mu) \in \mathbb{K}^2$, on a

$$\lambda\vec{x} + \mu\vec{y} = (\lambda x_1 + \mu y_1)\vec{e}_1 + \cdots + (\lambda x_n + \mu y_n)\vec{e}_n$$

et donc par définition $\vec{e}_k^*(\lambda\vec{x} + \mu\vec{y}) = \lambda x_k + \mu y_k = \lambda\vec{e}_k^*(\vec{x}) + \mu\vec{e}_k^*(\vec{y})$.

Montrons que $(\vec{e}_1^*, \dots \vec{e}_n^*)$ est une famille libre de E^* : en effet, si $\lambda_1\vec{e}_1^* + \cdots + \lambda_n\vec{e}_n^* = 0$, alors $(\lambda_1\vec{e}_1^* + \cdots + \lambda_n\vec{e}_n^*)(\vec{x}) = 0$ pour tout $\vec{x} \in E$. On prend $\vec{x} = \vec{e}_k$ et on obtient $\lambda_k = 0$. Ceci vaut pour tout $k = 1, \dots, n$, donc la famille est bien libre.

Montrons que $(\vec{e}_1^*, \dots \vec{e}_n^*)$ est une famille génératrice de E^* : Soit $\varphi \in E^*$, et $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \in E$. Alors,

$$\begin{aligned} \varphi(\vec{x}) &= \varphi(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n) \\ &= x_1\varphi(\vec{e}_1) + \cdots + x_n\varphi_n(\vec{e}_n) \\ &= \varphi(\vec{e}_1)\vec{e}_1^*(\vec{x}) + \cdots + \varphi(\vec{e}_n)\vec{e}_n^*(\vec{x}) \end{aligned}$$

Ceci est valable pour tout $\vec{x} \in E$, ce qui montre que $\varphi = \varphi(\vec{e}_1)\vec{e}_1^* + \cdots + \varphi(\vec{e}_n)\vec{e}_n^*$ et donc que φ est bien une combinaison linéaire de $\vec{e}_1^*, \dots, \vec{e}_n^*$. $\square$

Nous allons maintenant développer un point de vue un peu plus géométrique sur la dualité. En particulier nous allons démontrer que tout sous-espace de dimension k de E est l'intersection de $n-k$ hyperplans de E. Commençons par la définition des hyperplans, qui généralise la notion familière de plan dans l'espace usuel $\mathbb{R}^3$.

Définition 7.3.4
Un sous-espace de E de dimension $n-1$ (on dit aussi qu'il est de codimension 1 *dans E) est appelé un* hyperplan *de E.*

Proposition 7.3.5
Si $\varphi \in E^*$, $\varphi \neq 0$, alors son noyau $\mathrm{Ker}(\varphi)$ *est un hyperplan de E. Réciproquement, tout hyperplan de E est le noyau d'une forme linéaire non nulle de E.*

Preuve. Soit $\varphi \in E^*$, $\varphi \neq 0$. Son image est $\mathbb{K}$ tout entier (puisque $\mathbb{K}$ est de dimension 1, et que $\mathrm{Im}(\varphi) \neq \{\vec{0}\}$). Donc son noyau est, par le théorème du rang, un sous-espace vectoriel de E de dimension $n-1$.

Réciproquement, soit H un sous-espace vectoriel de E de dimension $n-1$. Choisissons $(h_1, \ldots, h_{n-1})$ une base de H. Par le théorème de la base incomplète, il existe un vecteur h_n de E tel que $(h_1, \ldots, h_{n-1}, h_n)$ soit une base de E. Alors, en prenant $\varphi = h_n^*$, on a bien $H = \mathrm{Ker}(\varphi)$. □

Exemple 7.3.6
Dans le cas où $E = \mathbb{K}^n$, en combinant la proposition 7.3.5 et l'expression des formes linéaires de $\mathbb{K}^n$ donnée dans l'exemple 7.3.1, on voit que les hyperplans de E sont exactement les sous-espaces définis par une équation linéaire de la forme $a_1x_1 + \cdots + a_nx_n = 0$ avec $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$:

$$H = \{(x_1, \ldots x_n) \in \mathbb{K}^n \mid a_1x_1 + \cdots + a_nx_n = 0\}.$$

Revenons au cas général d'un espace vectoriel E quelconque. On peut donc décrire un hyperplan soit en choisissant une base, c'est alors l'ensemble des combinaisons linéaires des éléments de la base choisie, soit comme le noyau d'une forme linéaire. On va généraliser cela aux sous-espaces de dimension quelconque dans la prochaine proposition.

Proposition 7.3.7
Soit $F \subset E$ un sous-espace vectoriel de E de dimension k. Alors, F est égal à l'intersection de $n-k$ hyperplans.

Preuve. Choisissons $(\vec{e}_1, \ldots, \vec{e}_k)$ une base de F, et complétons-la en une base de E $(\vec{e}_1, \ldots, \vec{e}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n)$. Alors il est facile de vérifier que

$$F = \mathrm{Ker}(\vec{e}_{k+1}^{\,*}) \cap \cdots \cap \mathrm{Ker}(\vec{e}_n^{\,*}).$$

□

Remarque 7.3.8
Dans le case où $E = \mathbb{K}^n$, on a donc montré que tout sous-espace vectoriel de $\mathbb{K}^n$ peut s'exprimer soit comme l'ensemble des combinaison linéaires d'une famille de vecteurs, soit comme l'ensemble des solutions d'un système linéaire. Ainsi, les deux exemples de sous-espaces vectoriels du chapitre 2 sont en fait les mêmes.