Notes du Cours
Algèbre linéaire
Math103

Guy Henniart & Thierry Ramond
Université Paris Sud
e-mail: guy.henniart@math.u-psud.fr, thierry.ramond@math.u-psud.fr

30 avril 2018

Table des matières

Chapitre 1

Systèmes linéaires

Les systèmes d'équations linéaires sont présents partout, en tous cas dès que l'on veut modéliser des phénomènes du monde réel où plusieurs paramètres entrent en jeu, éventuellement après linéarisation. Par exemple, les prévisions météorologiques sont obtenues par la résolution à l'aide d'ordinateurs d'énormes systèmes d'équations ou les inconnues sont la température, la pression, etc, en différents endroits du territoire que l'on considère.

Les mathématiques elles-mêmes regorgent de situations où l'on doit résoudre de tels systèmes. On commence modestement par des exemples issus de la géométrie, dans le plan ou dans l'espace. On trouvera dans l'appendice A un exposé rapide des notions essentielles sur les droites du plan et les droites et plans de l'espace.

1.1 Exemples de systèmes d'équations

1.1.1 En géométrie plane

Soit $(0, \vec{\imath}, \vec{\jmath})$ un repère du plan, et $A = (0, 1)$, $B = (1, 0)$, $C = (1, 2)$ et $D = (3, 1)$ quatre points. On cherche l'intersection des droites (AB) et (CD). Un point $M = (x, y)$ appartient à (AB) si et seulement si les vecteurs $\overrightarrow{AM}$ et $\overrightarrow{AB}$ sont colinéaires, c'est-à dire si et seulement si il existe $t_1 \in \mathbb{R}$ tel que $\overrightarrow{AM} = t_1\overrightarrow{AB}$ (cf. (A.2)). En termes de coordonnées, cela s'écrit

$$\left\{ \begin{array}{rcr} x - 0 & = & t_1 \times 1, \\ y - 1 & = & t_1 \times (-1). \end{array} \right.$$

Il s'agit d'un système de deux équations à trois inconnues (x, y et t_1). De même un point $M = (x, y)$ appartient à (CD) si et seulement si il existe $t_2 \in \mathbb{R}$ tel que

$$\left\{ \begin{array}{rcr} x - 1 & = & t_2 \times 2, \\ y - 2 & = & t_2 \times (-1). \end{array} \right.$$

Donc les droites (AB) et (CD) se coupent si et seulement si il existe $x, y, t_1, t_2 \in \mathbb{R}$ tels que

$$\left\{\begin{array}{ccccccccc} x & & & - & t_1 & & & = & 0, \\ & & y & + & t_1 & & & = & 1, \\ x & & & & & - & 2t_2 & = & 1, \\ & & y & & & + & t_2 & = & 2. \end{array}\right. \tag{1.1}$$

Il s'agit donc déterminer si le système de quatre équations à quatre inconnues ci-dessus admet des solutions, et, pour une réponse complète, donner les solutions s'il y en a. D'ailleurs on sait (c'est notre bon sens qui le dit, mais on va le prouver un peu plus loin) qu'il y a trois possibilités :

— ou bien (AB) et (CD) sont égales, et dans ce cas le système (1.1) a une infinité de solutions,
— ou bien (AB) et (CD) se coupent en un point, et le système (1.1) a une unique solution,
— ou bien (AB) et (CD) ne se coupent pas (elles sont distinctes et parallèles), et le système (1.1) n'a aucune solution.

On peut aussi partir des équations cartésiennes de (AB) et (CD) respectivement. On sait que (AB) admet une équation de la forme $ax + by = c$, et que A et B appartiennent à (AB). On doit donc avoir

$$\begin{aligned} & \left\{\begin{array}{l} a.x_A + b.y_A = c, \\ a.x_B + b.y_B = c, \end{array}\right. \\ \Longleftrightarrow & \left\{\begin{array}{l} a.0 + b.1 = c, \\ a.1 + b.0 = c, \end{array}\right. \\ \Longleftrightarrow & \ a = b = c. \end{aligned}$$

On est en présence d'un système de deux équations linéaires à trois inconnues qui a une infinité de solutions. On pouvait s'y attendre puisqu'une droite a une infinité d'équations cartésiennes, toutes multiples les unes des autres (cf. la proposition A.3.2). On peut choisir $a = b = c = 1$ et donc (AB) a pour équation cartésienne

$$(AB) : x + y = 1.$$

On notera que cette équation cartésienne est un système de une équation à deux inconnues, dont l'ensemble des solutions est par définition la droite (AB). De la même manière on voit que la droite (CD) a pour équation cartésienne

$$(CD) : x + 2y = 5.$$

Un point $M(x, y)$ appartient donc à l'intersection de (AB) et (CD) si et seulement si

$$\left\{\begin{array}{ccccc} x & + & y & = & 1, \\ x & + & 2y & = & 5. \end{array}\right.$$

Là encore, ce système de deux équations à deux inconnues peut avoir une infinité de solutions, ou bien une unique solution, ou bien pas de solution du tout.

1.1.2 En géométrie dans l'espace

Soit $(O,\vec{\imath},\vec{\jmath},\vec{k})$ un repère de l'espace, et $A=(0,1,1)$, $B=(1,0,1)$. Un point $M(x,y,z)$ appartient à la droite (AB) si et seulement si $\overrightarrow{AM}$ et $\overrightarrow{AB}$ sont colinéaires, c'est-à-dire si et seulement si il existe $t_1\in\mathbb{R}$ tel que $\overrightarrow{AM}=t\overrightarrow{AB}$. En coordonnées, cela donne le système de trois équations à quatre inconnues

$$\left\{\begin{array}{rcr} x-0 &=& t_1.1,\\ y-1 &=& t_1.(-1),\\ z-1 &=& t_1.0.\end{array}\right.$$

Autrement dit, la droite (AB) est l'ensemble des points de coordonnées (x,y,z) correspondant aux solutions (x,y,z,t_1) du système linéaire de trois équations à quatre inconnues

$$\left\{\begin{array}{ccccccc} x & & & - & t_1 & = & 0,\\ & y & & + & t_1 & = & 1,\\ & & z & & & = & 1.\end{array}\right.$$

Soit aussi $C=(1,2,2)$ et $D=(3,1,0)$. On s'intéresse maintenant à l'intersection des droites (AB) et (CD). On peut voir comme ci-dessus qu'un point $M(x,y,z)$ appartient à (CD) si et seulement si il existe $t_2\in\mathbb{R}$ tel que

$$\left\{\begin{array}{rcr} x-1 &=& 2t_2,\\ y-2 &=& -t_2,\\ z-2 &=& -2t_2.\end{array}\right.$$

Ainsi, le point $M(x,y,z)$ appartient à $(AB)\cap(CD)$ si et seulement si il existe $t_1,t_2\in\mathbb{R}$ tels que

$$\left\{\begin{array}{ccccccccc} x & & & - & t_1 & & & = & 0,\\ & y & & + & t_1 & & & = & 1,\\ & & z & & & & & = & 1,\\ x & & & & & - & 2t_2 & = & 1,\\ & y & & & & + & t_2 & = & 2,\\ & & z & & & + & 2t_2 & = & 2.\end{array}\right.$$

On arrive donc à un système linéaire de six équations à cinq inconnues. Sans donner de méthode générale pour résoudre ce genre de système (ce sera l'objet des sections suivantes !), on peut remarquer que la 3ème et la 6ème équation ensemble donnent $t_2=1/2$, puis que la 4ème équation donne $x=0$, que la première équation donne alors $t_1=0$, la deuxième $y=1$ et la cinquième $y=3/2$, ce qui est impossible. Donc (AB) et (CD) ne se coupent pas.

Ces droites sont-elles parallèles ? C'est le cas si et seulement si $\overrightarrow{AB}$ et $\overrightarrow{CD}$ sont colinéaires, c'est-à dire si et seulement si il existe $s\in\mathbb{R}$ tel que

$$\overrightarrow{AB}=s.\overrightarrow{CD}.$$

En coordonnées, cela donne encore un système linéaire, de trois équations à une inconnue :

$$\left\{\begin{array}{rcr} 1 &=& s.2,\\ -1 &=& s.(-1),\\ 0 &=& s.(-2).\end{array}\right.$$

Ce système n'a pas de solution, donc $\overrightarrow{AB}$ et $\overrightarrow{CD}$ ne sont pas colinéaires et (AB) et (CD) ne sont pas non plus parallèles : elles sont non-coplanaires.

En particulier les points A,B et C ne sont pas alignés, et on peut chercher l'équation cartésienne du plan (ABC) qu'ils définissent. On sait (cf. la section A.4.1) que l'équation cartésienne d'un plan dans l'espace est de la forme $ax + by + cz = d$, et que les coordonnées de A,B et C vérifient cette équation. On arrive donc au système de trois équations à quatre inconnues

$$\begin{cases} a.0 + b.1 + c.1 = d, \\ a.1 + b.0 + c.1 = d, \\ a.1 + b.2 + c.2 = d. \end{cases}$$

Comme pour une droite du plan, un plan dans l'espace admet une infinité d'équation cartésiennes, toutes multiples les unes des autres. On peut vérifier que $(a, b, c, d) = (1, 1, -2, -1)$ est une solution du système ci-dessus, donc que le plan (ABC) a pour équation cartésienne

$$(ABC) : x + y - 2z = -1.$$

Il s'agit d'un système de une équation à trois inconnues.

Pour finir, on veut déterminer l'intersection du plan (ABC) avec la droite (CD). On reprend les résultats ci-dessus : un point $M = (x, y, z)$ appartient à $(ABC) \cap (CD)$ si et seulement si il existe $t_2 \in \mathbb{R}$ tel que

$$\left\{\begin{array}{ccccccccc} x & & & & & - & 2t_2 & = & 1, \\ & & y & & & + & t_2 & = & 2, \\ & & & & z & + & 2t_2 & = & 2, \\ x & + & y & - & 2z & & & = & -1, \end{array}\right.$$

ce qui est cette fois un système linéaire de quatre équations à quatre inconnues. Il n'est pas difficile de résoudre ce système : les trois premières équations donnent x, y et z en fonction de t_2, et la dernière équation donne la valeur de t_2 si l'on y reporte ces expressions : on doit avoir

$$(2t_2 + 1) + (2 - t_2) - 2(2 - 2t_2) = -1,$$

c'est-à-dire $t_2 = 0$. Ce système a au plus une solution, et le plan (ABC) et la droite (CD) au plus un point commun. Comme C appartient aux deux, c'est leur intersection :

$$(ABC) \cap (CD) = \{C\}.$$

1.2 Solutions des systèmes linéaires

On formalise les exemples vu dans la section précédente. Il s'agit de pouvoir écrire des résultats généraux sur les systèmes linéaires et leurs solutions. On commence par quelques

1.2.1 Définitions

Définition 1.2.1 Une équation linéaire à p inconnues est une expression de la forme

$$a_1X_1 + a_2X_2 + \cdots + a_pX_p = b \tag{1.2}$$

où a_1, a_2, ..., a_p et b sont des réels fixés et X_1, X_2, ..., X_p des symboles.

Un solution de cette équation est un p-uplet $(x_1, x_2, \dots x_p)$ de réels tel que, si l'on remplace les symboles $X_1, X_2, \dots, X_p$ dans l'équation (1.2) par les valeurs $x_1, x_2, \dots x_p$, l'égalité correspondante est vraie.

Définition 1.2.2 Un système de n équations linéaires à p inconnues est une expression de la forme

$$\left\{\begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p & = & b_1 \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p & = & b_n \end{array}\right. \tag{1.3}$$

où les a_{ij} et les b_i pour $i \in \{1, \dots n\}$ et $j \in \{1, \dots p\}$ sont des réels fixés.

Une solution de ce système est un p-uplet $(x_1, x_2, \dots x_p)$ de réels tel que, si l'on remplace les symboles $X_1, X_2, \dots, X_p$ dans chacune des équations du système par les valeurs $x_1, x_2, \dots x_p$, les égalités correspondantes sont toutes vraies.

Les nombres $b_1, b_2, \dots, b_n$ sont les seconds membres du système (1.3). Lorsqu'ils sont tous nuls, on dit que le système (1.3) est homogène. Dans ce cas, le p-uplet $(0, 0, \dots, 0)$ est une solution du système.

Remarque 1.2.3 Chose un peu troublante au début, mais qui se révèle très pratique, on note la plupart du temps les symboles $X_1, X_2, \dots, X_p$ de la même manière que les solutions, c'est-à-dire $x_1, x_2, \dots, x_p$.

Définition 1.2.4 On dit que le système (1.3) est compatible (ou plutôt que les équations du système sont compatibles) lorsqu'il a au moins une solution. Sinon, on dit que le système (ou l'ensemble de ses équations) est incompatible.

1.2.2 Système homogène associé

Définition 1.2.5 On appelle système homogène associé au système (1.3), le système obtenu en remplaçant tous les second membres $b_1, b_2, \dots b_n$ par 0, c'est à dire

$$(1.4) \qquad \left\{ \begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p & = & 0 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p & = & 0 \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \dots + a_{n,p}X_p & = & 0 \end{array} \right.$$

où les a_{ij} sont ceux de (1.3).

Proposition 1.2.6 Soit $(x_1, x_2, \dots x_p)$ une solution du système (1.3). Le p-uplet $(x_1, x_2, \dots x_p) + (t_1, t_2, \dots, t_p)$ est aussi une solution de (1.3) si et seulement si $(t_1, t_2, \dots, t_p)$ est une solution du système homogène associé.

Preuve.— On rappelle d'abord que

$$(x_1, x_2, \dots x_p) + (t_1, t_2, \dots, t_p) = (x_1 + t_1, x_2 + t_2, \dots, x_p + t_p),$$

est le p-uplet obtenu en ajoutant les deux p-uplets composante par composante.

Or pour $k \in \{1, \dots n\}$, on a, puisque $a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kp}x_p = b_k$,

$$a_{k1}(x_1 + t_1) + a_{k2}(x_2 + t_2) + \dots + a_{kp}(x_p + t_p) = b_k + a_{k1}t_1 + a_{k2}t_2 + \dots + a_{kp}t_p$$

Donc $(x_1, x_2, \dots x_p) + (t_1, t_2, \dots, t_p)$ est une solution de (1.3) si et seulement si, pour tout $k \in \{1, \dots n\}$,

$$a_{k1}t_1 + a_{k2}t_2 + \dots + a_{kp}t_p = 0.$$

□

On en déduit par exemple le résultat suivant, qui ramène le problème de la détermination de toutes les solutions d'un système linéaire à celui de la détermination

— de toutes les solutions du système homogène associé,
— et d'une solution du système complet.

Corollaire 1.2.7 L'ensemble des solutions du système (1.3) s'obtient en ajoutant chacune des solutions du système homogène (1.4) associé à une solution fixée de (1.3).

1.2.3 Systèmes équivalents

Définition 1.2.8 On dit que deux systèmes linéaires à p inconnues sont équivalents lorsqu'ils ont le même ensemble de solutions.

Remarque 1.2.9 Attention ! De cette définition il découle que tous les systèmes incompatibles sont équivalents entre eux.

Bien entendu, l'ordre des lignes d'un système n'a pas d'importance : une solution du système n'est rien d'autre qu'une solution de toutes les lignes. Autrement dit, si l'on permute les lignes d'un système, on obtient un système équivalent.

Voici plusieurs moyens de transformer un système en un autre qui lui est équivalent mais, on l'espère, plus simple à résoudre. On adopte les notations suivantes : L_i désigne la i-ième ligne du système (1.3). D'autre part l'écriture $L_i \leftarrow e$ signifie qu'on considère désormais le système obtenu en remplaçant la i-ième ligne par l'équation e.

Proposition 1.2.10 Les opérations suivantes transforment un système en un autre qui lui est équivalent :

- Echange de deux lignes : $L_i \leftrightarrow L_j$, $i \neq j$.
- Multiplication des coefficients d'une ligne par un réel non-nul : $L_i \leftarrow \alpha L_i$, $\alpha \neq 0$.
- Remplacement d'une ligne par sa somme avec la multiple d'une autre : $L_i \leftarrow L_i + \alpha L_j$, $i \neq j$, $\alpha \in \mathbb{R}$

Preuve.— Le premier point est un cas particulier de permutation de lignes, et le second est vraiment simple. On prouve le troisième.

Soit $(x_1, x_2, \ldots, x_p)$ un p-uplet. S'il est solution du système (1.3), on a pour tout i et tout j,

$$\left\{\begin{array}{ccccccccc} a_{i1}x_1 & + & a_{i2}x_2 & + & \ldots & + & a_{ip}x_p & = & b_i \\ a_{j1}x_1 & + & a_{ji2}x_2 & + & \ldots & + & a_{jp}x_p & = & b_j \end{array}\right.$$

donc aussi

$$\left\{\begin{array}{ccccccccc} a_{i1}x_1 & + & a_{i2}x_2 & + & \ldots & + & a_{ip}x_p & = & b_i, \\ \alpha a_{j1}x_1 & + & \alpha a_{ji2}x_2 & + & \ldots & + & \alpha a_{jp}x_p & = & \alpha b_j, \end{array}\right.$$

En sommant membre à membre ces deux égalités, on obtient

$$(a_{i1} + \alpha a_{j1})x_1 + (a_{i2} + \alpha a_{j2})x_2 + \cdots + (a_{ip} + \alpha a_{jp})x_p = b_i + \alpha b_j,$$

ce qui montre que $(x_1, x_2, \ldots, x_p)$ est aussi solution de $L_i + \alpha L_j$. C'est donc bien une solution du système modifié.

Réciproquement, si $(x_1, x_2, \ldots, x_p)$ est une solution du système modifié, on a

$$\left\{\begin{array}{ccccccccc} a_{j1}x_1 & + & a_{j2}x_2 & + & \ldots & + & a_{jp}x_p & = & b_j, \\ (a_{i1} + \alpha a_{j1})x_1 & + & (a_{i2} + \alpha a_{j2})x_2 & + & \ldots & + & (a_{ip} + \alpha a_{jp})x_p & = & b_i + \alpha b_j. \end{array}\right.$$

En multipliant la première égalité par α et en la retranchant à la seconde, on trouve

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ip}x_p = b_i,$$

ce qui montre que $(x_1, x_2, \ldots, x_p)$ est une solution du système original. $\square$

1.3 Cas des systèmes échelonnés

1.3.1 Définitions

On considère à nouveau le système

$$\left\{ \begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p & = & b_1 \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p & = & b_n \end{array} \right. \tag{1.5}$$

Définition 1.3.1 On dit qu'une ligne de (1.5) est nulle lorsque tous les coefficients de son membre de gauche sont nuls. Lorsqu'une ligne n'est pas nulle, son coefficient non-nul le plus à gauche est appelé pivot de la ligne.

Exemple 1.3.2 Dans le système

$$\left\{ \begin{array}{ccccccccc} 2X_1 & + & X_2 & & & + & 3X_4 & = & 1, \\ & & & & aX_3 & & & = & 0, \\ & & X_2 & + & X_3 & - & X_4 & = & -1, \end{array} \right.$$

le pivot de la première ligne est $a_{11} = 2$, et le pivot de la troisième ligne est $a_{32} = 1$. La deuxième ligne est nulle si et seulement si $a = 0$. Pour $a \neq 0$, son pivot est $a_{23} = a$.

Définition 1.3.3 On dit qu'un système est échelonné lorsque
- — il n'y a pas de ligne non-nulle en dessous d'une ligne nulle, et
- — le pivot d'une ligne non-nulle est strictement plus à droite que celui de la ligne précédente (s'il y en a une).

Le système de l'exemple 1.3.2 n'est pas échelonné : si $a = 0$, la deuxième ligne est nulle mais la troisième ne l'est pas, et si $a \neq 0$ le pivot de la troisième ligne est $a_{3,2} = 1$ qui se trouve à gauche du pivot de la seconde ligne $a_{2,3} = a$.

Définition 1.3.4 Le rang d'un système échelonné est le nombre de pivots du système, autrement dit le nombre de lignes non-nulles. On dit qu'une inconnue d'un système échelonné est principale lorsque l'un de ses coefficients dans le système est un pivot. Les autres inconnues sont dites secondaires.

Remarque 1.3.5 On a bien sûr $r \leq n$: il y a au plus un pivot par ligne. Dans un système échelonné, on a aussi $r \leq p$ puisque le numero de colonne des pivots est strictement croissant, et qu'il y a p colonnes.

1.3.2 Solutions d'un système échelonné

Lorsqu'un système est échelonné, il est très simple de décrire l'ensemble de ses solutions. On commence par des exemples.

Exemple 1.3.6 Le système à deux équations et trois inconnues

$$\left\{\begin{array}{ccccccc} x_1 & + & 3x_2 & + & x_3 & = & 3, \\ & & & & x_3 & = & 1. \end{array}\right.$$

est échelonné, de rang 2. Le pivot de la première ligne est son premier coefficient $a_{1,1} = 1$ et celui de la deuxième ligne est son troisième coefficient $a_{2,3} = 1$. Les inconnues principales sont donc x_1 et x_3, et x_2 est une inconnue secondaire. La deuxième ligne fixe la valeur de l'inconnue principale $x_3 = 1$, et compte tenu de ce fait, la première équation s'écrit $x_1 = 3 - 3x_2 - 1 = 2 - x_2$. Autrement dit le système admet une solution pour chaque valeur de l'inconnue secondaire x_2. Appelons s cette valeur. Le système s'écrit

$$\left\{\begin{array}{ccccc} x_1 & = & 2 & - & 3s, \\ x_2 & = & & & s, \\ x_3 & = & 1. & & \end{array}\right.$$

L'ensemble des solutions est donc $((2 - 3s, s, 1),\ s \in \mathbb{R})$, ou encore l'ensemble $((2, 0, 1) + s(-3, 1, 0), s \in \mathbb{R})$. On reconnait dans cette dernière écriture le contenu du corollaire 1.2.7 : une solution du système complet s'écrit comme somme de la solution $(2, 0, 1)$ du système complet, et d'une solution quelconque du système homogène associé : $s(-3, 1, 0)$ pour un s dans $\mathbb{R}$.

Exemple 1.3.7 Le système

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

est un système échelonné de 3 équations à quatre inconnues, de rang 2. Ses deux inconnues principales sont x_1 et x_3, et x_2, x_4 sont des inconnues secondaires. La troisième ligne est nulle, mais pas son second membre : cette équation n'a pas de solution, et par conséquent le système non plus. Il est incompatible.

Ces deux exemples sont des cas particuliers du résultat totalement général suivant, qui est le résultat central de ce chapitre.

Proposition 1.3.8 Soit un système échelonné de n équations à p inconnues, de rang r.

Compatibilité : les $n-r \geq 0$ dernières lignes sont nulles, et on examine les seconds membres de ces lignes $b_{r+1}, ..., b_n$. Si l'un d'entre eux est non nul, le système est incompatible. Sinon il est compatible.

Solutions : Si le système est compatible, on peut attribuer des valeurs arbitraires aux $p-r \geq 0$ inconnues secondaires, et les r inconnues principales sont alors uniquement déterminées. On les trouve en résolvant la r-ème équation, puis la $r-1$-ème, etc.

Dans le cas $r = n$, il n'y a pas de ligne nulle, et pas de second membre à examiner : le système est compatible. Si de plus $p = r = n$, il n'y a pas d'inconnue secondaire et le système admet une unique solution. Pour prouver cette proposition, on va transformer un système échelonné en un système équivalent pour lequel l'énoncé est presque évident.

1.3.3 Système échelonné réduit

Définition 1.3.9 On dit qu'un système échelonné est réduit lorsque

i) Tous les pivots du système sont égaux à 1,

ii) Dans sa colonne, un pivot est le seul élément non-nul.

Exemple 1.3.10 Voici un exemple de système échelonné réduit,

$$\left\{\begin{array}{ccccccc} x_1 & + & 3x_2 & & & = & 3, \\ & & & & x_3 & = & 1. \end{array}\right.$$

et d'un système échelonné qui n'est pas réduit :

$$\left\{\begin{array}{ccccccc} x_1 & + & 3x_2 & + & 2x_3 & = & 3, \\ & & & & 3x_3 & = & 1. \end{array}\right.$$

En effet dans ce dernier système, les pivots sont $a_{11} = 1$ et $a_{2,3} = 3$ ce qui contredit le point $i)$ de la définition. De plus, dans la colonne de ce pivot, donc la 3ème, il y a un autre élément non-nul : $a_{1,3} = 2$, ce qui contredit aussi le point $ii)$.

Sur ce système, on peut effectuer l'opération

$$\mathbf{L_2} \leftarrow \frac{\mathbf{1}}{\mathbf{3}}\mathbf{L_2}.$$

On obtient le système équivalent :

$$\left\{\begin{array}{ccccccc} x_1 & + & 3x_2 & + & 2x_3 & = & 3, \\ & & & & x_3 & = & \frac{1}{3}, \end{array}\right.$$

pour lequel la première objection ci-dessus a disparu. On peut alors faire

$$\mathbf{L_1} \leftarrow \mathbf{L_1} - \mathbf{2L_2},$$

et on obtient le système :

$$\begin{cases} x_1 + 3x_2 & = \frac{7}{3}, \\ \qquad\qquad x_3 & = \frac{1}{3}, \end{cases}$$

qui est échelonné réduit et équivalent au système dont on est parti.

La méthode utilisée dans cet exemple est tout à fait générale, et permet de montrer la

Proposition 1.3.11 Pour un système échelonné donné, il existe un système échelonné réduit qui lui est équivalent.

Preuve.— Un système échelonné de n équations à p inconnues, de rang r, s'écrit nécessairement

$$\left\{\begin{array}{ccccccccccccccc} a_{11}X_1 & + & a_{12}X_2 & + & a_{13}X_3 & + & \dots & + & a_{1r}X_r & + & \dots & + & a_{1p}X_p & = & b_1 \\ & & a_{22}X_2 & + & a_{23}X_3 & + & \dots & + & a_{2r}X_r & + & \dots & + & a_{2p}X_p & = & b_2 \\ & & & & a_{33}X_3 & + & \dots & + & a_{3r}X_r & + & \dots & + & a_{3p}X_p & = & b_3 \\ & & & & & & & & \vdots & & & & \vdots & & \vdots \\ & & & & & & & & a_{rr}X_r & + & \dots & + & a_{r,p}X_p & = & b_r \\ & & & & & & & & & & & & 0 & = & b_{r+1} \\ & & & & & & & & & & & & \vdots & & \vdots \\ & & & & & & & & & & & & 0 & = & b_n \end{array}\right.$$

où certains des a_{ij} peuvent être nuls, mais pas tous ceux qui se trouvent sur une même ligne. Pour $i \in \{1, \dots, r\}$, on note $k(i)$ le numéro de colonne du pivot de la ligne i. On sait que $1 \leq k(i) < k(i+1) \leq r$. On aussi, sur la ligne i,

$$a_{i,j} = 0 \text{ pour } j < k(i), \text{ et } a_{i,k(i)} \neq 0. \tag{1.6}$$

Comme le système est échelonné, on sait aussi que, pour chaque i, sur la colonne $k(i)$,

$$a_{\ell,k(i)} = 0 \text{ pour } \ell > i. \tag{1.7}$$

Etape 1 : Pour tout $1 \leq i \leq r$, on effectue l'opération

$$\mathbf{L_i} \leftarrow \frac{\mathbf{1}}{\mathbf{a_{i,k(i)}}}\mathbf{L_i}.$$

On obtient un système équivalent, dont tous les pivots valent 1, et dont on note $a'_{i,j}$ les coefficients et L'_i les lignes. Il faut aussi noter que l'on a toujours (1.6) et (1.7) pour les $a'_{i,j}$, et de plus

$$a'_{i,k(i)} = 1 \text{ pour tout } i \in \{1, \dots, r\}.$$

Etape 2 : Pour chaque colonne $k(i)$ de chaque pivot, avec $i = r, r-1, \dots, 1$ (dans cet ordre), on "nettoie" la colonne : pour toutes les lignes au dessus de la ligne i dont le coefficient dans la colonne $k(i)$ n'est pas nul, c'est à dire pour tout $\ell < i$ tel que $a'_{\ell,k(i)} \neq 0$ on effectue l'opération

$$\mathbf{L}'_\ell \leftarrow \mathbf{L}'_\ell - \mathbf{a}'_{\ell,\mathbf{k(i)}}\mathbf{L}'_\mathbf{i},$$

ce qui donne un nouveau système équivalent dont on note $a''_{i,j}$ les coefficients et L''_i les lignes. Sur les nouvelles lignes L''_ℓ avec $\ell < i$, le coefficient sur la colonne $k(i)$ est, puisque $a'_{i,k(i)} = 1$,

$$a''_{\ell,k(i)} = a'_{\ell,k(i)} - a'_{\ell,k(i)}.a'_{i,k(i)} = 0.$$

De plus les coefficients $a'_{i,j}$ de la ligne L'_i sont tous nuls avant la colonne $j = k(i)$, donc on a, pour $\ell < i$ et $j < k(i)$,

$$a''_{\ell,j} = a'_{\ell,j} - a'_{\ell,k(i)}.a'_{i,j} = a'_{\ell,j} - a'_{\ell,k(i)}.0 = a'_{\ell,j}.$$

Autrement dit, cette opération n'a pas changé la ligne ℓ jusqu'à la colonne qui précède la $k(i)$-ième.

Par contre, les coefficients des colonnes qui suivent peuvent avoir changé sur la ligne $\ell < i$: pour $j > k(i)$ on a aussi

$$a''_{m,j} = a'_{m,j} - a'_{m,k(i)}.a'_{i,j}$$

mais rien ne dit cette fois que $a'_{i,j} = 0$ pour ces valeurs de j... sauf si la colonne j correspond à un pivot, pour lequel on a déjà fait le nettoyage nécessaire puisqu'on prend les colonnes dans l'ordre décroissant : cette opération ne modifie pas les colonnes des pivots à droite de celle que l'on considère, et donc les laisse avec pour seul élément non nul le pivot qui vaut 1.

□

Exemple 1.3.12 On considère le système échelonné

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

C'est un système de trois équations à 4 inconnues, de rang trois. Les pivots sont $a_{1,1} = 1$, $a_{2,2} = 2$, et $a_{3,3} = 1$, les inconnues principales x_1, x_2 et x_3, et x_4 est une inconnue secondaire. On applique la procédure ci-dessus.

1. Comme le pivot de la seconde ligne ne vaut pas 1 on effectue l'opération $\mathbf{L_2} \leftarrow \frac{\mathbf{1}}{\mathbf{2}}\mathbf{L_2}$. On obtient le système équivalent

$$\left\{\begin{array}{ccccccccc} x_1 & + & x_2 & - & x_3 & & & = & 0, \\ & & x_2 & - & \frac{1}{2}x_3 & - & \frac{1}{2}x_4 & = & 2, \\ & & & & x_3 & + & x_4 & = & 1. \end{array}\right.$$

2. On considère le pivot de la dernière ligne $a_{3,3} = 1$. Dans sa colonne, il y a deux coefficients non-nuls $a_{2,3} = -\frac{1}{2}$ et $a_{1,3} = -1$. On effectue donc les opérations $\mathbf{L_2} \leftarrow \mathbf{L_2} + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{L_3}$ et $\mathbf{L_1} \leftarrow \mathbf{L_1} + \mathbf{L_3}$. On obtient le système

$$\left\{\begin{array}{ccccccccc} x_1 & + & x_2 & & & + & x_4 & = & 1, \\ & & x_2 & & & & & = & \frac{5}{2}, \\ & & & & x_3 & + & x_4 & = & 1. \end{array}\right.$$

3. On considère le pivot de la deuxième ligne $a_{2,2} = 1$. Dans sa colonne, il y a un coefficient non-nul $a_{1,2} = 1$. On effectue l'opération $\mathbf{L_1} \leftarrow \mathbf{L_1} - \mathbf{L_2}$, et on obtient le système

$$\left\{\begin{array}{ccccccc} x_1 & & & + & x_4 & = & -\frac{3}{2}, \\ & x_2 & & & & = & \frac{5}{2}, \\ & & x_3 & + & x_4 & = & 1. \end{array}\right.$$

On peut maintenant facilement prouver la proposition 1.3.8.

Preuve de la Proposition 1.3.8.— Le système échelonné réduit associé à un système échelonné de n équations à p inconnues de rang r s'écrit, notant $k(i)$ le numéro de colonne du pivot de la ligne i,

$$\left\{\begin{array}{cccccccccccc} X_{k(1)} & & & & + & a_{1,k(r)+1}X_{k(r)+1} & + & \dots & + & a_{1p}X_p & = & b_1, \\ & X_{k(2)} & & & + & a_{2,k(r)+1}X_{k(r)+1} & + & \dots & + & a_{2p}X_p & = & b_2, \\ & & \ddots & & & \vdots & & & & \vdots & & \vdots \\ & & & X_{k(r)} & + & a_{r,k(r)+1}X_{k(r)+1} & + & \dots & + & a_{r,p}X_p & = & b_r, \\ & & & & & & & & & 0 & = & b_{r+1}, \\ & & & & & & & & & \vdots & & \vdots \\ & & & & & & & & & 0 & = & b_n. \end{array}\right.$$

Les r inconnues principales sont les $X_{k(i)}$ pour $i \in \{1, \dots r\}$, alors que les $X_{k(r)+1}, \dots, X_p$ sont les $p - r$ inconnues secondaires. Si l'un des seconds membres $b_{r+1}, \dots b_n$ des $n - r$ dernières équations n'est pas nul, alors le système est incompatible. Sinon, il n'y a aucune condition sur les inconnues secondaires, et chacun d'entre elles peut donc prendre n'importe quelle valeur. Après avoir fixées la valeur des inconnues secondaires, les r premières équations ont exactement une solution pour l'inconnue principale qu'elle contient. $\square$

1.4 Cas général : la méthode du pivot

On s'attaque maintenant à la résolution d'un système linéaire quelconque de n équations à p inconnues, qu'on ne suppose donc plus échelonné. Pour ce faire, on dispose de la méthode du pivot, souvent attribuée au mathématicien K.F. Gauss. Il s'agit d'un algorithme qui permet de transformer un système quelconque en un système échelonné réduit qui lui est équivalent. Compte tenu de ce qui précède, c'est donc bien un moyen d'obtenir l'ensemble des solutions de n'importe quel système linéaire.

Puisque l'on sait déjà transformer un système échelonné en un système échelonné réduit équivalent, il reste à prouver la

Proposition 1.4.1 Pour tout système linéaire de n équations à p inconnues, il existe un système échelonné qui lui est équivalent.

Preuve.— Algorithme du Pivot

On considère le système linéaire général

$$\left\{\begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p & = & b_1, \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p & = & b_2, \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p & = & b_n. \end{array}\right. \tag{1.8}$$

Etape 1 : On cherche la première colonne non-nulle. S'il n'y en pas, toutes les équations sont nulles, et le système est déjà échelonné. Sinon, dans cette colonne j_1, on cherche le premier (en partant du haut) élément non-nul, c'est-à dire le pivot de la colonne. Notant $i(j_1)$ le numéro de sa ligne, on effectue l'échange

$$\mathbf{L_1} \leftrightarrow \mathbf{L_{i(j_1)}}.$$

On obtient un système équivalent dont on note $a'_{i,j}$ les coefficients, b'_i les seconds membres et L'_i les lignes. On sait que $a'_{1,j_1} \neq 0$. Pour chaque ligne L'_i du système avec $i \geq 2$, on effectue l'opération

$$\mathbf{L'_i} \leftarrow \mathbf{L'_i} - \frac{\mathbf{a'_{i,j_1}}}{\mathbf{a'_{1,j_1}}}\mathbf{L'_1}.$$

On obtient un nouveau système équivalent dont on note $a''_{i,j}$ les coefficients, b''_i les seconds membres et L''_i les lignes. On sait que pour $i \geq 2$,

$$a''_{i,j} = a'_{i,j} - \frac{a'_{i,j_1}}{a'_{1,j_1}}a'_{1,j}$$

donc en particulier, pour tout $i \geq 2$,

$$a''_{i,j_1} = a'_{i,j_1} - \frac{a'_{i,j_1}}{a'_{1,j_1}}a'_{1,j_1} = 0,$$

autrement dit tous les coefficients de la j_1-ième colonne sont nuls sauf le premier, et le système est de la forme

$$\left\{\begin{array}{ccccccccccccc} 0X_1 & + & \ldots & + & a''_{1,j_1}X_{j_1} & + & a''_{1,j_1+1}X_{j_1+1} & + & \ldots & + & a''_{1,p}X_p & = & b''_1, \\ 0X_1 & + & \ldots & + & 0X_{j_1} & + & a''_{2,j_1+1}X_{j_1+1} & + & \ldots & + & a''_{2,p}X_p & = & b''_2, \\ \vdots & & & & \vdots & & \vdots & & & & \vdots & & \vdots \\ 0X_1 & + & \ldots & + & 0X_{j_1} & + & a''_{n,j_1+1}X_{j_1+1} & + & \ldots & + & a''_{n,p}X_p & = & b''_n. \end{array}\right.$$

Etape 2 : On oublie la j_1-ième colonne et la première ligne, et on applique l'étape 1 au sous-sytème obtenu (en bleu ci-dessus), tant qu'il reste des colonnes.

Le système qu'on obtient en appliquant cet algorithme est échelonné. □

On trouvera sur le web à l'adresse `http://www.math.u-psud.fr/~ramond/docs/codes/pivot.py` une implémentation en python de l'algorithme du pivot, qui se veut être très proche de la description mathématiques de l'algorithme. Elle n'est donc pas la plus rapide possible, mais donne un autre éclairage sur ce chapitre. Chose peut-être utile, le programme permet de résoudre n'importe quel système linéaire.

Pour un système quelconque, on a finalement aussi une notion de rang :

Définition 1.4.2 Le rang d'un système linéaire est le rang du système échelonné qu'on obtient en appliquant l'algorithme du pivot.

On verra plus loin (cf. la remarque **??**) que cette notion ne dépend en fait pas de la méthode qu'on utilise pour trouver un système échelonné équivalent.

Exemple 1.4.3 On veut résoudre le système linéaire de 3 équations à 3 inconnues

$$\left\{\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & 9, \\ 4x & + & 2y & + & z & = & 9, \\ -6x & - & y & + & 2z & = & 12. \end{array}\right.$$

1. On examine la première colonne. Elle n'est pas nulle, et son pivot est $a_{1,1} = 2$. On effectue donc les opérations

$$\mathbf{L_2} \leftarrow \mathbf{L_2} - \frac{\mathbf{4}}{\mathbf{2}}\mathbf{L_1} \quad \text{et} \quad \mathbf{L_3} \leftarrow \mathbf{L_3} - \frac{\mathbf{-6}}{\mathbf{2}}\mathbf{L_1},$$

   ce qui donne le système équivalent

$$\left\{\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & 9, \\ & & 4y & - & 5z & = & -9, \\ & & -4y & + & 11z & = & 39. \end{array}\right.$$

2. On applique la même procédure au sous-système obtenu en oubliant la première ligne et la première colonne (en bleu ci-dessus). Sa première colonne n'est pas nulle, et son pivot est $a_{2,2} = 4$. On effectue donc l'opération

$$\mathbf{L_3} \leftarrow \mathbf{L_3} - \frac{\mathbf{-4}}{\mathbf{4}}\mathbf{L_2},$$

   ce qui donne le système équivalent

$$\left\{\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & 9, \\ & & 4y & - & 5z & = & -9, \\ & & & & 6z & = & 30, \end{array}\right.$$

   et ce système est échelonné.

Exercice 1.4.4 Donner la forme échelonnée réduite de ce système.

Chapitre 2

Espaces vectoriels

2.1 Définition et exemples

Soit E un ensemble. On suppose que l'on sait additionner les éléments de E entre eux, et qu'on obtient alors un élément de E. On note

$$a + b$$

la somme des éléments a et b de E. On suppose aussi que l'on sait multiplier un élément de E quelconque par un réel, et que l'on obtient ainsi un élément de E. On note

$$\lambda.a$$

le produit du réel λ et de $a \in E$. Dans ces conditions, on dit que "$+$" est une loi de composition interne sur E, et que "." est une loi de composition externe sur $\mathbb{R} \times E$.

Exemple 2.1.1 Soit $E = \mathbb{R}$. On sait ajouter deux éléments de $\mathbb{R}$, et multiplier un élément de $\mathbb{R}$ par un élément de $\mathbb{R}$.

Exemple 2.1.2 Soit $E = \mathbb{R}^n$, c'est-à-dire l'ensemble des n-uplets $(x_1, x_2, \ldots, x_n)$ d'éléments de $\mathbb{R}$. Les réels x_1, x_2, …, x_n sont appelés composantes du n-uplet $(x_1, x_2, \ldots, x_n)$.

On sait ajouter deux n-uplets, suivant la définition

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

c'est-à-dire composante par composante. On sait aussi multiplier un n-uplets par un réel :

$$\lambda.(x_1, x_2, \ldots, x_n) := (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Il s'agit là encore de multiplier chacune des composantes par λ.

Exemple 2.1.3 Soit $\mathcal{V}$ l'ensemble des vecteurs du plan ou de l'espace. On se rappelle qu'un vecteur non-nul est la donnée d'une direction, d'un sens et d'une longueur. Si $\vec{v}$ est le vecteur nul, $\lambda.\vec{v}$ est le vecteur nul pour tout réel λ. Pour $\lambda \in \mathbb{R}$ et $\vec{v} \in \mathcal{V}$ non-nul, le vecteur $\lambda.\vec{v}$ est le vecteur nul si $\lambda = 0$, et pour $\lambda \neq 0$,

— le vecteur de même direction que $\vec{v}$,
— de même sens que $\vec{v}$ si $\lambda > 0$ et de sens opposé si $\lambda < 0$
— de longueur $|\lambda|$ fois la longueur de $\vec{v}$.

On sait aussi ajouter deux vecteurs $\vec{u}$ et $\vec{v}$ de $\mathcal{V}$: on trace un représentant de $\vec{v}$ en partant de l'extrémité d'un représentant de $\vec{u}$.

Exemple 2.1.4 Soit X un ensemble quelconque, et $\mathcal{F}(X, \mathbb{R})$ l'ensemble des fonctions de X dans $\mathbb{R}$. La somme $h = f + g$ de deux fonctions f et g est la fonction définie par

$$h : x \mapsto f(x) + g(x).$$

Le produit $k = \lambda.f$ de la fonction f par un réel λ est la fonction définie par

$$k : x \mapsto \lambda f(x).$$

Ces définitions sont les définitions usuelles dans le cas de $X = \mathbb{R}$.

Définition 2.1.5 Soit $(E, +, .)$ un triplet constitué d'un ensemble E, d'une loi de composition interne "+" sur E et d'une loi de composition externe "." sur E. On dit que $(E, +, .)$ est un espace vectoriel lorsque les lois "+" et "." vérifient les huit propriétés suivantes :

1 Pour tous a, b de E, $a + b = b + a$.
2 Pour tous a, b et c de E, $a + (b + c) = (a + b) + c$.
3 Il existe un élément e de E tel que pour tout a de E, $a + e = e + a = a$.
4 Pour tout a de E, il existe un élément b de E tel que $a + b = b + a = e$.
N Pour tout a de E, $1.a = a$.
A Pour tous réels λ et μ, et tout $a \in E$, $\lambda.(\mu.a) = (\lambda\mu).a$.
D_1 Pour tous réels λ et μ, et tout $a \in E$, $(\lambda + \mu).a = \lambda.a + \mu.a$.
D_2 Pour tous a, b de E et tout réel λ, $\lambda.(a + b) = \lambda.a + \lambda.b$.

Lorsque la propriété (1) est vérifiée, on dit que la loi '+' est commutative. Si la propriété (2) est vérifiée, on dit que la loi '+' est associative. L'élément e de la propriété (3) est appelé élément neutre pour la loi '+'. On le note la plupart du temps 0_E où même 0.

Remarque 2.1.6 Il ne peut y avoir qu'un seul élément neutre pour une loi interne : si e_1 et e_2 sont élément neutre, on a $e_1 = e_1 + e_2 = e_2$.

L'élément b de la propriété (4) s'il existe est lui aussi nécessairement unique : $b_2 = b_2 + e = b_2 + (a + b_1) = (b_2 + a) + b_1 = b_1$. Le cas échéant, on l'appelle symétrique de a, et on le note $b = -a$.

Définition 2.1.7 Lorsque les propriétés (1) à (4) sont vérifiées, on dit que $(E, +)$ est un groupe commutatif.

Lorsque la propriété (N) est vérifiée, on dit que le réel 1 est l'élement neutre pour la loi externe. On dit que la loi externe est associative lorsque (A) est vraie. Les deux dernières propriétés (D_1) et (D_2) sont des propriétés de distributivité.

On reprend les exemples précédents, et on montre que l'on a bien affaire à des espaces vectoriels.

Exemple 2.1.8 Dans $E = \mathbb{R}$, l'addition et a multiplication usuelles possèdent bien les huit propriétés de la définition : $(\mathbb{R}, +, .)$ est un espace vectoriel.

Exemple 2.1.9 Dans $E = \mathbb{R}^n$, les deux lois de compositions définies ci-dessus ont aussi les huit propriétés de la définition. En effet les opérations sont définies composante par composante, donc il suffit de vérifier leurs propriétés composante par composante, et l'on se retrouve alors dans $(\mathbb{R}, +, .)$. On notera que l'élément neutre pour l'addition dans $\mathbb{R}^n$ est $(0, 0, \dots, 0)$.

Exemple 2.1.10 L'ensemble $\mathcal{F}(X, \mathbb{R})$ muni des deux opérations définies dans l'exemple 2.1.4 est un espace vectoriel. Là encore on se ramène aux propriétés équivalentes dans $(\mathbb{R}, +, .)$ qui est cette fois l'espace d'arrivée des fonctions de $\mathcal{F}(X, \mathbb{R})$. L'élément neutre pour l'addition dans $\mathcal{F}(X, \mathbb{R})$ est la fonction nulle $0 : x \mapsto 0$.

Si F est un espace vectoriel, on peut aussi définir sur l'ensemble $\mathcal{F}(X, F)$ des fonctions de X dans F les deux opérations de somme et de produit par un réel en utilisant les opérations dans F. On montre facilement, comme dans le cas de $\mathcal{F}(X, \mathbb{R})$, que muni de ces opérations, $\mathcal{F}(X, F)$ est un espace vectoriel. L'élément neutre pour l'addition dans $\mathcal{F}(X, F)$ est la fonction nulle $0 : x \mapsto 0_F$.

Exemple 2.1.11 Soit $\mathcal{S}$ l'ensemble des suites de nombres réels. On définit la somme $(u_n) + (v_n)$ de (u_n) et (v_n) comme étant la suite $(u_n + v_n)$, et le produit $\lambda.(u_n)$ de la suite (u_n) par le réel λ comme étant la suite $(\lambda.u_n)$. Comme dans les deux exemples ci-dessus, il est facile de voir que $(\mathcal{S}, +, .)$ est un espace vectoriel en se ramenant aux propriétés de l'addition et de la multiplication dans $\mathbb{R}$. Son élément neutre pour l'addition est la suite nulle $(0, 0, \dots)$.

2.2 Sous-espaces vectoriels

Définition 2.2.1 Soit $(E, +, .)$ un espace vectoriel, et V une partie de E. On dit que V est un sous-espace vectoriel de E lorsque les trois propriétés suivantes sont vérifiées :

- Le vecteur nul 0_E de E appartient à V.
- Pour tout réel λ et tout élément v de V, $\lambda.v$ est aussi un élément de V.
- Pour tous éléments a et b de V, $a + b$ est aussi un élément de V.

Exemple 2.2.2 *i)* Dans $(\mathbb{R}, +, .)$, il n'y a pas beaucoup de sous-espaces vectoriels différents. Soit en effet un sous-espace vectoriel F de $\mathbb{R}$. Si F contient un élément non nul a, alors pour tout $\lambda \in \mathbb{R}$, on sait que λa appartient à F. Comme n'importe quel élément x de $\mathbb{R}$ peut s'écrire λa (prendre $\lambda = x/a$), on voit que F contient $\mathbb{R}$ tout entier, donc $F = \mathbb{R}$.

D'autre part le singleton $(F = \{0\}, +, .)$ est un sous-espace vectoriel de $\mathbb{R}$: les propriétés de la définition 2.2.1 sont très faciles à prouver !

ii) L'ensemble $F = \{(x, y) \in \mathbb{R}^2,\ x + y = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$. En effet :
- — $(0, 0) \in F$.
- — Si $\lambda \in \mathbb{R}$ et $(x, y) \in F$ on a $\lambda x + \lambda y = \lambda.(x + y) = 0$ ce qui prouve que $\lambda.(x, y) = (\lambda x + \lambda y) \in F$.
- — Si (x, y) et (x', y') appartiennent à F alors $(x + x') + (y + y') = (x + y) + (x' + y') = 0 + 0 = 0$, ce qui prouve que $(x, y) + (x', y') = (x + x', y + y')$ appartient à F.

iii) Attention, l'ensemble $G = \{(x, y) \in \mathbb{R}^2,\ x + y = 1\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En fait cet ensemble ne vérifie aucune des propriétés de la définition 2.2.1.

iv) L'ensemble $\mathcal{C}^0$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$. En effet la fonction nulle est continue (comme toutes les fonctions constantes). De plus la somme de deux fonctions continues est continue, et le produit d'une fonction continue par un réel est encore une fonction continue.

v) L'ensemble $\mathcal{S}_0$ des suites de nombres réels qui convergent est un sous-espace vectoriel de $\mathcal{S}$. En effet la suite nulle est une suite qui converge (vers 0 !), la somme de deux suites convergentes est convergente (vers la somme des limites de chacune des suites), et le produit d'une suite convergente par un réel λ est encore une suite convergente (vers le produit de λ par la limite de la suite).

vi) L'ensemble $\mathcal{S}_\infty$ des suites de réels qui divergent n'est pas un sous-espace vectoriel de $\mathcal{S}$. Non seulement il ne contient pas la suite nulle, mais en plus la somme de deux suites divergentes peut être convergente.

Remarque 2.2.3 L'ensemble $F = \{0_E\}$ est un sous-espace vectoriel de n'importe quel espace vectoriel E. En effet $0_E + 0_E = 0_E$ et $\lambda.0_E = 0_E$ pour tout $\lambda \in \mathbb{R}$. On parle du sous-espace vectoriel nul.

Exercice 2.2.4 Montrer que l'ensemble des solutions d'une équation linéaire homogène à n inconnues est un sous-espace vectoriel de $\mathbb{R}^n$.

Remarque 2.2.5 Pour montrer qu'un ensemble muni de deux lois est un espace vectoriel, on a toujours intérêt à démontrer que c'est un sous-espace vectoriel d'un espace vectoriel connu qui le contient. Il n'y a alors que trois propriétés à démontrer, au lieu des huit propriétés de la définition 2.1.5.

2.3 Combinaisons linéaires

Définition 2.3.1 Soit $u_1, u_2, \ldots u_n$ des vecteurs d'un espace vectoriel $(E, +, .)$. On dit que $v \in E$ est une combinaison linéaire de $u_1, u_2, \ldots u_n$ lorsqu'il existe des réels $\lambda_1, \lambda_2, \ldots \lambda_n$ tels que

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n.$$

Il est parfois commode d'utiliser une notation plus compacte pour écrire les combinaisons linéaires, ou de manière plus générale, les sommes. Pour des vecteurs (ou des réels) u_1, u_2, ...,

u_n, on écrit

$$\sum_{j=1}^{n} u_j = u_1 + u_2 + \cdots + u_n.$$

Le symbole $\sum$ est la lettre majuscule grecque sigma (voir l'appendice B) qui a donné le "S" de "Somme" en alphabet latin. Avec cette notation la combinaison linéaire ci-dessus s'écrit

$$v = \sum_{j=1}^{n} \lambda_j u_j.$$

Exemple 2.3.2 Dans $\mathbb{R}^3$, le vecteur $u = (3,3,1)$ est-il combinaison linéaire de $v_1 = (1,1,0)$ et $v_2 = (1,1,1)$? Il s'agit de savoir s'il existe deux réels λ_1 et λ_2 tels que $\lambda_1.(1,1,0) + \lambda_2.(1,1,1) = (3,3,1)$. On est donc amené à résoudre le système

$$\left\{\begin{array}{ccccc} \lambda_1 & + & \lambda_2 & = & 3, \\ \lambda_1 & + & \lambda_2 & = & 3, \\ & & \lambda_2 & = & 1. \end{array}\right.$$

Ce système a une (unique) solution $(\lambda_1, \lambda_2) = (2,1)$, donc $u = 2v_1 + v_2$ est bien combinaison linéaire de v_1 et v_2.

Exemple 2.3.3 A quelle condition le vecteur $w = (a,b,c)$ est-il une combinaison linéaire de $u = (1,2,-1)$ et $v = (6,4,2)$? C'est le cas si et seulement si il existe deux réels λ_1 et λ_2 tels que $\lambda_1.(1,2,-1) + \lambda_2.(6,4,2) = (a,b,c)$. Il s'agit donc de savoir à quelle condition sur a, b et c le système

$$\left\{\begin{array}{ccccc} \lambda_1 & + & 6\lambda_2 & = & a, \\ 2\lambda_1 & + & 4\lambda_2 & = & b, \\ -\lambda_1 & + & 2\lambda_2 & = & c, \end{array}\right.$$

admet des solutions. On applique la méthode du pivot à ce système $(\mathbf{L_2} \leftarrow \mathbf{L_2} - \mathbf{2L_1}$, $\mathbf{L_3} \leftarrow \mathbf{L_3} + \mathbf{L_1}$, $\mathbf{L_3} \leftarrow \mathbf{L_3} + \mathbf{L_1})$, et on obtient le système équivalent

$$\left\{\begin{array}{ccccc} \lambda_1 & + & 6\lambda_2 & = & a, \\ & & -8\lambda_2 & = & b-2a, \\ & & 0 & = & c-a+b. \end{array}\right.$$

La proposition 1.3.8 dit que ce système a une solution si et seulement si le dernier second membre est nul, c'est-à-dire $c - a + b = 0$. On vient de trouver une équation cartésienne de l'ensemble des solutions du système.

2.4 Familles génératrices

Proposition 2.4.1 Soit $u_1, u_2, \dots u_n$ des vecteurs d'un espace vectoriel $(E, +, .)$. L'ensemble de toutes les combinaisons linéaires de $u_1, u_2, \dots u_n$ est un sous-espace vectoriel de E. On le note $\mathrm{Vect}(u_1, u_2, \dots u_n)$ et on parle du sous-espace vectoriel engendré par la famille $(u_1, u_2, \dots u_n)$.

Preuve.— D'abord le vecteur nul de E peut s'écrire

$$0_E = 0.u_1 + 0.u_2 + \cdots + 0.u_n,$$

donc appartient à $\mathrm{Vect}(u_1, u_2, \ldots u_n)$. Soit ensuite v et w des vecteurs de $\mathrm{Vect}(u_1, u_2, \ldots u_n)$. On peut écrire

$$v = a_1.u_1 + a_2.u_2 + \cdots + a_n.u_n \quad \text{et} \quad w = b_1.u_1 + b_2.u_2 + \cdots + b_n.u_n,$$

donc

$$v + w = a_1.u_1 + \cdots + a_n.u_n + b_1.u_1 + \cdots + b_n.u_n = (a_1 + b_1).u_1 + \cdots + (a_n + b_n).u_n,$$

ce qui prouve que $v + w \in \mathrm{Vect}(u_1, u_2, \ldots u_n)$. Enfin pour $\lambda \in \mathbb{R}$, on a

$$\lambda.v = \lambda.(a_1.u_1 + \cdots + a_n.u_n) = (\lambda a_1).u_1 + \cdots + (\lambda a_n).u_n,$$

donc $\lambda.v$ appartient aussi à $\mathrm{Vect}(u_1, u_2, \ldots u_n)$. □

Remarque 2.4.2 Le sous-espace vectoriel $\mathrm{Vect}(u_1, u_2, \ldots, u_n)$ est le plus petit (pour l'inclusion) des sous-espaces vectoriels de E qui contient $u_1, \ldots, u_n$. En effet si F est un tel sous-espace vectoriel de E, comme il est stable par combinaison linéaire, il contient toutes les combinaisons linéaires de $u_1, \ldots, u_n$, donc $\mathrm{Vect}(u_1, u_2, \ldots, u_n)$.

Définition 2.4.3 Soit $(E, +, .)$ un espace vectoriel et G un sous-espace vectoriel de E. On dit que la famille de vecteurs $(u_1, u_2, \ldots, u_n)$ engendre G, ou encore est une famille génératrice de G, lorsque $G = \mathrm{Vect}(u_1, u_2, \ldots, u_n)$.

Exemple 2.4.4 Soit $n \in \mathbb{N}$ et $\mathbb{R}_n[X]$ l'ensemble des polynômes à coefficients réels de degré inférieur ou égal à n. C'est un sous-ensemble de $\mathcal{F}$, et même un sous-espace vectoriel pour l'addition des fonctions et la multiplication d'une fonction par un réel. En effet la fonction nulle est par convention un polynôme de degré $-\infty$, donc (par abus de langage) inférieur ou égal à n, et une combinaison linéaire de polynômes de degré inférieur ou égal à n est bien un polynôme de degré inférieur ou égal à n.

Mais au fait qu'est-ce qu'un polynôme de degré inférieur ou égal à n ? C'est une combinaison linéaire des monômes $1, X, X^2, \ldots X^n$. Autrement dit $\mathbb{R}_n[X]$ est le sous-espace vectoriel de $\mathcal{F}$ engendré par la famille $(1, X, \ldots X^n)$:

$$\mathbb{R}_n[X] = \mathrm{Vect}(1, X, \ldots X^n).$$

Exemple 2.4.5 Soit $a \in \mathbb{R}$. Dans $\mathbb{R}^3$, la famille $((1, 1, 2), (-1, 0, 1), (2, 1/3, -1))$ est-elle génératrice ? Pour le savoir, on se donne un vecteur (x, y, z) de $\mathbb{R}^3$, et on cherche à l'écrire comme combinaison linéaire des trois vecteurs de la famille. Autrement dit on cherche trois réels λ_1, λ_2 et λ_3 tels que

$$\lambda_1.(1, 1, 2) + \lambda_2.(-1, 0, 1) + \lambda_3.(2, 1/3, -1) = (x, y, z).$$

Il s'agit donc de savoir si le système

$$\left\{\begin{array}{rcrcrcl} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & x, \\ \lambda_1 & & & + & \frac{1}{3}\lambda_3 & = & y, \\ 2\lambda_1 & + & \lambda_2 & - & \lambda_3 & = & z. \end{array}\right.$$

d'inconnues λ_1, λ_2 et λ_3 admet une solution.

Pour cela on utilise l'algorithme du pivot : les opérations $\mathbf{L_2} \leftarrow \mathbf{L_2} - \mathbf{L_1}$ et $\mathbf{L_3} \leftarrow \mathbf{L_3} - \mathbf{2L_1}$ donnent le système équivalent

$$\left\{\begin{array}{rcrcrcl} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & x, \\ & & \lambda_2 & - & \frac{5}{3}\lambda_3 & = & y-x, \\ & & 3\lambda_2 & - & 5\lambda_3 & = & z-2x. \end{array}\right.$$

Enfin l'opération $\mathbf{L_3} \leftarrow \mathbf{L_3} - \mathbf{3L_2}$ donne le système échelonné

$$\left\{\begin{array}{rcrcrcl} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & x, \\ & & \lambda_2 & - & \lambda_3 & = & y-x, \\ & & & & 0 & = & z-3y+x. \end{array}\right.$$

Ce système n'est compatible que si $z-3y+x=0$. Donc la famille $((1,1,2),(-1,0,1),(2,1/3,-1))$ n'est pas génératrice de $\mathbb{R}^3$: si $z-3y+x \neq 0$, le vecteur (x,y,z) de $\mathbb{R}^3$ ne s'écrit pas comme combinaison linéaire des vecteurs de cette famille.

Au passage, on récupère une équation cartésienne de $F = \mathrm{Vect}((1,1,2),(-1,0,1),(2,1/3,-1))$: $(x,y,z) \in F \iff z-3y+x=0$.

2.5 Familles libres

Définition 2.5.1 On dit que la famille $(u_1, u_2, \ldots, u_n)$ de vecteurs d'un espace vectoriel $(E,+,.)$ est libre lorsque :

$$\lambda_1 u_1 + \lambda_2 + \cdots + \lambda_n u_n = 0_E \Longrightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

Exemple 2.5.2 Dans $\mathbb{R}^n[X]$, la famille $(1, X, \ldots X^n)$ est libre. En effet supposons que

$$\lambda_0.1 + \lambda_1.X + \cdots + \lambda_n.X^n = 0.$$

Cela signifie que le polynôme $P(X) = \lambda_0 + \lambda_1 X + \cdots + \lambda_n X^n$ est la fonction nulle, donc a une infinité de racines. Ce n'est possible que si tous les coefficients sont nuls, i.e.

$$\lambda_0 = \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

Exemple 2.5.3 Dans l'espace vectoriel $\mathcal{F}$ des fonction de $\mathbb{R}$ dans $\mathbb{R}$, on considère la famille $(\cos, \sin)$ constituée des deux fonctions trigonométriques cosinus et sinus bien connues, et on se demande si elle est libre. On suppose donc que pour deux réels λ_1, λ_2, on a

$$\lambda_1. \cos + \lambda_2. \sin = 0. \tag{2.1}$$

Attention ! cette égalité est une égalité dans $\mathcal{F}$, autrement dit entre fonctions ; en particulier le "0" qui figure dans le membre de droite est la fonction nulle, pas le nombre 0. L'équation (2.1) signifie :

$$\forall x \in \mathbb{R},\ \lambda_1 \cos(x) + \lambda_2 \sin(x) = 0.$$

En particulier en prenant $x = 0$, on trouve $\lambda_1 = 0$, et pour $x = \pi/2$, on trouve $\lambda_2 = 0$. Donc la famille $(\cos, \sin)$ est libre dans $\mathcal{F}$.

Exemple 2.5.4 Soit $a \in \mathbb{R}$. Dans $\mathbb{R}^3$, la famille $((1, 1, 2), (-1, 0, 1), (2, a, -1))$ est-elle libre ? Pour le savoir, on suppose qu'il existe trois réels λ_1, λ_2 et λ_3 tels que

$$\lambda_1.(1, 1, 2) + \lambda_2.(-1, 0, 1) + \lambda_3.(2, a, -1) = 0.$$

Cela signifie que $(\lambda_1, \lambda_2, \lambda_3)$ est une solution du système

$$\left\{ \begin{array}{ccccccc} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & 0, \\ \lambda_1 & & & + & a\lambda_3 & = & 0, \\ 2\lambda_1 & + & \lambda_2 & - & \lambda_3 & = & 0. \end{array} \right. \tag{2.2}$$

Il s'agit donc de connaître l'ensemble des solutions de (2.2). Pour cela on utilise l'algorithme du pivot : les opérations $\mathbf{L_2} \leftarrow \mathbf{L_2} - \mathbf{L_1}$ et $\mathbf{L_3} \leftarrow \mathbf{L_3} - \mathbf{2L_1}$ donnent le système équivalent

$$\left\{ \begin{array}{ccccccc} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & 0, \\ & & \lambda_2 & + & (a-2)\lambda_3 & = & 0, \\ & & 3\lambda_2 & - & 5\lambda_3 & = & 0. \end{array} \right.$$

Enfin l'opération $L_3 \leftarrow L_3 - 3L_2$ donne le système échelonné

$$\left\{ \begin{array}{ccccccc} \lambda_1 & - & \lambda_2 & + & 2\lambda_3 & = & 0, \\ & & \lambda_2 & + & (a-2)\lambda_3 & = & 0, \\ & & & & (1-3a)\lambda_3 & = & 0. \end{array} \right.$$

Ou bien $a \neq 1/3$ et le système a pour unique solution $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$, ou bien $a = 1/3$ et le sytème a une infinité de solutions autres que $(0, 0, 0)$.

Conclusion : la famille $((1, 1, 2), (-1, 0, 1), (2, a, -1))$ est libre si et seulement si $a \neq 1/3$.

Remarque 2.5.5 Il est important de noter que pour déterminer si une famille est génératrice, on s'est posé la question de **l'existence de solutions** pour un système linéaire, et que d'autre part pour savoir si une famille est génératrice, il s'est agit de savoir si un système linéaire admet une **unique solution** ou bien plusieurs solutions.

2.6 Bases d'un espace vectoriel

2.6.1 Définition, exemples

Définition 2.6.1 On dit que la famille finie $\mathcal{B} = (u_1, u_2, \dots, u_n)$ de vecteurs d'un espace vectoriel $(E, +, .)$ est une base de E lorsque $\mathcal{B}$ est libre et génératrice.

Exemple 2.6.2 Dans $\mathbb{R}^n$, on note $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, …, $e_n = (0, \dots, 0, 1)$. La famille $(e_1, \dots, e_n)$ est une base, qu'on appelle base canonique de $\mathbb{R}^n$. Prouvons que cette famille est génératrice. Soit $(x_1, x_2, \dots, x_n)$ un vecteur de $\mathbb{R}^n$. On cherche $\lambda_1, \dots, \lambda_n$ dans $\mathbb{R}$ tels que

$$\lambda_1.e_1 + \dots + \lambda_n.e_n = (x_1, x_2, \dots, x_n).$$

Puisque $\lambda_1.e_1 + \dots + \lambda_n.e_n = (\lambda_1, \dots, \lambda_n)$, il suffit de prendre $\lambda_j = x_j$ pour tout $j \in \{1, \dots, n\}$. On montre maintenant que $(e_1, \dots, e_n)$ est une famille libre. Supposons que

$$\lambda_1.e_1 + \dots + \lambda_n.e_n = 0 = (0, \dots, 0).$$

Encore une fois puisque $\lambda_1.e_1 + \dots + \lambda_n.e_n = (\lambda_1, \dots, \lambda_n)$, cette égalité entraine $\lambda_1 = \dots = \lambda_n = 0$.

Exemple 2.6.3 Dans $\mathbb{R}^n[X]$, on a déjà vu que la famille $(1, X, \dots X^n)$ est génératrice (c'est la définition de $\mathbb{R}^n[X]$), et libre. C'est donc une base de $\mathbb{R}^n[X]$, appelée elle aussi base canonique, de $\mathbb{R}^n[X]$ cette fois.

2.6.2 Dimension d'un espace vectoriel

On commence par un résultat dont la preuve est un peu longue, mais qui a beaucoup de conséquences importantes.

Proposition 2.6.4 Soit $(E, +, .)$ un espace vectoriel. On suppose que E admet
- une famille génératrice $(u_1, u_2, \dots, u_n)$,
- une famille libre $(v_1, v_2, \dots v_p)$.

Alors nécessairement $p \leq n$. De plus, si $p = n$, alors $(u_1, u_2, \dots, u_n)$ et $(v_1, v_2, \dots v_p)$ sont des bases de E.

Preuve.— On veut savoir quels vecteurs de E peuvent s'écrire comme combinaison linéaire des vecteurs de $(v_1, v_2, \dots v_p)$. Il s'agit donc de déterminer pour quels vecteurs v de E l'équation d'inconnues $\lambda_1, \lambda_2 \dots, \lambda_p$,

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_p v_p = v. \tag{2.3}$$

a des solutions.

Puisque $(u_1, u_2, \ldots, u_n)$ une famille génératrice de E, il existe des réels $a_{i,j}$, $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots, n\}$ et des réels $x_1, \ldots x_n$ tels que

$$\begin{cases} v_1 &= a_{1,1}u_1 + a_{1,2}u_2 + \cdots + a_{1,n}u_n, \\ v_2 &= a_{2,1}u_1 + a_{2,2}u_2 + \cdots + a_{2,n}u_n, \\ \vdots & \qquad\qquad \vdots \\ v_p &= a_{p,1}u_1 + a_{p,2}u_2 + \cdots + a_{p,n}u_n. \end{cases}$$

et

$$v = x_1u_1 + x_2u_2 + \cdots + x_nu_n.$$

Donc l'équation s'écrit

$$\begin{aligned} x_1u_1 + x_2u_2 + \cdots + x_nu_n &= v \\ &= \lambda_1v_1 + \lambda_2v_2 + \cdots + \lambda_pv_p \\ &= \sum_{i=1}^{p} \lambda_i(a_{i,1}u_1 + a_{i,2}u_2 + \cdots + a_{i,n}u_n) \\ &= (\sum_{i=1}^{p} \lambda_i a_{i,1})u_1 + (\sum_{i=1}^{p} \lambda_i a_{i,2})u_2 + \cdots + (\sum_{i=1}^{p} \lambda_i a_{i,n})u_n \end{aligned}$$

On note alors que si le système

$$\begin{cases} a_{1,1}\lambda_1 + a_{1,2}\lambda_2 + \cdots + a_{1,n}\lambda_n &= x_1, \\ a_{2,1}u_1 + a_{2,2}u_2 + \cdots + a_{2,n}u_n &= x_2, \\ \qquad\vdots & \quad\vdots \\ a_{p,1}u_1 + a_{p,2}u_2 + \cdots + a_{p,n}u_n &= x_n, \end{cases} \tag{2.4}$$

admet une solution $(\lambda_1, \ldots, \lambda_n)$, c'est aussi une solution de l'équation (2.3). Puisque $(v_1, v_2, \ldots v_p)$ est une famille libre, l'équation (2.3) a au plus une solution : s'il y en avait deux distinctes, disons $\lambda_1, \ldots, \lambda_p$ et $\mu_1, \ldots, \mu_p$, on aurait

$$0_E = v - v = (\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_p - \mu_p)v_p,$$

donc $\lambda_i = \mu_i$ pour tout $i \in \{1, \ldots, p\}$, ce qui serait absurde.

Ainsi le système (2.4) a-t-il au plus une solution. D'après la proposition 1.3.8, cela signifie qu'il n'y a pas d'inconnues secondaires, donc le rang de ce système est p. Comme le rang d'un système est inférieur ou égal au nombre d'équations, on obtient bien

$$p \leq n.$$

Enfin si $n = p$, le système a une unique solution pour tout $(x_1, x_2, \ldots x_n)$, ce qui signifie que $(v_1, v_2, \ldots v_p)$ est une famille génératrice, donc une base. De plus la famille $(u_1, u_2, \ldots, u_n)$ est alors libre : s'il existait x_1,..., x_n non tous nuls tels que

$$x_1u_1 + \cdots + x_nu_n = 0,$$

alors la solution $(\lambda_1, \ldots, \lambda_n)$ correspondante du système (2.4) donnerait une solution non-nulle à l'équation (2.3), ce qui n'est pas possible puisque la famille $v_1, \ldots, v_p$ est libre. □

Proposition 2.6.5 Si E admet une base, alors toutes les bases de E ont le même nombre d'éléments.

Preuve.— Soit $(u_1, u_2, \dots, u_n)$ une base de E ; c'est donc une partie génératrice à n éléménts, et une partie libre à n éléménts. Si $(v_1, v_2, \dots, v_p)$ est une autre base de E, c'est une partie libre donc, d'après la proposition 2.6.4 $p \leq n$, et c'est une partie génératrice, donc, encore avec la proposition 2.6.4, $p \geq n$. □

Définition 2.6.6 Soit $(E, +, .)$ un espace vectoriel non nul. S'il existe une base de E, on appelle dimension de E, et on note $\dim E$, le nombre commun d'éléments de toutes les bases de E. Sinon on dit que E est de dimension infinie, et on note $\dim E = +\infty$.

Remarque 2.6.7 L'espace vectoriel $E = \{0_E\}$ n'admet pas de famille libre puisque toute famille de E contient 0_E. Par convention, on dit pourtant que E a pour dimension 0.

Par analogie avec la situation dans $E = \mathbb{R}^3$, les sous-espaces vectoriels de dimension 1 sont appelés "droites", et les les sous-espaces vectoriels de dimension 2 sont appelés "plans". Les sous-espaces vectoriels de dimension $n - 1$ d'un espace de dimension n sont appelés hyperplans.

Proposition 2.6.8 (Dimension et inclusion) Soit E un espace vectoriel de dimension finie. Soit V et W deux sous-espaces vectoriels de E, tels que $V \subset W$. Alors V et W sont de dimension finie et $\dim V \leq \dim W$. De plus $V = W$ si et seulement si $\dim V = \dim W$.

Preuve.— On montre d'abord qu'un sous-espace vectoriel V d'un espace E de dimension finie est de dimension finie. Si $V = \{0_E\}$, V est bien de dimension finie. Sinon V admet des familles libres. Comme $V \subset E$, ce sont aussi des familles libres de E, donc, d'après la proposition 2.6.4, elles ont au plus $n = \dim E$ éléments. Soit alors $k \leq n$ le plus grand nombre d'éléments des familles libres de V, et $(v_1, \dots, v_k)$ une famille libre de V à k éléments. Si $(v_1, \dots, v_k)$ n'est pas génératrice pour V, il existe $v \in V$ qui ne s'écrit pas comme combinaison linéaire des v_j, $j \in \{1, \dots, k\}$. Autrement dit $(v_1, \dots, v_k, v)$ est une famille libre d'éléments de v, ce qui est absurde. Donc $(v_1, \dots, v_k)$ est une base de V, et V est de dimension finie $k \leq n$.

Supposons maintenant que $V \subset W$ soient des sous-espaces vectoriels de E. On vient de prouver qu'ils sont de dimension finie, donc ils admettent des bases $(v_1, \dots v_k)$ et $(w_1, \dots, w_\ell)$. Comme $(v_1, \dots v_k)$ est une partie libre de V, c'est une partie libre de W. Puisque $(w_1, \dots, w_\ell)$ est une partie génératrice de W, la proposition 2.6.4 donne $k \leq \ell$. Si $k = \ell$, c'est-à-dire $\dim V = \dim W$, la proposition 2.6.4 donne aussi $V = W$. □

2.6.3 Comment extraire une base d'une famille génératrice ?

Exemple 2.6.9 Soit $v_1 = (1, 2, 3, 0)$, $v_2 = (0, 2, 2, 4)$ et $v_3 = (-1, a, 2a - 3, 4a)$ trois vecteurs de $\mathbb{R}^4$, et $F = \mathrm{Vect}(v_1, v_2, v_3)$. Par définition, la famille (v_1, v_2, v_3) est une famille génératrice de

F, et on veut trouver une base de F.

Si la famille (v_1, v_2, v_3) est libre, c'est une base. Or l'équation

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0,$$

équivaut au système

$$\left\{\begin{array}{ccccccc} \lambda_1 & & & - & \lambda_3 & = & 0, \\ 2\lambda_1 & + & 2\lambda_2 & + & a\lambda_3 & = & 0, \\ 3\lambda_1 & + & 2\lambda_2 & + & (2a-3)\lambda_3 & = & 0, \\ & & 4\lambda_2 & + & 4a\lambda_3 & = & 0. \end{array}\right.$$

Avec l'algorithme du pivot, on obtient le système échelonné équivalent :

$$\left\{\begin{array}{ccccccc} \lambda_1 & & & - & \lambda_3 & = & 0, \\ & & 2\lambda_2 & + & (a+2)\lambda_3 & = & 0, \\ & & & & (a-2)\lambda_3 & = & 0. \end{array}\right.$$

Donc si $a \neq 2$, le système est de rang 3, et admet une unique solution $(0, 0, 0)$. Donc (v_1, v_2, v_3) est une base dans ce cas.

Si $a = 2$, le système est de rang 2 et admet une infinité de solutions, donc (v_1, v_2, v_3) n'est pas une famille libre. Mais on peut dire plus : λ_3 est la seule inconnue secondaire, et le système peut s'écrire

$$\left\{\begin{array}{ccccc} \lambda_1 & & & = & s, \\ & \lambda_2 & & = & -2s, \\ & & \lambda_3 & = & s. \end{array}\right.$$

L'ensemble des solutions est $\mathrm{Vect}((1, -2, 1))$, et en particulier $v_1 - v_2 + v_3 = 0$, ou encore $v_3 = -v_1 + 2v_2$. Donc (v_1, v_2) est une famille génératrice de F : si $v \in F$, il existe $\lambda_1, \lambda_3, \lambda_3$ tels que $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. Mais alors

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3(-v_1 + 2v_2) = (\lambda_1 - 3\lambda_3)v_1 + (\lambda_2 + 2\lambda_3)v_2,$$

ce qui prouve que v s'écrit comme combinaison linéaire de v_1 et v_2. Enfin (v_1, v_2) est une famille libre de F : dans le cas $s = 0$, la seule solution du système ci-dessus est $\lambda_1 = \lambda_2 = 0$.

Dans cet exemple, on a obtenu une base en gardant les vecteurs correspondants aux inconnues principales. C'est un fait général :

Proposition 2.6.10 Soit $V = \mathrm{Vect}(v_1, \dots v_n)$ un sous-espace vectoriel de $\mathbb{R}^p$. Pour extraire de la famille génératrice $(v_1, \dots v_n)$ une base de V, il suffit d'échelonner le système correspondant à l'équation

$$x_1 v_1 + \cdots + x_n v_n = 0.$$

Les vecteurs $v_{j_1}, \dots, v_{j_r}$ correspondant aux inconnues principales constituent alors une base de V. En particulier la dimension de V est égale au rang r du système.

Preuve.— Si l'on donne la valeur -1 à l'une des inconnues secondaires λ_{j_0}, et la valeur 0 à toutes les autres, on trouve une solution $\lambda_{j_1}, \lambda_{j_2}, \dots, \lambda_{j_r}$ du système, de sorte que

$$v_{j_0} = \lambda_{j_1} v_{j_1} + \cdots + \lambda_{j_r} v_{j_r}.$$

Donc la famille $(v_{j_1}, \dots v_{j_r})$, obtenue en retirant les vecteurs correspondants aux inconnues secondaires reste une famille génératrice.

Ensuite, si l'on donne la valeur 0 à toutes les inconnues secondaires, le système restant admet $(0, \dots, 0)$ comme seule solution : la famille $(v_{j_1}, \dots v_{j_r})$ est libre. □

Proposition 2.6.11 (Théorème de la base incomplète) Soit $(E, +, .)$ un espace vectoriel de dimension finie. Toute famille libre de E peut être complétée en une base de E en lui adjoignant des vecteurs d'une base de E.

Preuve.— Soit $(v_1, v_2, \dots v_k)$ une famille libre de E, et $(w_1, \dots, w_n)$ une base de E. On considère la famille $\mathcal{F} = (v_1, v_2, \dots v_k, w_1, \dots, w_n)$. Elle est génératrice puisqu'elle contient la sous-famille génératrice $(w_1, \dots, w_n)$. On considère alors le système homogène de n équations à $k+n$ inconnues dont les coefficients sont les vecteurs de $\mathcal{F}$:

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k + y_1 w_1 + \cdots + y_n w_n = 0_E,$$

et on l'échelonne avec la méthode du pivot, de manière à obtenir une base de E. Comme les v_j sont libres, les inconnues correspondantes sont principales, donc la base obtenue contient les v_j. □

2.6.4 Comment trouver une base d'un espace vectoriel donné par un système d'équations ?

On sait que l'ensemble des solutions d'un système homogène de n équations à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. On se pose la question d'en trouver une base, et de déterminer sa dimension.

Exemple 2.6.12 Soit $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3,\ 2x_1 + 3x_2 = 0, x_1 + x_2 + x_3 = 0\}$. V est l'ensemble des solutions du système

$$\left\{ \begin{array}{ccccccc} 2x_1 & + & 3x_2 & & & = & 0, \\ x_1 & + & x_2 & + & x_3 & = & 0. \end{array} \right.$$

On l'échelonne avec la méthode du pivot. On obtient $(L_2 \leftarrow L_2 - \frac{1}{2} L_1)$ le système équivalent

$$\left\{ \begin{array}{ccccccc} 2x_1 & + & 3x_2 & & & = & 0, \\ & - & \frac{1}{2} x_2 & + & x_3 & = & 0. \end{array} \right.$$

Les inconnues principales sont x_1 et x_2. On peut les exprimer en fonction de l'inconnue secondaire x_3 en écrivant le système sous forme échelonnée réduite $(L_1 \leftarrow L_1 + 6L_2,\ L_1 \leftarrow \frac{1}{2} L_1,\ L_2 \leftarrow -2L_2)$:

$$\left\{ \begin{array}{ccccc} x_1 & & + & 3x_3 & = & 0, \\ & x_2 & - & 2x_3 & = & 0. \end{array} \right.$$

Autrement dit, $V = \{(-3s, 2s, s),\ s \in \mathbb{R}\}$, ou encore $V = \mathrm{Vect}((-3, 2, 1))$

De manière générale, on a la

Proposition 2.6.13 Soit V un sous-espace vectoriel de $\mathbb{R}^p$ donné comme l'ensemble des solutions d'un système homogène à p inconnues. Les vecteurs dont les coordonnées sont les coefficients des inconnues secondaires dans la forme échelonnée réduite du système forment une base de V. En particulier, la dimension de V est $p - r$, où r est le rang du système.

2.7 Intersection et somme de sous espaces vectoriels

2.7.1 Intersection

On sait que l'ensemble des solutions d'un système linéaire homogène à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. Cela vaut quel que soit le nombre d'équations : en particulier l'ensemble des solutions d'une équation linéaire homogène à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. La résolution d'un système de plusieurs équations consiste donc à trouver l'intersection des espaces vectoriels définis par chacune. Dans le cas de $\mathbb{R}^p$, on a donc pratiquement déjà prouvé la

Proposition 2.7.1 Soit $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. Leur intersection de $V \cap W$ est un sous-espace vectoriel de E.

Preuve.— Voici une preuve directe. Puisque V est un sous-espace vectoriel, $0_E \in V$. De même $0_E \in W$, donc $0_E \in V \cap W$. Il reste à montrer que $V \cap W$ est stable par combinaison linéaire. Soit donc u_1 et u_2 deux vecteurs de $V \cap W$, et λ_1, λ_2 deux réels. Puisque u_1 et u_2 appartiennent à V, $\lambda_1 u_1 + \lambda_2 u_2$ appartient à V. De même $\lambda_1 u_1 + \lambda_2 u_2$ appartient à W, donc finalement $\lambda_1 u_1 + \lambda_2 u_2$ appartient à $V \cap W$. $\square$

Proposition 2.7.2 Soit $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. On a

$$\dim(V \cap W) \leq \min(\dim V, \dim W).$$

De plus s'il y a égalité, c'est que l'un des sous-espaces est inclus dans l'autre.

Preuve.— Puisque $V \cap W \subset V$ et $V \cap W \subset W$, l'inégalité découle directement de la proposition 2.6.8. Supposons que $\dim(V \cap W) = \min(\dim V, \dim W)$, par exemple que $\dim(V \cap W) = \dim V$. Comme on a $V \cap W \subset V$, la proposition 2.6.8 donne aussi $V \cap W = V$, donc $V \subset W$. $\square$

2.7.2 Somme

Définition 2.7.3 Soit $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. On appelle somme de V et W, et on note $V + W$ le sous-ensemble de E défini par

$$V + W = \{x + y, x \in V, y \in W\}.$$

Proposition 2.7.4 Si V, W deux sous-espaces vectoriels de E, alors $V + W$ est aussi un sous-espace vectoriel de E. De plus si F est un sous-espace vectoriel qui contient V et W, alors F contient $V + W$.

Preuve.— Puisque $0_E \in V$, $0_E \in W$ et $0_E = 0_E + 0_E$, le vecteur nul 0_E s'écrit bien comme somme d'un élément de V et d'un élément de W, donc appartient à $V + W$. Soit ensuite u_1 et u_2 deux éléments de $V + W$. Il existe x_1, x_2 dans V et y_1, y_2 dans W tels que

$$\begin{cases} u_1 = x_1 + y_1, \\ u_2 = x_2 + y_2. \end{cases}$$

Ainsi

$$u_1 + u_2 = x_1 + y_1 + x_2 + y_2 = (x_1 + x_2) + (y_1 + y_2)$$

et $u_1 + u_2 \in V + W$. Enfin pour $\lambda \in \mathbb{R}$ et $u = x + y \in V + W$, on a $\lambda u = \lambda x + \lambda y$ donc $\lambda u \in V + W$. Donc $V + W$ est bien un sous-espace vectoriel de E.

Supposons maintenant que F soit un un sous-espace vectoriel qui contient V et W. Puisque F est stable pour l'addition, pour n'importe quel $x \in V$ et n'importe quel $y \in W$, on a $x \in F$ et $y \in F$ donc $x + y \in F$. Cela signifie que $V + W \subset F$. $\square$

Exemple 2.7.5 Dans $E = \mathbb{R}^3$, on note $a = (0, 1, 2)$, $b = (1, 0, 1)$ et $c = (1, 1, 1)$. Soit aussi $V = \mathrm{Vect}(a, b)$ et $W = \mathrm{Vect}(c)$. L'espace vectoriel $V + W$ est constitué des vecteurs v qui s'écrivent

$$v = (\lambda a + \mu b) + \nu c = \lambda a + \mu b + \nu c.$$

Donc $V + W$ est engendré par la famille (a, b, c). D'ailleurs puisque cette famille est libre, $V + W = \mathbb{R}^3$: l'espace $\mathbb{R}^3$ s'écrit comme somme du plan V et de la droite W.

Proposition 2.7.6 Si $V = \mathrm{Vect}(v_1, v_2, \dots v_k)$ et $W = \mathrm{Vect}(w_1, w_2, \dots w_\ell)$, alors $V + W = \mathrm{Vect}(v_1, v_2, \dots v_k, w_1, w_2, \dots w_\ell)$. En particulier, $\dim(V + W) \leq \dim V + \dim W$.

Preuve.— Notons $Y = \mathrm{Vect}(v_1, v_2, \dots v_k, w_1, w_2, \dots w_\ell)$. On va montrer successivement les deux inclusions $V + W \subset Y$ et $Y \subset V + W$.

— $V + W \subset Y$: en effet Y contient $v_1, v_2, \dots v_k$, donc contient V. De même Y contient W, et donc Y contient $V + W$ d'après la proposition 2.7.4.

— $Y \subset V + W$: en effet $V + W$ contient V donc les v_j, et contient W donc les w_j. Donc $E + F$ contient la famille $(v_1, v_2, \dots v_k, w_1, w_2, \dots w_\ell)$, donc le sous-espace qu'elle engendre, c'est à dire Y.

□

2.7.3 Somme directe

Proposition 2.7.7 Soit $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. Tout vecteur y de $V + W$ s'écrit de manière unique $y = v + w$ avec $v \in V$ et $w \in W$ si et seulement si $V \cap W = \{0_E\}$.

Preuve.— Soit $y \in V + W$. Supposons qu'il existe $v_1, v_2 \in V$ et $w_1, w_2 \in W$ tels que $y = v_1 + w_1$ et $y = x_2 + w_2$. On soustrayant ces deux égalités on obtient

$$0_E = (v_1 + w_1) - (v_2 + w_2) = (v_1 - v_2) + (w_1 - w_2)$$

et donc $v_1 - v_2 = w_2 - w_1$. Dans cette égalité, le membre de gauche appartient à V, et le membre de droite à W. Donc les deux vecteurs appartiennent à $V \cap W = \{0_E\}$. Ainsi $v_1 = v_2$ et $w_1 = w_2$. On a bien montré que la décomposition d'un élément de $V + W$ en somme d'un élément de V et d'un élément de W est unique.

Réciproquement, supposons que $x \in V \cap W$. Par hypothèse, on peut écrire de manière unique $x = v + w$ avec $v \in V$ et $w \in W$. Comme $x \in V$ et $0_E \in W$, cette écriture est nécessairement $x = x + 0_E$, c'est-à dire que $w = 0_E$. De même on a aussi $x = 0_E + x$ donc $v = 0_E$ et $x = 0_E + 0_E = 0_E$. □

Définition 2.7.8 Soit $(E, +, .)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E.

i) On dit que V et W sont en somme directe lorsque $V \cap W = \{0_E\}$. Dans ce cas on note $F = V \oplus W$ l'espace $F = V + W$.

ii) Si de plus $V \oplus W = E$, on dit que V et W sont supplémentaires dans E.

2.7.4 Formule de la dimension

Proposition 2.7.9 Soit $(E, +, .)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E. On a toujours

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

Preuve.— Soit $\mathcal{B}_0 = (u_1, u_2, \dots u_k)$ une base de l'espace vectoriel $V \cap W$. On note au passage que $\dim V \cap W = k$. Les vecteurs de $\mathcal{B}_0$ forment un système libre de V. D'après le théorème de la base incomplète, il existe des vecteurs $v_1, \dots, v_\ell$ de V tels que $\mathcal{B}_V = (u_1, u_2, \dots u_k, v_1, \dots v_\ell)$ soit une base de V. Encore une fois, au passage, on a $\dim V = k + \ell$.

De même, il existe des vecteurs $w_1, \dots, w_m$ de W tels que $\mathcal{B}_W = (u_1, u_2, \dots u_k, w_1, \dots w_m)$ soit une base de W. On a encore $\dim W = k + m$.

D'après la proposition 2.7.6, la famille

$$(u_1, u_2, \dots u_k, v_1, \dots v_\ell, u_1, u_2, \dots u_k, w_1, \dots, w_m)$$

engendre $V + W$. Bien sûr, c'est aussi le cas de $\mathcal{B} = (u_1, u_2, \dots u_k, v_1, \dots v_\ell, w_1, \dots, w_m)$.

Montrons que $\mathcal{B}$ est libre. Pour cela on suppose que

$$\lambda_1 u_1 + \dots \lambda_k u_k + \mu_1 v_1 + \dots + \mu_\ell v_\ell + \nu_1 w_1 + \dots + \nu_m w_m = 0_E, \tag{2.5}$$

et on note $v = \lambda_1 u_1 + \dots \lambda_k u_k + \mu_1 v_1 + \dots + \mu_\ell v_\ell \in V$ et $w = -(\nu_1 w_1 + \dots + \nu_m w_m) \in W$. On a $v = w$, donc les deux sont dans $V \cap W$. Comme $\mathcal{B}_0$ est une base de $V \cap W$, il existe $\lambda'_1, \dots, \lambda'_k$ tels que

$$v = w = \lambda'_1 u_1 + \dots + \lambda'_k u_k.$$

Du coup $v - (\lambda'_1 u_1 + \dots + \lambda'_k u_k) = 0_E$, c'est-à-dire

$$(\lambda_1 - \lambda'_1) u_1 + \dots (\lambda_k - \lambda'_k) u_k + \mu_1 v_1 + \dots + \mu_\ell v_\ell = 0_E,$$

et comme $\mathcal{B}_V$ est une base de V, tous les coefficients sont nuls, en particulier $\mu_1 = \dots = \mu_l = 0$. De la même manière, en partant de $w - (\lambda'_1 u_1 + \dots + \lambda'_k u_k) = 0_E$, on obtient $\nu_1 = \dots = \nu_m = 0$. Du coup (2.5) devient

$$\lambda_1 u_1 + \dots \lambda_k u_k = 0_E,$$

ce qui entraine $\lambda_1 = \dots = \lambda_k = 0$.

On a donc prouvé que $\mathcal{B}$ est une base de $V + W$, et

$$\dim(V + W) = k + \ell + m = (k + \ell) + (k + m) - k = \dim V + \dim W - \dim(V \cap W).$$

$\square$

Chapitre 3

Matrices

3.1 L'espace vectoriel des matrices

Définition 3.1.1 Une matrice à n lignes et p colonnes est un tableau de nombres de la forme

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix}$$

Les nombres a_{ij} sont appelés coefficients de la matrice A, et on note souvent $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ ou même $A = (a_{ij})$ quand il n'y a pas de confusion possible. Les nombres n et p sont appelés dimensions de la matrice.

On note $\mathcal{M}_{n,p}(\mathbb{R})$ l'ensemble des matrices à n lignes et p colonnes de nombres réels. Si les dimensions d'une matrice sont égales, i.e. s'il y a autant de lignes que de colonnes, on dit que la matrice est carrée, et on note simplement $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices à n lignes et n colonnes.

Exemple 3.1.2 *i)* On appelle matrice nulle la matrice dont tous les coefficients sont nuls :

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

ii) La matrice d'un système linéaire à n équations et p inconnues

$$\left\{\begin{array}{ccc} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p & = & b_1 \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p & = & b_n \end{array}\right.$$

est le tableau à n lignes et p colonnes dont les coefficients sont les a_{ij} :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix}$$

La matrice complète du système est la matrice à n lignes et $p+1$ colonnes définie par

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1p} & b_1 \\ a_{21} & a_{22} & \dots & a_{2p} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} & b_p \end{array}\right)$$

On notera la présence de la barre verticale pour séparer les seconds membres du reste de la matrice. Celle-ci n'est pas obligatoire, mais souvent pratique.

iii) On appelle matrice élémentaire toute matrice dont tous les coefficients sont nuls sauf l'un d'entre eux qui vaut 1. On note

$$E_{ij} = \begin{pmatrix} & & \vdots & \\ & & \vdots & \\ \dots & \dots & 1 & \dots \\ & & \vdots & \end{pmatrix}$$

la matrice élémentaire dont l'élément non nul est sur la ligne i et la colonne j.

iv) Lorsqu'une matrice n'a qu'une ligne, on parle de matrice-ligne. Lorsqu'elle n'a qu'une colonne, on parle de matrice-colonne. Par exemple on identifie souvent un vecteur v de $\mathbb{R}^n$ avec la matrice colonne de ses coordonnées dans la base canonique $(e_1, e_2, \dots, e_n)$ de $\mathbb{R}^n$: si $v = (x_1, x_2, \dots x_n)$, on a $v = x_1e_1 + x_2e_2 + \dots x_ne_n$ et on écrit

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Définition 3.1.3 Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de $\mathcal{M}_{n,p}(\mathbb{R})$. On note $C = A + B$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par $C = (a_{ij} + b_{ij})$. Plus explicitement

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n,1} & b_{n2} & \dots & b_{n,p} \end{pmatrix} =$$
$$\begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1p} + b_{1p} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2p} + b_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} + b_{n,1} & a_{n2} + b_{n2} & \dots & a_{n,p} + b_{n,p} \end{pmatrix}$$

On a ainsi défini une loi de composition interne "$+$" sur l'ensemble $\mathcal{M}_{n,p}(\mathbb{R})$. Comme il s'agit d'additionner les coefficients terme à terme (donc de faire des additions de nombres réels), il est très facile de vérifier que cette loi "$+$" vérifie les quatre premières propriétés de la définition 2.1.5. En particulier l'élément neutre pour l'addition des matrices est la matrice nulle, et la matrice symétrique d'une matrice $A = (a_{ij})$ est la matrice, notée $-A$, définie par

$$-A = \begin{pmatrix} -a_{11} & -a_{12} & \dots & -a_{1p} \\ -a_{21} & -a_{22} & \dots & -a_{2p} \\ \vdots & \vdots & & \vdots \\ -a_{n,1} & -a_{n2} & \dots & -a_{n,p} \end{pmatrix}$$

Pour résumer :

Proposition 3.1.4 $(\mathcal{M}_{n,p}(\mathbb{R}), +)$ est un groupe commutatif.

On définit maintenant une loi de composition externe "." sur l'ensemble $\mathcal{M}_{n,p}(\mathbb{R})$.

Définition 3.1.5 Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ et λ un réel. On note $C = \lambda.A$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par $C = (\lambda a_{ij})$. Plus explicitement

$$\lambda \,.\, \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1p} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2p} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n,1} & \lambda a_{n2} & \dots & \lambda a_{n,p} \end{pmatrix}$$

Comme il s'agit de multiplier chaque coefficient par λ, il est également facile de vérifier que cette loi "." vérifie les 4 dernières propriétés de la définition 2.1.5. On a donc prouvé la première partie de la

Proposition 3.1.6 $(\mathcal{M}_{n,p}(\mathbb{R}), +)$ est un espace vectoriel, de dimension np.

Preuve.— Il reste à montrer que la dimension est np. Or l'ensemble des matrices élémentaires $(E_{ij}, 1 \leq i \leq n, 1 \leq j \leq p)$ est une base de $\mathcal{M}_{n,p}(\mathbb{R})$: si $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{R})$, on a

$$A = \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq p} a_{ij} E_{ij},$$

donc c'est une famille génératrice. De plus

$$\sum_{1 \leq i \leq n} \sum_{1 \leq j \leq p} \lambda_{ij} E_{ij} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1p} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2p} \\ \vdots & \vdots & & \vdots \\ \lambda_{n,1} & \lambda_{n2} & \dots & \lambda_{n,p} \end{pmatrix}$$

Donc la combinaison linéaire $\sum_{1 \leq i \leq n} \sum_{1 \leq j \leq p} \lambda_{ij} E_{ij}$ est la matrice nulle si et seulement si tous les coefficients λ_{ij} sont nuls. □

Remarque 3.1.7 On notera que la famille $(E_{1,1},\ E_{1,2},\ \dots,\ E_{n,p-1},\ E_{n,p})$ n'est pas indéxée par une partie de $\mathbb{N}$ mais par une partie de $\mathbb{N} \times \mathbb{N}$. Ce n'est qu'une façon de l'écrire : l'énumération $(E_{1,1},\ E_{1,2},\ \dots,\ E_{n,p-1},\ E_{n,p})$ fournit une nouvelle numérotation de la famille pour des indices entre 1 et np. De manière générale n'importe quel ensemble fini peut servir à énumerer une base.

3.2 Produit de matrices

3.2.1 Définition

On définit maintenant le produit de deux matrices. Il est naturel de penser multiplier terme à terme les coefficients de chacune, mais cette multiplication là ne sert à rien. Ce que l'on veut, c'est retrouver le membre de gauche des équations du système

$$\left\{ \begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p & = & b_1 \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p & = & b_n \end{array} \right.$$

en multipliant sa matrice par la matrice colonne des symboles $X_1, X_2, \dots X_p$. Autrement dit, on veut que

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix} \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\ \vdots \\ a_{n,1}X_1 + a_{n2}X_2 + \cdots + a_{n,p}X_p \end{pmatrix},$$

et c'est cela que l'on prend comme définition du produit d'une matrice par une matrice-colonne. Il est essentiel de noter que le nombre de lignes de la matrice colonne doit être égal au nombre de colonnes de la matrice.

On étend ensuite cette définition au cas du produit d'une matrice A par une matrice B quelconque, en construisant la matrice dont les colonnes sont les produits successifs de A par chacune des colonnes de B. Le produit de deux matrices se fait donc "ligne par colonne".

Exemple 3.2.1 Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 3 \end{pmatrix}$. Le produit $A \times B$ est bien défini puisque A a trois colonnes et B trois lignes. Comme A a deux lignes et B a 2 colonnes, le produit $A \times B$ sera une matrice à 2 lignes et 2 colonnes.

Le produit de A par la première colonne de B est

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1*1 & + & 2*0 & + & 3*2 \\ 0*1 & + & (-1)*0 & + & 1*2 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

Celui de A par la seconde colonne de B est

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1*(-1) & + & 2*1 & + & 3*3 \\ 0*(-1) & + & (-1)*1 & + & 1*3 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix}$$

On obtient donc

$$A \times B = \begin{pmatrix} 7 & 10 \\ 2 & 2 \end{pmatrix}.$$

On peut formaliser cette définition. C'est l'objet de la

Définition 3.2.2 Soit $A = (a_{ij})$ une matrice à n lignes et **p colonnes**, et $B = (b_{ij})$ une matrice à **p lignes** et m colonnes. On note $C = A \times B$ la matrice à n lignes et m colonnes dont les coefficients c_{ij} sont donnés par

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ip} b_{pj}.$$

Remarque 3.2.3 Le produit de deux matrices n'est pas commutatif. Autrement dit, en général $A \times B \neq B \times A$. Par exemple

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

D'ailleurs l'un des produit peut être bien défini alors que l'autre ne l'est pas en fonction des dimensions de A et de B.

Dans la pratique, il est commode pour calculer le produit $C = A \times B$ de deux matrices de disposer les calculs de la manière suivante

$$\begin{array}{c|c}
 & \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1m} \\ b_{21} & \dots & b_{2j} & \dots & b_{2m} \\ \vdots & & \vdots & & \vdots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pm} \end{pmatrix} \\
\hline
\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} \end{pmatrix} & \begin{pmatrix} & \vdots & \\ \dots & \displaystyle\sum_{k=1}^{p} a_{ik} b_{kj} & \dots \\ & \vdots & \end{pmatrix}
\end{array}$$

Proposition 3.2.4 Le produit des matrices est associatif : si $A \in \mathcal{M}_{mn}(\mathbb{R})$, $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et $C \in \mathcal{M}_{pq}(\mathbb{R})$, on a

$$A \times (B \times C) = (A \times B) \times C.$$

Preuve.— On note $M = (m_{ij}) = A \times (B \times C)$, et $N = (n_{ij}) = (A \times B) \times C$. Soit aussi $D = (d_{ij}) = B \times C \in \mathcal{M}_{nq}(\mathbb{R})$. On sait que

$$m_{ij} = \sum_{k=1}^{n} a_{ik} d_{kj} = \sum_{k=1}^{n} a_{ik} \big(\sum_{\ell=1}^{p} b_{k\ell} c_{\ell j} \big).$$

La distributivité de ma multiplication dans $\mathbb{R}$ par rapport à l'addition donne

$$m_{ij} = \sum_{k=1}^{n} \sum_{\ell=1}^{p} a_{ik} b_{k\ell} c_{\ell j}.$$

De la même manière, on trouve

$$n_{ij} = \sum_{\ell=1}^{p} \big(\sum_{k=1}^{n} a_{ik} b_{k\ell} \big) c_{\ell j} = \sum_{\ell=1}^{p} \sum_{k=1}^{n} a_{ik} b_{k\ell} c_{\ell j},$$

ce qui montre que $m_{ij} = n_{ij}$. □

Au passage on note que pour $\lambda \in \mathbb{R}$, $\lambda A = (\lambda I) \times A$ et

$$A \times (\lambda B) = A \times ((\lambda I) \times B) = (A \times (\lambda I)) \times B = ((\lambda I) \times A) \times B = \lambda A \times B.$$

Proposition 3.2.5 Le produit des matrices est distributif par rapport à l'addition, à gauche et à droite : si $A \in \mathcal{M}_{n,p}(\mathbb{R})$, $B \in \mathcal{M}_{p,m}(\mathbb{R})$ et $C \in \mathcal{M}_{p,m}(\mathbb{R})$, on a

$$A \times (B + C) = A \times B + A \times C.$$

et, si $B \in \mathcal{M}_{n,p}(\mathbb{R})$, $C \in \mathcal{M}_{n,p}(\mathbb{R})$ et $A \in \mathcal{M}_{p,m}(\mathbb{R})$,

$$(B + C) \times A = B \times A + C \times A.$$

3.2.2 Produit par une matrice élémentaire

Proposition 3.2.6 Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$, et $E_{k\ell} = (e_{ij}) \in \mathcal{M}_{n,n}((\mathbb{R})$ la matrice élémentaire dont le seul élément non nul est $e_{k\ell} = 1$. On a

$$E_{kl} \times A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ a_{\ell 1} & \dots & a_{\ell p} \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{ligne } k \\ \\ \\ \\ \\ \end{array}$$

Preuve.— Notant $C = (c_{ij}) = E_{kl} \times A$, on sait que

$$c_{ij} = \sum_{m=1}^{n} e_{im} a_{mj}.$$

Donc pour $i \neq k$, $c_{ij} = 0$ puisque tous les e_{im} sont nuls dans ce cas. Ensuite

$$c_{kj} = \sum_{m=1}^{n} e_{km} a_{mj} = e_{k\ell} a_{\ell j} = a_{\ell j}.$$

□

Ainsi, multiplier une matrice A par $E_{k\ell}$ permet d'extraire la ℓ-ième ligne de A et de la placer sur la ligne k. En particulier

Remarque 3.2.7 L'opération $L_i \leftarrow L_i + \alpha L_j$, que l'on peut effectuer sur une matrice comme sur un système, correspond à remplacer A par $A + \alpha E_{ij} \times A$. L'opération $L_i \leftarrow \alpha L_i$ correspond à remplacer A par $\alpha E_{ii} \times A$

3.3 Inverse d'une matrice carrée

3.3.1 La matrice identité

Dans les espaces de matrices carrées $\mathcal{M}_n(\mathbb{R})$, la multiplication est une loi interne. Ce n'est bien sûr pas le cas dans $\mathcal{M}_{n,p}(\mathbb{R})$ quand $n \neq p$: la multiplication de deux élément n'est alors même pas définie. Dans $\mathcal{M}_n(\mathbb{R})$, la multiplication a un élément neutre :

Proposition 3.3.1 Soit $I \in \mathcal{M}_n(\mathbb{R})$ la matrice dont les éléments sont tous nuls, sauf ceux qui se trouvent sur la diagonale qui valent 1 :

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

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, on a

$$A \times I = I \times A = A.$$

Cette matrice I est appelée matrice identité. On se rappelle qu'il ne peut y avoir qu'un seul élément neutre pour une loi interne (cf la remarque 2.1.6).

Exercice 3.3.2 Soit J la matrice carrée dont seuls les éléments antidiagonaux ne sont pas nuls et valent 1, c'est à dire

$$J = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \iddots & & \vdots \\ 1 & \dots & 0 & 0 \end{pmatrix}$$

Calculer $A \times J$.

3.3.2 Definition

On se pose maintenant la question de l'existence pour une matrice $A \in \mathcal{M}_n(\mathbb{R})$ donnée, d'un symétrique pour la loi $\times$.

Définition 3.3.3 Soit $A \in \mathcal{M}_n(\mathbb{R})$. On dit que A admet la matrice $Q \in \mathcal{M}_n(\mathbb{R})$ pour inverse lorsque

$$A \times Q = Q \times A = I.$$

Par exemple la matrice I admet elle-même pour inverse puisque $I \times I = I$.

On a déjà montré dans la remarque 2.1.6 qu'un élément ne peut avoir plus d'un symétrique pour une loi interne associative. Ayant en vue la Proposition 3.3.5 ci-dessous, on reprend la preuve dans le cadre actuel.

Proposition 3.3.4 Une matrice A admet au plus une matrice inverse.

Preuve.— Supposons que A admette Q_1 et Q_2 pour inverses. On a

$$Q_2 = Q_2 \times I = Q_2 \times (A \times Q_1) = (Q_2 \times A) \times Q_1 = I \times Q_1 = Q_1.$$

□

Dans cette preuve, on a seulement utilisé que Q_1 est inverse à droite de A, c'est-à-dire $A \times Q_1 = I$ et que Q_2 est inverse à gauche de A, c'est-à-dire $Q_2 \times A = I$. On a donc démontré la

Proposition 3.3.5 Si A admet une matrice inverse à droite Q_1 et une matrice inverse à gauche Q_2, alors $Q_1 = Q_2$ et A admet $Q_1 = Q_2$ comme inverse.

Lorsqu'une matrice A admet une matrice inverse, on dit que A est inversible, et note A^{-1} la matrice inverse de A. Reste à trouver un moyen de savoir si une matrice donnée est inversible, et le cas échéant, de calculer son inverse.

3.3.3 Existence

Une manière de voir ce problème est de considérer le système linéaire $A \times X = B$ associé à la matrice A avec un second membre $B \in \mathcal{M}_{n,1}(\mathbb{R})$ quelconque, c'est-à-dire l'équation d'inconnue X qui s'écrit

$$A \times X = B.$$

S'il existe Q telle que $Q \times A = I$, cette équation entraine

$$X = (Q \times A) \times X = Q \times (A \times X) = Q \times B,$$

donc la seule solution possible est $X = QB$. Si on a aussi $A \times Q = I$, alors

$$A \times (Q \times B) = (A \times Q) \times B = B,$$

ce qui prouve que $X = QB$ est bien une solution du système.

Autrement dit : si A est inversible, pour tout second membre $B \in \mathcal{M}_{n,1}(\mathbb{R})$ le système de n équations à n inconnues $AX = B$ a une unique solution.

Examinons la réciproque : supposons que pour tout second membre $B \in \mathcal{M}_{n,1}(\mathbb{R})$ le système $AX = B$ a une unique solution. C'est un particulier le cas si B est (le vecteur colonne associé à) l'un des vecteurs B_j de la base canonique, c'est à dire l'un des vecteurs

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

On note Q_1, Q_2,. . .,Q_n la solution correspondante du système, c'est-à dire l'unique vecteur tel que

$$AQ_j = B_j,$$

et Q la matrice dont les colonnes sont $Q_1, Q_2, \ldots, Q_n$. Par définition du produit de A par Q, on a

$$A \times Q = I,$$

donc Q est inverse à droite de A.

D'autre part, puisque le système $AX = B$ a une unique solution pour n'importe quel second membre B, toutes ses inconnues sont principales, donc sa forme échelonnée réduite est

$$I \times X = C.$$

On passe du système $AX = B$ à sa forme échelonnée réduite en effectuant des opérations du type $L_i \leftarrow L_i + \alpha L_j$ ou $L_i \leftarrow \alpha L_i$, c'est-à-dire compte tenu de la remarque 3.2.7, en multipliant A par des matrices de la forme $(I + \alpha E_{ij})$. Notant R le produit de toutes ces matrices, on a,

$$R \times A = I,$$

donc R est un inverse à gauche de A. On a démontré que si pour tout second membre $B \in \mathcal{M}_{n,1}(\mathbb{R})$ le système $AX = B$ a une unique solution, alors A est inversible. En résumé, on a obtenu la

Proposition 3.3.6 Soit $A \in \mathcal{M}_n(\mathbb{R})$. Les conditions suivantes sont équivalentes :

i) La matrice A est inversible.

ii) Le système $AX = B$ a une unique solution pour tout second membre $B \in \mathcal{M}_{n,1}(\mathbb{R})$.

iii) La matrice de la forme échelonnée réduite du système est la matrice identité I.

3.3.4 Calcul pratique de l'inverse

La discussion qui précède donne deux procédés pour calculer l'inverse d'une matrice inversible A :

- Par résolution de n systèmes linéaires $A \times X = B_j$, $j = 1, \ldots, n$, où les B_j sont les n vecteurs de la base canonique de $\mathbb{R}^n$.
- Par transformation du système en sa forme échelonnée réduite à l'aide de la méthode du pivot, qu'on a réinterprété comme la multiplication de A par des matrices de la forme $(I + \alpha E_{ij})$ ou αE_{ii}.

En fait ces deux procédés conduisent exactement aux mêmes calculs. Commençons par le second : si l'on effectue les mêmes substitutions de lignes sur la matrice identité que celle que l'on fait sur A pour l'amener à sa forme échelonnée réduite, on multiplie la matrice identité par A^{-1}, et donc on récupère l'inverse A^{-1} de A.

Voici un exemple d'un tel calcul. On place à droite de A la matrice identité, et on effectue sur les lignes de la grande matrice obtenue toutes les opérations nécessaires pour obtenir sa forme échelonnée réduite. Si cette forme est la matrice identité, c'est que A est inversible et dans ce cas, la matrice de droite est A^{-1}.

Exemple 3.3.7 Pour savoir si la matrice $A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$ est inversible et calculer son inverse éventuel, on écrit

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right),$$

et on effectue successivement les opérations $L_3 \leftarrow L_3 - \frac{1}{2}L_1$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 3/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right),$$

$L_3 \leftarrow L_3 - \frac{3}{2}L_2$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -2 & -1/2 & -3/2 & 1 \end{array}\right),$$

$L_3 \leftarrow -\frac{1}{2}L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_2 \leftarrow L_2 - L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_1 \leftarrow L_1 + L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 0 & 5/4 & 3/4 & -1/2 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_1 \leftarrow L_1 - L_2$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & 1/2 & -1 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

et enfin $L_1 \leftarrow \frac{1}{2}L_1$, ce qui donne

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & 1/4 & -1/2 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right).$$

On en conclut que A est inversible, d'inverse $A^{-1} = \begin{pmatrix} 3/4 & 1/4 & -1/2 \\ -1/4 & 1/4 & 1/2 \\ 1/4 & 3/4 & -1/2 \end{pmatrix}$.

Le lien avec la résolution des n systèmes linéaires est clair si l'on remarque que les colonnes de la matrices identité sont les B_j. Les substitutions de lignes ci-dessus sont exactement les opérations nécessaires à la résolution par la méthode du pivot d'un système $A \times X = B$ pour un second membre quelconque. Effectuer ces opérations sur chacune des colonnes de la matrice identité, c'est en fait résoudre simultanément les n systèmes $A \times X = B_j$.

Chapitre 4

Applications linéaires

4.1 Généralités

Rappelons qu'une fonction $f : A \to B$ est un procédé qui à chaque élément de l'ensemble de départ A associe 0 ou 1 élément (on dit aussi "au plus un élément") de l'ensemble d'arrivée B. L'ensemble de définition de la fonction f est le sous-ensemble de A constitué des éléments qui ont 1 image par f. On peut penser par exemple à la fonction $f : \mathbb{R} \to \mathbb{R}$ qui a un nombre x associe son inverse. Son ensemble de définition est $\mathcal{D}_f = \mathbb{R}^*$. Une fonction $f : A \to B$ est appelée application lorsque son ensemble de définition est A tout entier : $\mathcal{D}_f = A$.

4.1.1 Définition

Définition 4.1.1 Soit $(E, +, .)$ et $(F, +, .)$ deux espaces vectoriels. On dit qu'une application $f : E \to F$ est linéaire lorsque

$$f(a + b) = f(a) + f(b) \text{ et } f(\lambda.x) = \lambda.f(x)$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F.

Les applications linéaires sont aussi appelées "morphismes d'espace vectoriel". Comme pour les sous-espaces vectoriels, on peut vérifier les deux propriétés de la définition en une fois : une application $f : E \to F$ est linéaire si et seulement si pour tous réels λ, μ et tous x, y dans E, on a

$$f(\lambda.x + \mu.y) = \lambda.f(x) + \mu.f(y).$$

Une conséquence importante de cette définition est que l'image du vecteur nul de E par une application linéaire $f : E \to F$ est toujours 0_F. En effet

$$f(0_E) + 0_F = f(0_E) = f(0_E + 0_E) = f(0_E) + f(0_E),$$

d'où, en ajoutant l'opposé de $f(0_E)$, $0_F = f(0_E)$ comme annoncé.

Remarque 4.1.2 Dans la définition, on a noté '+' et '.' les lois de compositions sur E et sur F. Il faut prendre garde au fait que les ensembles E et F contiennent des objets à priori différents, et que les opérations sur ces objets différents n'ont à priori pas de rapport. Il est utile de se demander de quelle loi il s'agit quand on lit la définition ci-dessus.

4.1.2 Exemples

i) On prend $E = F = \mathbb{R}$ munis de l'addition et de la multiplication externe $(\ldots)$ usuelles. Si $f : \mathbb{R} \to \mathbb{R}$ est une application linéaire, alors pour tout $x \in \mathbb{R}$, on a

$$f(x) = f(x.1) = x.f(1).$$

Notant $a = f(1)$, on voit que $f : x \mapsto ax$.

ii) Maintenant $E = F$ est l'espace vectoriel des vecteurs du plan. L'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(\vec{u}) = \lambda.\vec{u}$ est linéaire. Il s'agit de l'homothétie vectorielle de rapport λ. Si $\vec{v}$ est un vecteur fixé non-nul, la translation de vecteur $\vec{v}$, définie par $f(\vec{u}) = \vec{u}+\vec{v}$ n'est pas linéaire : par exemple $f(\vec{u}_1+\vec{u}_2) = \vec{u}_1+\vec{u}_2+\vec{v}$ alors que $f(\vec{u}_1)+f(\vec{u}_2) = \vec{u}_1+\vec{u}_2+2\vec{v}$. On aurait aussi pu noter que $f(\vec{0}) = \vec{v} \neq \vec{0}$.

iii) Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ une matrice. L'application $f : \mathbb{R}^p \to \mathbb{R}^n$ définie par $f(X) = A \times X$ est une application linéaire. En effet d'après la proposition 3.2.5 et la remarque qui la précède, on a

$$f(\lambda_1 X_1+\lambda_2 X_2) = A\times(\lambda_1 X_1+\lambda_2 X_2) = A\times(\lambda_1 X_1)+A\times(\lambda_2 X_2) = \lambda_1 f(X_1)+\lambda_2 f(X_2).$$

Dans le paragraphe 4.3, on verra que, dans des espaces de dimension finie, une application linéaire est toujours de cette forme.

iv) Voici un exemple d'application linéaire qui sort du cadre des espaces vectoriels de dimension finie. On note $\mathcal{C}^0([0,1])$ l'espace vectoriel des fonctions continues de $[0,1]$ dans $\mathbb{R}$. On sait qu'une telle fonction est intégrable sur $[0,1]$ et que

$$\int_0^1 \lambda f(x) + \mu g(x) dx = \lambda \int_0^1 f(x)dx + \mu \int_0^1 g(x)dx.$$

Autrement dit, l'application $\varphi : \mathcal{C}^0([0,1]) \to \mathbb{R}$ définie par

$$\varphi(f) = \int_0^1 f(x)dx$$

est une application linéaire.

4.1.3 Somme et composition d'applications linéaires

Proposition 4.1.3 L'ensemble $\mathcal{L}(E,F)$ des applications linéaires de E dans F, muni de l'addition des fonctions et de la multiplication d'une fonction par un nombre, est un espace vectoriel.

Preuve.— On va montrer que $\mathcal{L}(E, F)$ est un sous-espace vectoriel de l'espace $\mathcal{F}(E, F)$ des application de E dans F (voir l'exemple 2.1.10). D'abord on vérifie facilement que l'application nulle $f : E \to F$ est linéaire, donc appartient à $\mathcal{L}(E, F)$. Il reste à montrer que $\mathcal{L}(E, F)$ est stable par combinaison linéaire : soit donc f_1 et f_2 deux applications linéaires et $\lambda_1, \lambda_2 \in \mathbb{R}$. Pour $x, y \in E$, on a

$$\begin{aligned}(\lambda_1 f_1 + \lambda_2 f_2)(x + y) &= \lambda_1 f_1(x + y) + \lambda_2 f_2(x + y)\\ &= \lambda_1 f_1(x) + \lambda_1 f_1(y) + \lambda_2 f_2(x) + \lambda_2 f_2(y)\\ &= \lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_1 f_1(y) + \lambda_2 f_2(y)\\ &= (\lambda_1 f_1 + \lambda_2 f_2)(x) + (\lambda_1 f_1 + \lambda_2 f_2)(y)\end{aligned}$$

De la même manière, pour $a \in \mathbb{R}$,

$$\begin{aligned}(\lambda_1 f_1 + \lambda_2 f_2)(ax) &= \lambda_1 f_1(ax) + \lambda_2 f_2(ax)\\ &= \lambda_1 a f_1(x) + \lambda_2 a f_2(x)\\ &= a(\lambda_1 f_1 + \lambda_2 f_2)(x).\end{aligned}$$

□

On peut aussi composer des applications linéaires :

Proposition 4.1.4 Soit E, F et G trois espaces vectoriels. Soit $f : E \to F$ et $g : F \to G$ deux applications linéaires. Alors $g \circ f : E \to G$ est une application linéaire.

Preuve.— Soit $x_1, x_2 \in E$ et $\lambda_1, \lambda_2 \in \mathbb{R}$. On a, en utilisant la linéarité de f

$$g \circ f(\lambda_1 x_1 + \lambda_2 x_2) = g(f(\lambda_1 x_1 + \lambda_2 x_2)) = g(\lambda_1 f(x_1) + \lambda_2 f(x_2)).$$

En utilisant maintenant la linéarité de g, on obtient

$$g\circ f(\lambda_1 x_1+\lambda_2 x_2) = g(\lambda_1 f(x_1)+\lambda_2 f(x_2) = \lambda_1 g(f(x_1))+\lambda_2 g(f(x_2)) = \lambda_1 g\circ f(x_1)+\lambda_2 g\circ f(x_2).$$

□

4.2 Noyau et Image

4.2.1 Injection, surjection, bijection

Les notions que l'on rappelle maintenant ne sont pas liées à la linéarité.

Définition 4.2.1 On dit qu'une application $f : E \to F$ est injective lorsque

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

Exemple 4.2.2 L'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas injective, puisque $f(-1) = f(1)$. Par contre l'application $f : \mathbb{R}^+ \to \mathbb{R}$ définie par $f(x) = x^2$ est injective.

Définition 4.2.3 On dit qu'une application $f : E \to F$ est surjective pour tout élément y de F, il existe au moins un élément x de E tel que $f(x) = y$.

Exemple 4.2.4 L'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas surjective, puisqu'il n'existe pas de réel x tel que $f(x) = -1$. Par contre l'application $f : \mathbb{R} \to \mathbb{R}^+$ définie par $f(x) = x^2$ est surjective.

Définition 4.2.5 On dit qu'une application $f : E \to F$ est bijective lorsque pour tout élément y de F, il existe exactement un élément x de E tel que $f(x) = y$.

Exemple 4.2.6 L'application $f : \mathbb{R}^+ \to \mathbb{R}^+$ définie par $f(x) = x^2$ est bijective, puisque pour tout $y \in \mathbb{R}^+$, il existe un unique x tel que $f(x) = y$: il s'agit de $x = \sqrt{y}$.

Proposition 4.2.7 Une application $f : E \to F$ est bijective si et seulement si elle est injective et surjective.

Preuve.— Si f bijective, f est clairement surjective. Soit $y \in E$ et $z = f(y)$. L'équation $f(x) = z$ a une seule solution, qui est y. Donc $f(x) = f(y)$ entraine $x = y$ et f est aussi injective. Reciproquement, supposons f injective et surjective. Pour $y \in F$, l'équation $f(x) = y$ a au moins une solution. Si elle en avait deux distinctes, disons x_1 et x_2, on aurait $f(x_1) = f(x_2)$ donc $x_1 = x_2$, ce qui est absurde. Donc f est bien bijective. □

Définition 4.2.8 Soit $f : E \to F$ une application bijective. L'application $f^{-1} : F \to E$ qui à y dans F associe l'unique $x \in E$ tel que $f(x) = y$ est appelée bijection réciproque de f.

On est allé un peu vite : il faut justifier le fait que f^{-1} est bien une bijection :

Proposition 4.2.9 Soit $f : E \to F$ une application bijective, et $f^{-1} : F \to E$ sa bijection réciproque.

i) f^{-1} est bijective, et sa bijection réciproque est f.

ii) On a $f^{-1} \circ f = I_E$ et $f \circ f^{-1} = I_F$.

Preuve.— On veut montrer que $f^{-1} : F \to E$ est bijective. D'abord $f^{-1} : F \to E$ est surjective : pour $x \in E$, notant $y = f(x) \in F$, on a $f^{-1}(y) = x$. Au passage, on remarque que $f^{-1}(f(x)) = x$, ce qui prouve la première partie de *ii)*. De plus $f^{-1} : F \to E$ est injective : si $y_1, y_2 \in F$ vérifient $f^{-1}(y_1) = f^{-1}(y_2) = x$, on a $f(x) = y_1$ et $f(x) = y_2$. Comme f est une fonction, cela entraine $y_1 = y_2$. Enfin, puisque $f^{-1} : F \to E$ est bijective, pour tout $x \in E$, l'équation $f^{-1}(y) = x$ d'inconnue $y \in F$ admet une unique solution. Comme $f(x)$ convient, on a l'équivalence

$$f^{-1}(y) = x \Longleftrightarrow y = f(x),$$

ce qui montre que $(f^{-1})^{-1} = f$ et termine la preuve de *i)*. La deuxième assertion de *ii)* s'obtient à partir de la première en échangeant le rôle de f et f^{-1}. □

Dans le cas des applications linéaires, on a un peu de vocabulaire spécifique :

Définition 4.2.10 Les applications linéaires bijectives de E dans F sont appelés isomorphismes. Les applications linéaires de E dans E sont appelés endomorphismes. Les endomorphismes bijectifs sont appelés automorphismes.

Plutôt que $\mathcal{L}(E, E)$, on note simplement $\mathcal{L}(E)$ l'ensemble des endomorphisme de E.

4.2.2 Noyau d'une application linéaire

Définition 4.2.11 Soit $f : E \to F$ une application linéaire. Le sous-ensemble de E des éléments dont l'image est 0_F est appelé noyau de f. On le note

$$\operatorname{Ker} f = \{x \in E,\ f(x) = 0_F\}.$$

Proposition 4.2.12 Soit $f : E \to F$ une application linéaire. Le noyau de f est un sous-espace vectoriel de E.

Preuve.— On a déjà vu que $f(0_E) = 0_F$, donc $0_E \in \operatorname{Ker} f$. De plus pour $\lambda_1, \lambda_2 \in \mathbb{R}$ et $x_1, x_2 \in \operatorname{Ker} f$,

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

ce qui prouve que $\operatorname{Ker} f$ est stable par combinaison linéaire. □

Proposition 4.2.13 Une application linéaire $f : E \to F$ est injective si et seulement si son noyau ne contient que le vecteur nul de E.

Preuve.— Supposons f injective. Si $x \in \operatorname{Ker} f$ on a $f(x) = 0_E = f(0_E)$, donc $x = O_E$. Réciproquement, supposons que $\operatorname{Ker} f = \{0_E\}$. Si $f(x) = f(y)$, on a $f(x-y) = f(x) - f(y) = 0_E$. Donc $x - y \in \operatorname{Ker} f$, i.e. $x = y$. □

Exercice 4.2.14 Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $f(x, y, z) = (y+z, x+z, x-y)$. Déterminer son noyau. L'application f est-elle injective ?

On a $(x, y, z) \in \operatorname{Ker} f$ si et seulement si (x, y, z) est solution du système

$$\left\{ \begin{array}{ccccccc} & & y & + & z & = & 0 \\ x & & & + & z & = & 0 \\ x & - & y & & & = & 0. \end{array} \right.$$

En appliquant la méthode du pivot par exemple, on voit que ce système est de rang 2, et que l'ensemble de ses solutions est $((-s, -s, s),\ s \in \mathbb{R}\} = \operatorname{Vect}((-1, -1, 1))$. Donc $\operatorname{Ker} f$ n'est pas réduit au vecteur nul, et f n'est pas injective.

4.2.3 Image d'une application linéaire

Définition 4.2.15 Le sous-ensemble de F constitué des images par f des éléments de E est appelé image de f. On le note

$$\operatorname{Im} f = \{f(x),\ x \in E\}.$$

Autrement dit $\operatorname{Im} f$ est pour une application linéaire ce que l'on note $f(E)$ pour une application quelconque. Par définition donc, une application linéaire $f : E \to F$ est surjective si et seulement si $\operatorname{Im} f = F$.

Proposition 4.2.16 Soit $f : E \to F$ une application linéaire. L'image de f est un sous-espace vectoriel de E.

Preuve.— On a d'abord $0_F \in \operatorname{Im} f$ puisque $f(0_E) = 0_F$. Enfin, si $\lambda_1, \lambda_2 \in \mathbb{R}$ et $y_1, y_2 \in \operatorname{Im} f$, il existe x_1 et x_2 dans E tels que $f(x_1) = y_1$ et $f(x_2) = y_2$. Ainsi

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

ce qui prouve que $\operatorname{Im} f$ est stable par combinaison linéaire. □

Exercice 4.2.17 Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $f(x,y,z) = (y+z, x+z, x-y)$. Déterminer son image. L'application f est-elle surjective ?

Un vecteur $(x', y', z') \in \mathbb{R}^3$ est dans l'image de f si et seulement si il existe $(x,y,z) \in \mathbb{R}^3$ tel que $f(x,y,z) = (x', y', z')$, autrement dit si et seulement si le système

$$\left\{\begin{array}{ccccccc} & & y & + & z & = & x', \\ x & & & + & z & = & y', \\ x & - & y & & & = & z'. \end{array}\right.$$

admet au moins une solution. Par la méthode du pivot, c'est équivalent à l'existence de solutions pour le système

$$\left\{\begin{array}{ccccccc} x & & & + & z & = & y', \\ & & y & + & z & = & x', \\ & & & & 0 & = & x'-y'+z'. \end{array}\right.$$

Or ce système est compatible si et seulement si $x' - y' + z' = 0$. Donc $\operatorname{Im} f = \{(x', y', z') \in \mathbb{R}^3,\ x' - y' + z' = 0\} \neq \mathbb{R}^3$. L'application f n'est donc pas surjective. On peut donner une base de $\operatorname{Im} f$: ce sous-espace vectoriel est donné par le système à une équation $x' - y' + z' = 0$. Comme l'équation est non-nulle, son rang est 1. La seule inconnue principale est x', et y', z' sont les inconnues secondaires. Le système s'écrit donc

$$\left\{\begin{array}{ccccc} x' & = & s & - & t \\ y' & = & s & & \\ z' & = & & & t. \end{array}\right.$$

L'ensemble des solutions est $((s-t, s, t), (s,t) \in \mathbb{R}^2\} = \operatorname{Vect}((1,1,0), (-1,0,1))$.

4.2.4 Applications linéaires bijectives

Voici une particularité des applications bijectives qui sont linéaires :

> **Proposition 4.2.18** Soit $f : E \to F$ une application linéaire. Si f est bijective, alors son application réciproque $f^{-1} : F \to E$ est aussi linéaire.

Preuve.— Soit $y_1, y_2 \in F$ et $\alpha_1, \alpha_2 \in \mathbb{R}$. Soit aussi $x_1 = f^{-1}(y_1)$ et $x_2 = f^{-1}(y_2)$. On a

$$f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2) = \alpha_1 y_1 + \alpha_2 y_2.$$

Donc

$$f^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1 + \alpha_2 x_2 = \alpha_1 f^{-1}(y_1) + \alpha_2 f^{-1}(y_2),$$

ce qui démontre la linéarité de f^{-1}. □

La proposition suivante, un peu technique, donne un caractérisation de l'injectivité et de la surjectivité d'une application linéaire f en termes de l'image d'une base par f. On retiendra en particulier qu'il ne peut pas y avoir d'application linéaire bijective entre deux espaces de dimension finie différentes, et qu'une application linéaire est bijective si et seulement si elle transforme toute base de l'espace de départ en une base de l'espace d'arrivée :

Proposition 4.2.19 Soit E un espace vectoriel de dimension finie, et F un espace vectoriel quelconque. Soit $\mathcal{B}_E$ une base de E et $f \in \mathcal{L}(E,F)$ une application linéaire.

i) f est injective si et seulement si $f(\mathcal{B}_E)$ est une famille libre de F ; en particulier, si f est injective on a $\dim E \leq \dim F$.

ii) f est surjective si et seulement si $f(\mathcal{B}_E)$ est une famille génératrice de F ; en particulier, si f est surjective on a $\dim F \leq \dim E$.

iii) f est bijective si et seulement si $f(\mathcal{B}_E)$ est une base de F ; en particulier si f est bijective on a $\dim E = \dim F$.

Preuve.— On note $\mathcal{B}_E = (u_1, u_2, \dots, u_p)$, de sorte que $f(\mathcal{B}_E) = (f(u_1), f(u_2), \dots, f(u_p))$.

i) Supposons f injective. Soit $\alpha_1, \alpha_2, \dots, \alpha_n$ des réels tels que

$$\alpha_1 f(u_1) + \cdots + \alpha_p f(u_p) = 0_F.$$

Par linéarité, on a $f(\alpha_1 u_1 + \cdots + \alpha_p u_p) = 0_F$, ie $\alpha_1 u_1 + \cdots + \alpha_p u_p \in \operatorname{Ker} f$. Puisque f est injective, cela entraine $\alpha_1 u_1 + \cdots + \alpha_p u_p = 0_E$, et puisque $(u_1, u_2, \dots, u_p)$ est une famille libre, $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. Donc $(f(u_1), f(u_2), \dots, f(u_p))$ est une famille libre.

Réciproquement, supposons que $(f(u_1), f(u_2), \dots, f(u_p))$ est une famille libre. Pour $u \in \operatorname{Ker} f$, on peut écrire $u = \alpha_1 u_1 + \cdots + \alpha_p u_p$, où les α_j sont les coordonnées de u dans la base $\mathcal{B}_E$. Mais alors

$$0_F = f(u) = f(\alpha_1 u_1 + \cdots + \alpha_p u_p) = \alpha_1 f(u_1) + \cdots + \alpha_p f(u_p),$$

donc $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$ et $u = 0_E$. Le noyau de f ne contient donc que 0_E et f est injective.

Pour f injective, l'inégalité $\dim E \leq \dim F$ provient du lemme 2.6.4 : le nombre d'éléments de la famille libre $f(\mathcal{B}_E)$ est $\dim E$, et il est nécessairement inférieur au nombre d'éléments d'une base de F, c'est-à dire $\dim F$.

ii) Supposons f surjective. Soit $y \in F = \operatorname{Im} f$. Il existe $u \in E$ tel que $y = f(u)$. Comme $\mathcal{B}_E$ est une base de E, on peut écrire $u = \alpha_1 u_1 + \cdots + \alpha_p u_p$, où les α_j sont les coordonnées de u dans la base $\mathcal{B}_E$. Mais alors

$$y = f(u) = f(\alpha_1 u_1 + \cdots + \alpha_p u_p) = \alpha_1 f(u_1) + \cdots + \alpha_p f(u_p),$$

donc $f(B_E)$ engendre bien F.

Réciproquement, si $f(B_E)$ engendre F, tout $y \in F$ peut s'écrire

$$y = \alpha_1 f(u_1) + \cdots + \alpha_p f(u_p) = f(\alpha_1 u_1 + \cdots + \alpha_p u_p),$$

c'est à dire comme un élément de $\operatorname{Im} f$. Donc $F = \operatorname{Im} f$ et f est surjective.

Pour f surjective, l'inégalité $\dim E \geq \dim F$ provient elle aussi du lemme 2.6.4 : le nombre d'éléments de la famille génératrice $f(\mathcal{B}_E)$ est $\dim E$, et il est nécessairement supérieur au nombre d'éléments d'une base de F, c'est-à dire $\dim F$.

iii) découle directement de *i)* et *ii)*. $\square$

4.2.5 Le théorème du rang

Définition 4.2.20 Soit $f \in \mathcal{L}(E, F)$. On appelle rang de f la dimension de $\operatorname{Im} f$, et on note

$$\operatorname{rang}(f) = \dim \operatorname{Im} f.$$

Proposition 4.2.21 Soit E un espace vectoriel de dimension finie, et $f \in \mathcal{L}(E, F)$. On a

$$\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f.$$

Preuve.— On utilise le théorème de la base incomplète (cf. la proposition 2.6.11). Soit $\mathcal{K} = (u_1, \ldots, u_k)$ une base de $\operatorname{Ker} f$. Comme E est de dimension finie - appelons la n, on peut completer $\mathcal{K}$ en une famille $\mathcal{B} = (u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ qui est une base de E.

Si $v \in \operatorname{Im} f$, il existe $u \in E$ tel que $v = f(u)$. Puisque $\mathcal{B}$ est une base de E, on peut écrire

$$u = \alpha_1 u_1 + \cdots + \alpha_k u_k + \alpha_{k+1} u_{k+1} + \cdots + \alpha_n u_n.$$

Ainsi, tout élément v de $\operatorname{Im} f$ peut s'écrire

$$\begin{aligned} v &= f(\alpha_1 u_1 + \cdots + \alpha_k u_k + \alpha_{k+1} u_{k+1} + \cdots + \alpha_n u_n) \\ &= \alpha_{k+1} f(u_{k+1}) + \cdots + \alpha_n f(u_n), \end{aligned}$$

puisque $f(u_j) = 0$ pour $j \in \{1, \ldots, k\}$. Donc la famille $(f(u_{k+1}), \ldots, f(u_n))$ est une famille génératrice de $\operatorname{Im} f$.

Montrons que c'est une famille libre de F : supposons que

$$0_F = \alpha_{k+1} f(u_{k+1}) + \cdots + \alpha_n f(u_n).$$

Alors $f(\alpha_k u_k + \cdots + \alpha_{k+1} u_{k+1}) = 0_F$, donc $\alpha_k u_k + \cdots + \alpha_{k+1} u_{k+1} \in \operatorname{Ker} f$. Puisque $\mathcal{K}$ est une base de $\operatorname{Ker} f$, cela signifie qu'il existe $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tels que

$$\alpha_k u_k + \cdots + \alpha_{k+1} u_{k+1} = \alpha_1 u_1 + \cdots + \alpha_k u_k,$$

ou encore

$$\alpha_1 u_1 + \cdots + \alpha_k u_k - \alpha_k u_k - \cdots - \alpha_{k+1} u_{k+1} = 0_E.$$

Comme $\mathcal{B} = (u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ est une base de E, tous les coefficients de cette combinaison linéaire sont nuls, en particulier $\alpha_{k+1}, \ldots, \alpha_n$.

En conclusion, la famille $(f(u_{k+1}), \ldots, f(u_n))$ est une base de $\operatorname{Im} f$, qui est donc un sous-espace vectoriel de F de dimension $n - k$, où k est la dimension de $\operatorname{Ker} f$. On a donc bien

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = k + (n - k) = n = \dim E.$$

□

Voici une des conséquences très utiles de cette proposition :

Corollaire 4.2.22 Soit $f \in \mathcal{L}(E, F)$ avec E et F dimension finie. Si $\dim E = \dim F$, on a l'équivalence
$$f \text{ bijective } \Leftrightarrow f \text{ surjective } \Leftrightarrow f \text{ injective.}$$

Preuve.— Supposons f injective. On a $\operatorname{Ker} f = \{0_E\}$, donc $\dim \operatorname{Ker} f = 0$, et, d'après le théorème du rang, $\dim \operatorname{Im} f = \dim E = \dim F$. Comme $\operatorname{Im} f \subset F$, cela entraine $\operatorname{Im} f = E$, donc f est surjective, et bijective.

Supposons maintenant f surjective. On a $\operatorname{Im} f = E$, donc $\dim \operatorname{Im} f = \dim E$, et le théorème du rang donne $\dim \operatorname{Ker} f = 0$. Ainsi $\operatorname{Ker} f = \{0_E\}$ et f est injective, donc aussi bijective. □

On a aussi le

Corollaire 4.2.23 Soit $f \in \mathcal{L}(E, F)$ avec E et F dimension finie. Si $\dim E > \dim F$, f ne peut pas être injective. De même si $\dim E < \dim F$, f ne peut pas être surjective.

Preuve.— Comme $\dim F \geq \dim \operatorname{Im} f$, si $\dim E > \dim F$, le théorème du rang donne $\dim \operatorname{Ker} f = \dim E - \dim \operatorname{Im} f > 0$, donc f n'est pas injective. Supposons $\dim E < \dim F$, comme $\dim \operatorname{Ker} f \geq 0$, le théorème du rang dit que $\dim \operatorname{Im} f \leq \dim E$, donc $\dim \operatorname{Im} f < \dim F$, et f n'est pas surjective. □

Exemple 4.2.24 Soit $f : \mathbb{R}^n[X] \to \mathbb{R}^n[X]$ définie par $f(P) = \lambda P + P'$ avec $\lambda \neq 0$. On veut montrer que f est un isomorphisme sur $\mathbb{R}^n[X]$. D'après ce qui précède, il suffit de montrer que $\operatorname{Ker} f = \{0\}$. Or $f(P) = 0$ si et seulement si $P = Ce^{-\lambda x}$, ce qui, dans $\mathbb{R}_n[X]$, entraine $C = 0$. Il est au moins aussi intéressant de considérer le cas $\lambda = 0$ (**exercice !**).

4.3 Matrices d'une application linéaire

4.3.1 Définition

Dans ce paragraphe, on considère deux espaces vectoriels E et F de dimension finie, respectivement p et n. On choisit une base $\mathcal{B}_E = (u_1, u_2, \ldots, u_p)$ de E, et une base $\mathcal{B}_F = (v_1, v_2, \ldots, v_n)$ de F. On commence par la

Proposition 4.3.1 Soit $f \in \mathcal{L}(E, F)$ une application linéaire. La donnée de l'image par f des vecteurs de la base $\mathcal{B}_E$ détermine entièrement l'application f.

Preuve.—Il s'agit de montrer que si l'on connait l'image par f des vecteurs de la base $\mathcal{B}_E$, alors on peut déterminer l'image par f de n'importe quel vecteur x de E. Or pour $x \in E$, on peut écrire, de manière unique,
$$x = x_1 u_1 + \cdots + x_p u_p.$$

Donc, par linéarité de f,

$$f(x) = f(x_1u_1 + \cdots + x_pu_p) = x_1f(u_1) + \cdots + x_pf(u_p),$$

et l'on voit que $f(x)$ est parfaitement connu à partir des coordonnées de x dans la base $\mathcal{B}_E$. $\square$

Bien entendu, les vecteurs $f(u_1), \ldots, f(u_p)$ sont eux parfaitement déterminés par la donnée de leurs coordonnées dans la base $\mathcal{B}_F$. Autrement dit, la donnée du tableau à n lignes et p colonnes constituées des coordonnées des $f(u_j)$ dans la base $(v_1, v_2, \ldots, v_n)$ de F determine entièrement l'application linéaire f. Cela conduit à la

Définition 4.3.2 Soit $f \in \mathcal{L}(E, F)$ une application linéaire. On appelle matrice de f dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$ la matrice $A \in \mathcal{M}_{p,n}(\mathbb{R})$ dont les p colonnes sont les coordonnées des images par f des vecteurs de $\mathcal{B}_E$ dans la base $\mathcal{B}_F$. On la note

$$A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) = \begin{array}{c} \begin{array}{cccc} f(u_1) & f(u_2) & \ldots & f(u_p) \end{array} \\ \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,p} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,p} \end{pmatrix} \end{array} \begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{array}$$

avec, pour $j \in \{1, \ldots, p\}$,

$$f(u_j) = a_{1,j}v_1 + a_{2,j}v_2 + \cdots + a_{n,j}v_n = \sum_{i=1}^{n} a_{i,j}v_i.$$

Exemple 4.3.3 On considère l'espace vectoriel $\mathbb{R}^n$, et $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ une base quelconque de $\mathbb{R}^n$. La matrice de l'homothétie $h_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ de rapport λ dans les bases $\mathcal{B}$ et $\mathcal{B}$ est

$$H_\lambda = \mathrm{Mat}_{\mathcal{B} \leftarrow \mathcal{B}}(h_\lambda) = \begin{array}{c} \begin{array}{cccc} h_\lambda(e_1) & h_\lambda(e_2) & \ldots & h_\lambda(e_n) \end{array} \\ \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ \\ e_n \end{array} = \lambda.I$$

Exemple 4.3.4 On note $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle θ dans $\mathbb{R}^2$, définie par

$$f_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

La matrice de f_θ dans la base canonique $\mathcal{B}$ de $\mathbb{R}^2$ est donc

$$A = \mathrm{Mat}_{\mathcal{B} \leftarrow \mathcal{B}}(f_\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Exemple 4.3.5 On note $f_\theta : \mathbb{R}^3 \to \mathbb{R}^3$ la rotation d'axe $(0,0,1)$ et d'angle θ dans $\mathbb{R}^3$, définie par

$$f_\theta(x,y,z) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z).$$

La matrice de f_θ dans la base canonique $\mathcal{B}$ de $\mathbb{R}^3$ est donc

$$A = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}}(f_\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exemple 4.3.6 On considère l'espace vectoriel $\mathbb{R}_n[X]$ des polynômes de degré inférieur ou égal à n. La famille $\mathcal{B} = (1, X, X^2, \ldots, X^n)$ est une base de $\mathbb{R}_n[X]$. Soit alors $\varphi : \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ l'application qui a un polynôme $P \in \mathbb{R}_n[X]$ associe son polynôme dérivée P'. L'application φ est linéaire, et $\varphi(1) = 0$, $\varphi(X) = 1$, ..., $\varphi(X^n) = nX^{n-1}$. La matrice de φ dans les bases $\mathcal{B}$ et $\mathcal{B}$ est donc

$$A = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}}(\varphi) = \begin{array}{c} \begin{array}{ccccc} \varphi(1) & \varphi(X) & \varphi(X^2) & \ldots & \varphi(X^n) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ \vdots \\ X^{n-1} \\ X^n \end{array}$$

Remarque 4.3.7 La matrice d'une application linéaire dans des bases $\mathcal{B}_E$ et $\mathcal{B}_F$ dépend en général des bases $\mathcal{B}_E$ et $\mathcal{B}_F$. Comme on l'a vu plus haut, ce n'est cependant pas le cas pour les homothéties, et donc en particulier pour l'application identité (qu'on peut écrire h_λ avec $\lambda = 1$).

4.3.2 Liens entre une application linéaire et sa matrice dans des bases données

Soyons plus précis : on a vu que "la donnée du tableau à n lignes et p colonnes constituées des coordonnées des $f(u_j)$ dans la base $(v_1, v_2, \ldots, v_n)$ de F détermine entièrement l'application linéaire f ", mais comment ?

Proposition 4.3.8 Soit $f \in \mathcal{L}(E,F)$ une application linéaire. La matrice $A \in \mathcal{M}_{n,p}(\mathbb{R})$ est la matrice de f dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$, si et seulement si pour tout $x \in E$, de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ dans la base $\mathcal{B}_E$, les coordonnées $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ de son image $y = f(x)$ dans la base $\mathcal{B}_F$ vérifient $Y = A \times X$.

Preuve.— Supposons que A soit la matrice de f dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$. Par hypothèse, on a $X = \sum_{j=1}^{p} x_j u_j$. Donc, par linéarité de f,

$$f(X) = \sum_{j=1}^{p} x_j f(u_j) = \sum_{j=1}^{p} x_j \sum_{i=1}^{n} a_{i,j} v_i = \sum_{i=1}^{n} \big(\sum_{j=1}^{p} a_{i,j} x_j\big) v_i,$$

où les $a_{i,j}$ sont les coefficients de la matrice A. Autrement dit, la coordonnée de $f(X)$ selon le vecteur v_i est $\sum_{j=1}^{p} a_{i,j}x_j$ c'est à dire la i-ième ligne de $A \times X$.

Réciproquement si les coordonnées Y de $f(X)$ sont données par $Y = A \times X$ pour tout $X \in E$, c'est le cas pour les vecteurs de $\mathcal{B}_E$. Or les coordonnées X_j de u_j dans la base $\mathcal{B}_E$ sont $X_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, de sorte que $A \times X_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$. Cela prouve que la j-ième colonne de A est bien constituée des coordonnées de $f(u_j)$ dans la base $\mathcal{B}_F$, et donc que $A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f)$. □

Remarque 4.3.9 On a déjà parlé de l'application linéaire associée à une matrice $A \in \mathcal{M}_{p,n}(\mathbb{R})$ donnée. Il est maintenant clair qu'il s'agit de l'application linéaire f de $\mathbb{R}^n$ dans $\mathbb{R}^p$ dont la matrice dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est A.

Proposition 4.3.10 Soit $f : E \to F$ une application linéaire, et $A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f)$ sa matrice dans $\mathcal{B}_E$ et $\mathcal{B}_F$. Le rang de f est le rang du système linéaire associé à la matrice A.

Preuve.— Si $y \in \mathrm{Im}\, f$, il existe $x = \alpha_1 u_1 + \cdots + \alpha_p u_p$ tel que $y = f(x) = \alpha_1 f(u_1) + \cdots + \alpha_p f(u_p)$. Donc $f(\mathcal{B}_E) = (f(u_1), \dots f(u_p))$ est une famille génératrice de $\mathrm{Im}\, f$. Trouver le rang de f, c'est trouver le nombre d'éléments d'une base de $\mathrm{Im}\, f$. Or on sait construire une base d'un espace vectoriel quand on dispose d'une famille génératrice (voir la section 2.6.3) : on doit échelonner le système obtenu en écrivant coordonnée par coordonnée l'équation

$$x_1 f(u_1) + x_2 f(u_2) + \dots x_p f(u_p) = 0.$$

Il s'agit bien du système associé à la matrice A si l'on prend les coordonnées des $f(u_j)$ dans la base $\mathcal{B}_F$. □

Remarque 4.3.11 Dans la pratique, pour calculer le rang d'une application linéaire f, on applique l'algorithme du pivot à une matrice A représentant f. Le rang est alors le nombre de lignes non-nulles de la matrice échelonnée obtenue.

Remarque 4.3.12 On vient de prouver un point important. Le rang d'un système linéaire a été défini (cf. la définition 1.4.2) comme le nombre de lignes non-nulles du système échelonné obtenue en appliquant l'algorithme du pivot au sytème en question. Obtiendrait-on le même rang si l'on utilisait un autre procédé pour trouver un système échelonné équivalent au premier ? La réponse est oui : ce que l'on calcule, c'est la dimension de l'image de l'application linéaire f associée au système, et ce nombre ne dépend pas de la méthode utilisée pour le calculer.

Proposition 4.3.13 Soit E, F et G des espaces vectoriels de dimension finie, et deux applications linéaires $f : E \to F$ et $g : F \to G$. Soit aussi $\mathcal{B}_E$, $\mathcal{B}_F$ et $\mathcal{B}_G$ des bases respectives de E, F et G, et

$$A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f), \quad B = \mathrm{Mat}_{\mathcal{B}_G \leftarrow \mathcal{B}_F}(g).$$

Alors la matrice C de l'application linéaire $g \circ f : E \to G$ dans les bases $\mathcal{B}_E$ et $\mathcal{B}_G$ est donnée par le produit $B \times A$:

$$C = \mathrm{Mat}_{\mathcal{B}_G \leftarrow \mathcal{B}_E}(g \circ f) = \mathrm{Mat}_{\mathcal{B}_G \leftarrow \mathcal{B}_F}(g) \times \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) = B \times A.$$

Preuve.— D'après la proposition précédente, il suffit de prouver que pour tout $x \in E$, de coordonnées X dans la base $\mathcal{B}_E$, le vecteur Z des coordonnées de $g \circ f(X)$ dans la base $\mathcal{B}_G$ est donné par $(B \times A) \times X$. Or $f(x)$ a pour coordonnées $A \times X$ dans la base $\mathcal{B}_F$, et $g(f(x))$ a pour coordonnées $B \times (A \times X)$ dans la base $\mathcal{B}_G$. La proposition découle donc simplement de l'associativité du produit de matrices $B \times (A \times X) = (B \times A) \times X$. □

Exemple 4.3.14 Soit $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ et $f_{\theta'} : \mathbb{R}^2 \to \mathbb{R}^2$ les rotations d'angle θ et θ' respectivement. La matrice de $f_\theta \circ f'_\theta$ dans la base canonique est

$$\begin{aligned}
&\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \times \begin{pmatrix} \cos\theta' & \sin\theta' \\ -\sin\theta' & \cos\theta' \end{pmatrix} \\
&\qquad = \begin{pmatrix} \cos\theta\cos\theta' - \sin\theta\sin\theta' & \cos\theta\sin\theta' + \sin\theta\cos\theta' \\ -\sin\theta\cos\theta' - \cos\theta\sin\theta' & -\sin\theta\sin\theta' + \cos\theta\cos\theta' \end{pmatrix} \\
&\qquad = \begin{pmatrix} \cos(\theta+\theta') & \sin(\theta+\theta') \\ -\sin(\theta+\theta') & \cos(\theta+\theta') \end{pmatrix}
\end{aligned}$$

Donc $f_\theta \circ f'_\theta$ est la rotation d'angle $\theta + \theta'$.

Proposition 4.3.15 Soit E et F de dimension finie, respectivement p et n. Soit $f \in \mathcal{L}(E, F)$ une application linéaire, et $A \in \mathcal{M}_{n,p}(\mathbb{R})$ la matrice de f dans les bases $\mathcal{B}_E$ de E et $\mathcal{B}_F$ de F. L'application f est bijective si et seulement si la matrice A est une matrice carrée inversible. Dans ce cas

$$\mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}_F}(f^{-1}) = A^{-1}.$$

Preuve.— Supposons f bijective. D'après le point *iii)* de la proposition 4.2.19, on a $\dim E = \dim F$, et $A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f)$ est une matrice carrée de $\mathcal{M}_n(\mathbb{R})$. Notons $B = \mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}_F}(f^{-1})$. La matrice de $f \circ f^{-1} : F \to F$ dans les bases $\mathcal{B}_F$ et $\mathcal{B}_F$ est $B \times A$. Comme $f \circ f^{-1}$ est l'application identité, on a $B \times A = I$. De la même manière on obtient $A \times B = I$, donc A est inversible et $A^{-1} = B = \mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}_F}(f^{-1})$.

Réciproquement, supposons la matrice carrée $A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f)$ inversible, d'inverse A^{-1}. On note $g : F \to E$ l'application linéaire dont la matrice dans les bases $\mathcal{B}_F$ et $\mathcal{B}_E$ est A^{-1}. Pour $x \in E$, notant $X \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne des coordonnées de x dans la base $\mathcal{B}_E$, on sait que les coordonnées Y de $g \circ f(x)$ dans la base $\mathcal{B}_E$ sont données par

$$Y = (\mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}_F}(g) \times \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f)) \times X = A^{-1} \times A \times X,$$

donc $g \circ f(x) = x$ pour tout x, ou encore $g \circ f = I_E$. On montre de la même manière que $f \circ g = I_F$, et donc que f est bijective. □

Exercice 4.3.16 Soit $a \in \mathbb{R}$ et $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire définie par $f(x, y) = (x + 2y, 2x + ay)$. L'application f est-elle bijective ? Si oui donner f^{-1}.

4.4 Projections et symétries vectorielles

On examine en détail deux types d'applications linéaires qui jouent un rôle important. Comme leur nom l'indique, il est utile d'avoir en tête l'exemple des projections et symétries du plan ou de l'espace.

4.4.1 Projections

Définition algébrique

Définition 4.4.1 Soit $(E, +, .)$ un espace vectoriel. On dit qu'un endomorphisme $p \in \mathcal{L}(E)$ est une projection lorsque $p \circ p = p$.

Exemple 4.4.2 Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire dont A est la matrice dans la base canonique de $\mathbb{R}^2$. D'après la proposition 4.3.13, la matrice de $f \circ f$ est

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

Donc $f \circ f = f$, et f est une projection.

Exemple 4.4.3 Soit $p \in \mathcal{L}(E, E)$ une projection. On a

$$(I - p) \circ (I - p) = I \circ (I - p) - p \circ (I - p) = I \circ I - I \circ p - p \circ I + p \circ p = I - p.$$

Donc $I - p$ est aussi une projection.

Proposition 4.4.4 Si $p \in \mathcal{L}(E)$ est une projection, alors $\operatorname{Ker} p$ et $\operatorname{Im} p$ sont des sous-espaces vectoriels supplémentaire de E :

$$E = \operatorname{Ker} p \oplus \operatorname{Im} p.$$

De plus la restriction de p à $\operatorname{Ker} p$ est l'application nulle, et celle de p à $\operatorname{Im} p$ est l'application identité.

Preuve.— Soit $y \in \operatorname{Ker} p \cap \operatorname{Im} p$. Comme $y \in \operatorname{Im} p$, il existe $x \in E$ tel que $p(x) = y$. Comme $p \circ p = p$ on a aussi $p(y) = p(p(x)) = p(x) = y$. Or $p(y) = 0_E$ puisque $y \in \operatorname{Ker} p$. Donc $y = 0_E$ et $\operatorname{Ker} p \cap \operatorname{Im} p = \{0_E\}$. La dimension du sous-espace vectoriel $\operatorname{Ker} p + \operatorname{Im} p$ de E est donné par la formule de la dimension (cf. la proposition 2.7.9) :

$$\dim(\operatorname{Ker} p + \operatorname{Im} p) = \dim \operatorname{Ker} p + \dim \operatorname{Im} p - \dim \operatorname{Ker} p \cap \operatorname{Im} p = \dim \operatorname{Ker} p + \dim \operatorname{Im} p.$$

Enfin le théorème du rang dit que

$$\dim \operatorname{Ker} p + \dim \operatorname{Im} p = \dim E,$$

et donc que $E = \operatorname{Ker} p \oplus \operatorname{Im} p$. Enfin si $x \in \operatorname{Ker} p$ on a $p(x) = 0$ et si $x \in \operatorname{Im} p$, il existe $u \in E$ tel que $x = p(u)$. Mais alors $p(x) = p \circ p(u) = p(u) = x$. □

Définition géométrique

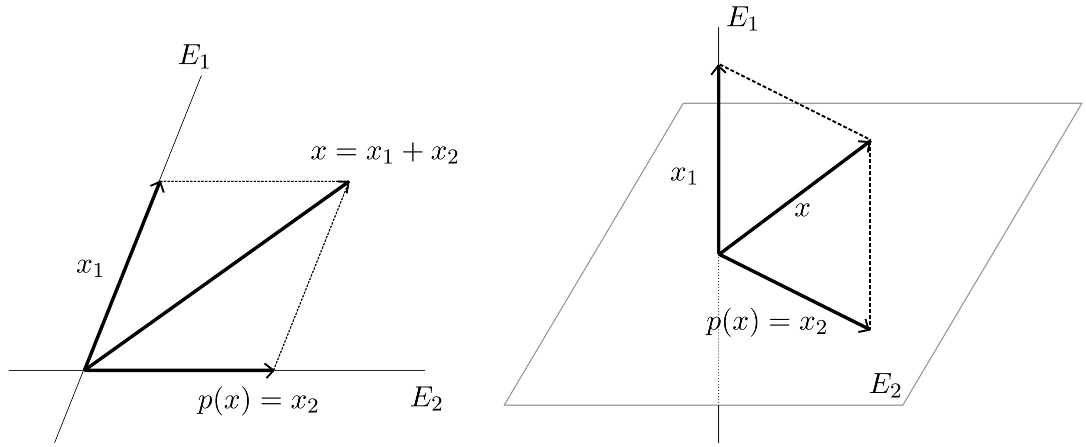

FIGURE 4.1 – Projections vectorielles dans le plan et dans l'espace

Soit $p \in \mathcal{L}(E)$ une projection. La proposition 4.4.11 dit que $E = \operatorname{Ker} p \oplus \operatorname{Im} p$, c'est-à dire que tout vecteur $x \in E$ s'écrit de manière unique $x = u + v$ avec $u \in \operatorname{Ker} p$ et $v \in \operatorname{Im} p$. On dit que u est la composante de x sur $\operatorname{Ker} p$ et que v est la composante de x sur $\operatorname{Im} p$.

Avec ces notations, on a $p(x) = p(u) + p(v) = p(v)$. De plus, comme $v \in \operatorname{Im} p$, il existe $w \in E$ tel que $v = p(w)$, donc

$$p(v) = p(p(w)) = p^2(w) = p(w) = v.$$

Finalement, on a donc

$$x = u + v, u \in \operatorname{Ker} p, v \in \operatorname{Im} p \Rightarrow p(x) = v.$$

Cette remarque va nous permettre d'interpréter géométriquement la notion de projection :

> **Proposition 4.4.5** Soit E_1 et E_2 deux sous-espaces vectoriels supplémentaires de l'espace vectoriel E. Soit $p : E \to E$ l'application qui à $x \in E$ associe sa composante sur E_2. Alors p est une projection ; son image est E_2 et son noyau est E_1. On dit que p est la projection sur E_2 dans la direction de E_1.

Preuve.— On montre d'abord que p est linéaire. Soit $x \in E$ et $\lambda \in \mathbb{R}$. On note x_1 et x_2 les composantes de x dans E_1 et E_2 respectivement, de sorte que $x = x_1 + x_2$ et $p(x) = x_2$. Alors $\lambda x = \lambda x_1 + \lambda x_2$, avec $\lambda x_1 \in E_1$ et $\lambda x_2 \in E_2$ puisque E_1 et E_2 sont des sous-espaces vectoriels. Donc

$$p(\lambda x) = \lambda x_2 = \lambda p(x).$$

Soit aussi $y \in E$, de composantes y_1 et y_2 dans E_1 et E_2 respectivement. De la même manière on a

$$p(x + y) = p((x_1 + x_2) + (y_1 + y_2)) = p((x_1 + y_1) + (x_2 + y_2)) = x_2 + y_2 = p(x) + p(y),$$

puisque $x_1 + y_1 \in E_1$ et $x_2 + y_2 \in E_2$.

On calcule maintenant p^2. Soit $x \in E$, de composantes x_1 et x_2 dans E_1 et E_2 respectivement. On a

$$p^2(x) = p(p(x)) = p(p(x_1 + x_2)) = p(x_2) = p(0 + x_2) = x_2 = p(x).$$

Donc p est bien une projection, de noyau E_1 et d'image E_2. □

Dans le cas où E est l'ensemble des vecteurs du plan, et E_1, E_2 deux droites vectorielles distinctes, comme dans le cas où E est l'ensemble des vecteurs de l'espace, E_1 une droite et E_2 un plan ne contenant pas E_1, la projection sur E_2 dans la direction de E_1 est bien ce que l'on pense, comme l'illustre la figure 4.1.

Matrice d'une projection

Si A est la matrice d'une projection p dans des bases données, on a $A^2 = A \times A = A$, puisque que $A \times A$ est la matrice dans les mêmes bases de $p \circ p = p$. On peut choisir les bases de manière à ce que la matrice de p dans celles-ci soit très simple. Pour le voir, on doit commencer par la

Proposition 4.4.6 Soit E_1 et E_2 deux sous-espaces vectoriels de E, et $\mathcal{B}_1 = (u_1, \dots u_k)$, $\mathcal{B}_2 = (v_1, \dots, v_{k_2})$ des bases de E_1 et E_2 respectivement. Si $E = E_1 \oplus E_2$, alors $\mathcal{B} = \mathcal{B}_1 \bigcup \mathcal{B}_2 = (u_1, \dots u_k, v_1, \dots, v_{k_2})$ est une base de E.

Preuve.— Soit $x \in E = E_1 \oplus E_2$. Il existe $x_1 \in E_1$ et $x_2 \in E_2$ tel que $x = x_1 + x_2$. Comme $\mathcal{B}_1$ est une base de E_1, il existe $\alpha_1, \dots, \alpha_{k_1}$ tels que $x_1 = \alpha_1 u_1 + \dots + \alpha_{k_1} x_{k_1}$. De même, il existe $\beta_1, \dots, \beta_{k_2}$ tels que $x_2 = \beta_1 v_1 + \dots + \beta_{k_2} v_{k_2}$. Donc

$$x = x_1 + x_2 = \alpha_1 u_1 + \dots + \alpha_{k_1} x_{k_1} + \beta_1 v_1 + \dots + \beta_{k_2} v_2,$$

ce qui montre que $\mathcal{B}$ est une famille génératrice de E.

Supposons maintenant que $\alpha_1 u_1 + \dots + \alpha_{k_1} x_{k_1} + \beta_1 v_1 + \dots + \beta_{k_2} v_{k_2} = 0_E$. Alors

$$E_1 \ni \alpha_1 u_1 + \dots + \alpha_{k_1} x_{k_1} = -(\beta_1 v_1 + \dots + \beta_{k_2} v_{k_2}) \in E_2.$$

Puisque l'intersection de E_1 et E_2 est réduite à (0_E), on obtient

$$\alpha_1 u_1 + \dots + \alpha_{k_1} x_{k_1} = \beta_1 v_1 + \dots + \beta_{k_2} v_{k_2} = 0_E.$$

Comme $(u_1, \dots u_k)$ d'une part et $(v_1, \dots v_{k_2})$ d'autre part sont des familles libres, cela entraine $\alpha_1 = \dots = \alpha_{k_1} = \beta_1 = \dots = \beta_{k_2} = 0$. Donc $\mathcal{B}$ est une famille libre. □

Soit donc $p : E \to E$ une projection, $\mathcal{B}_1 = (u_1, \dots, u_{k_1})$ une base de $\operatorname{Im} p$ et $\mathcal{B}_2 = (v_1, \dots, v_{k_2})$ une base de $\operatorname{Ker} p$. D'après ce qui précède, $\mathcal{B} = \mathcal{B}_1 \cap \mathcal{B}_2$ est une base de E. Notant A la matrice de p dans les bases $\mathcal{B}$ et $\mathcal{B}$, puisque $p(u_j) = u_j$ et $p(v_j) = 0$, on a

$$A = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}}(p) = \begin{array}{c} \begin{array}{ccc|ccc} p(u_1) & \dots & p(u_{k_1}) & p(v_1) & \dots & p(v_{k_2}) \end{array} \\ \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right) \end{array} \begin{array}{c} \\ u_1 \\ \\ u_{k_1} \\ v_1 \\ \\ v_{k_2} \end{array}$$

On pourra rapprocher ce résultat de l'exemple 4.4.2.

4.4.2 Symétries

Définition algébrique

Définition 4.4.7 Soit $(E, +, .)$ un espace vectoriel. On dit qu'un endomorphisme $s \in \mathcal{L}(E)$ est une symétrie lorsque $s \circ s = I$.

On notera qu'une symétrie est une application linéaire bijective, qui est sa propre inverse.

Exemple 4.4.8 Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et $s : \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire dont A est la matrice dans la base canonique de $\mathbb{R}^2$. La matrice de $s \circ s$ est

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

Donc $s \circ s = I$, et s est une symétrie.

Exemple 4.4.9 Soit $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +, .)$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. A une fonction $f \in \mathcal{F}$ on associe la fonction $s(f)$ définie par

$$s(f) : x \mapsto f(-x).$$

Il est simple de voir que $s : \mathcal{F} \to \mathcal{F}$ est linéaire. De plus on a $s \circ s(f)(x) = s(s(f))(x) = s(f)(-x) = f(x)$, donc $s \circ s = I$ et s est une symétrie.

Proposition 4.4.10 Soit $p \in \mathcal{L}(E)$ une projection. L'application $s = I - 2p$ est une symétrie.

Preuve.— On calcule $s \circ s$:

$$(I - 2p) \circ (I - 2p) = I \circ (I - 2p) - 2p \circ (I - 2p) = I \circ I - 2I \circ p - 2p \circ I + 4p \circ p = I.$$

Donc $s = I - 2p$ est une symétrie. $\square$

Puisque s est une bijection, on a $\operatorname{Ker} s = \{0_E\}$ et $\operatorname{Im} s = E$, donc en particulier encore $E = \operatorname{Ker} s \oplus \operatorname{Im} s$, mais cette égalité est sans intérêt pour une symétrie. Par contre on a aussi la

Proposition 4.4.11 Si $s \in \mathcal{L}(E)$ est une symétrie, alors $\operatorname{Ker}(s+I)$ et $\operatorname{Ker}(s-I)$ sont des sous-espaces vectoriels supplémentaires de E :

$$E = \operatorname{Ker}(s+I) \oplus \operatorname{Ker}(s-I).$$

Preuve.— Soit $p = \frac{1}{2}(I-s)$. On a

$$p \circ p = \frac{1}{2}(I-s) \circ \frac{1}{2}(I-s) = \frac{1}{4}(I-s)\circ(I-s) = \frac{1}{4}(I - I\circ s - s\circ I + s\circ s) = \frac{1}{4}(2I-2s) = p,$$

donc p est une projection, et

$$E = \operatorname{Ker} p \oplus \operatorname{Im} p.$$

Si $y \in \operatorname{Im} p$, il existe $x \in E$ tel que $y = p(x) = \frac{1}{2}(x - s(x))$, donc

$$s(y) + y = \frac{1}{2}(s(x) - s(s(x))) + \frac{1}{2}(x - s(x)) = \frac{1}{2}(s(x) - x)) + \frac{1}{2}(x - s(x)) = 0_E,$$

ce qui montre que $\operatorname{Im} p \subset \operatorname{Ker}(s+I)$. Réciproquement, si $y \in \operatorname{Ker}(s+I)$, on a $-y = s(y)$ et

$$p(y) = \frac{1}{2}(y - s(y)) = \frac{1}{2}(y+y) = y,$$

donc $y \in \operatorname{Im} p$. On a prouvé que $\operatorname{Im} p = \operatorname{Ker}(s+I)$ et puisque $\operatorname{Ker} p = \operatorname{Ker}(s-I)$, on a bien la relation $E = \operatorname{Ker}(s+I) \oplus \operatorname{Ker}(s-I)$. $\square$

Exemple 4.4.12 On reprend l'exemple 4.4.9. On a vu que l'application $s : \mathcal{F} \to \mathcal{F}$ définie par $s(f) : x \to f(-x)$ est une symétrie. Une fonction $f \in \mathcal{F}$ appartient à $\operatorname{Ker}(s+I)$ si et seulement si pour tout $x \in \mathbb{R}$, on a $(s(f)+f)(x) = f(-x) + f(x) = 0$. Autrement dit $\operatorname{Ker}(s+I)$ est le sous-espace des fonction impaires. De la même manière, $f \in \operatorname{Ker}(s-I)$ si et seulement si $(s(f)-f)(x) = f(-x) - f(x) = 0$, donc $\operatorname{Ker}(s-I)$ est le sous-espace des fonctions paires. Dans ce cas, la proposition précédente dit donc que toute fonction de $\mathbb{R}$ dans $\mathbb{R}$ s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire.

Définition géométrique

Soit $s : E \to E$ une symétrie, et $x \in E$. On peut écrire $x = x_1 + x_2$ avec $x_1 \in \operatorname{Ker}(s-I)$ et $x_2 \in \operatorname{Ker}(s+I)$, de sorte que $s(x) = s(x_1 + x_2) = s(x_1) + s(x_2) = x_1 - x_2$.

Proposition 4.4.13 Soit E_1 et E_2 deux sous-espaces vectoriels supplémentaires de l'espace vectoriel E. Soit $s : E \to E$ l'application qui à $x \in E$ associe $x_1 - x_2$ où x_1 et x_2 sont définis par $x = x_1 + x_2$ avec $x_1 \in E_1$ et $x_2 \in E_2$. Alors s est une symétrie. On dit que s est la symétrie par rapport à E_2 dans la direction de E_1.

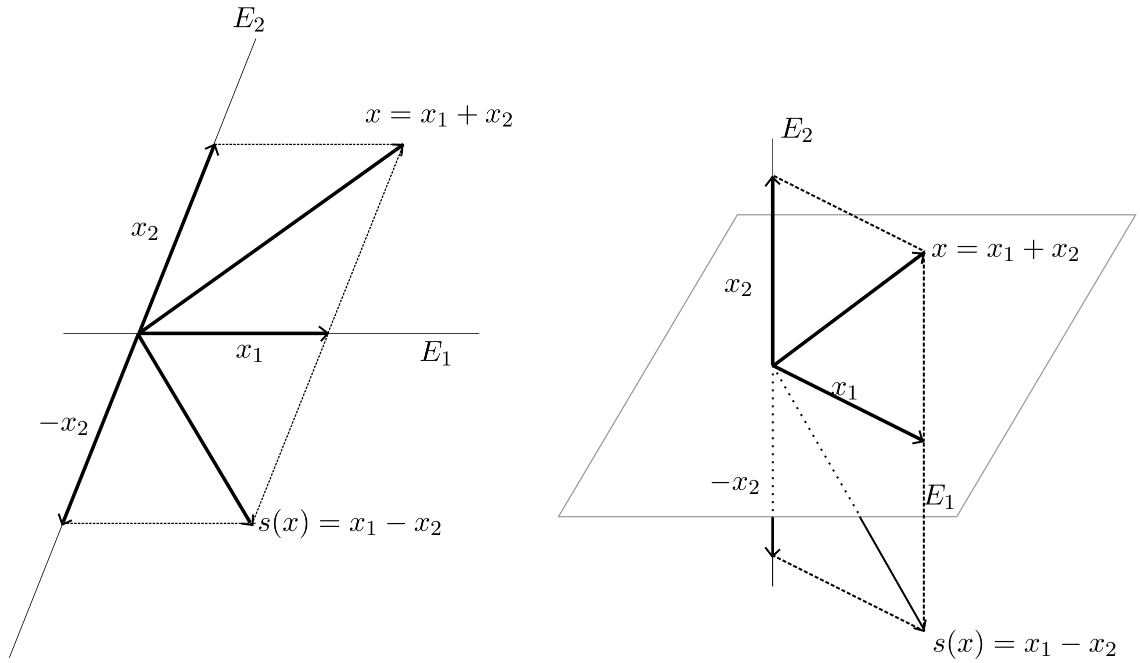

FIGURE 4.2 – Symétries vectorielles dans le plan et dans l'espace

Preuve.— Soit $x \in E$ et $x_1 \in E_1$, $x_2 \in E_2$ tels que $x = x_1 + x_2$. On a $s(x) = x_1 - x_2 = x_1 + (-x_2)$, avec $x_1 \in E_1$, $-x_2 \in E_2$. Donc

$$s(s(x)) = s(x_1 - x_2) = s(x_1 + (-x_2)) = x_1 - (-x_2) = x_1 + x_2 = x,$$

donc $s \circ s = I$. □

Dans le cas où E est l'ensemble des vecteurs du plan, et E_1, E_2 deux droites vectorielles distinctes, comme dans le cas où E est l'ensemble des vecteurs de l'espace, E_1 un plan et E_2 une droite qui n'est pas contenue dans E_1, la symétrie par rapport à E_1 dans la direction de E_2 est bien ce que l'on pense, comme l'illustre la figure 4.2.

Matrice d'une symétrie

Si A est la matrice d'une symétrie s dans des bases données, on a $A^2 = A \times A = I$, puisque que $A \times A$ est la matrice dans les mêmes bases de $s \circ s = I$. Pour une symétrie aussi, on peut trouver des bases dans lesquelles la matrice de A est très simple.

Soit E_1 et E_2 deux sous-espaces supplémentaires de E, et $s : E \to E$ la symétrie par rapport à E_1 dans la direction de E_2. Soit aussi $\mathcal{B}_1 = (u_1, \ldots, u_{k_1})$ une base de E_1 et $\mathcal{B}_2 = (v_1, \ldots, v_{k_2})$ une base de E_2. On sait que $\mathcal{B} = \mathcal{B}_1 \cap \mathcal{B}_2$ est une base de E. Notant A la matrice de s dans les bases $\mathcal{B}$ et $\mathcal{B}$, puisque $s(u_j) = u_j$ et $s(v_j) = -v_j$, on a

$$A = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}}(s) = \begin{array}{c} \begin{array}{cccc|ccc} s(u_1) & \dots & s(u_{k_1}) & & s(v_1) & \dots & s(v_{k_2}) \end{array} \\ \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & -1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 \end{array}\right) \end{array} \begin{array}{c} \\ u_1 \\ \\ u_{k_1} \\ v_1 \\ \\ v_{k_2} \end{array}$$

Là encore, on pourra rapprocher ce résultat de l'exemple 4.4.8.

4.5 Changement de base

La matrice d'une application linéaire f donnée dépend en général des bases dans lesquelles on la définit. On s'intéresse dans cette section aux relations entre les matrices de f dans des bases différentes.

4.5.1 Matrice de passage d'une base dans une autre

Soit E un espace vectoriel de dimension n. Soit $\mathcal{B} = (u_1, u_2, \dots, u_n)$ une base de E, et $\mathcal{B}' = (u'_1, u'_2, \dots, u'_n)$ une "nouvelle base" de E. En général, les vecteurs de $\mathcal{B}' = (u'_1, u'_2, \dots, u'_n)$ sont donnés par leurs coordonnées dans "l'ancienne base" $\mathcal{B}$, et il est naturel et très simple de former la matrice P des coordonnées des vecteurs de la nouvelle base $\mathcal{B}'$ dans l'ancienne base $\mathcal{B}$:

$$P = \begin{array}{c} \begin{array}{cccc} u'_1 & u'_2 & \dots & u'_n \end{array} \\ \begin{pmatrix} p_{1,1} & p_{1,2} & & p_{1,n} \\ p_{2,1} & p_{2,2} & & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & p_{n,2} & & p_{n,n} \end{pmatrix} \end{array} \begin{array}{c} \\ u_1 \\ u_2 \\ \\ u_n \end{array}$$

Cette matrice a un nom, pour l'instant un peu troublant :

Définition 4.5.1 La matrice P des coordonnées des vecteurs de la nouvelle base $\mathcal{B}'$ dans l'ancienne base $\mathcal{B}$ est appelée matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. On la note $P_{\mathcal{B},\mathcal{B}'}$.

Attention ! à l'ordre des bases dans cette définition. $P_{\mathcal{B},\mathcal{B}'}$ est la matrice dont les colonnes sont les coordonnées dans la base $\mathcal{B}$ des vecteurs de la base $\mathcal{B}'$.

Exemple 4.5.2 Soit $\mathcal{B} = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $e'_1 = (1, -1)$, $e'_2 = (0, 2)$. Les vecteurs e'_1 et e'_2 ne sont pas colinéaires, donc $\mathcal{B}' = (e'_1, e'_2)$ est une base de $\mathbb{R}^2$. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice

$$P_{\mathcal{B},\mathcal{B}'} = \begin{array}{c} \begin{array}{cc} e'_1 & e'_2 \end{array} \\ \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

Proposition 4.5.3 Soit E un espace vectoriel de dimension $n \in \mathbb{N}^*$. Soit $\mathcal{B} = (u_1, u_2, \dots, u_n)$ et $\mathcal{B}' = (u'_1, u'_2, \dots, u'_n)$ deux bases de E. La matrice de de passage de $\mathcal{B}$ dans $\mathcal{B}'$ est la matrice de l'application identité de E dans les bases $\mathcal{B}'$ et $\mathcal{B}$. Autrement dit

$$P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}'}(I).$$

Preuve.— Il suffit de remarquer que

$$P_{\mathcal{B},\mathcal{B}'} = \begin{array}{c} \begin{array}{cccc} u'_1 = I(u'_1) & u'_2 = I(u'_2) & \dots & u'_n = I(u'_n) \end{array} \\ \begin{pmatrix} p_{1,1} & p_{1,2} & & p_{1,n} \\ p_{2,1} & p_{2,2} & & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & p_{n,2} & & p_{n,n} \end{pmatrix} \end{array} \begin{array}{c} \\ u_1 \\ u_2 \\ \\ u_n \end{array} = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}'}(I)$$

□

Remarque 4.5.4 Une matrice de passage est toujours inversible, puisque l'application linéaire identité l'est (cf. la proposition 4.3.15).

De la définition ci-dessus découle l'importante

Proposition 4.5.5 Une matrice $P \in \mathcal{M}_n(\mathbb{R})$ est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ si et seulement si, pour tout vecteur $x \in E$, ses coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}$ et $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ dans la base $\mathcal{B}'$ vérifent $X = P \times X'$.

Preuve.— Supposons que $X = PX'$ pour tout vecteur $x \in E$. C'est en particulier le cas pour tous les vecteurs $u'_1, \dots, u'_n$ de la base $\mathcal{B}'$. Les coordonnées du vecteur u'_k dans la base $\mathcal{B}'$ sont $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, où le 1 figure sur la k-ième ligne. Or $P \times \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ est la k-ième colonne de P. Donc cette colonne est bien constituée des coordonnées de u'_k dans la base $\mathcal{B}$.

Réciproquement supposons que $P = P_{\mathcal{B},\mathcal{B}'}$. Il suffit de vérifier la proposition pour les vecteurs de la base $\mathcal{B}'$. Le résultat pour les autres vecteurs en découlera par linéarité. Les coordonnées du vecteur u'_k dans la base $\mathcal{B}'$ sont $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, où le 1 figure sur la k-ième ligne. On calcule donc

$P_{\mathcal{B},\mathcal{B}'} \times \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} p_{1,k} \\ p_{2,k} \\ \vdots \\ p_{n,k} \end{pmatrix}$, ce qui est bien, par définition de $P_{\mathcal{B},\mathcal{B}'}$, le vecteur colonne des coordonnées de u'_k dans la base $\mathcal{B}$.

□

Exemple 4.5.6 On reprend les vecteurs de l'exemple 4.5.2 : $\mathcal{B} = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $\mathcal{B}' = (e'_1, e'_2)$ avec $e'_1 = (1, -1)$, $e'_2 = (0, 2)$. Si $x = (x_1, x_2) \in \mathbb{R}^2$, on a $x = x_1e_1 + x_2e_2$ et $x = x'_1e'_1 + x'_2e'_2$ avec

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}.$$

On vérifie en effet que

$$x_1.e_1 + x_2.e_2 = (1.x'_1 + 0.x'_2)e_1 + (-1.x'_1 + 2.x'_2)e_2 = x'_1.(e_1 - e_2) + x'_2.(2e_2) = x'_1e'_1 + x'_2e'_2.$$

Proposition 4.5.7 Si P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

Preuve.— Soit $x \in E$, X le vecteur colonne des coordonnées de x dans $\mathcal{B}$ et X' celui des coordonnées de x dans $\mathcal{B}'$. D'après la proposition 4.5.5 on a $X = P_{\mathcal{B},\mathcal{B}'}X'$, donc compte tenu de la remarque 4.5.4, $X' = (P_{\mathcal{B},\mathcal{B}'})^{-1}X$. Donc, d'après la proposition 4.5.5, $P_{\mathcal{B}',\mathcal{B}} = (P_{\mathcal{B},\mathcal{B}'})^{-1}$.
□

4.5.2 Matrices semblables

Proposition 4.5.8 Soit $f \in \mathcal{L}(E, F)$, et A la matrice de f dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$. Soit aussi $\mathcal{B}'_E$ et $\mathcal{B}'_F$ des nouvelles bases de E et F respectivement. Alors la matrice A' de f dans les bases $\mathcal{B}'_E$ et $\mathcal{B}'_F$ est donnée par

$$A' = (P_{\mathcal{B}_F,\mathcal{B}'_F})^{-1} \times A \times P_{\mathcal{B}_E,\mathcal{B}'_E}$$

On notera que cette relation s'écrit

$$\mathrm{Mat}_{\mathcal{B}'_F \leftarrow \mathcal{B}'_E}(f) = (P_{\mathcal{B}_F,\mathcal{B}'_F})^{-1} \times \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) \times P_{\mathcal{B}_E,\mathcal{B}'_E}$$

ou encore

$$\mathrm{Mat}_{\mathcal{B}'_F \leftarrow \mathcal{B}'_E}(f) = \mathrm{Mat}_{\mathcal{B}'_F \leftarrow \mathcal{B}_F}(I) \times \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) \times \mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}'_E}(I).$$

Preuve.— Soit x un vecteur de E, de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathcal{B}_E$ et de coordonnées $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ dans la base $\mathcal{B}'_E$. On a $X = P_{\mathcal{B}_E,\mathcal{B}'_E} X'$, donc $A \times X = A \times P_{\mathcal{B}_E,\mathcal{B}'_E} \times X'$ est le vecteur colonne des coordonnées de $f(x)$ dans la base $\mathcal{B}_F$. On obtient les coordonnées de ce vecteur dans la base $\mathcal{B}'_F$ en le multipliant par $P_{\mathcal{B}'_F,\mathcal{B}_F} = (P_{\mathcal{B}_F,\mathcal{B}'_F})^{-1}$. □

Corollaire 4.5.9 Soit $f \in \mathcal{L}(E)$ un endomorphisme, et A la matrice de f dans la base $\mathcal{B}$. Soit aussi $\mathcal{B}'$ une nouvelle base de E. Alors la matrice A' de f dans la base $\mathcal{B}'$ vérifie

$$A' = P^{-1}AP,$$

où $P = P_{\mathcal{B},\mathcal{B}'}$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

Définition 4.5.10 On dit que deux matrices A et B sont semblables lorsqu'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$ telle que $A = PBP^{-1}$.

De manière équivalente, deux matrices sont semblables si et seulement si elles représentent la même application linéaire dans des bases différentes.

Annexe A

Un peu de géométrie

A.1 Droites et vecteurs

Une droite du plan $\mathcal{P}$ ou de l'espace $\mathcal{E}$ est déterminée par la donnée de deux points distincts. Si $A \neq B$ sont deux points d'une droite, on la note (AB).

Un vecteur du plan ou de l'espace est la donnée d'une direction (une droite), d'un sens (un des deux sens de parcours sur la droite), et d'une longueur, un réel positif. Si la longueur est nulle, on parle de vecteur nul. Dans ce cas la direction et le sens ne comptent pas : il y a un seul vecteur nul, qu'on note $\vec{0}$.

Deux points A et B distincts définissent un vecteur non-nul : sa direction est (AB), son sens est "de A vers B" et sa longueur la distance entre A et B. On le note $\overrightarrow{AB}$.

Dit autrement, si $\vec{u}$ est un vecteur et A un point du plan, il existe un unique point B du plan tel que $\overrightarrow{AB} = \vec{u}$. On le trouve en traçant la droite qui passe par A dont la direction est celle de $\vec{u}$, et en reportant depuis A la longueur de $\vec{u}$ dans la direction de $\vec{u}$. On dit que le couple (A, B) est un représentant du vecteur $\vec{u}$.

On peut ajouter deux vecteurs $\vec{u}$ et $\vec{v}$: on choisit un point A, puis on note B le point tel que $\overrightarrow{AB} = \vec{u}$ et C le point tel que $\overrightarrow{BC} = \vec{v}$. On note alors $\overrightarrow{AC} = \vec{u} + \vec{v}$ et l'on peut voir qu'on obtiendrait le même vecteur (direction, sens et longueur) en partant d'un autre point que A. Par définition de l'addition des vecteurs, on a donc la relation de Chasles :

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

Si t est un réel non-nul et $\vec{u}$ un vecteur du plan, on note $t.\vec{u}$ le vecteur de même direction que $\vec{u}$, de même sens si $t > 0$ et de sens opposé si $t < 0$, enfin de longueur $|t|$ fois celle de $\vec{u}$. Pour $t = 0$, $t.\vec{u} = \vec{0}$. En particulier $\vec{u} + (-1).\vec{u} = \vec{0}$, et on note $-\vec{u}$ le vecteur $(-1).\vec{u}$.

En termes de vecteurs, la droite (AB) est l'ensemble des points M de $\mathcal{P}$ tels que $\overrightarrow{AM}$ et $\overrightarrow{AB}$ ont la même direction (la droite (AB) !). Autrement dit, il s'agit des points $M \in \mathcal{P}$ tels que

$$\overrightarrow{AM} = t\overrightarrow{AB} \text{ pour un réel } t.$$

On note que A et B appartiennent bien à la droite (AB) : prendre $t = 0$ et $t = 1$.

Au passage on note qu'une droite est déterminé par la donnée d'un point A et d'un vecteur non-nul $\vec{u}$: il suffit de définir le point B tel que $\overrightarrow{AB} = \vec{u}$ pour se retrouver dans la situation précédente. D'ailleurs, on dit que $\vec{u}$ est un vecteur directeur de la droite $\mathcal{D}$ lorsque $\mathcal{D}$ est la direction de $\vec{u}$.

A.2 Plans de l'espace

Un plan de l'espace $\mathcal{E}$ est déterminé par la donnée de trois points non-alignés, c'est-à dire n'appartenant pas à la même droite. Si A, B et C sont trois points non-alignés d'un plan, on le note (ABC). Le plan (ABC) est l'ensemble des points M tels que

$$\overrightarrow{AM} = s\overrightarrow{AB} + t\overrightarrow{AC} \text{ pour des réels } \ s \text{ et } t.$$

Pour $s = 0, t = 0$, $s = 1, t = 0$ et $s = 0, t = 1$ respectivement on retrouve que A, B et C appartiennent à (ABC).

Comme pour les droites, on note qu'un plan est déterminé par la donnée d'un point A et de deux vecteurs non-colinéaires $\vec{u}$ et $\vec{v}$: il suffit de définir B et C les points tels que $\overrightarrow{AB} = \vec{u}$ et $\overrightarrow{AC} = \vec{v}$ pour se retrouver dans la situation précédente.

Si A et B sont deux points distincts d'un plan $\mathcal{P}$ de $\mathcal{E}$, alors la droite (AB) est incluse dans le plan $\mathcal{P}$: il suffit de prendre $t = 0$ dans l'équation ci-dessus.

On dit que des points sont coplanaires lorsqu'ils sont contenus dans un même plan. Cette notion n'est intéressante que lorsqu'il s'agit de 4 points ou plus, puisque trois points sont toujours coplanaires : c'est le cas s'ils sont alignés, et sinon ils définissent un plan.

Des vecteurs $\vec{u}$, $\vec{v}$ et $\vec{w}$ sont dits coplanaires si, étant donné un point A de l'espace, et les points B, C et D tels que $\overrightarrow{AB} = \vec{u}$, $\overrightarrow{AC} = \vec{v}$ et $\overrightarrow{AD} = \vec{w}$, les points A, B, C et D sont coplanaires. On peut voir que cette définition ne dépend pas du choix de A.

A.3 Avec des coordonnées, dans le plan

On fixe un repère $(O, \vec{\imath}, \vec{\jmath})$ du plan $\mathcal{P}$, c'est à dire un triplet constitué d'un point et de deux vecteurs non-colinéaires (ie. qui n'ont pas la même direction). Un point A de $\mathcal{P}$ est identifié au couple (x_A, y_A) de coordonnées du vecteur $\overrightarrow{OA}$ définies par

$$\overrightarrow{OA} = x_A\vec{\imath} + y_A\vec{\jmath}.$$

Autrement dit, l'ensemble $\mathcal{P}$ des points du plan est identifé à l'ensemble $\mathbb{R}^2$ des couples de nombres réels au moyen du repère $(O, \vec{\imath}, \vec{\jmath})$.

Voyons ce que cela donne pour les droites du plan.

A.3.1 Equation paramétrique

Le point $M = (x_M, y_M)$ appartient à la droite (AB) si et seulement si il existe un réel t tel que

$$\overrightarrow{AM} = t\overrightarrow{AB} \tag{A.1}$$

Si $A = (x_A, y_A)$ et $B = (x_B, y_B)$, les coordonnées du vecteur $\overrightarrow{AB}$ dans le repère $(O, \vec{\imath}, \vec{\jmath})$ sont donnés par la relation de Chasles

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (x_B - x_A)\vec{\imath} + (y_B - y_A)\vec{\jmath}.$$

De même, le vecteur $\overrightarrow{AM}$ a pour coordonnées $(x_M - x_A, y_M - y_A)$, et la relation (A.1) ci-dessus devient

$$M \in (AB) \iff \exists t \in \mathbb{R}, \left\{ \begin{array}{rcl} x_M - x_A & = & t(x_B - x_A), \\ y_M - y_A & = & t(y_B - y_A), \end{array} \right. \tag{A.2}$$

ou même

$$M \in (AB) \iff \exists t \in \mathbb{R}, \left\{ \begin{array}{rcl} x_M & = & x_A + t(x_B - x_A), \\ y_M & = & y_A + t(y_B - y_A). \end{array} \right.$$

Ce dernier système d'équation est appelé "représentation paramétrique" de la droite (AB).

A.3.2 Equation cartésienne

Comme $A \neq B$, le vecteur $\overrightarrow{AB}$ n'est pas le vecteur nul. Donc l'une au moins de ses coordonnées $x_B - x_A$ ou $y_B - y_A$ n'est pas nulle.

Supposons que $x_B - x_A \neq 0$, autrement dit que $\overrightarrow{AB}$ et $\vec{\jmath}$ ne sont pas colinéaires. L'équivalence (A.2) peut s'écrire

$$M \in (AB) \iff \exists t \in \mathbb{R}, \left\{ \begin{array}{rcl} t & = & \dfrac{x_M - x_A}{x_B - x_A}, \\ y_M - y_A & = & t(y_B - y_A). \end{array} \right.$$

ou encore, puisque la condition sur t est toujours vérifiée,

$$M \in (AB) \iff y_M - y_A = \frac{y_B - y_A}{x_B - x_A}(x_M - x_A).$$

Le nombre $m = \dfrac{y_B - y_A}{x_B - x_A}$ est appelé coefficient directeur de la droite (AB) et $p = y_A - mx_A$ est appelé "ordonnée à l'origine" de (AB). Avec ces notations, on obtient

$$M \in (AB) \iff y_M = mx_M + p, \tag{A.3}$$

et l'équation à droite est appelée "équation cartésienne" de la droite (AB). En prenant $x_M = 0$ on voit que le point $(0, p)$ est le point de (AB) d'abscisse nulle, ce qui justifie la terminologie pour p.

Dans le cas où $x_B - x_A = 0$, autrement dit quand $\overrightarrow{AB}$ et $\vec{\jmath}$ sont colinéaires, l'équivalence (A.2) devient, puisque $y_B - y_A \neq 0$,

$$M \in (AB) \iff \exists t \in \mathbb{R}, \begin{cases} x_M - x_A &= 0, \\ t &= \dfrac{y_M - y_A}{y_B - y_A}, \end{cases}$$

ou simplement

$$M \in (AB) \iff x_M = x_A. \tag{A.4}$$

On a démontré pour l'essentiel le résultat suivant :

Proposition A.3.1 Soit $(O, \vec{\imath}, \vec{\jmath})$ un repère du plan $\mathcal{P}$, et $\mathcal{D}$ une droite de $\mathcal{P}$. Il existe trois nombres réels a, b et c avec $(a, b) \neq (0, 0)$ tels que

$$M = (x, y) \in \mathcal{D} \iff ax + by = c.$$

Preuve.— Soit $\mathcal{D}$ une droite du plan. Dans le cas où $\vec{\jmath}$ est un vecteur directeur de $\mathcal{D}$, il suffit de prendre $a = 1$, $b = 0$ et $c = x_A$ où A est un point quelconque de $\mathcal{D}$: on retrouve alors (A.4). Sinon, on peut prendre $a = -m$, $b = 1$ et $c = p$ où m et p sont définis comme ci-dessus à partir de la donnée de deux points distincts de $\mathcal{D}$, et l'on retrouve (A.3). □

L'équation $ax + by = c$ de la proposition est appelée équation cartésienne de la droite $\mathcal{D}$.

Remarque A.3.2 Si $ax + by = c$ est une équation cartésienne de $\mathcal{D}$, alors, pour tout $\alpha \neq 0$, l'ensemble des points (x, y) qui vérifient l'équation $\alpha ax + \alpha by = \alpha c$ est aussi $\mathcal{D}$. Autrement dit pour tout $\alpha \neq 0$, $\alpha ax + \alpha by = \alpha c$ est aussi une équation cartésienne de $\mathcal{D}$.

Cette proposition admet une réciproque :

Proposition A.3.3 Soit $\mathcal{D}$ l'ensemble de points (x, y) du plan tels que $ax + by = c$ où a, b et c sont trois réels donnés tels que $(a, b) \neq (0, 0)$. Alors $\mathcal{D}$ est une droite.

Preuve.— Supposons que $b \neq 0$. Les points $A(0, \frac{c}{b})$ et $B(b, \frac{c}{b} - a)$ sont deux points distincts de $\mathcal{D}$. Le vecteur $\overrightarrow{AB} = (b, -a)$ n'est pas nul, et si $M = (x, y)$ appartiennent à $\mathcal{D}$, on a, puisque $ax + by = c$,

$$\overrightarrow{AM} = (x, y - \frac{c}{b}) = \frac{x}{b}\overrightarrow{AB}.$$

Autrement dit $\mathcal{D}$ est la droite (AB).

Lorsque $b = 0$, on sait que a n'est pas nul. Dans ce cas $A = (\frac{c}{a}, 0)$ et $B = (\frac{c}{a}, 1)$ sont deux points distincts de $\mathcal{D}$. Si $M = (x, y)$ appartiennent à $\mathcal{D}$, on a $\overrightarrow{AB} = (0, 1)$ et, puisque $ax = c$,

$$\overrightarrow{AM} = (x - \frac{c}{a}, y) = y\overrightarrow{AB},$$

ce qui prouve encore que $\mathcal{D}$ est la droite (AB). □

A.4 Dans l'espace, avec des coordonnées

On fixe un repère $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ de l'espace $\mathcal{E}$, c'est à dire un quadruplet constitué d'un point et de trois vecteurs non-coplanaires. Un point A de $\mathcal{E}$ est identifié au triplet (x_A, y_A, z_A) de coordonnées du vecteur $\overrightarrow{OA}$ définies par

$$\overrightarrow{OA} = x_A\vec{\imath} + y_A\vec{\jmath} + z_A\vec{k}.$$

Autrement dit, l'ensemble $\mathcal{E}$ des points de l'espace est identifié à l'ensemble $\mathbb{R}^3$ des triplets de nombres réels au moyen du repère $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$.

A.4.1 Plans

Equation paramétrique

Soit $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ et $C = (x_C, y_C, z_C)$ trois points non-alignés de $\mathcal{E}$. Le point $M = (x, y, z)$ appartient au plan (ABC) si et seulement si il existe deux réels s et t tels que

$$\overrightarrow{AM} = s\overrightarrow{AB} + t\overrightarrow{AC} \text{ pour des réels } s \text{ et } t.$$

En écrivant les coordonnées de chacun des vecteurs ci-dessus, on obtient

$$M = (x, y, z) \in (ABC) \Longleftrightarrow \left\{ \begin{array}{rcrcl} x - x_A & = & s(x_B - x_A) & + & t(x_C - x_A), \\ y - y_A & = & s(y_B - y_A) & + & t(y_C - y_A), \\ z - z_A & = & s(z_B - z_A) & + & t(z_C - z_A). \end{array} \right.$$

C'est ce que l'on appelle une équation paramétrique (ou un système d'équations paramétriques) du plan (ABC).

Equation cartésienne

Comme pour les droites du plan, on a aussi une notion d'équation cartésienne.

Proposition A.4.1 Soit P un plan de l'espace. Il existe quatre réels a, b, c, d tels que $M = (x, y, z)$ appartient à P si et seulement si $ax+by+cz = d$. On dit que l'équation $ax+by+cz = d$ est une équation cartésienne de P.

Réciproquement si a, b, c, d sont quatre réels donnés, avec $(a, b, c) \neq (0, 0, 0)$, l'ensemble des points $M = (x, y, z)$ de l'espace qui vérifient l'équation $ax + by + cz = d$ est un plan.

La preuve de cette proposition est plus délicate que dans le cas des droites du plan, mais l'idée est la même : on part de l'équation paramétrique, écrite sous la forme d'un système de 3 équations à cinq inconnues (x, y, z, s, t),

$$\left\{ \begin{array}{l} x - s(x_B - x_A) - t(x_C - x_A) = x_A, \\ y - s(y_B - y_A) - t(y_C - y_A) = y_A, \\ z - s(z_B - z_A) - t(z_C - z_A) = z_A. \end{array} \right.$$

On élimine les deux inconnues s et t dans l'une des équations, et l'équation qu'on obtient est une condition nécessaire et suffisante pour que le système ait une solution : c'est l'équation cartésienne recherchée. En fait, le but du premier chapitre de ces notes est de montrer comment faire cela et de prouver que c'est possible, bien sûr dans un cadre beaucoup plus général.

A.4.2 Droites

Soit $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ deux points distincts de l'espace. Le point $M = (x, y, z)$ appartient à la droite (AB) si et seulement si il existe un réel s tel que

$$\overrightarrow{AM} = s\overrightarrow{AB}.$$

En coordonnées cela donne le système d'équations paramétrique de la droite (AB) :

$$M = (x, y, z) \in (AB) \Longleftrightarrow \left\{ \begin{array}{lcl} x - x_A & = & s(x_B - x_A), \\ y - y_A & = & s(y_B - y_A), \\ z - z_A & = & s(z_B - z_A). \end{array} \right.$$

On ne sera pas surpris par la forme, données ci-dessous, des équations cartésiennes de droites dans l'espace si l'on pense qu'une droite de l'espace est l'intersection de deux plans.

Proposition A.4.2 Soit D une droite de l'espace. Il existe huit nombres réels $a_1, a_2, b_1, b_2, c_1, c_2, d_1$ et d_2 tels que

$$M = (x, y, z) \in D \Longleftrightarrow \left\{ \begin{array}{l} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2. \end{array} \right.$$

Réciproquement, si $a_1, a_2, b_1, b_2, c_1, c_2, d_1$ et d_2 sont huit nombres réels donnés tels que $(a_1, b_1, c_1) \neq (0, 0, 0)$ et $(a_2, b_2, c_2) \neq (0, 0, 0)$, alors l'ensemble des points $M = (x, y, z)$ de l'espace qui vérifient le système

$$\left\{ \begin{array}{l} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2. \end{array} \right.$$

est une droite.

La preuve là est plus facile : c'est la même que pour les droites du plan. On exprime l'inconnue s en fonction de l'une des trois autres inconnues x, y ou z (c'est possible parce que $A \neq B$...) et on remplace s par l'expression obtenue dans les deux autres équations.

Annexe B

L'alphabet grec ancien

minuscule	**majuscule**	**nom**	**équivalent latin**
α	A	alpha	a
β	B	beta	b
γ	Γ	gamma	g
δ	Δ	delta	d
ε	E	épsilon	é
ζ	Z	dzeta	z
η	H	èta	è
θ	Θ	théta	th
ι	I	iota	i
κ	K	kappa	k
λ	Λ	lambda	l
μ	M	mu	m
ν	N	nu	n
ξ	Ξ	xi	x
o	O	omicron	o
π	Π	pi	p
ρ	P	rho	r
σ	Σ	sigma	s
τ	T	tau	t
υ	Y	upsilon	u
φ	Φ	phi	f
χ	X	chi	kh
ψ	Ψ	psi	ps
ω	Ω	oméga	ô

Annexe C

Bugs corrigés

version 1.02 $\rightarrow$ 1.03

— Le calcul de l'exemple 3.2.1 contenait plusieurs erreurs. Corrigées.