Notes de cours
L1 — LM 125

Sophie Chemla

9 septembre 2010

Table des matières

Dans ce cours $(K, +, *)$ désignera $(\mathbb{R}, +, *)$, $(\mathbb{C}, +, *)$ ou $(\mathbb{Q}, +, *)$ mais la théorie développée reste valable pour tout corps commutatif $(K, +, *)$.

Ce cours reprend des parties des cours de

• LM120 écrit par H. Ledret

• LM125 2004-2009 qui reposait sur des modules d'algèbre de l'université en ligne (UeL) adaptés au programme de l'UE. Ces modules sont dûs à une équipe de collègues de l'université Bordeaux 1 animée par J Queyrut.

Chapitre 1

Matrices

Les matrices sont des tableaux de nombres. La résolution d'un certain nombre de problèmes d'algèbre linéaire se ramène à des manipulations sur les matrices. Comme nous le verrons dans le deuxième chapitre, cela est vrai pour la résolution des sytèmes linéaires.

1.1 Matrices : définitions, opérations

1.1.1 Définitions

Définition 1 *Soit deux entiers n et p supérieurs ou égaux à 1. On appelle matrice de type (n,p) à coefficients dans K, un tableau rectangulaire à n lignes et p colonnes d'éléments de K.*

Terminologie et notations :

Un tel tableau est représenté de la manière suivante :

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

Les éléments $a_{i,j}$ de K sont appelés les coefficients de la matrice A. L'élement $a_{i,j}$ désigne le coefficient du tableau situé à l'intersection de la ligne i et de la colonne j. On écrira aussi $A = (a_{i,j})_{(i,j)\in[1,n]\times[1,p]}$ ou s'il n'y a pas d'ambiguité $A = (a_{i,j})$. On dira que $a_{i,j}$ est le terme général de la matrice A.

L'ensemble des matrices à n lignes et p colonnes à coefficients dans K est noté $M_{n,p}(K)$. Les éléments de $M_{n,p}(\mathbb{R})$ (respectivement $M_{n,p}(\mathbb{C})$) sont appelées matrices réelles (respectivement complexes). Les inclusions $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ entrainent les inclusions $M_{n,p}(\mathbb{Q}) \subset M_{n,p}(\mathbb{R}) \subset M_{n,p}(\mathbb{C})$.

Exemple 1 La matrice $A = \begin{pmatrix} 1 & \sqrt{2} \\ 1 & 0 \\ 3 & 1 \\ 0,5 & 0 \end{pmatrix}$ est un élément de $M_{4,2}(\mathbb{R})$.

Définition 2 (Définition de l'egalité de deux matrices) *Soient n, p, n', p' quatre entiers. On considère $A = (a_{i,j})$ une matrice appartenant à $M_{n,p}(K)$ et $A' = (a'_{i,j})$ une matrice appartenant à $M_{n',p'}(K)$. On dit que ces matrices sont égales si et seulement si*

$$\begin{array}{l} n = n', p = p' \\ \forall (i,j) \in [1,n] \times [1,p],\ \ a_{i,j} = a'_{i,j} \end{array}$$

Matrices particulières

Soient n et p deux entiers supérieurs ou égaux à 1.

Une matrice qui a une seule ligne est appelée matrice ligne. Si elle a p colonnes, on la note

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \end{pmatrix}$$

De même, une matrice qui a une seule colonne est appelée matrice colonne. Si elle a n lignes, on la note

$$A = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix}$$

Définition 3 (Matrice carrée) *Une matrice qui a le même nombre de lignes et de colonnes est appelée matrice carrée. Si ce nombre est l'entier n, on dit que la matrice est d'ordre n et on note $M_n(K)$ au lieu de $M_{n,n}(K)$, l'ensemble des matrices carrées d'ordre n à coefficients dans K.*

Sur une matrice carrée, on a la notion de diagonale principale.

Définition 4 (Définition de la diagonale principale d'une matrice carrée) *Soit*

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

une matrice carrée d'ordre n. Sa diagonale principale est la diagonale $(a_{1,1}, a_{2,2} \dots, a_{n,n})$.

Exemple 2 la matrice réelle $\begin{pmatrix} 1 & 0 & 9 \\ 2 & \sqrt{2} & 2 \\ 3 & 5 & 9 \end{pmatrix}$ est carrée d'ordre 3. Les termes de sa diagonale principale sont $a_{1,1} = 1, a_{2,2} = \sqrt{2}, a_{3,3} = 9$.

Il y a un certain nombre de cas particuliers de matrices carrées intéressants.

Définition 5 (Matrices triangulaires) *Une matrice carrée d'ordre n de terme général $a_{i,j}$ est triangulaire supérieure si pour tout entier $i \in [1,n]$ et tout entier j tel que $1 \leq j < i$, $a_{i,j} = 0$.*

Une matrice carrée d'ordre n de terme général $a_{i,j}$ est triangulaire inférieure si pour tout entier $i \in [1,n]$ et tout entier j tel que $i < j \leq n$, $a_{i,j} = 0$.

Une matrice carrée d'ordre n de terme général $(a_{i,j})$ est diagonale si, pour tout couple $(i,j) de [1,n] \times [1,n]$ tel que $i \neq j$, on a $a_{i,j} = 0$.

Une matrice triangulaire supérieure (respectivement inférieure) est une matrice carrée dont tous les termes "en dessous" (respectivement "au dessus") de la diagonale principale sont nuls.

Une matrice diagonale est une matrice carrée dont les termes situés hors de la diagonale principale sont tous nuls.

Exemple 3 La matrice réelle $\begin{pmatrix} 1 & 0 & 9 \\ 0 & \sqrt{2} & 2 \\ 0 & 0 & 9 \end{pmatrix}$ est triangulaire supérieure.

La matrice réelle $\begin{pmatrix} 1 & 0 & 0 \\ 2 & \sqrt{2} & 0 \\ 3 & 5 & 9 \end{pmatrix}$ est triangulaire inférieure.

La matrice réelle $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 9 \end{pmatrix}$ est diagonale.

Cas particulier important :

La matrice diagonale d'ordre n dont les termes de la diagonale principale sont tous égaux à 1 est appelée matrice unité et est notée I_n

Soit

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

Définition 6 (Transposée d'une matrice) *Soit $A = (a_{i,j})$ un élément de $M_{n,p}(K)$. On appelle transposée de A et on note A^T la matrice à p lignes et n colonnes de terme général $b_{k,l}$ défini par :*

$$\forall k,\ 1 \leq k \leq p,\ \forall l, 1 \leq l \leq n\ \ b_{k,l} = a_{l,k}.$$

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

Remarque 1 La $i$ème ligne de A devient la $i$ème colonne de A^T.

Notation : La transposée de la matrice A se note aussi ${}^tA$.

1.1.2 Opérations sur les matrices

Commençons par définir la somme de deux matrices.

Soient n et p deux entiers supérieurs ou égaux à 1. On ne considère que des matrices de même type appartenant à $M_{n,p}(K)$. Si l'on a des matrices de types différents, parler de leur somme n'a aucun sens !

Définition 7 *Soient $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices appartenant à $M_{n,p}(K)$. On appelle somme des matrices A et B, et l'on note $A+B$, la matrice appartenant à $M_{n,p}(K)$ de terme général la somme des termes généraux de A et B. Autrement dit on a $A+B = (c_{i,j})_{(i,j)\in[1,n]\times[1,p]}$ avec*

$$\forall (i,j) \in [1,n]\times[1,p],\ \ c_{i,j} = a_{i,j} + b_{i,j}.$$

Exemple 5 Dans $M_{2,3}$, on a la somme

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

Proposition 1 (Propriété de la somme de deux matrices) *Soient A, B et C trois éléments de $M_{n,p}$*

1) On a $A+B = B+A$ (on dit que l'addition est commutative).

2) On a $(A+B)+C = A+(B+C)$ (on dit que l'addition est associative).

3) Si on note $0_{n,p}$ la matrice, élément de $M_{n,p}(K)$, dont tous les coefficients sont nuls, on a $A+0_{n,p} = A$.

4) Si $A = (a_{i,j})$ et $A' = (-a_{i,j})$, on a $A+A' = 0_{n,p}$.

Démonstration. 1) Si $A = (a_{i,j})$ et $B = (b_{i,j})$, $A+B$ est la matrice de terme général $a_{i,j}+b_{i,j}$ et $B+A$ est la matrice de terme général $b_{i,j}+a_{i,j}$. Comme on a l'égalité $a_{i,j}+b_{i,j} = b_{i,j}+a_{i,j}$ dans K, on en déduit $A+B = B+A$.

2) La justification est semblable à la précédente : c'est une conséquence directe de la propriété $a+(b+c) = (a+b)+c$ vraie pour tout élément a, b, et c de K.

3) La justification est semblable à la précédente : c'est une conséquence directe de la propriété $a+0 = a$, vraie pour tout élément a de K.

4) La justification est semblable à la précédente : c'est une conséquence directe de la propriété $a+(-a) = 0$ vraie pour tout élément a de K.

Remarque 2 *Lorsqu'il n'y aura pas d'ambiguité, la matrice $0_{n,p}$ sera notée 0.*

Proposition 2 *La transposée de la somme de deux matrices est la somme des matrices transposées.*

La démonstration est laissée au lecteur. Elle se fonde sur la définition de la transposée d'une matrice et sur celle de l'addition de deux matrices. □

Nous allons introduire une nouvelle opération sur $M_{n,p}(K)$: la multiplication d'une matrice par un scalaire.

Définition 8 (Multiplication d'une matrice par un scalaire) *Soient $A = (a_{i,j})$ une matrice appartenant à $M_{n,p}(K)$ et un élément α de K. On désigne par αA la matrice appartenant à $M_{n,p}(K)$ dont le terme général est le produit par α du terme général de A. On a donc $\alpha A = (\alpha a_{i,j})$. On dit que est αA est le produit de la matrice A par le scalaire α.*

Remarque 3 Il est clair que le fait que le produit de deux éléments de K soit encore un élément de K, est essentiel dans cette définition.

Exemple 6 Dans $M_{2,3}(\mathbb{R})$, on a l'égalité :

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

Dans $M_3(\mathbb{C})$, on a

$$i\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} i & i & i \\ i & i & i \\ i & i & i \end{pmatrix}$$

Énonçons maintenant quelques propriétés de la multiplication par un scalaire.

Proposition 3 *Soit $A = (a_{i,j})$ une matrice quelconque appartenant à $M_{n,p}(K)$*
a) On a $1A = A$.
b) Soient α et β deux éléments quelconques de K, on a $\alpha(\beta A) = (\alpha\beta)A$
c) Si α est un élément quelconque de K, on a $(\alpha A)^T = \alpha A^T$.

Démonstration. a) provient de l'égalité dans K, valable pour tout a de K : $1a = a$.
b) Cela provient de l'égalité dans K, valable pour tout a : $\alpha(\beta a) = (\alpha\beta)a$
c) Cela provient immédiatement des définitions de la transposée d'une matrice et du produit d'une matrice par un scalaire.

Énonçons maintenant deux propriétés liant ces deux relations :

Proposition 4 *Soient $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices quelconques appartenent à $M_{n,p}(K)$. Soient α et β deux scalaires. On a*
a) $(\alpha+\beta)A = \alpha A + \beta A$.
b) $\alpha(A+B) = \alpha A + \alpha B$

Démonstration. a) En effet, le terme général de $(\alpha+\beta)A$ est égal à $(\alpha+\beta)a_{i,j}$. D'après les règles de calcul dans K, $(\alpha+\beta)a_{i,j}$ est égal à $\alpha a_{i,j} + \beta a_{i,j}$ qui est le terme général de la matrice $\alpha A + \beta A$.
b) La démonstration de b) est semblable à la précédente.

1.2 Produit de matrices

Soient A et B deux matrices. On suppose que le nombre de colonnes de A est égal au nombre de lignes de B. Dans ces conditions, on va pouvoir définir le produit AB. Soient n, p et q trois entiers supérieurs ou égaux à 1.

Définition 9 *Soit $A = (a_{i,j})$ une matrice à n lignes et p colonnes et soit $B = (b_{i,j})$ une matrice à p lignes et q colonnes. Alors AB est la matrice à n lignes et q colonnes dont le terme général $c_{i,j}$ est donné par la formule*

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

Remarque 4 On peut écrire le coefficient de façon plus développée, à savoir : .

$$c_{i,k} = a_{i,1}b_{1,k} + a_{i,2}b_{2,k} + \cdots + a_{i,j}b_{j,k} + \cdots + a_{i,p}b_{p,k}.$$

Sous cette forme, on comprend mieux la contrainte imposée sur les nombres de colonnes et de lignes de A et B respectivement. Donc, pour avoir l'élément de la i-ième ligne k-ième colonne de AB, on prend la i-ième ligne de la matrice qui est à gauche c'est-à-dire A, la k-ième colonne de la matrice qui est à droite c'est-à-dire B ; on fait le produit du premier élément de la ligne par le premier élément de la colonne, puis du deuxième élément de la ligne par le deuxième élément de la colonne et ainsi de suite jusqu'au produit du p-ième élément de la ligne par le p-ième élément de la colonne et on fait la somme de tous ces produits. Il est commode quand on débute de disposer les calculs de la façon suivante.

$$\begin{array}{cc} & \begin{pmatrix} & & \times & \\ & & \times & \\ & & \times & \\ & & \times & \end{pmatrix} \leftarrow B \\ A \rightarrow \begin{pmatrix} & & & \\ \times & \times & \times & \times \\ & & & \end{pmatrix} & \begin{pmatrix} & & & | & & \\ & & & | & & \\ - & - & - & | & - & - \\ & & & | & & \end{pmatrix} \leftarrow AB \end{array}$$

On a fait apparaître une ligne générique de A et une colonne générique de B avec les coefficients qui doivent être multipliés les uns avec les autres (représentés par des $\times$ dans l'ordre de gauche à droite dans A et de haut en bas dans B) puis additionnés pour donner le coefficient de AB situé à l'intersection de cette ligne et de cette colonne.

Avec un peu plus de pratique, on pose directement l'opération en ligne comme dans l'exemple ci-dessous.

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 5 & 11 \end{pmatrix}.$$

Remarque 5 Écrivons $B = (b_1 b_2 \ldots b_q)$ où b_i désigne la i-ème colonne de B. Le produit Ab_i est défini et est élément de $M_{n,1}(K)$. On a

$$AB = (Ab_1 Ab_2 \ldots Ab_q).$$

Remarque 6 *Deux erreurs grossières à éviter.* Les règles du calcul des produits de matrices diffèrent de celles des produits dans un corps par d'autres aspects.

i) Si $AB = AC$, on ne peut pas simplifier par A pour en déduire que $B = C$ même si A n'est pas nulle. C'est faux en général comme le montre l'exemple ci-dessous.

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

ii) Si $AB = 0$, on ne peut pas en déduire que soit $A = 0$ soit $B = 0$. C'est faux en général.

Si on a un exemple de ii), on a aussi un exemple de i) puisque $0 = A \times 0$. Il suffit de prendre

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \text{ Alors, } AB = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

mais ni A ni B n'est nulle. □

Proposition 5 (Propriétés du produit des matrices) *a) Associativité du produit : Soient $A \in M_{n,p}$, $B \in M_{p,q}$ et $C \in M_{q,r}$. Alors les produits AB, $(AB)C$, BC, $A(BC)$ ont un sens et l'on a l'égalité suivante dans $M_{n,r}(K)$:*

$$(AB)C = A(BC).$$

b) Distributivité à droite du produit par rapport à la somme : Soient $A \in M_{n,p}(K)$, $B \in M_{n,p}(K)$ et $C \in M_{p,q}(K)$. Alors $A+B$ et les produits $(A+B)C$, AC, BC ont un sens et on a l'égalité dans $M_{n,q}(K)$

$$(A+B)C = AC + BC.$$

c) Distributivité à gauche du produit par rapport à la somme : Soient $A \in M_{n,p}(K)$, $B \in M_{n,p}(K)$ et $C \in M_{p,q}(K)$. Alors $A+B$ et les produits $(A+B)C$, AC, BC ont un sens et on a l'égalité dans $M_{n,q}(K)$

$$A(B+C) = AB + AC.$$

d) Comportement du produit des matrices par rapport au produit par un scalaire : Soient , $A \in M_{n,p}(K)$, $B \in M_{p,q}$ deux matrices et λ un scalaire. alors les produits AB, $(\lambda A)B$, $A(\lambda B)$ et $\lambda(AB)$ Alors les produits ont un sens et on a les égalités dans :

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

Démonstration. Posons $A = (a_{i,j})$, $B = (b_{i,j})$ et $C = (c_{i,j})$.

a) Le terme d'indice (i,j) de la matrice AB est $x_{i,k} = \sum_{l=1}^{p} a_{i,l} b_{l,k}$. Le terme d'indice (i,j) de la matrice $(AB)C$ est donc

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

Le terme d'indice (l,j) de la matrice BC est $y_{l,j} = \sum_{k=1}^{p} b_{l,k} c_{k,j}$. Le terme d'indice (i,j) de la matrice $A(BC)$ est

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

Comme la multiplication est distributive sur l'addition et associative dans K, les termes généraux de $(AB)C$ et $A(BC)$ coincident.

Les démonstrations de b) et c) et d) se font comme celle de l'associativité. □

Proposition 6 (Produit d'une matrice par une matrice unité) *Notons I_r la matrice unité d'ordre r. Si $A \in M_{n,p}(K)$, on a les propriétés suivantes :*

$$AI_p = A \text{ et } I_n A = A.$$

En particulier, si A est une matrice carrée d'ordre n, on a

$$AI_n = I_n A = A.$$

Démonstration. Soit $A \in M_{n,p}(K)$ de terme général $a_{i,j}$. La matrice unité d'ordre p est telle que tous les éléments de la diagonale principale sont égaux à 1, les autres étant tous nuls. On peut formaliser cela en introduisant le symbole de Kronecker.

Si r et s sont deux entiers, on appelle symbole de Kronecker et on note $\delta_{r,s}$ le réel qui vaut 0 si r est différent de s, et 1 si r est égal à s. Donc

$$\begin{array}{l} \delta_{r,s} = 0 \text{ si } r \neq s \\ \delta_{r,s} = 1 \text{ si } r = s \end{array}$$

Alors on peut dire que le terme général de la matrice carrée d'ordre p, I_p, est $\delta_{r,s}$ avec r et s entiers, compris entre 1 et p.

Alors la matrice produit AI_p est une matrice appartenant à $M_{n,p}(K)$ dont le terme général $c_{i,l}$ est donné par la formule $c_{i,l} = \sum_{j=1}^{p} a_{i,j}\delta_{j,l}$. Dans cette somme i et l sont fixés et j prend toutes les valeurs comprises entre 1 et p. Si j est différent de l, $\delta_{j,l} = 0$, et si j est égal à l, $\delta_{l,l} = 1$.

Donc dans la somme qui définit $c_{i,l}$, tous les termes correspondant à des valeurs de j différentes de l sont nuls et il reste donc $c_{i,l} = a_{i,l}\delta_{l,l} = a_{i,l}1 = a_{i,l}$. Donc les matrices AI_p et A ont le même terme général et sont donc égales.

L'égalité $I_nA = A$ se démontre de la même façon. □

Proposition 7 (Produit d'une matrice par la matrice nulle) *Soit A un élément de $M_{n,p}(K)$, et soit $0_{p,q}$ (respectivement $0_{r,n}$) la matrice de $M_{p,q}(K)$ (respectivement $M_{r,n}(K)$) dont tous les éléments sont nuls. On a les propriétés $A0_{p,q} = 0_{n,q}$ et $0_{r,n}A = 0_{r,p}$.*

La démonstration de cette proposition est laissée au lecteur.

Proposition 8 *Si le produit AB est défini, alors le produit B^TA^T est aussi défini et l'on a $(AB)^T = B^TA^T$.*

Démonstration. Soit $A = (a_{ij}) \in M_{mn}(K)$ et $B = (b_{jk}) \in M_{np}(K)$, d'où $AB \in M_{mp}(K)$. On voit donc que $B^T \in M_{pn}(K)$ et $A^T \in M_{nm}(K)$. Par conséquent, B^TA^T est bien défini et de la même taille que $(AB)^T$.

Utilisons la formule générale ci-dessus.

$$(B^TA^T)_{ik} = \sum_{j=1}^{n} (B^T)_{ij}(A^T)_{jk} = \sum_{j=1}^{n} b_{ji}a_{kj} = \sum_{j=1}^{n} a_{kj}b_{ji} = (AB)_{ki} = ((AB)^T)_{ik}$$

d'où le résultat. □

Remarque 7 ATTENTION ! Le produit matriciel *n'est pas commutatif.* En effet, il peut se faire que AB soit défini mais pas BA, ou que AB et BA soient tous deux définis mais pas de la même taille. Mais même dans le cas où AB et BA sont définis et de la même taille, on a en général $AB \neq BA$. Considérons l'exemple suivant.

$$\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 3 \\ -2 & -6 \end{pmatrix} \text{ mais } \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 29 & -2 \end{pmatrix}.$$

C'est là la situation générale. L'ordre des facteurs dans un produit matriciel ne doit donc *jamais* être modifié, sous peine de fausser le résultat (sauf si l'on sait que l'on est dans un cas particulier où deux matrices *commutent*, c'est-à-dire sont telles que $AB = BA$. Mais c'est rare...).

En fait, le produit matriciel est le premier exemple que l'on rencontre de produit non commutatif. □

1.3 Matrices carrées, matrices carrées inversibles

Dans l'ensemble $M_n(K)$ des matrices carrées d'ordre n, (c'est à dire celles qui ont n lignes et n colonnes) à coefficients dans K, la multiplication des matrices est une opération interne. De plus, si I_n désigne la matrice carrée unité d'ordre n , on a :

$$\forall A \in M_n(K),\ \ AI_n = I_nA = A.$$

On peut aussi définir les puissances successives d'une matrice.

Définition 10 *Pour tout $A \in M_n(K)$, on définit les puissances successives de A par $A^0 = I_n$ et $A^{p+1} = AA^p = A^pA$ pour tout $p \in \mathbb{N}$.*

Exemple 7 On cherche à calculer A^p avec $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. On calcule A^2 et A^3 et on obtient

$$A^2 = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix} = 3A,\ \ A^3 = 3A.A = 3^2A.$$

L'observation de ces premières puissances permet de penser que la formule est : $A^{p+1} = 3^pA = \begin{pmatrix} 3^p & 3^p & 3^p \\ 3^p & 3^p & 3^p \\ 3^p & 3^p & 3^p \end{pmatrix}$. Démontrons ce résultat par récurrence.

Il est vrai pour $p = 0$. On le suppose vrai pour un entier p et on va le démontrer par récurrence pour $p+1$. On a, d'après la définition,

$$A^{p+2} = A^{p+1}A = \begin{pmatrix} 3^p & 3^p & 3^p \\ 3^p & 3^p & 3^p \\ 3^p & 3^p & 3^p \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3^{p+1} & 3^{p+1} & 3^{p+1} \\ 3^{p+1} & 3^{p+1} & 3^{p+1} \\ 3^{p+1} & 3^{p+1} & 3^{p+1} \end{pmatrix}.$$

Donc la propriété est démontrée.

Remarque 8 Ce n'est pas toujours simple de calculer la puissance d'une matrice car la formule de récurrence n'est pas toujours aussi apparente que dans l'exemple qui vient d'être traité. Il existe des méthodes plus systématiques, mais qui sortent du cadre de ce cours. Nous verrons cependant des méthodes qui marchent dans des cas favorables.

Remarque 9 Comme la multiplication n'est pas commutative, les identités binomiales usuelles sont fausses. En particulier, $(A+B)^2 \neq A^2 + 2AB + B^2$, mais bien $(A+B)^2 = A^2 + AB + BA + B^2$. □

On a cependant :

Proposition 9 (Calcul de $(A+B)^n$ lorsque $AB = BA$) *Soient A et B deux éléments de $M_n(K)$ qui commutent c'est à dire tels que $AB = BA$. Alors, pour tout entier m, supérieur ou égal à 1, on a la formule*

$$(A+B)^m = \sum_{k=0}^{m} C_m^k A^{m-k} B^k$$

où C_m^k désigne le coefficient du binôme.

La démonstration se fait par récurrence en utilisant les propriétés bien connues du binôme, à savoir $C_{m-1}^{k-1}+C_{m-1}^{k}=C_m^k$ et $C_m^k=C_m^{m-k}$.

Exemple 8 Soit $M=\begin{pmatrix}1&1&1&1\\0&1&2&1\\0&0&1&3\\0&0&0&1\end{pmatrix}$. On pose $N=M-I=\begin{pmatrix}0&1&1&1\\0&0&2&1\\0&0&0&3\\0&0&0&0\end{pmatrix}$. La matrice N est nilpotente (c'est à dire $\exists k\in\mathbb{N}\mid N^k=0$) comme le montrent les calculs suivants :

$$N^2=\begin{pmatrix}0&0&2&4\\0&0&0&6\\0&0&0&0\\0&0&0&0\end{pmatrix},N^3=\begin{pmatrix}0&0&0&6\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}\ \text{ et }\ N^4=0$$

Comme on a $M=I+N$ et les matrices N et I commutent, on peut appliquer la formule du binôme de Newton. Si k est supérieur ou égale à 3, en utilisant le fait que $N^k=0$ si $k\geq 4$, on obtient

$$\begin{aligned}M^k &= \sum_{l=0}^{k}C_k^lN^lI_n^{k-l}\\ &= \sum_{l=0}^{3}C_k^lN^l\\ &= I_4+kN+\tfrac{k(k-1)}{2!}N^2+\tfrac{k(k-1)(k-2)}{3!}N^3\end{aligned}$$

D'où

$$M^k=\begin{pmatrix}1&k&k^2&k(k^2-k+1)\\0&1&2k&k(3k-2)\\0&0&1&3k\\0&0&0&1\end{pmatrix}.$$

On vérifie immédiatement que cette formule est aussi vraie pour $k=0,1$ ou 2.

Définition 11 *On dit que $A\in M_n(K)$ est* inversible *si et seulement si il existe une matrice $A'\in M_n(K)$ telle que $AA'=A'A=I_n$.*

Notation :

• On note $A'=A^{-1}$, et, plus généralement, $A^{-p}=(A^{-1})^p$ pour tout $p\in\mathbb{N}$ quand A est inversible.

• L'ensemble des matrices inversibles de $M_n(K)$ est noté $GL_n(K)$.

Exemple 9 Soit I_n la matrice carrée unité d'ordre n. C'est une matrice inversible (immédiat à partir de l'égalité $I_nI_n=I_n$).

Exemple 10 La matrice nulle , d'ordre n avec n quelconque, n'est pas inversible. En effet on sait que, pour tout matrice M de $M_n(K)$, on a $M0_n=0_nM=0_n$. Comme la matrice nulle est différente de la matrice unité, on peut conclure.

Exemple 11 Soit $A=\begin{pmatrix}1&1\\0&0\end{pmatrix}$. Etudier si A est inversible, c'est étudier l'existence d'une matrice $B=\begin{pmatrix}a&c\\b&d\end{pmatrix}$ à coefficients dans K, telle que $AB=BA=I_2$. Or $AB=I_2$ équivaut à l'égalité :

$$\begin{pmatrix}1&1\\0&0\end{pmatrix}\begin{pmatrix}a&c\\b&d\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

qui équivaut à :

$$\begin{pmatrix} a+b & c+d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Or les matrices $\begin{pmatrix} a+b & c+d \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ne sont pas égales, puisque les coefficients de la deuxième colonne, deuxième ligne sont différents.

Donc il n'existe pas de matrice B telle que $AB = BA = I_2$ et A n'est pas inversible.

Exemple 12 Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Etudier si A est inversible, c'est étudier l'existence d'une matrice $B = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ à coefficients dans K, telle que $AB = BA = I_2$. Or $AB = I_2$ équivaut à l'égalité :

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Cette égalité équivaut au système : .

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

Sa résolution est immédiate : $a = 1$, $b = 0$, $c = -1$, $d = 1$

Il n'y a donc qu'une seule matrice possible $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. Pour prouver qu'elle convient, il faut montrer l'égalité $BA = I_2$ dont la vérification est laissée au lecteur.

On a donc trouvé une matrice carrée d'ordre 2, $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, telle que $AB = BA = I_2$. La matrice A est donc inversible.

On a remarqué, en cours de calcul, qu'il n'y avait qu'une seule solution possible. En fait c'est une propriété générale.

Remarque 10 La méthode de Gauss fournira une méthode pour calculer l'inverse d'une matrice.

Proposition 10 *Si A est inversible, alors son inverse est unique.*

Démonstration. La méthode classique pour mener à bien une telle démonstration est de supposer l'existence de deux matrices B_1 et B_2 satisfaisant aux conditions imposées et de démontrer que $B_1 = B_2$.

Soit donc B_1 telle que $AB_1 = B_1A = I_n$ et B_2 telle que $AB_2 = B_2A = I_n$. Calculons $B_2(AB_1)$. D'une part, comme $AB_1 = I_n$, on a $B_2(AB_1) = B_2$. D'autre part, comme le produit des matrices est associatif, on a $B_2(AB_1) = (B_2A)B_1 = I_nB_1 = B_1$. Donc $B_1 = B_2$. □

Pour ceux qui connaissent la théorie des groupes, les points a) et b) de la proposition suivante montre que l'ensemble non vide $GL(n,K)$ est un groupe pour la multiplication.

Proposition 11 *Soient A et B deux matrices de $GL_n(K)$.*
a) A^{-1} est inversible et on a $(A^{-1})^{-1} = A$.
b) AB est inversible et on a $(AB)^{-1} = B^{-1}A^{-1}$.
c) A^T est inversible et on a et $(A^T)^{-1} = (A^{-1})^T$.

Démonstration. a) découle de l'identité $AA^{-1} = A^{-1}A = I_n$.
b) En utilisant l'associativité de la multiplication des matrices, on montre

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

La formule pour l'inverse du produit a lieu dans n'importe quel groupe.
c) Pour la transposée, on remarque que comme $I_n = AA^{-1}$ et que $I_n^T = I_n$, on a

$$I_n = (AA^{-1})^T = (A^{-1})^T A^T.$$

De même, le fait que $I_n = A^{-1}A$ implique que $I_n = A^T(A^{-1})^T$, donc A^T est inversible et on a pour inverse la transposée de A^{-1}. □

Si M est une matrice quelconque de $M_n(K)$, nous avons vu que la relation $MA = MB$ où A, et B sont des éléments de $M_n(K)$ n'entraîne pas forcément l'égalité $A = B$. En revanche, si M est une matrice inversible, on a la proposition suivante :

Proposition 12 (Simplification par une matrice inversible) *Soient A et B deux matrices de $M_n(K)$ et M une matrice inversible de $M_n(K)$. Alors l'égalité $MA = MB$ implique l'égalité $A = B$. On dit que M est un élément régulier de $M_n(K)$.*

Démonstration. Ce résultat est immédiat : si on multiplie à gauche l'égalité $MA = MB$ par M^{-1}, on obtient l'égalité : $M^{-1}(MA) = M^{-1}(MB)$. Soit en utilisant l'associativité du produit des matrices $(M^{-1}M)A = (M^{-1}M)B$, ce qui donne d'après la définition de l'inverse $I_nA = I_nB$ d'où $A = B$.

1.4 Algorithme de Gauss sur les matrices

L'algorithme de Gauss (ou encore du pivot de Gauss) est fondé sur les notions de matrices échelonnées réduites et d'opérations élémentaires sur les lignes.

Définition 12 *Une matrice A est dite* échelonnée *si et seulement si elle a les deux propriétés suivantes*
1) *Si une ligne est entièrement nulle, toutes les lignes situées en dessous sont également entièrement nulles.*
2) *Dans chaque ligne non entièrement nulle (à partir de la deuxième), le premier coefficient non nul en comptant à partir de la gauche est situé strictement à droite du premier coefficient non nul de la ligne précédente.*
On dit qu'une matrice est échelonnée réduite *si et seulement elle a en plus les deux propriétés suivantes*
3) *Le premier coefficient non nul d'une ligne en comptant à partir de la gauche vaut* 1.
4) *Et c'est le seul élément non nul de sa colonne.*

Remarque 11 Grâce à 1), on voit que 2) a un sens : si une ligne contient un élément non nul, alors la ligne précédente contient aussi un élément non nul, sinon cela contredirait 1). Par ailleurs, toujours à cause de 2) et de 1), on voit que tous les coefficients situés dans la même colonne qu'un tel premier élément non nul d'une ligne et en dessous de cet élément, sont nuls. □

Définition 13 *Soit U une matrice échelonnée réduite. Les* positions de pivot *de U sont les emplacements (au sens du couple (numéro de ligne, numéro de colonne)) des coefficients valant* 1 *du point 3) de la définition 12.*

Exemple 13

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

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

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

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

Définition 14 *On appelle* opérations élémentaires sur les lignes *les trois opérations suivantes :*

i) Échanger deux lignes (échange).

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

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

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

Exemple 14 Considérons la matrice

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

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

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

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

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

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

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

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

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

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

Notation Soient A et B deux matrices de même taille $m\times n$. Si B se déduit de A par une suite finie d'opérations élémentaires, on notera $A\sim B$.

Proposition 13 *1. Si A est un élément de $M_{m,n}$, on a $A\sim A$ (on dit que la relation $\sim$ est reflexive).*

2. Soient A, B et C trois éléments de $M_{m,n}$. Si $A\sim B$ et $B\sim C$ alors $A\sim C$ (on dit que la relation $\sim$ est transitive).

3. Si $A\sim B$ alors $B\sim A$ (on dit que la relation $\sim$ est symétrique).

Remarque 12 Une relation binaire possédant les trois propriétés, réflexive, symétrique et transitive est appelée relation d'équivalence. Donc la relation binaire $\sim$ définie sur $M_{m,n}(K)$ est une relation d'équivalence sur $M_{m,n}(K)$.

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

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

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

Théorème 1 *Étant donnée une matrice A, il existe une unique matrice échelonnée réduite U obtenue à partir de A par des opérations élémentaires sur les lignes.*

Démonstration. Ce théorème est en deux parties, une partie d'existence (il existe U échelonnée réduite obtenue à partir de A par des opérations élémentaires sur les lignes) et une partie unicité (c'est la seule).

Commençons par l'existence, laquelle se démontre grâce à l'algorithme de Gauss proprement dit. L'idée générale de l'algorithme de Gauss consiste à utiliser des substitutions de lignes pour placer des zéros là où il faut de façon à créer d'abord une forme échelonnée, puis une forme échelonnée réduite. Soit A une matrice $m\times n$ quelconque.

Passage à une forme échelonnée.

Étape 1 : Choix du pivot. On commence par inspecter la première colonne. Soit elle ne contient que des zéros, auquel cas on passe directement à l'étape 3, soit elle contient au moins un terme non nul. On choisit alors un tel terme, que l'on appelle le *pivot.* Si c'est le terme a_{11} on passe directement à l'étape 2, si c'est un terme a_{i1} avec $i \neq 1$, on échange les lignes 1 et i et on passe à l'étape 2.

Au terme de l'étape 1, on a obtenu une matrice de la forme

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

dans le premier cas, ou bien

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

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

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

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

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

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

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

dont la première colonne est bien celle d'une matrice échelonnée. On va donc conserver cette première colonne. Si $a_{11}^1 \neq 0$, on conserve aussi la première ligne, et l'on va boucler sur l'étape 1 en l'appliquant à la sous-matrice $(m-1)\times(n-1)$ qui reste

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

Si $a_{11}^1 = 0$, on boucle sur l'étape 1 en l'appliquant à la sous-matrice $m\times(n-1)$

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

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

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

et ainsi de suite.

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

Passage à une forme échelonnée réduite.

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

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

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

Nous admettrons la partie unicité du théorème.

Remarque 13 Le théorème précédent nous dit que l'on ne peut obtenir qu'une seule matrice échelonnée réduite à partir de A par des opérations élémentaires sur les lignes.

En revanche, on peut obtenir une infinité de matrices échelonnées à partir de A par des opérations élémentaires sur les lignes.

Exemple 15 Soit

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

Passage à une forme échelonnée.

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

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

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

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

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

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

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

Passage à une forme échelonnée réduite.

Étape 1, homothéties. On multiplie la ligne 1 par 1, la ligne 2 par $1/2$ et la ligne 3 par $-1/2$ et l'on obtient

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

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

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

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

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

qui est bien échelonnée réduite.

Le théorème 1 permet d'étendre un certain nombre de définitions aux matrices quelconques.

Définition 15 *Soit A une matrice quelconque et U l'unique matrice échelonnée réduite obtenue à partir de A par des opérations élémentaires sur les lignes. Les* positions, colonnes et lignes de pivot *de A sont les positions, colonnes et lignes de pivot de U.*

Attention ! Les positions de pivot ne sont en général pas apparentes sur la matrice A. Pour les déterminer, Il faut effectivement calculer la matrice U, ou au moins une matrice échelonnée obtenue à partir de A par des opérations élémentaires sur les lignes. Ainsi, dans l'exemple 15, on voit trois positions de pivot : $(1,1)$, $(2,2)$ et $(3,4)$ sur la matrice échelonnée réduite que l'on ne pouvait pas deviner sur la matrice A elle-même.

1.5 Interprétation matricielle de la méthode de Gauss

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

Définition 16 *On appelle* matrice élémentaire *toute matrice qui résulte de l'application d'une opération élémentaire sur les lignes à la matrice identité I_m.*

Exemple 16 Dans le cas 3×3,

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

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

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

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

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

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

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

Et ainsi de suite. □

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

Proposition 14 *Soit A une matrice $m \times n$ et E une matrice élémentaire. La matrice EA est celle qui résulte de l'application de la même opération élémentaire à la matrice A.*

Démonstration. Écrivons les matrices A et E comme une ligne de n matrices colonnes soit $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ et $E = \begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{pmatrix}$. On sait que $EA = \begin{pmatrix} Ea_1 & Ea_2 & \cdots & Ea_n \end{pmatrix}$. Il suffit par conséquent de vérifier quel est l'effet de la multiplication par E sur une matrice-colonne x. Soit e_i la matrice colonne ayant un 1 sur la ième ligne et des 0 ailleurs. On a $I_m = (e_1, \ldots, e_m)$, donc par définition d'une matrice élémentaire, Ee_i est le vecteur obtenu par l'opération élémentaire considérée appliquée au vecteur e_i. Soit $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$. On a $x = \sum_{i=1}^{m} x_i e_i$. On a

$$Ex = E\left(\sum_{i=1}^{m} x_i e_i\right) = \sum_{i=1}^{m} x_i E e_i = \sum_{i=1}^{m} x_i \varepsilon_i.$$

Ex n'est autre que la matrice colonne obtenue par l'opération élémentaire considérée appliquée à la matrice colonne x. □

Proposition 15 *Une matrice élémentaire est inversible.*

Démonstration. Nous avons trois cas à traiter.

Premier cas : La matrice élémentaire est obtenue à partir de la matrice I_m par une homothétie. Soit λ un réel non nul. Notons $E(L_i \times \lambda)$ la matrice élémentaire obtenue à partir de I_m en multipliant la ième ligne par λ. En appliquant la proposition précédente à la matrice $A = E(L_i \times \lambda)$, on obtient la relation :

$$E(L_i \times \frac{1}{\lambda})E(L_i \times \lambda) = I_m.$$

La relation ci dessus appliquée à $\frac{1}{\lambda}$ nous donne

$$E(L_i \times \lambda)E(L_i \times \frac{1}{\lambda}) = I_m.$$

On en déduit que $E(L_i \times \lambda)$ est inversible d'inverse $E(L_i \times \frac{1}{\lambda})$.

Deuxième cas : La matrice élémentaire est obtenue à partir de la matrice I_m par un échange.

Notons $E(L_i \longleftrightarrow L_j)$ la matrice élémentaire obtenue à partir de I_m en échangeant les ième et jème lignes. En appliquant la proposition précédente à la matrice $A = E(L_i \longleftrightarrow L_j)$, on obtient la relation :

$$E(L_i \longleftrightarrow L_j)E(L_j \longleftrightarrow L_i) = I_m.$$

On en déduit que $E(L_i \longleftrightarrow L_j)$ est inversible d'inverse elle-même.

Troisième cas : La matrice élémentaire est obtenue à partir de la matrice I_m par une substitution. Soit λ un réel. Notons $E(L_i \longleftarrow L_i + \lambda L_j)$ la matrice élémentaire obtenue à partir de I_m en ajoutant à la ième ligne la $j$ième ligne multipliée par λ. En appliquant la proposition précédente à la matrice $A = E(L_i \longleftarrow L_i - \lambda L_j)$, on obtient la relation :

$$E(L_i \longleftarrow L_i + \lambda L_j)E(L_i \longleftarrow L_i - \lambda L_j) = I_m.$$

La relation ci dessus appliquée à $-\lambda$ nous donne

$$E(L_i \longleftarrow L_i - \lambda L_j)E(L_i \longleftarrow L_i + \lambda L_j) = I_m.$$

On en déduit que $E(L_i \longleftarrow L_i + \lambda L_j)$ est inversible d'inverse $E(L_i \longleftarrow L_i - \lambda L_j)$.□

Théorème 2 *Soit $A \in M_{mn}(K)$ et $U \in M_{mn}(K)$ l'unique matrice échelonnée réduite obtenue à partir de A par des opérations élémentaires sur les lignes. Alors il existe une matrice $M \in GL_m(K)$ telle que*

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

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

1ère étape : $A_1 = E_1A$.

2ème étape : $A_2 = E_2A_1 = E_2(E_1A) = (E_2E_1)A$.

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

$p$ème étape : $U = A_p = E_pA_{p-1} = (E_pE_{p-1}\cdots E_2E_1)A$.

On pose donc $M = E_pE_{p-1}\cdots E_2E_1$. Comme chacune des opérations élémentaires est inversible, chaque matrice élémentaire E_k appartient à $GL_m(K)$, d'où $M \in GL_m(K)$. □

Théorème 3 *Soit $A \in M_n(K)$ et U l'unique matrice échelonnée réduite obtenue à partir de A par des opérations élémentaires sur les lignes. La matrice A est inversible si et seulement si U est égale à I_n.*

Démonstration. On conserve les notations du théorème précédent. Si $U = I_n$ alors $A = M^{-1}$ est inversible puique $M \in GL_n(K)$. Si $U \neq I_n$, la dernière ligne de U est nulle. Donc, pour toute matrice carrée V, la dernière ligne de UV est nulle. On n'aura donc jamais $UV = I_n$. Donc U n'est pas inversible. Alors, A n'est pas inversible non plus car, si A était inversible, on aurait $U = MA$ et U serait inversible comme produit de matrices inversibles. □

Remarque 14 *A est une matrice carrée inversible si et seulement si $U = I_n$. On a alors $M = A^{-1}$. On retrouve donc le calcul de A^{-1} par la méthode de Gauss en utilisant $\tilde{A} = \begin{pmatrix} A & I_m \end{pmatrix}$. En effet, $\tilde{U} = M\tilde{A} = \begin{pmatrix} MA & MI_n \end{pmatrix} = \begin{pmatrix} I_n & A^{-1} \end{pmatrix}$.* □

Exemple 17 Dans le calcul qui suit, on adopte les notations suivantes pour les opérations élémentaires :

• un échange entre la $i$ème ligne et la $j$ième ligne de la matrice sera noté $L_i \longleftrightarrow L_j$

• si on multiplie la $i$ème ligne par le facteur non nul λ, on écrit $L_i \times \lambda$ (homothétie).

• si on ajoute λ fois la $i$ème ligne à la $j$ième ligne, on notera $L_j \longleftarrow L_j + \lambda L_i$ (substitution).

Considérons la matrice $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Montrons qu'elle est inversible et calculons son inverse.

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) (L_1 \longleftrightarrow L_2)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array}\right) (L_3 \longleftarrow L_3 - L_1)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{array}\right) (L_3 \longleftarrow L_3 - L_2)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right) (L_3 \times \tfrac{1}{2})$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right) (L_1 \longleftarrow L_1 - L_3 \text{ et } \mathrm{L}_2 \longleftarrow \mathrm{L}_2 - \mathrm{L}_3)$$

Donc A est inversible et

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

1.6 Matrices semblables

Les matrices considérées dans ce paragraphe sont des matrices carrées, éléments de $M_n(K)$.

1.6.1 Définition et propriétés

Définition 17 *Soient A et B deux matrices de $M_n(K)$. On dit que A est semblable à B si et seulement si il existe une matrice inversible P appartenant à $M_n(K)$ telle que $A = PBP^{-1}$.*

Remarque 15 On en déduit immédiatement que si A est une matrice quelconque de $M_n(K)$ et P une matrice inversible de $M_n(K)$, A est semblable à $P^{-1}AP$.

Proposition 16 *La relation binaire " être semblable à ...", définie sur $M_n(K)$, est appelée relation de similitude. Elle possède les propriétés suivantes :*

1.Si A est une matrice de $M_n(K)$, A est semblable à elle même (on dit que la relation est réflexive).

2. Soient A et B deux matrices de $M_n(K)$. Si A est semblable à B, alors B est semblable à A (on dit que la relation est symétrique).

3. Soient A, B et C trois matrices de $M_n(K)$. Si A est semblable à B, et B est semblable à C, alors A est semblable à C (on dit que la relation est transitive).

Autrement dit, la relation binaire définie sur $M_n(K)$ " A est semblable à B " est une relation d'équivalence sur $M_n(K)$.

Vocabulaire :

Compte tenu de ces propriétés, on peut dire indifféremment que la matrice A est semblable à la matrice B ou que les matrices A et B sont semblables.

Démonstration. Les démonstrations sont basées sur les propriétés des matrices inversibles.

1. Comme la matrice unité I_n est inversible, d'inverse $I_n^{-1} = I_n$, on peut écrire $I_nAI_n = A$, ce qui prouve que A est semblable à elle-même ($P = I_n$).

2. Soient A et B deux matrices de $M_n(K)$. Si A est semblable à B, il existe une matrice inversible P de $M_n(K)$ telle que $A = PBP^{-1}$. Si on multiplie les deux membres de

cette égalité, à gauche par P^{-1} et à droite par P, on obtient l'égalité $B = P^{-1}AP$. Comme P^{-1} est inversible d'inverse P, on a $B = P^{-1}A(P^{-1})^{-1}$. Cela permet de conclure.

3. Soient A, B et C trois matrices de $M_n(K)$. Si A est semblable à B, et B est semblable à C, il existe deux matrices inversibles P et Q de $M_n(K)$ telles que $A = PBP^{-1}$ et $B = QCQ^{-1}$. Alors on a $A = PQCQ^{-1}P^{-1}$. Or on a vu, dans les propriétés des matrices inversibles, que si P et Q sont des matrices inversibles, la matrice PQ l'est aussi et $(PQ)^{-1} = Q^{-1}P^{-1}$. L'égalité précédente peut donc s'écrire $A = (PQ)C(PQ)^{-1}$; cela prouve que A et C sont semblables.

1.6.2 Application au calcul des puissances d'une matrice

La notion de matrices semblables a aussi une utilisation intéressante pour le calcul des puissances de matrices en partant du constat que plus une matrice est simple (c'est-à-dire avec beaucoup de zéros) plus le calcul est facile.

Pour s'en convaincre on peut considérer l'exemple des matrices diagonales et établir le résultat suivant

Proposition 17 (Puissances d'une matrice diagonale) *Soit*

$$D = \begin{pmatrix} \alpha_1 & 0 & \dots & \dots & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \alpha_n \end{pmatrix}$$

une matrice diagonale d'ordre n. Alors, pour tout entier positif p, on a

$$D^p = \begin{pmatrix} \alpha_1^p & 0 & \dots & \dots & 0 \\ 0 & \alpha_2^p & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n-1}^p & 0 \\ 0 & \dots & \dots & 0 & \alpha_n^p \end{pmatrix}.$$

On démontre cette formule en faisant une démonstration par récurrence sur p.

Comme annoncé, on va donner une formule liant les puissances de deux matrices semblables.

Théorème 4 (Relation entre les puissances de deux matrices semblables) *Soient A et B deux matrices semblables, c'est-à-dire telles qu'il existe une matrice inversible P telle que $A = P^{-1}BP$. Alors pour tout entier positif p, on a $A^p = P^{-1}B^pP$, et donc A^p et B^p sont semblables.*

Démonstration. On démontre encore cette formule par récurrence sur p.

Si $p = 1$, la formule est triviale (c'est la formule traduisant le fait que A et B sont semblables).

Supposons la propriété vraie pour $p = k-1$ c'est-à-dire $A^{k-1} = PB^{k-1}P^{-1}$. Alors on a

$$A^k = A^{k-1}A = (PB^{k-1}P^{-1})(PBP^{-1}) = PB^{k-1}(P^{-1}P)BP^{-1} = PB^kP^{-1}$$

et la propriété est vraie pour $p = k$.

Remarque 16 On a vu dans le paragraphe consacré aux matrices carrées que l'on avait un procédé simple pour calculer les puissances d'une matrice de la forme $\lambda I + N$ où N est une matrice nilpotente de $M_n(K)$.

Par conséquent on aura une méthode systématique pour calculer les puissances d'une matrice semblable à une matrice de la forme $\lambda I + N$ où N est une matrice nilpotente de $M_n(K)$. Ce résultat est très utile.

Dans le dernier chapitre de ce cours, nous apprendrons à caractériser les matrices semblables à une matrice diagonale et nous donnerons aussi un exemple de calcul des puissances d'une telle matrice.

Chapitre 2

Systèmes linéaires

Les systèmes linéaires interviennent dans de nombreux contextes d'applications de l'algèbre linéaire (sciences de l'ingénieur, météorologie, économie, mais aussi codes de transmission d'information et cryptographie). Pour ce qui concerne les mathématiques, ils forment la base calculatoire de l'algèbre linéaire. Ils permettent également de traiter une bonne partie de la théorie de l'algèbre linéaire en dimension finie.

2.1 Définition

Soit $n \in \mathbb{N}^*$ un entier naturel supérieur ou égal à 1. Une *équation linéaire à n inconnues* $x_1, x_2, \ldots, x_n$ est une équation de la forme

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

où $a_1, a_2, \ldots, a_n$ et b sont des éléments de K donnés.

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

Définition 18 *Un* système de m équations linéaires à n inconnues*, ou système linéaire, est une liste de m équations linéaires.*

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

Exemple 18 Voici un système de 2 équations à 3 inconnues à coefficients dans $\mathbb{R}$.

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

On aurait pu l'écrire tout aussi bien

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

□

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

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

Les nombres a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, sont les *coefficients* du système. Ce sont des données. Les nombres b_i, $i = 1, \ldots, m$, constituent le *second membre* du système et sont également des données.

Il convient de bien observer comment on a rangé le système en lignes (une ligne par équation) numérotées de 1 à m par l'indice i, et en colonnes : les termes correspondant à une même inconnue x_j sont alignés verticalement les uns sous les autres. L'indice j varie de 1 à n. Il y a donc n colonnes à gauche des signes d'égalité, plus une colonne supplémentaire à droite pour le second membre. La notation avec double indice a_{ij} correspond à ce rangement : le *premier* indice (ici i) est le numéro de *ligne* et le *second* indice (ici j) est le numéro de *colonne.* Il est extrêmement important de toujours respecter cette convention.

Dans l'exemple 18, on a $m = 2$ (nombre d'équations = nombre de lignes), $n = 3$ (nombre d'inconnues = nombre de colonnes à gauche du signe =) et $a_{11} = 2$, $a_{12} = -1$, $a_{13} = 3/2$, $a_{21} = 1$, $a_{22} = 0$, $a_{23} = -4$, $b_1 = 8$ et $b_2 = -7$.

Définition 19 *Une solution du système linéaire est une liste de n nombres réels* $(s_1, s_2, \ldots, s_n)$ *(un n-uplet) tels que si l'on substitue* s_1 *pour* x_1*,* s_2 *pour* x_2*, etc., dans le système linéaire, on obtient une égalité. L'*ensemble des solutions du système *est l'ensemble de tous ces n-uplets.*

Ainsi, $(5, 13/2, 3)$ est une solution du système linéaire de l'exemple 18. En règle générale, on s'attache à déterminer l'ensemble des solutions d'un système linéaire. C'est ce que l'on appelle *résoudre* le système linéaire. Ceci amène à poser la définition suivante.

Définition 20 *On dit que deux systèmes linéaires sont* équivalents *s'ils ont le même ensemble de solutions.*

À partir de là, le jeu pour résoudre un système linéaire donné consistera à le transformer en un système équivalent dont la résolution sera plus simple que celle du système de départ. Nous verrons plus loin comment procéder de façon systématique pour arriver à ce but.

Remarque 17 Deux systèmes équivalents ont toujours visiblement le même nombre d'inconnues. Par contre, ils n'ont pas forcément le même nombre d'équations. Dans ce dernier cas, on peut toujours ajouter au système avec le moins d'équations le nombre manquant à l'aide d'équations triviales

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

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

Exemple 19 Résolution dans le cas d'un système 2×2 à coefficients réels. Considérons le système suivant

$$\begin{cases} x_1 & -2x_2 & = & -1, \\ -x_1 & +3x_2 & = & 3. \end{cases}$$

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

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

$$\begin{cases} x_1 & -2x_2 & = & -1, \\ -x_1 & +2x_2 & = & 3. \end{cases}$$

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

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

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

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

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

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

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

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

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

2.2 Notation matricielle

En réfléchissant un petit peu, on se rend compte que dans la donnée d'un système linéaire, seuls comptent les coefficients du système et le second membre. Écrire les

équations avec les inconnues permet de visualiser le système, mais n'est pas autrement utile. Il est donc naturel d'introduire la matrice suivante :

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

appelée *la matrice du système linéaire*. Elle a m lignes et n colonnes, c'est une *matrice* $m \times n$ (à coefficients dans K). Si on pose $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, le système s'écrit matriciellement $Ax = b$.

On introduit aussi

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

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

Exemple 20 Il est très facile de passer d'un système linéaire à sa matrice augmentée et vice-versa : il suffit de lire les coefficients au bon endroit. Considérons l'exemple du système 3×3 suivant

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

Sa matrice est

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

et sa matrice augmentée

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

□

2.3 Systèmes échelonnés réduits

Il se trouve que les systèmes linéaires dont la matrice augmentée est échelonnée réduite — appelés systèmes échelonnés réduits pour aller plus vite — sont particulièrement simples à résoudre. Commençons par deux exemples.

Exemple 21 Considérons le système suivant dont la résolution est triviale

$$\left\{\begin{array}{ccccccc} x_1 & & & & & = & b_1, \\ & x_2 & & & & = & b_2, \\ & & \ddots & & & \vdots & \vdots \\ & & & & x_n & = & b_n. \end{array}\right.$$

Sa matrice augmentée

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

est échelonnée réduite.

Exemple 22 Supposons que

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

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

$$\left\{\begin{array}{ccccccc} x_1 & & +2x_3 & & & = & 25, \\ & x_2 & -2x_3 & & & = & 16, \\ & & & & x_4 & = & 1. \end{array}\right.$$

Ce système se résout trivialement en

$$\left\{\begin{array}{ccc} x_1 & = & 25-2x_3, \\ x_2 & = & 16+2x_3, \\ x_4 & = & 1. \end{array}\right.$$

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

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

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

L'exemple qui précède montre que les inconnues d'un système échelonné réduit ne jouent pas toutes le même rôle. Rappelons la définition suivante.

Soit U une matrice échelonnée réduite. Les positions de pivot *de U sont les emplacements (au sens du couple (numéro de ligne, numéro de colonne)) des coefficients valant* 1 *du point 3) de la définition 12.*

Ainsi, dans l'exemple 22, on voit trois positions de pivot : $(1,1)$, $(2,2)$ et $(3,4)$. Le coefficient 1 situé en position $(3,5)$ n'est pas un pivot car il n'est pas le premier élément non nul de sa ligne.

Dans une matrice échelonnée réduite, on appelle *colonnes de pivot* les colonnes qui contiennent une position de pivot et *lignes de pivot* les lignes qui contiennent une position de pivot. D'après le point 3) de la définition 12, on voit qu'il y a au plus une position de pivot par ligne, et d'après le point 4), au plus une position de pivot par colonne. Par conséquent, le nombre de colonnes de pivot est égal au nombre de lignes de pivot, tous deux étant égaux au nombre de positions de pivot.

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

Définition 21 *Les inconnues correspondant à une colonne de pivot sont appelées* inconnues ou variables essentielles. *Les autres sont appelées* inconnues ou variables libres.

Remarquons qu'un système échelonné a toujours au moins une variable essentielle, mais qu'il n'a pas forcément de variables libres, voir le tout premier exemple de cette section. Nous pouvons maintenant résoudre les systèmes échelonnés réduits dans tous les cas.

Théorème 5 *Un système échelonné réduit est compatible si et seulement si sa matrice augmentée ne contient aucune ligne de la forme*

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

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

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

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

Cette ligne correspond à l'équation linéaire

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

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

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

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

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

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

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

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

fournit une représentation paramétrique de l'ensemble des solutions, les variables libres parcourant indépendamment K. □

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

Corollaire 6 *Dans le cas d'un système échelonné réduit $m \times n$ on a l'alternative suivante.*

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

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

b) Soit il y a une solution unique s'il n'y a pas de telle ligne ni de variables libres.

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

2.4 Résolution des systèmes par l'Algorithme de Gauss

À partir de maintenant, la stratégie pour résoudre un système général sera de se ramener à un système échelonné réduit qui lui soit équivalent. On va pour cela raisonner uniquement sur les matrices et utiliser l'algorithme de Gauss présenté dans le chapitre 1.

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

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

La seconde raison est que dans les applications pratiques de l'algèbre linéaire, lesquelles sont extrêmement nombreuses et importantes, les systèmes à résoudre sont énormes (des milliers, voire des millions d'équations et d'inconnues) et qu'il n'est pas question d'effectuer les calculs à la main. Ce sont des ordinateurs qui s'en chargent, et

ces derniers ont besoin de programmes, lesquels sont la traduction en tel ou tel langage d'un algorithme.

La notion d'équivalence de matrices est directement liée à celle d'équivalence des systèmes linéaires de la Définition 20.

Proposition 18 *Si les matrices augmentées de deux systèmes linéaires sont équivalentes, alors les systèmes linéaires sont équivalents.*

Démonstration. Il suffit de le vérifier sur les opérations élémentaires, l'équivalence des systèmes se propageant visiblement de proche en proche à chacune d'entre elles. Soient donc deux systèmes linéaires dont les matrices augmentées $A = (a_{ij})$ et $A' = (a'_{ij})$, $i = 1,\dots,n$, $j = 1,\dots,n+1$ (on note $a_{i,n+1}$ la colonne correspondant au second membre pour simplifier la notation), diffèrent par une opération élémentaire. Notons S_A l'ensemble des solutions du système associé à A et $S_{A'}$ l'ensemble des solutions du système associé à A'. Il faut distinguer suivant les trois cas possibles.

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

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

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

Multipliant les deux membres par λ, on voit que

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

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

Le cas de la substitution est très semblable : $a'_{ij} = a_{ij} + \lambda a_{kj}$ pour un certain i, un certain k et tous $j = 1,\dots,n+1$. Soit $(s_1,s_2,\dots,s_n) \in S_A$. Ce n-uplet vérifie en particulier les équations numéros i et k

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

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

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

On factorise le membre de gauche

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

qui n'est autre que

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

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

Nous avons vu que toute matrice A est équivalente à une unique matrice échelonnée réduite U. Les opérations élémentaires appliquées aux matrices augmentées produisant des systèmes équivalents entre eux, on va s'en servir pour se ramener à un système échelonné réduit. Rappelons que les *positions, colonnes et lignes de pivot* de A sont les positions, colonnes et lignes de pivot de U. Si A est la matrice augmentée d'un système linéaire, alors les inconnues correspondant à une colonne de pivot sont appelées *inconnues ou variables essentielles*. Les autres sont appelées *inconnues ou variables libres*.

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

Théorème 7 *Un système linéaire est compatible si et seulement si la matrice échelonnée réduite équivalente à sa matrice augmentée ne contient aucune ligne de la forme*

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

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

De même,

Corollaire 8 *Soit un système linéaire $m \times n$ quelconque, A sa matrice augmentée et U l'unique matrice échelonnée réduite équivalente à A. On a l'alternative suivante.*

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

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

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

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

Remarque 18 On n'a décrit qu'un seul algorithme de résolution, l'algorithme de Gauss. Or cet algorithme est bien insuffisant pour résoudre numériquement, c'est-à-dire sur ordinateur, les énormes systèmes linéaires rencontrés dans la pratique. L'analyse numérique matricielle est l'étude d'algorithmes qui généralisent celui de Gauss, ou qui sont de nature totalement différente, dans le but de résoudre effectivement et efficacement de tels systèmes. C'est un vaste champ de recherche toujours très actif de nos jours. □

Exemple 23 Soient a, b et c trois nombres réels. On considère le système suivant :

$$\begin{cases} x + y - z = a \\ -x + + 2z = b \\ 2y + 2z = c \end{cases}$$

La matrice augmentée du système est

$$\tilde{A} = \begin{pmatrix} 1 & 1 & -1 & a \\ -1 & 0 & 2 & b \\ 0 & 2 & 2 & c \end{pmatrix}.$$

En faisant les opérations sur les lignes suivantes : $L_2 \longleftarrow L_2 + L_1$, $L_3 \longleftarrow L_3 - 2L_2$ puis $L_1 \longleftarrow L_1 - L_2$, on obtient

$$\tilde{A} \sim \begin{pmatrix} 1 & 1 & -1 & a \\ 0 & 1 & 1 & a+b \\ 0 & 2 & 2 & c \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & a \\ 0 & 1 & 1 & a+b \\ 0 & 0 & 0 & c-2a-2b \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 & -b \\ 0 & 1 & 1 & a+b \\ 0 & 0 & 0 & c-2a-2b \end{pmatrix}.$$

Le système est compatible si et seulement si $c-2a-2b=0$.

Si $c-2a-2b \neq 0$, alors le système n'admet pas de solution.

si $c-2a-2b=0$, alors l'ensemble des solutions du système est

$$S = \{(2z-b, -z+a+b, z) \mid z \in K\}.$$

Remarque 19 Dans le cas d'un système homogène (tous les termes du second membre sont nuls), on peut appliquer l'algorithme de Gauss directement sur la matrice du système (au lieu de la matrice augmentée).

Proposition 19 *Soit un système linéaire donné sous sa forme matricielle $SX = b$ où S est un élément de $M_n(K)$. Il admet une solution unique si et seulement si S est inversible.*

Démonstration. Si S est inversible, alors le système admet une solution unique à savoir $x = S^{-1}b$. Réciproquement, supposons que le système admette une solution unique. Soit A la matrice augmentée du système. On a $A = (S \;\; b)$. D'après le corollaire 8, par une succession d' opérations élémentaires sur les lignes, on peut transformer $A = (S \;\; b)$ en $U = (I \;\; u)$ où

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

Il existe donc p matrices élémentaires $E_1, E_2, \ldots E_p$ telles que

$$E_p E_{p-1} \ldots E_1 A = (I_n \; u).$$

Or $E_p E_{p-1} \ldots E_1 A = (E_p E_{p-1} \ldots E_1 S \;\; E_p E_{p-1} \ldots E_1 b)$. Donc $E_p E_{p-1} \ldots E_1 S = I_n$. Comme chaque E_i est inversible, on en déduit $S = E_1^{-1} E_2^{-1} \ldots E_p^{-1}$. La matrice S est donc inversible comme produit de matrices inversibles.

Remarque 20 *Si S est inversible, l'unique solution du système linéaire $SX = b$ est $S^{-1}b$.*

Chapitre 3

Espaces vectoriels et applications linéaires

3.1 Cours sur les espaces vectoriels (généralités)

L'ensemble des vecteurs du plan ou de l'espace est muni d'une loi de composition interne (à savoir la somme de deux vecteurs) et d'une loi de composition externe (à savoir la multiplication d'un vecteur par un scalaire). De plus, ces deux lois satisfont un certains nombres de propriétés. Les espaces vectoriels généralisent cette situation.

3.1.1 Définition d'un espace vectoriel

Définition 22 *Un K-espace vectoriel est un ensemble non vide E muni*

• *d'une loi de composition interne c'est à dire d'une application de $E \times E$ dans E*

$$\begin{array}{rcl} E \times E & \to & E \\ (v, v') & \mapsto & v + v' \end{array}$$

• *d'une loi de composition externe de domaine d'opérateurs un corps commutatif, c'est à dire d'une application de $K \times E$ dans E*

$$\begin{array}{rcl} K \times E & \to & E \\ (\alpha, v) & \mapsto & \alpha \cdot v \end{array}$$

vérifiant trois groupes d'axiomes :
1) Axiomes relatifs à la loi interne
2) Axiomes relatifs à la loi externe
3) Axiomes liant les deux lois : double distributivité.

Axiomes

Dans la description des axiomes, la loi externe sur E sera notée $\cdot$ alors que la multiplication dans K sera notée $\times$. La loi de composition interne dans E et la somme dans K seront toutes les deux notées $+$ mais le contexte permettra de déterminer aisément de quelle loi il s'agit.

1) Axiomes relatifs à la loi interne

a) **Associativité** , c'est à dire que pour tous éléments u,v et w de E

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

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

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

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

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

Cet élément v' de E est noté $-v$.

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

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

Remarque 21 S'il existe un élément neutre 0_E vérifiant les axiomes b ci-dessus. il est unique.

Démonstration. Soient 0_E et $0'_E$ deux éléments vérifiant la définition de l'élément neutre. On a alors : pour tout élément v de E

$$\begin{array}{c} v+0_E = 0_E+v = v \\ v+0'_E = 0'_E+v = v \end{array}$$

Alors, la première propriété utilisée avec $v = 0'_E$ donne

$$0'_E+0_E = 0_E+0'_E = 0'_E.$$

La deuxième propriété utilisée avec $v = 0_E$ donne

$$0_E+0'_E = 0'_E+0_E = 0_E.$$

En comparant ces deux résultats, il vient $0_E = 0'_E$.

Remarque 22 De même, si v est un élément de E et s'il existe un élément v' de E vérifiant l'axiome c, il est unique.

Démonstration. Supposons qu'il existe deux symétriques de v notés v et v'. On a :

$$\begin{array}{c} v+v' = v'+v = 0_E \\ v+v'' = v''+v = 0_E \end{array}$$

Caculons $v'+(v+v'')$ de deux façons différentes en utilisant l'associativité de la loi + et les relations précédentes.

$$\begin{array}{l} \bullet v'+(v+v'') = v'+0_E = v' \\ \bullet v'+(v+v'') = (v'+v)+v'' \text{ d'où } v'+(v+v'') = 0_E+v'' = v''. \end{array}$$

On en déduit $v' = v''$.

Remarque 23 Les étudiants connaissant la théorie des groupes reconnaitront dans les axiomes a, b, c et d ci-dessus, les axiomes caractérisant les groupes abéliens.

2) Axiomes relatifs à la loi externe

a) Pour tous éléments λ et μ de K et pour tout élément v de E, on a

$$(\lambda \times \mu) \cdot v = \lambda \cdot (\mu \cdot v).$$

b) Soit 1 l'élément neutre de la multiplication de K. Pour tout élément v de E, on a

$$1 \cdot v = v.$$

3) Axiomes liant les deux lois : double distributivité

a) Distributivité par rapport à l'addition des scalaires : Pour tous λ et μ de K et pour tout élément v de E, on a :

$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v.$$

b) Distributivité par rapport à l'addition des vecteurs : Pour tout élément λ de K et pour tous éléments u et v de E, on a

$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v.$$

La loi interne et la loi externe doivent donc satisfaire huit axiomes pour que $(E, +, \cdot)$ soit un espace vectoriel sur K.

Terminologie et notations

• Au lieu de K-espace vectoriel, on dit aussi espace vectoriel sur K.

• Les élements du corps K sont appelés des scalaires et les éléments de E des vecteurs. Le corps K est appelé le corps des scalaires.

• La loi de composition interne sur E (notée usuellement +) est appelée couramment l'addition et $v + v'$ est appelée somme des vecteurs v et v'. La loi de composition externe sur E est appelée couramment multiplication par un scalaire. La multiplication du vecteur v par le scalaire α sera notée αv.

• 0_E est l'élément neutre de la loi interne de E et est appelé vecteur nul. Il ne doit pas être confondu avec l'élément 0 de K. Lorsqu'il n'y aura pas de risque de confusion, 0_E sera aussi noté 0.

Somme de n vecteurs

Il est possible de définir, par récurrence, l'addition de n vecteurs, $n \geq 2$. La structure d'espace vectoriel permet de définir l'addition de deux vecteurs, ce qui initialise la démonstration. Si la somme de $n-1$ vecteurs est définie, alors la somme de n vecteurs $v_1, \ldots, v_n$ est définie par

$$v_1 + v_2 + \cdots + v_n = (v_1 + v_2 + \cdots + v_{n-1}) + v_n.$$

Notation : $v_1 + v_2 + \cdots + v_n = \sum_{i=1}^{n} v_i$.

L'associativité de la loi + nous permet de ne pas mettre de parenthèses dans la somme $v_1 + \cdots + v_n$.

3.1.2 Exemples

Dans tous les exemples qui suivent, la vérification des axiomes se fait simplement et est laissée au soin des étudiants. Seules seront indiquées, dans chaque cas, les valeurs de l'élément neutre de la loi interne et du symétrique d'un élément.

Il est important de remarquer que les règles de calcul proviennent de l'addition et de la multiplication des éléments du corps K qui est sous-jacent dans tous les exemples.

Exemple 24 : Le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$

Le produit cartésien $\mathbb{R} \times \mathbb{R}$ est noté $\mathbb{R}^2$. C'est l'ensemble des couples (x,y) avec x élément de $\mathbb{R}$ et y élément de $\mathbb{R}$. Ceci s'écrit

$$\mathbb{R}^2 = \{(x,y) \mid x \in \mathbb{R}, y \in \mathbb{R}\}.$$

Remarque : L'écriture (x,y) traduit un ordre sur les élément x et y : x est la première composante du couple (x,y) et y est la seconde. Le couple $(1,2)$ est différent du couple $(2,1)$.

• Définition de la loi interne :
Si (x,y) et (x',y') sont deux éléments de $\mathbb{R}^2$,

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

•Définition de la loi externe :
Si α est un réel et (x,y) est élément de $\mathbb{R}^2$,

$$\alpha(x,y) = (\alpha x, \alpha y).$$

L'élément neutre de la loi interne est $(0,0)$. Le symétrique de (x,y) est $(-x,-y)$.

Exemple 25 : Le $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$

Cet exemple généralise le précédent. Soit n un entier supérieur ou égal à 1. Le produit cartésien de n ensembles égaux à $\mathbb{R}$, $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ est noté $\mathbb{R}^n$. C'est l'ensemble des n-uplets $(x_1,\ldots,x_n)$ avec $x_1,\ldots,x_n$ éléments de $\mathbb{R}$. Ceci s'écrit

$$\mathbb{R}^n = \{(x_1,\ldots,x_n) \mid \forall i, 1 \leq i \leq n, x_i \in \mathbb{R}\}.$$

De même que dans l'exemple précédent, l'écriture $(x_1,\ldots,x_n)$ traduit un ordre sur les élément x_i : x_i est la $i$ème composante du n-uplet $(x_1,\ldots,x_n)$

• Définition de la loi interne :
Si $(x_1,\ldots,x_n)$ et $(x'_1,\ldots,x'_n)$ sont deux éléments de $\mathbb{R}^n$,

$$(x_1,\ldots,x_n)+(x'_1,\ldots,x'_n) = (x_1+x'_1,\ldots,x_n+x'_n)$$

•Définition de la loi externe :
Si α est un réel et $(x_1,\ldots,x_n)$ est élément de $\mathbb{R}^n$,

$$\alpha(x_1,\ldots,x_n) = (\alpha x_1,\ldots,\alpha x_n).$$

L'élément neutre de la loi interne est $(0,0,\ldots,0)$. Le symétrique de $(x_1,\ldots,x_n)$ est $(-x_1,\ldots,-x_n)$.

Dans le cas particulier où $n = 1$, nous avons défini une structure de $\mathbb{R}$-espace vectoriel sur $\mathbb{R}$. Dans ce cas particulier, la loi interne est la somme sur $\mathbb{R}$ et la loi externe est la multiplication sur $\mathbb{R}$.

De manière analogue, on peut définir le $\mathbb{C}$-espace vectoriel $\mathbb{C}^n$ et plus généralement le K-espace vectoriel K^n.

Exemple 26 : Le $\mathbb{R}$-espace vectoriel $A(\mathbb{R},\mathbb{R})$

L'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}$ est noté $A(\mathbb{R},\mathbb{R})$. Il peut être muni d'une structure de $\mathbb{R}$-espace vectoriel de la manière suivante.

• Définition de la loi interne :

Soient f et g deux éléments de $A(\mathbb{R},\mathbb{R})$. L'application $f+g$ est donc définie en donnant l'image de tout élément réel x par $f+g$, soit :

$$\forall x \in \mathbb{R},\ \ (f+g)(x) = f(x) + g(x)$$

où le signe $+$ désigne la loi interne de $A(\mathbb{R},\mathbb{R})$ dans le membre de gauche et l'addition dans $\mathbb{R}$ dans le membre de droite.

• Définition de la loi externe :

Nous désignerons par $\cdot$ la loi externe de $A(\mathbb{R},\mathbb{R})$ et par $\times$ la multiplication dans $\mathbb{R}$. De même, si α est un nombre réel et f un élément de $A(\mathbb{R},\mathbb{R})$, la fonction $\alpha \cdot f$ est définie par l'image de tout réel x comme suit :

$$\forall x \in \mathbb{R},\ (\alpha \cdot f)(x) = \alpha \times f(x).$$

L'élément neutre pour la loi interne est l'application de $\mathbb{R}$ dans $\mathbb{R}$ définie par

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

C'est la fonction nulle qu'il est difficile de noter 0 (car alors, on serait en droit d'écrire $0(0) = 0$, ce qui est difficile à décoder !). Le symétrique de l'élément f de $A(\mathbb{R},\mathbb{R})$ est l'application g de $\mathbb{R}$ dans $\mathbb{R}$ définie par

$$\forall x \in \mathbb{R},\ \ g(x) = -f(x).$$

Le symétrique de f est noté $-f$.

Exemple 27 : Le $\mathbb{R}$-espace vectoriel des suites réelles

L'ensemble des suites réelles, noté $S = A(\mathbb{N},\mathbb{R})$ est l'ensemble des applications de $\mathbb{N}$ dans $\mathbb{R}$.

• Définition de la loi interne

Soient $U = (U_n)_{n\in\mathbb{N}}$ et $V = (V_n)_{n\in\mathbb{N}}$ deux éléments de S, La suite $U+V$ est la suite $W = (W_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N},\ \ W_n = U_n + V_n$$

où $U_n + V_n$ désigne la somme de U_n et de V_n dans $\mathbb{R}$.

• Définition de la loi externe :

Si α est un nombre réel et $U = (U_n)_{n\in\mathbb{N}}$ un élément de S, αU est la suite $T = (T_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N},\ \ T_n = \alpha \times U_n$$

où $\times$ désigne la multiplication dans $\mathbb{R}$.

L'élément neutre de la loi interne est la suite dont tous les termes sont nuls. Le symétrique de la suite $U = (U_n)_{n\in\mathbb{N}}$ est la suite $U' = (U'_n)_{n\in\mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N},\ \ U'_n = -U_n.$$

Elle est notée $-U$.

Exemple 28 : Le $\mathbb{R}$-espace vectoriel $\mathbb{C}$

L'ensemble $\mathbb{C}$ des nombres complexes peut être aussi muni naturellement d'une structure de $\mathbb{R}$-espace vectoriel.

• **Définition de la loi interne :** La loi interne est la somme dans $\mathbb{C}$.

• **Définition de la loi externe :** La loi externe est définie de la façon suivante :

$$\begin{array}{rcl} \mathbb{R}\times\mathbb{C} & \to & \mathbb{C} \\ (\alpha,\nu) & \mapsto & \alpha\nu \end{array}$$

où $\alpha\nu$ désigne la multiplication de α par ν dans le corps des complexes $\mathbb{C}$.

L'élement neutre est 0 et le symétrique du nombre complexe z est $-z$.

Exemple 29 : K-espace vectoriel des matrices

Soient n et p deux entiers naturels strictement positifs. L'ensemble $M_{n,p}(K)$ des matrices à n lignes et p colonnes à coefficients dans K est muni d'une structure de K-espace vectoriel.

• **Définition de la loi interne** : La loi interne est la somme de deux matrices.

• **Définition de la loi externe** : La loi externe est le produit d'une matrice par un scalaire :

$$\begin{array}{rcl} K\times M_{p,n}(K) & \to & M_{p,n}(K) \\ (\alpha,A) & \mapsto & \alpha A \end{array}$$

l' élément neutre pour la loi interne est la matrices à n lignes et p colonnes dont tous les coefficients sont nuls. Le symétrique de la matrice $A = (a_{i,j})$ est la matrice $(-a_{i,j})$.

Exemple 30 : Applications d'un ensemble dans un espace vectoriel

L'exemple suivant généralise à la fois les exemples $\mathbb{R}^n$, $A(\mathbb{R},\mathbb{R})$, $A(\mathbb{N},\mathbb{R})$, $M_{n,p}(K)$.

Soit X un ensemble et E un K-espace vectoriel. L'ensemble $A(X,E)$ des applications de X dans E peut être muni d'une structure d'espace vectoriel comme suit.

Définition de la loi interne : Soient f et g deux applications de X dans E, l'application $f+g$ est définie de la façon suivante :

$$\begin{array}{rcl} f+g : X & \to & E \\ x & \mapsto & f(x)+g(x) \end{array}$$

où $f(x)+g(x)$ est la somme des éléments $f(x)$ et $g(x)$ dans E.

Définition de la loi externe : Si α est un élément de K et f un élément de $A(X,E)$, on définit l'application $\alpha \cdot f$ comme suit :

$$\begin{array}{rcl} \alpha \cdot f : X & \to & E \\ x & \mapsto & \alpha f(x) \end{array}$$

L'élément neutre pour la loi interne est l'application nulle, c'est à dire celle qui, à tout élément x de X, associe 0_E. Si f est un élément de $A(X,E)$, son symétrique par l'application interne est l'application

$$\begin{array}{rcl} X & \to & E \\ x & \mapsto & -f(x) \end{array}$$

où $-f(x)$ est le symétrique de $f(x)$ dans E.

Si $X=\mathbb{R}$, $K=\mathbb{R}$ et $E=\mathbb{R}$, on retrouve l'exemple $A(\mathbb{R},\mathbb{R})$. Si $X=\mathbb{N}$, $K=\mathbb{R}$ et $E=\mathbb{R}$, on retrouve l'exemple $A(\mathbb{N},\mathbb{R})$. Si $X=\{1,\ldots,n\}$, $K=\mathbb{R}$ et $E=\mathbb{R}$, on retrouve l'exemple $\mathbb{R}^n$. En effet, l'application naturelle

$$\begin{array}{rcl} A(X,\mathbb{R}) & \mapsto & \mathbb{R}^n \\ f & \mapsto & (f(1),\ldots,f(n)) \end{array}$$

est une bijection qui identifie $A(X,\mathbb{R})$ à $\mathbb{R}^n$. On peut en fait voir que c'est un isomorphisme (voir définition plus loin). Si $X=\{1,\ldots,n\}\times\{1,\ldots,p\}$, $E=K$, on retrouve l'exemple $M_{n,p}(K)$. En effet, l'application naturelle

$$\begin{array}{rcl} A(X,\mathbb{R}) & \mapsto & M_{n,p}(K) \\ f & \mapsto & (f(i,j)) \end{array}$$

est une bijection qui identifie $A(X,K)$ à $M_{n,p}(K)$. On peut en fait voir que c'est un isomorphisme.

3.1.3 Règles de calcul, combinaisons linéaires

Proposition 20 *Soit E un espace vectoriel sur un corps K. On notera par 0_E l'élément neutre pour la loi interne de E pour le distinguer de l'élément 0 de K. Les propriétés suivantes sont satisfaites :*

1) L'addition est régulière : Si u,v et w sont des vecteurs tels que

$$u+v=u+w, \text{ alors } v=w.$$

2) Pour tout vecteur v de E, $0\cdot v=0_E$.

3) Pour tout scalaire α, $\alpha 0_E=0_E$.

4) Pour tout vecteur v de E, $(-1)v=-v$.

5) L'opération $(v,w)\mapsto v+(-w)$ s'appelle la soustraction ; Le vecteur $v+(-w)$ est noté $v-w$. Les propriétés suivantes sont satisfaites :

a) Pour tout scalaire α et tous vecteurs v et w, $\alpha(v-w)=\alpha v-\alpha w$.

b) Pour tous scalaire α et β et tout vecteur v, $(\alpha-\beta)v=\alpha v-\beta v$.

6) Si λ est un scalaire et v un vecteur tels $\lambda v=0_E$, alors

$$\text{soit } \lambda=0, \text{ soit } v=0_E.$$

SYNTHÈSE : Les propriétés 2,3 et 6 peuvent êre résumées par le **résultat fondamental** suivant

$$\lambda v = 0_E \Longleftrightarrow \lambda = 0 \text{ ou } v = 0_E.$$

Démonstration. Les démonstrations des propriétés sont des manipulations sur les axiomes définissant les espaces vectoriels.

Démonstration de la propriété 1 : En ajoutant aux deux membres de l'inégalité $u+v = u+w$ le symétrique de u, soit $-u$, on obtient l'égalité

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

ce qui, en utilisant l'associativité de l'addition des vecteurs, permet d'obtenir

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

Or $(-u)+u = 0_E$ d'après la définition du symétrique, l'égalité devient donc

$$0_E + v = 0_E + w$$

d'où $v = w$ d'après la définition de l'élément neutre de l'addition dans E.

Démonstration de la propriété 2 : Le point de départ de la démonstration est l'égalité dans K

$$0+0 = 0.$$

D'où, pour tout vecteur de E, l'égalité

$$(0+0)v = 0v.$$

Soit, en utilisant la distributivité de la loi externe par rapport à la loi interne et la définition de l'élément neutre

$$0v+0v = 0v = 0v+0_E.$$

Ce qui permet d'obtenir l'égalité souhaitée $0_K v = 0_E$ grâce à la propriété 1.

Démonstration de la propriété 3 : La preuve est semblable en partant de l'égalité $0_E + 0_E = 0_E$.

Démonstration de la propriété 4 : Compte tenu de la définition du symétrique pour l'addition d'un élément de E, il suffit, pour justifier la propriété, de calculer l'expression $v+(-1)v$. On a, en utilisant la propriété de la multiplication par 1 puis la distributivité de l'addition des scalaires et enfin a propriété 2 :

$$v+(-1)v = 1v+(-1)v = (1+(-1))\,v = 0v = 0_E$$

La loi interne étant commutative, on en déduit que $(-1)v$ est le symétrique du vecteur v.

Démonstration de la propriété 5 :

Preuve du a :

Avec la notation introduite dans l'énoncé, $\alpha v - \alpha w$ est égal à $\alpha v + (-\alpha w)$. Alors,

$$\begin{aligned}
\alpha(v-w)+\alpha w &= \alpha[(v+(-w))+w] \text{ (distributivité de l'addition dans E)} \\
&= \alpha[v+((-w)+w)] \text{ (associativité de l'addition dans E)} \\
&= \alpha[v+0_E] \text{ (définition du symétrique)} \\
&= \alpha v \text{(définition de l'élément neutre)}
\end{aligned}$$

ce qui donne le résultat en rajoutant à chaque membre de l'égalité le symétrique de αv.

Preuve du b :
Elle est du même type. Le point de départ en est le calcul de $(\alpha-\beta)v+\beta v$.

Démonstration de la propriété 6 :
Soit λ un scalaire et v un vecteur tels que $\lambda v = 0_E$. Supposons λ différent de 0. Alors λ est inversible pour le produit dans le corps K. Soit λ^{-1} son inverse. En multipliant par λ^{-1} les deux membres de l'égalité, il vient :

$$\lambda^{-1}(\lambda v) = \lambda^{-1}0_E.$$

D'où en utilisant les propriétés de la multiplication par un scalaire

$$(\lambda^{-1}\lambda)v = 0_E$$

et donc

$$1v = 0_E.$$

D'où $v = 0_E$.

Combinaisons linéaires d'éléments dans un espace vectoriel

Définition 23 *Soit n un entier supérieur ou égal à 1 et $v_1, v_2, \ldots, v_n$ n vecteurs d'un espace vectoriel E. Tour vecteur de la forme $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ où $\alpha_1, \alpha_2, \ldots, \alpha_n$sont des éléments de K est appelé* **combinaison linéaire** *des vecteurs $v_1, v_2, \ldots, v_n$. Les scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont appelés coefficients de la combinaison linéaire.*

Remarque : Si $n = 1$, on dit que w est colinéaire à v_1.

Exemples :
•Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$, $(3,3,1)$ est combinaison linéaire des vecteurs $(1,1,0)$ et $(1,1,1)$ car on a l'égalité

$$(3,3,1) = 2(1,1,0) + (1,1,1).$$

• Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$, le vecteur $v = (2,1)$ n'est pas combinaison linéaire du vecteur $v_1 = (1,1)$ car s'il l'était, il existerait un réel α tel que $v = \alpha v_1$, ce qui équivaudrait à l'égalité $(2,1) = (\alpha, \alpha)$.

• Soit $E = A(\mathbb{R}, \mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions réelles. Soient f_0, f_1, f_2 et f_3 les fonctions définies par :

$$\forall x \in \mathbb{R},\ \ f_0(x) = 1,\ \ f_1(x) = x,\ \ f_2(x) = x^2, f_3(x) = x^3$$

Alors la fonction f définie par

$$\forall x \in \mathbb{R},\ \ f(x) = x^3 - 2x^2 - 7x - 4$$

est combinaison linéaire des fonctions f_0, f_1, f_2, f_3 puisque l'on a l'égalité

$$f = f_3 - 2f_2 - 7f_1 - 4f_0.$$

• Dans $M_{2,3}(\mathbb{R})$, on considère $A = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 3 \end{pmatrix}$. On peut écrire A naturellement sous la forme suivante :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On a démontré que A est combinaison linéaire des matrices de type 2×3 dont tous les éléments sont nuls sauf un, égal à 1.

3.2 Cours sur les espaces vectoriels (constructions)

3.2.1 Sous-espaces vectoriels

Définition 24 *Soit E un K-espace vectoriel. Une partie F de E telle que*
- *F est non vide*
- *F est stable pour l'addition : $\forall u \in F,\ \forall v \in F,\ u + v \in F$.*
- *F est stable pour la multiplication par un scalaire : $\forall \lambda \in K, \forall u \in F, \lambda u \in F$*

est appelé sous espace vectoriel de E.

Théorème 9 *Soit E un K-espace vectoriel et F un sous espace vectoriel de K. Les lois de composition interne et externe sur E induisent des lois de composition internes et externes sur F. Muni de ces deux lois, F a une structure de K-espace vectoriel.*

Démonstration. La stabilité de F pour les deux lois permet de munir cet ensemble d'une loi de composition interne et d'une loi de composition externe à opérateurs dans K, en restreignant à F les opérations définies dans E. Les propriétés de commutativité et d'associativité de l'addition, ainsi que les quatre axiomes relatifs à la loi externe sont vérifiés, car ils sont satisfaits dans E donc en particulier dans F, qui est inclus dans E.

Il reste à montrer l'existence d'un élément neutre, et d'un symétrique pour tout élément de F :

L'espace vectoriel E possède un élément neutre 0_E. Cet élément appartient à F car pour u élément de F (l'hypothèse F non vide est ici essentielle) $0u$ appartient à F (stabilité de F pour la loi externe), or $0u = 0_E$, donc 0_E appartient à F. De plus F étant inclus dans E, cet élément est tel que :

$$\forall u \in F,\ u + 0_E = 0_E + u = u.$$

L'élément neutre de l'addition dans F est donc 0_E.

De même F étant inclus dans E, pour tout élément u de F, il existe un élément de E, noté $-u$, tel que $u + (-u) = 0_E$; il faut donc montrer que $-u$ appartient à F. Comme u est élément de F, $(-1)u$ appartient à F, d'après la stabilité de F pour la loi externe. Or $(-1)u = -u$. Donc le symétrique de u dans F est égal au symétrique de u dans E.

Remarque 24 Pour les étudiants connaissant la théorie des groupes, on peut noter que $(F, +)$ est un sous groupe de $(E, +)$.

Remarque 25 *1- La démonstration précédente fait ressortir les deux points suivants :*
•$0_E = 0_F$.
• *Le symétrique de u calculé dans E est le même que le symétrique de u calculé dans F.*
2- $\{0_E\}$ *et E sont des sous-espaces vectoriels de E.*
3- Un sous-espace vectoriel de E contient nécessairement à 0_E*. Ceci donne une méthode simple pour prouver qu'un sous-ensemble n'est pas un sous-espace vectoriel : si* 0_E *n'appartient pas à F alors F n'est pas un sous-espace vectoriel de E.*

Méthodologie :
1- Pour répondre à une question du type " le sous-ensemble F de l'espace vectoriel E est-il un sous-espace vectoriel de E ? ", il est judicieux de vérifier que 0_E appartient à F :
Si 0_E appartient à F, cela prouve que F est non vide et on peut poursuivre en étudiant la stabilité de F pour les lois de E. Sinon on peut alors affirmer que F n'est pas un sous-espace vectoriel de E.
2- Pour montrer qu'un ensemble F est un espace vectoriel sur K, on peut chercher un espace vectoriel E qui contient F, puis prouver que F est un sous-espace vectoriel de E.

Exemples immédiats
• L'ensemble $F = \{(x,y) \in \mathbb{R}^2 \mid x = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.
• L'ensemble $F = \{(x,y) \in \mathbb{R}^2 \mid x = 2\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.
• L'ensemble des fonctions continues sur $\mathbb{R}$ est un sous-espace vectoriel de l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$.
• L'ensemble des suites réelles convergentes est un sous-espace vectoriel de l'espace vectoriel des suites réelles.
• Notons $\mathcal{P}$ l'ensemble des fonctions paires et I l'ensemble des fonctions impaires. On a

$$\mathcal{P} = \{f \in A(\mathbb{R},\mathbb{R}) \mid \forall x \in \mathbb{R}, f(-x) = f(x)\}$$
$$I = \{f \in A(\mathbb{R},\mathbb{R}) \mid \forall x \in \mathbb{R}, f(-x) = -f(x)\}$$

$\mathcal{P}$ et I sont des sous espaces vectoriels de $A(\mathbb{R},\mathbb{R})$;

Théorème 10 (Caractérisation d'un sous-espace par la notion de combinaison linéaire)
Soit E un K-espace vectoriel et F une partie de E. F est un sous-espace vectoriel de E si et seulement si :
• *F est non vide*
•*Toute combinaison linéaire de deux éléments de F appartient à F :*

$$\forall (u,v) \in F^2,\ \forall (\alpha,\beta) \in K^2,\quad \alpha u + \beta v \in F.$$

Démonstration. Il suffit de démontrer que la deuxième propriété est équivalente à la stabilité de F pour les deux lois. Il est clair que si F est stable pour l'addition et la multiplication par un scalaire alors toute combinaison linéaire de deux vecteurs de F est dans F. Pour établir la réciproque il suffit de choisir convenablement les coefficients α et β : $\alpha = \beta = 1$ donne la stabilité de F pour l'addition.
α quelconque, élément de K, et $\beta = 0$ donne la stabilité de F pour la loi externe. □

On peut légèrement simplifier la caractérisation précédente de la façon suivante

Théorème 11 *Soit E un K-espace vectoriel et F une partie de E. F est un sous-espace vectoriel de E si et seulement si :*
- *F est non vide*
- $\forall (u,v) \in F^2,\ \forall \alpha \in K,\quad \alpha u + v \in F.$

Exemple 31 Une fonction polynôme sur $\mathbb{R}$ est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle qu'il existe un entier k et $k+1$ éléments $a_0,\dots,a_k$ de $\mathbb{R}$ tels que :

$$\forall x \in \mathbb{R},\ f(x) = a_0 + a_1 x + \cdots + a_k x^k.$$

On définit le degré de f comme étant le $\mathrm{Sup}\{k, a_k \neq 0\}$ On note $P(\mathbb{R})$ l'ensemble des fonctions polynômes sur $\mathbb{R}$. C'est un sous-espace vectoriel de $A(\mathbb{R},\mathbb{R})$, l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$. L'ensemble $P_n(\mathbb{R})$ des fonctions polynômes de degré inférieur ou égal à n est un sous-espace vectoriel de $P(\mathbb{R})$, donc de $A(\mathbb{R},\mathbb{R})$. En revanche, pour $n \geq 1$, l'ensemble des fonctions polynômes de degré exactement égal à n n'est pas un sous-espace vectoriel de $P(\mathbb{R})$. En effet ce n'est pas un ensemble stable pour l'addition des fonctions : par exemple les fonctions f et g définies par $f(x) = x+1$ et $g(x) = -x+1$ sont des fonctions polynômes de degré 1, mais leur somme ne l'est pas.

On définit de la même façon les fonctions polynômes sur K.

3.2.2 Sous-espace engendré par une partie finie-Intersection

Théorème 12 (Théorème de structure de l'ensemble des combinaisons linéaires) *Soit $\{v_1,\dots,v_n\}$ une partie finie du K-espace vectoriel E, alors l'ensemble des combinaisons linéaires des vecteurs $\{v_1,\dots,v_n\}$ est un sous-espace vectoriel de E. C'est le plus petit sous-espace vectoriel de E (au sens de l'inclusion) contenant les vecteurs $\{v_1,\dots,v_n\}$: autrement dit, il est inclus dans tout sous-espace vectoriel contenant $\{v_1,\dots,v_n\}$.*

Démonstration. On appelle F l'ensemble des combinaisons linéaires des vecteurs $\{v_1,\dots,v_n\}$. Cet ensemble est non vide, car il contient la combinaison linéaire particulière $0v_1 + \cdots + 0v_n$ qui vaut 0_E. On peut également vérifier que $v_1,\dots,v_n$ appartiennent à F, en effet pour tout k compris entre 1 et n, v_k est combinaison linéaire de $v_1,\dots v_n$ (il suffit de considérer la combinaison linéaire où tous les coefficients sont nuls sauf le kième qui vaut 1).

Il s'agit maintenant de prouver que F est stable par combinaison linéaire de deux vecteurs.

Soit u et w deux vecteurs de F et deux scalaires α et β . Comme u est élément de F, il existe des scalaires $\lambda_1,\dots,\lambda_n$ tels que

$$u = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

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

$$w = \mu_1 v_1 + \cdots + \mu_n v_n.$$

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

$$\alpha u + \beta v = (\alpha\lambda_1 + \beta\mu_1)v_1 + \cdots + (\alpha\lambda_2 + \beta\mu_2)v_2 + \cdots + (\alpha\lambda_n + \beta\mu_n)v_n.$$

C'est une combinaison linéaire des vecteurs $v_1,\dots,v_n$ donc un élément de F.

Si G est un sous-espace vectoriel contenant $\{v_1,\dots,v_n\}$ alors il est stable par combinaison linéaire ; il contient donc toute combinaison linéaire des vecteurs $\{v_1,\dots,v_n\}$. Par conséquent F est inclus dans G : F est le plus petit sous-espace (au sens de l'inclusion) contenant $\{v_1,\dots,v_n\}$.

Notation

Ce sous-espace vectoriel est appelé **sous-espace engendré par** $v_1,\dots,v_n$ et est noté $\mathrm{vect}(v_1,\dots,v_n)$. On a donc

$$u \in \mathrm{vect}(v_1,\dots,v_n) \Longleftrightarrow \exists(\lambda_1,\dots,\lambda_n) \in K^n \mid u = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Exemple 32 E étant un K-espace vectoriel, et u un élément quelconque de E, l'ensemble $F = \{\alpha u \mid \alpha \in K\}$ est le sous-espace vectoriel de E engendré par u. Il est souvent noté Ku.

Exemple 33 Soit E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$ et e_0, e_1, e_2 les applications définies par :

$$\forall x \in \mathbb{R},\ e_0(x) = 1,\ e_1(x) = x \text{ et } e_2(x) = x^2.$$

Le sous-espace vectoriel de E engendré par $\{e_0, e_1, e_2\}$ est l'espace vectoriel des fonctions polynômes de degré inférieur ou égal à 2, c'est-à-dire de la forme

$$f : x \mapsto ax^2 + bx + c.$$

Méthodologie :

On peut démontrer qu'une partie non vide F d'un espace vectoriel E est un sous-espace vectoriel de E en montrant que F est égal à l'ensemble des combinaisons linéaires d'un nombre fini de vecteurs de E.

Exemple 34 Soit $F = \{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\}$. Un triplet de $\mathbb{R}^3$ est élément de F si et seulement si $x = y + z$. Donc u est élément de F si et seulement s'il peut s'écrire $u = (y+z, y, z)$. Or, on a l'égalité

$$(y+z, y, z) = y(1,1,0) + z(1,0,1).$$

Donc F est l'ensemble des combinaisons linéaires de $\{(1,1,0),(1,0,1\}$. C'est le sous-espace vectoriel engendré par $\{(1,1,0),(1,0,1\}$..

Proposition 21 (Propriété de transitivité) *Soit F un sous-espace engendré par n vecteurs . On suppose qu'il existe p vecteurs $w_1,\dots,w_p$ appartenant à F tels que pour tout i compris entre 1 et n, , v_i soit une combinaison linéaire de $w_1,\dots,w_p$. Alors F est engendré par $w_1,\dots,w_p$.*

La démonstration est laissée à titre d'exercice.

Remarque 26 Plus généralement, on peut définir le sous-espace vectoriel engendré par une partie quelconque (non nécessairement finie) d'un espace vectoriel.

Proposition 22 (Intersection de deux sous-espaces) *Soit E un K-espace vectoriel. L'intersection de deux sous-espaces vectoriels de E est un sous-espace vectoriel de E.*

Démonstration. Soit F_1 et F_2 deux sous-espaces vectoriels de E. L'intersection $F_1 \cap F_2$ n'est pas vide car 0_E appartient à F_1 et F_2 (car ce sont des sous-espaces vectoriels de E).

Il suffit de montrer que $F_1 \cap F_2$ est stable par combinaison linéaire de deux vecteurs : Soient u et v deux vecteurs de $F_1 \cap F_2$ et α, β deux scalaires. Comme u et v sont éléments de F_1 et F_1 est un sous-espace vectoriel de E, le vecteur $\alpha u + \beta v$ appartient à F_1.

De même u et v appartenant à F_2, le vecteur $\alpha u + \beta v$ appartient donc à F_2. □

Exemple 35 Soit F le sous-ensemble de $\mathbb{R}^3$ défini par :

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=0 \text{ et } x+y+2z=0\}.$$

L'ensemble F est l'intersection de F_1 et F_2, les sous-ensembles de $\mathbb{R}^3$ définis par :

$$\begin{aligned} F_1 &= \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=0\} \\ F_2 &= \{(x,y,z) \in \mathbb{R}^3 \mid x+y+2z=0\} \end{aligned}$$

Ce sont des sous-espaces de $\mathbb{R}^3$ donc $F = F_1 \cap F_2$ est un sous-espace vectoriel de $\mathbb{R}^3$.

Remarque 27 On démontre de même que l'intersection d'une famille quelconque de sous espaces vectoriels de E est un sous espace vectoriel de E.

Remarque 28 La réunion de deux sous-espaces vectoriels de E n'est pas en général un sous-espace de E.

3.2.3 Somme de sous espaces vectoriels

Comme la réunion de deux sous-espaces vectoriels F_1 et F_2 n'est pas en général un sous-espace vectoriel, il est utile de connaître les sous-espaces vectoriels qui contiennent à la fois les deux sous-espaces vectoriels F_1 et F_2, et en particulier le plus petit d'entre eux (au sens de l'inclusion).

Définition 25 (Définition de la somme de deux sous-espaces) *Si F et G sont deux sous espaces vectoriels d'un K-espace vectoriel E, l'ensemble de tous les éléments $x+y$ où x est un élément de F et y un élément de G, est appelé somme des sous-espaces vectoriels F et G . Cette somme est notée $F+G$. On a donc*

$$F+G = \{z \in E \mid \exists x \in F, \exists y \in G, z = x+y\}.$$

Proposition 23 *Si F et G sont deux sous-espaces vectoriels du K-espace vectoriel E, alors $F+G$ est un sous-espace vectoriel de E.*

Remarque 29 L'ensemble $F+G$ contient F et contient G : en effet tout élément x de F s'écrit $x = x+0$ avec x appartenant à F et 0 appartenant à G (puisque G est un sous-espace vectoriel), donc x appartient à $F+G$. De même pour un élément de G. On peut montrer que $F+G$ est le plus petit sous espace vectoriel contenant F et G.

Démonstration. • $F+G$ est non vide car il contient F et G.

• Soient u et u' des éléments de $F+G$. Comme u est dans $F+G$, il existe x dans F et y dans G tels que $u = x+y$. Comme u' est dans $F+G$, il existe x' dans F et y' dans

G tels que $u' = x' + y'$. Soient λ et λ' des scalaires. En utilisant les axiomes des espaces vectoriels, on obtient :

$$\lambda u + \lambda' u' = (\lambda x + \lambda' x') + (\lambda y + \lambda' y').$$

Comme F et G sont des sous-espaces vectoriels, $\lambda x + \lambda' x'$ est dans F et $\lambda y + \lambda' y'$ est dans G. Donc $\lambda u + \lambda' u'$ est dans $F + G$. □

Exemple 36 Déterminons $F + G$ dans le cas où F et G sont les sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\} \text{ et } G = \{(x,y,z) \in \mathbb{R}^3 \mid x = z = 0\}$$

Un élément u de $F + G$ s'écrit $u = v + w$ où v est un élément de F et w un élément de G. Donc il existe deux nombres réels x et y tels que $v = (x,0,0)$ et $w = (0,y,0)$. Donc $u = (x,y,0)$. Réciproquement un tel élément $u = (x,y,0)$ est la somme de $(x,0,0)$ et de $(0,y,0)$. Donc $F + G = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}$. On voit que, dans cet exemple, tout élément de $F + G$ s'écrit de façon unique comme la somme d'un élément de F et d'un élément de G.

Exemple 37 Soient F' et G' les deux sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F' = \{(x,y,z) \in \mathbb{R}^3 \mid x = 0\} \text{ et } G' = \{(x,y,z) \in \mathbb{R}^3 \mid y = 0\}$$

Dans cet exemple, montrons que $F' + G' = \mathbb{R}^3$.

Par définition de $F' + G'$, tout élément de $F' + G'$ est dans $\mathbb{R}^3$. Mais réciproquement si $u = (x,y,z)$ est un élément quelconque de $\mathbb{R}^3$:

$$u = (x,y,z) = (0,y,z) + (x,0,0)$$

donc u appartient à $F' + G'$.

Remarquons que, dans cet exemple, un élément de $\mathbb{R}^3$ ne s'écrit pas de façon unique comme la somme d'un élément de F' et d'un élément de G'. Par exemple,

$$(1,2,3) = (0,2,3) + (1,0,0) = (0,2,0) + (1,0,3).$$

Exemple 38 Dans le $\mathbb{R}$-espace vectoriel $A(\mathbb{R},\mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$, on considère le sous-espace vectoriel des fonctions paires $\mathcal{P}$ et le sous espace vectoriel des fonctions impaires I. Montrons que $\mathcal{P} + I = A(\mathbb{R},\mathbb{R})$. L'inclusion $\mathcal{P} + I \subset A(\mathbb{R},\mathbb{R})$ est évidente. Montrons l'inclusion inverse.

Soit f une application de $\mathbb{R}$ dans $\mathbb{R}$. Cherchons une fonction paire α et une fonction impaire β telles que $f = \alpha + \beta$. Si α et β existent, on a pour tout x de $\mathbb{R}$:

$$f(x) = \alpha(x) + \beta(x)$$

et

$$f(-x) = \alpha(-x) + \beta(-x) = \alpha(x) - \beta(x).$$

Donc, si α et β existent, elles sont uniques et on a nécessairement

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x) + f(-x)}{2} \end{array}$$

et

$$\begin{array}{rcl} \beta : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x)-f(-x)}{2}. \end{array}$$

Considérons les fonctions α et β définies ci dessus. La fonctions α est dans $\mathcal{P}$ et la fonction β est dans I. Il est facile de voir que l'on a $f = \alpha + \beta$. Donc $\mathcal{P} + I = A(\mathbb{R},\mathbb{R})$. Nous avons montré que tout élément de $A(\mathbb{R},\mathbb{R})$ s'écrit de façon unique comme la somme d'un élément de $\mathcal{P}$ et d'un élément de I.

Dans cette démonstration, on a montré l'unicité de α et β avant leur existence car cela nous a permis d'avoir une expression pour α et β.

La notion de somme de deux sous-espaces vectoriels d'un K-espace vectoriel se généralise en la notion de somme de plusieurs sous-espaces.

Définition 26 (Définition de la somme de n sous-espaces vectoriels) *Si $F_1, F_2, \ldots, F_n$ sont n sous-espaces vectoriels d'un K-espace vectoriel E, l'ensemble de toutes les sommes $x_1 + x_2 + \cdots + x_n$ où, pour tout entier p compris entre 1 et n, l'élément x_p appartient à F_p est appelé somme des sous espaces $F_1, F_2, \ldots, F_n$ et est noté $F_1 + F_2 + \cdots + F_n$ ou $\sum_{p=1}^{n} F_p$. On a donc*

$$F_1 + F_2 + \cdots + F_n = \{x \in E \mid \exists (x_1, x_2, \ldots, x_n) \in F_1 \times F_2 \times \cdots \times F_n, x = x_1 + x_2 + \cdots + x_n\}$$

Théorème 13 (Théorème de structure de la somme de n sous-espaces vectoriels) *La somme $F_1 + F_2 + \cdots + F_n$ des sous-espaces F_p, pour p compris entre 1 et n, est un sous-espace vectoriel de E.*

La démonstration est analogue au cas n = 2.

Exemple 39 Considérons dans $\mathbb{R}^4$ les trois sous-espaces vectoriels F, G et H, engendrés respectivement par $(1,0,0,0), (0,1,0,0), (0,0,1,0)$, alors tout élément de la somme $F + G + H$ s'écrit sous la forme $(x, y, z, 0)$ et donc :

$$F + G + H = \{(x, y, z, t) \mid t = 0\}.$$

Dans le deuxième exemple, nous avons vu que $F' + G' = \mathbb{R}^3$ mais qu'un élément de $\mathbb{R}^3$ ne s'écrit pas de façon unique comme la somme d'un élément de F' et d'un élément de G'. En revanche, dans le premier exemple, un élement de $F + G$ s'écrit de façon unique comme la somme d'un élement de F et d'un élément de G. De même, dans le troisième exemple, un élément de $A(\mathbb{R},\mathbb{R})$ s'écrit de façon unique comme la somme d'un élément de $\mathcal{P}$ et d'un élément de I. Nous sommes donc amenés à introduire la définition suivante.

Définition 27 (Définition de la somme directe de deux sous-espaces) *Etant donnés deux sous-espaces vectoriels F et G de E, la somme $F + G$ des sous-espaces F et G est dite directe et s'écrit $F \oplus G$ si et seulement si tout élément de $F + G$ s'écrit d'une manière unique comme la somme d'un élément de F et d'un élément de G.*

La somme $F \oplus G$ est appelée somme directe de F et G.

Notation : La notation $F + G = F \oplus G$ signifie que la somme $F + G$ est directe.

Remarque 30 On dit que l'élément u de $F+G$ s'écrit d'une manière unique comme somme d'un élément de F et d'un élément de G lorsque la propriété suivante est vérifiée : Si u s'écrit $u=v+w$ et $u=v'+w'$ avec v,v' éléments de F et w,w' éléments de G alors $v=v'$ et $w=w'$.

Proposition 24 *Une condition nécessaire et suffisante pour que la somme de deux sous-espaces vectoriels F et G soit directe est que 0_E s'écrive de manière unique comme la somme d'un élément de F et d'un élément de G. La seule façon d'écrire 0_E comme la somme d'un élément de F et d'un élément de G est alors $0_E = 0_E + 0_E$*

Démonstration. Il est évident que la condition est nécessaire. Montrons qu'elle est suffisante. Supposons que 0_E s'écrive de manière unique comme la somme d'un élément de F et d'un élément de G et montrons qu'il en est de même pour tout élément de $F+G$. Soit u un élément de $F+G$. Supposons que u s'écrive $u=v+w$ et $u=v'+w'$ avec v,v' éléments de F et w,w' éléments de G. On a alors $(v-v')+(w-w')=0_E$, ce qui implique $v-v'=0_E$ et $w-w'=0_E$. On en déduit $v=v'$ et $w=w'$, ce qu'il fallait démontrer. □

Proposition 25 (Propriété caractéristique) *Une condition nécessaire et suffisante pour que la somme de deux sous-espaces vectoriels F et G soit directe est que l'intersection de F et de G soit réduite au vecteur nul.*

$$F+G=F\oplus G \Longleftrightarrow F\cap G=\{0\}.$$

Démonstration. Supposons que $F+G=F\oplus G$. Si u est un élément quelconque de $F\cap G$, il peut s'écrire des deux manières suivantes comme somme d'un élément de F et d'un élément de G :

$$u=0+u \text{ et } u=u+0.$$

L'élément u étant un élément de $F\cap G$, est donc un élément de $F+G$. D'après l'unicité de l'écriture d'un élément de $F+G$, cela entraîne : u=0 . Donc $F\cap G=\{0\}$.

Réciproquement supposons que $F\cap G=\{0\}$. Soit u un élément de $F+G$. Si u s'écrit des deux manières comme la somme d'un élément de F et d'un élément de G :

$$u=v+w; \text{ et } u'=v'+w'$$

où v et v' sont des éléments de F et w et w' des éléments de G, alors $v-v'=w'-w$. Mais $v-v'$ élément de F et $w'-w$ est un élément de G (puisque F et G sont des sous-espaces vectoriels) donc $v-v'=w'-w$ est un élément de $F\cap G$, c'est donc l'élément nul. On en déduit que $v=v'$ et $w=w'$. L'écriture de u comme somme d'un élément de F et d'un élément de G est donc unique.

La notion de somme directe de deux sous-espaces vectoriels d'un K-espace vectoriel E se généralise au cas de plusieurs sous-espaces.

Définition 28 (Définition de la somme directe de n sous-espaces vectoriels) *La somme de n sous-espaces vectoriels $F_1,F_2,\ldots,F_n$ d'un K-espace vectoriel E est dite directe et s'écrit $F_1\oplus F_2\oplus\cdots\oplus F_n$ si et seulement si tout élément de $F_1+F_2+\cdots+F_n$ s'écrit d'une manière unique comme somme d'éléments de $F_1,F_2,\ldots,F_n$. Ce qui s'écrit avec les quantificateurs :*

$$\forall x\in F_1+F_2+\cdots+F_n, \exists!(x_1,\ldots,x_n)\in F_1\times F_2\times\ldots F_n,\ x=x_1+x_2+\cdots+x_n.$$

La notation $\exists!$ signifie : il existe un unique.

Exemple 40 Soit dans $\mathbb{R}^4$ les trois sous-espaces vectoriels F, G et H, engendrés respectivement par $(1,0,0,0)$, $(0,1,0,0)$ et $(0,0,1,0)$. Alors tout élément u de la somme $F+G+H$ s'écrit sous la forme

$$u = \alpha(1,0,0,0) + \beta(0,1,0,0) + \gamma(0,0,1,0)$$

donc $u = (\alpha, \beta, \gamma, 0)$. Si u s'écrivait aussi

$$u = \alpha'(1,0,0,0) + \beta'(0,1,0,0) + \gamma'(0,0,1,0)$$

alors on aurait $u = (\alpha', \beta', \gamma', 0)$. Mais dans $\mathbb{R}^4$:

$$(\alpha, \beta, \gamma, 0) = (\alpha', \beta', \gamma', 0) \Longrightarrow \alpha = \alpha', \beta = \beta', \gamma = \gamma'$$

donc l'écriture d'un élément de $F+G+H$ comme somme d'éléments de F, G et H est unique :

$$F+G+H = F \oplus G \oplus H.$$

Proposition 26 *La somme de n sous-espaces vectoriels $F_1, F_2, \ldots, F_n$ d'un K-espace vectoriel E est directe si et seulement 0_E s'écrit d'une manière unique comme somme d'éléments de $F_1, F_2, \ldots, F_n$.*

Démonstration. La démonstration est identique à celle donnée dans le cas de deux sous espaces vectoriels. □

Remarque 31 ATTENTION : Dans le cas de plusieurs sous-espaces vectoriels, le fait que les sous-espaces aient deux à deux une intersection réduite au vecteur nul n'est pas une condition suffisante pour que la somme soit directe

Contre exemple : Dans l'espace vectoriel $\mathbb{R}^2$, soit F le sous-espace vectoriel engendré par $(1,0)$, G le sous-espace vectoriel engendré par $(0,1)$ et H le sous-espace vectoriel engendré par $(1,1)$. Il est immédiat que $F \cap G = \{0\}$, $G \cap H = \{0\}$ et $F \cap H = \{0\}$ et pourtant la somme $F+G+H$ n'est pas directe. En effet l'élément $(1,1)$ de $F+G+H$ se décompose en somme d'éléments de F, G et H de la manière suivante :

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

mais aussi de la manière suivante :

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

donc il n'y a pas unicité de l'écriture.

Définition 29 *1) Deux sous-espaces vectoriels F et G d'un K-espace vectoriel E sont des* **sous-espaces vectoriels supplémentaires** *de E si leur somme est directe et est égale à l'espace vectoriel E tout entier :*

$$E = F \oplus G.$$

2) Si F et G sont des sous-espaces vectoriels supplémentaires du K-espace vectoriel E, on dit que F est un supplémentaire de G, ou que G est un supplémentaire de F.

Propriétés caractéristiques :

Deux sous-espaces vectoriels F et G d'un K-espace vectoriel E sont des sous-espaces vectoriels supplémentaires de E si et seulement si tout élément de E s'écrit d'une manière unique comme la somme d'un élément de F et d'un élément de G.

Deux sous-espaces vectoriels F et G d'un K-espace vectoriel E sont des sous-espaces vectoriels supplémentaires de E si et seulement si $E = F + G$ et $F \cap G = \{0\}$.

Remarque 32 L'existence d'un supplémentaire d'un sous-espace vectoriel sera prouvée dans le cadre des espaces vectoriels de type fini.

Remarque 33 Il n'y a pas unicité du supplémentaire d'un sous-espace vectoriel donné (voir exemple suivant).

Exemple 41 L'espace vectoriel $\mathbb{R}^2$ est la somme directe du sous-espace vectoriel F engendré par $(1,0)$ et du sous-espace G engendré par $(0,1)$, donc F et G sont des sous-espaces vectoriels supplémentaires de $\mathbb{R}^2$. Mais l'espace vectoriel $\mathbb{R}^2$ est aussi la somme directe du sous-espace vectoriel F engendré par $(1,0)$ et du sous-espace vectoriel H engendré par $(1,1)$, donc F et H sont aussi des sous-espaces supplémentaires de $\mathbb{R}^2$.

Soit $u = (x,y)$ un élément de $\mathbb{R}^2$, cherchons deux éléments $v \in F$ et $w \in H$ tels que $u = v + w$.

$$\begin{array}{l} v \in F \Longleftrightarrow \exists \alpha \in \mathbb{R}, v = (\alpha, 0) \\ w \in H \Longleftrightarrow \exists \beta \in \mathbb{R}, w = (\beta, \beta) \\ u = v + w \Longleftrightarrow (x,y) = (\alpha+\beta, \beta) \Longleftrightarrow \beta = y \text{ et } \alpha = x - y \end{array}$$

Ceci prouve que pour tout élément $u = (x,y)$ de $\mathbb{R}^2$, il existe un unique élément $w = (y,y)$ de H et un unique élément $v = (x-y,0)$ de F tels que $u = v + w$. On a bien $\mathbb{R}^2 = F \oplus H$.

Exemple 42 Soient les sous-espaces vectoriels F'' et G'' de $\mathbb{R}^3$ suivants :

$$F'' = \{(x,y,z) \in \mathbb{R}^3 \mid x - y - z = 0\} \text{ et } G'' = \{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\}$$

Les sous-espaces vectoriels F'' et G'' sont des sous-espaces de $\mathbb{R}^3$ supplémentaires. Montrons le :

Il est immédiat de vérifier que $F'' \cap G'' = \{0\}$. En effet si l'élément $u = (x,y,z)$ appartient à l'intersection de F'' et de G'', alors les coordonnées de u vérifient : $x - y - z = 0$ (car u appartient à F''), et $y = z = 0$ (car u appartient à G''), donc $u = (0,0,0)$.

Il reste à démontrer que $F'' + G'' = \mathbb{R}^3$.. Soit donc $u = (x,y,z)$ un élément quelconque de $\mathbb{R}^3$; il faut déterminer des éléments u_1 de F'' et u_2 de G'' dont la somme soit égale à u : . L'élément u_1 doit être de la forme $(y_1 + z_1, y_1, z_1)$ et l'élément u_2 de la forme $u_2 = (x_2, 0, 0)$. On a $u = u_1 + u_2$ si et seulement si $y_1 = y$, $z_1 = z$, $x_2 = x - y - z$. On a donc

$$(x,y,z) = (y+z, y, z) + (x-y-z, 0, 0)$$

avec $(y+z, y, z)$ dans F'' et $(x-y-z, 0, 0)$ dans G''.

Remarque 34 La vérification du 1) était inutile puisque la recherche de u_1 et de u_2 montre leur unicité. Mais lorsque l'on ne sait pas si la somme est directe ou ne l'est pas, il est souvent plus facile de commencer par vérifier si l'intersection est nulle ou ne l'est pas.

3.3 Applications linéaires

3.3.1 Définition et premières propriétes

Définition 30 *Soient E et F deux K-espaces vectoriels. Une application f de E dans F est appelée application linéaire si elle satisfait aux deux conditions suivantes :*
(1) Pour tous vecteurs u et v de E, $f(u+v) = f(u) + f(v)$.
(2) Pour tout vecteur u de E et pour tout scalaire λ de K, $f(\lambda u) = \lambda f(u)$

Autrement dit : une application est linéaire si elle " respecte " les deux lois d'un espace vectoriel.

Notation :
L'ensemble des applications linéaires de E dans F est noté $L(E,F)$ ou $L_K(E,F)$.

Proposition 27 *Soient E et F deux K-espaces vectoriels. Si f est une application linéaire de E dans F alors*
- $f(0_E) = 0_F$
- *Pour tout vecteur u de E, $f(-u) = -f(u)$.*

Démonstration. Il suffit d'appliquer la propriété (2) de linéarité avec $\lambda = 0$ puis avec $\lambda = -1$. □

Remarque 35 *La nécessité que E et F soient des espaces vectoriels sur le même corps K apparaît clairement dans ces calculs.*

Méthodologie
Soit f une application d'un espace vectoriel E dans un espace vectoriel F. Lorsqu'on cherche à répondre à la question suivante : "f est-elle linéaire ? ", on peut rapidement déterminer $f(0_E)$:
si $f(0_E) \neq 0_F$, alors on peut conclure que f n'est pas linéaire. Si $f(0_E) = 0_F$, on ne peut rien conclure et il faut alors vérifier que f satisfait à chacune des deux propriétés de linéarité.

Pour démontrer qu'une application est linéaire, on peut aussi utiliser une propriété plus "concentrée" donnée par la caractérisation suivante :

Proposition 28 (Caractérisation d'une application linéaire) *Soient E et F deux K-espaces vectoriels et f une application de E dans F. L'application f est linéaire si et seulement si, pour tous vecteurs u et v de E et pour tous scalaires α et β de K,*

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v).$$

Démonstration : Soient f une application linéaire de E dans F, u et v deux vecteurs de E, et deux éléments α et β de K. En utilisant la propriété (1) puis la propriété (2) de la linéarité de K, on a

$$\begin{aligned} f(\alpha u+\beta v) &= f(\alpha u)+f(\beta v) \\ &= \alpha f(u)+\beta f(v) \end{aligned}$$

Montrons la réciproque. Soit f une application de E dans F telle que, pour tous vecteurs u et v de E et pour tous scalaires α et β de K, $f(\alpha u+\beta v)=\alpha f(u)+\beta f(v)$. Alors, pour tous vecteurs u et v de E, on a $f(u+v)=f(u)+f(v)$. (égalité (3) dans le cas particulier où $\alpha=\beta=1$). Pour tout vecteur u de E et pour tout scalaire α de K, on a$f(\alpha u)=\alpha f(u)$ (égalité (3) dans le cas particulier où $\beta=0$).

On a une caractérisation un peu plus simple des applications linéaires.

Proposition 29 *Soient E et F deux K-espaces vectoriels et f une application de E dans F. L'application f est linéaire si et seulement si, pour tous vecteurs u et v de E et pour tout scalaire α de K,*

$$f(\alpha u+v)=\alpha f(u)+f(v).$$

Proposition 30 (Image d'une combinaison linéaire) *Soient E et F deux K-espaces vectoriels et f une application linéaire de E dans F, alors*

$$\begin{aligned} &\forall n\in N^*, \forall(\lambda_1,\ldots,\lambda_n)\in K^n, \forall(u_1,\ldots,u_n)\in E^n \\ &f\left(\sum_{i=1}^{n}\lambda_i u_i\right)=\sum_{i=1}^{n}\lambda_i f(u_i) \end{aligned}$$

Cette proposition se démontre par récurrence sur n.

Vocabulaire

Soient E et F deux K-espaces vectoriels.

Une application linéaire de E dans F est aussi appelée homomorphisme d'espaces vectoriels.

L'ensemble des applications linéaires de E dans F est noté $L(E,F)$.

Une application linéaire bijective de E sur F est appelée isomorphisme d'espaces vectoriels.

Une application linéaire de E dans E est appelée endomorphisme de E.

L'ensemble des endomorphismes de E est noté L(E).

Un endomorphisme bijectif de E est appelé automorphisme de E.

L'ensemble des automorphismes de E est noté $GL(E)$.

Exemple 43 L'application f définie par

$$\begin{aligned} f:\mathbb{R}^3 &\to \mathbb{R}^2 \\ (x,y,z) &\mapsto (x,y+z) \end{aligned}$$

est linéaire. En effet, soient $u=(x,y,z)$ et $v=(x',y',z')$ deux éléments de $\mathbb{R}^3$ et λ un réel.

$$\begin{aligned} f(\lambda u+v) &= f(\lambda x+x',\lambda y+y',\lambda z+z') \\ &= (\lambda x+x',\lambda y+y'+\lambda z+z') \\ &= \lambda(x,y+z)+(x',y'+z') \\ &= \lambda f(u)+f(v) \end{aligned}$$

Exemple 44 Soient $D(\mathbb{R},\mathbb{R})$ le sous-espace vectoriel de $A(\mathbb{R},\mathbb{R})$ composé des fonctions dérivables sur $\mathbb{R}$. L'application d de $D(\mathbb{R},\mathbb{R})$ dans $A(\mathbb{R},\mathbb{R})$ définie par

$$\begin{array}{rcl} d : D(\mathbb{R},\mathbb{R}) & \to & A(\mathbb{R},\mathbb{R}) \\ f & \mapsto & f' \end{array}$$

est linéaire. En effet, soient f et g deux fonctions dérivables sur $\mathbb{R}$ et soit α un réel, on a

$$d(\alpha f+g) = (\alpha f+g)' = \alpha f' + g' = \alpha d(f) + d(g).$$

Exemple 45 Considérons l'application de $M_{n,p}(K)$ dans $M_{p,n}(K)$ donnée par la transposition.

$$\begin{array}{rcl} T : M_{n,p}(K) & \to & M_{p,n}(K) \\ A & \mapsto & A^T \end{array}$$

T est linéaire car pour tous éléments de $M_{n,p}(K)$ et tout scalaire α, on a vu que

$$(\alpha A+B)^T = (\alpha A)^T + B^T = \alpha A^T + B^T.$$

Exemple 46 Soient E un K-espace-vectoriel, w un vecteur non nul de E et f l'application définie par

$$\begin{array}{rcl} f : E & \to & E \\ u & \mapsto & u+w \end{array}$$

f est appelée translation de vecteur w. $f(0_E) = w$ d'où $f(0_E) \neq 0_E$. L'application f n'est donc pas linéaire.

Exemple 47 Soit f l'application définie par :

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 \end{array}$$

$f(1) = 1$ et $f(2) = 4$. Donc $2f(1) \neq f(2)$. Donc f n'est pas linéaire.

3.3.2 L'espace vectoriel L(E,F)

Soient E et F deux K-espaces vectoriels. Rappelons que l'ensemble des applications de E dans F, noté $A(E,F)$, est muni d'une loi de composition interne + et d'une loi de composition externe définies de la façon suivante :

f,g étant deux éléments de $A(E,F)$, et λ étant un élément de K, pour tout vecteur u de E

$$(f+g)(u) = f(u)+g(u) \text{ et } (\lambda f)(u) = \lambda f(u).$$

F étant un K-espace vectoriel, l'ensemble des applications de E dans F, noté $A(E,F)$ est un K-espace vectoriel.

Proposition 31 *Soient E et F deux K-espaces vectoriels. L'ensemble des applications linéaires de E dans F, noté $L(E,F)$, muni des deux lois définies précédemment, est un K-espace vectoriel.*

Démonstration. L'ensemble $L(E,F)$ est inclus dans le K-espace vectoriel $A(E,F)$. Pour montrer que $L(E,F)$ est un K-espace vectoriel, il suffit donc de montrer que $L(E,F)$ est un sous-espace vectoriel de $A(E,F)$:

L'application nulle appartient à $L(E,F)$, donc $L(E,F)$ est non vide. Soient f,g deux éléments de $L(E,F)$, et λ un élément de K. Pour tous vecteurs u et v de E et pour tous scalaires α, β de K,

$$\begin{aligned}
(f+g)(\alpha u+\beta v) &= f(\alpha u+\beta v)+g(\alpha u+\beta v)(\text{définition de f}+\text{g})\\
&= \alpha f(u)+\beta f(v)+\alpha g(u)+\beta g(v)(\text{linéarité de f et g})\\
&= \alpha\left(f(u)+g(u)\right)+\beta\left(f(v)+g(v)\right)(\text{propriétés des lois de F})\\
&= \alpha(f+g)(u)+\beta(f+g)(v)(\text{définition de f}+\text{g})
\end{aligned}$$

$f+g$ est donc linéaire et $L(E,F)$ est stable pour l'addition.

$$\begin{aligned}
(\lambda f)(\alpha u+\beta v) &= \lambda f(\alpha u+\beta v)\ \ (\text{définition de }\lambda\text{f})\\
&= \lambda\left(\alpha f(u)+\beta f(v)\right)\ \ (\text{linéarité de f})\\
&= \alpha\lambda f(u)+\beta\lambda f(v)(\text{propriétés des lois de } F)\\
&= \alpha(\lambda f)(u)+\beta(\lambda f)(v)(\text{définition de } \lambda f)
\end{aligned}$$

λf est donc linéaire et $L(E,F)$ est stable pour la loi externe.

$L(E,F)$ est donc un sous-espace vectoriel de $A(E,F)$. En particulier $L(E)$ est un sous-espace vectoriel de $A(E,E)$.

Proposition 32 (Composée de deux applications linéaires) *Soient E,F,G trois K-espaces vectoriels, f une application linéaire de E dans F et g une application linéaire de F dans G, alors $g\circ f$ est une application linéaire de E dans G.*

Remarque 36 En particulier, le composé de deux endomorphismes de E est un endomorphisme de E. Autrement dit $\circ$ est une loi de composition interne sur $L(E)$

Démonstration. Soient u et v deux vecteurs de E, et α et β deux éléments de K.

$$\begin{aligned}
(g\circ f)(\alpha u+\beta v) &= g\left(f(\alpha u+\beta v)\right)\ \ (\text{définition de } g\circ f)\\
&= g\left(\alpha f(u)+\beta f(v)\right)\ \ (\text{linéarité de f})\\
&= \alpha g\left(f(u)\right)+\beta g\left(f(v)\right)\ \ (\text{linéarité de g})
\end{aligned}$$

Attention ! Si les espaces vectoriels E et G sont distincts, on ne peut pas définir l'application $f\circ g$.

Proposition 33 (Propriétés de la composition d'applications linéaires) *Soient E, F, G trois K-espaces vectoriels.*

1. $\forall(f_1,f_2)\in L(E,F)\times L(E,F),\ \ \forall g\in L(F,G),\ \ g\circ(f_1+f_2)=g\circ f_1+g\circ f_2$

2. $\forall f\in L(E,F),\ \forall(g_1,g_2)\in L(F,G)\times L(F,G),\ \ (g_1+g_2)\circ f=g_1\circ f+g_2\circ f$

3. $\forall\alpha\in K,\forall f\in L(E,F),\forall g\in L(F,G),\ \ (\alpha g)\circ f=g\circ(\alpha f)=\alpha(g\circ f)$

Démonstration. 1. Pour tout vecteur u de E, on a

$$\begin{aligned}(g\circ(f_1+f_2))(u) &= g((f_1+f_2)(u))\\ &= g((f_1(u)+f_2(u))\\ &= g(f_1(u))+g(f_2(u))\end{aligned}$$

La dernière égalité utilise la linéarité de g. Les autres égalités se déduisent de la définition de la loi $\circ$ et de la loi + .

$$\begin{aligned}[(g\circ f_1)+(g\circ f_2)](u) &= (g\circ f_1)(u)+(g\circ f_2)(u)\\ &= g(f_1(u))+g(f_2(u))\end{aligned}$$

Remarque : Cette démonstration utilise la linéarité de g, mais pas celles de f_1 et de f_2.

2. Pour tout vecteur u de E,

$$\begin{aligned}((g_1+g_2)\circ f)(u) &= (g_1+g_2)(f(u))\\ &= g_1(f(u))+g_2(f(u))\\ &= (g_1\circ f+g_2\circ f)(u)\end{aligned}$$

Ces égalités se déduisent de la définition de la loi $\circ$ et de la loi +.

Remarquons que la démonstration de 2. n'utilise pas de linéarité.

3. La preuve de la troisième propriété est laissée au lecteur. La démonstration de la formule $(\alpha g)\circ f=\alpha(g\circ f)$ n'utilise pas de linéarité. La démonstration de la formule $g\circ(\alpha f)=\alpha(g\circ f)$ utilise la linéarité de g.

Proposition 34 (Linéarité de l'application réciproque d'un isomorphisme) *Soient E et F deux K-espaces vectoriels, si f est un isomorphisme de E sur F, alors f^{-1} est un isomorphisme de F sur E.*

Démonstration. f étant une application bijective de E sur F, f^{-1} est une application bijective de F sur E. Il reste donc à prouver que f^{-1} est bien linéaire. Soient u' et v' deux vecteurs de F et soient α et β deux éléments de K, on pose $f^{-1}(u')=u$ et $f^{-1}(v')=v$ et on a alors $f(u)=u'$ et $f(v)=v'$. Comme f est linéaire, on a

$$f^{-1}(\alpha u'+\beta v')=f^{-1}(\alpha f(u)+\beta f(v))=f^{-1}(f(\alpha u+\beta v))=\alpha u+\beta v$$

car $f^{-1}\circ f=Id_E$ (où Id_E désigne l'application identique de E dans E) donc

$$f^{-1}(\alpha u'+\beta v')=\alpha f^{-1}(u')+\beta f^{-1}(v')$$

f^{-1} est donc linéaire.

Vocabulaire

La proposition précédente prouve donc que s'il existe un isomorphisme de E sur F, alors il existe aussi un isomorphisme de F sur E. Les deux espaces vectoriels E et F sont dits isomorphes.

3.3.3 Exemples d'endomorphismes : homothétie, projection

•Homothétie

Soient E un K-espace-vectoriel, et λ un élément de K. On définit l'application f_λ par :

$$\begin{array}{rcl} f_\lambda : E & \to & E \\ u & \mapsto & \lambda u \end{array}$$

f_λ est linéaire. En effet, soient u et v deux vecteurs de E, et α, β deux scalaires de K. En utilisant les propriétés des lois de l'espace vectoriel E et la définiton de f_λ

$$\begin{array}{rcl} f_\lambda(\alpha u+\beta v) & = & \lambda(\alpha u+\beta v) \\ & = & \lambda(\alpha u)+\lambda(\beta v) \\ & = & (\lambda\alpha)u+(\lambda\beta)v \\ & = & (\alpha\lambda)u+(\beta\lambda)v \\ & = & \alpha(\lambda u)+\beta(\lambda v) \\ & = & \alpha f_\lambda(u)+\beta f_\lambda(v) \end{array}$$

f_λ est appelée homothétie de rapport λ. Dans le cas particulier où $\lambda = 1$, f_λ est l'application identité. Dans le cas particulier où $\lambda = 0$, f_λ est l'application nulle.

Si $k \neq 0$, f_λ est une bijection de E sur E (tout élément v de E admet un antécédent unique $u = \dfrac{1}{\lambda}$) donc c'est un automorphisme de E.

• **Projection**

Soient E un K-espace-vectoriel et F et G deux sous-espaces vectoriels supplémentaires dans E. Tout vecteur u de E s'écrit de façon unique $u = v+w$ avec v élément de F et w élément de G. L'unicité de la décomposition précédente permet de définir l'application p de E dans E telle que $p(u) = v$. L'application p est appelée projection sur F parallèlement à G. C'est une application linéaire.

En effet, soient deux vecteurs u et u' de E, et deux scalaires α , β deux scalaires de K, le vecteur u s'écrit de façon unique $u = v+w$ avec v élément de F et w élément de G et, par définition de p, $p(u) = v$. De même, le vecteur u' s'écrit de façon unique $u' = v'+w'$ avec v' élément de F et w' élément de G et, par définition de p, $p(u') = v'$.

$$\alpha u+\beta u' = (\alpha v+\beta v')+(\alpha w+\beta w').$$

F est un sous-espace vectoriel de E, il est donc stable par combinaison linéaire et donc le vecteur $\alpha v+\beta v'$appartient à F. De même le vecteur $\alpha w+\beta w'$ appartient à G et, d'après la définition de p,on a

$$p(\alpha u+\beta u') = \alpha v+\beta v' = \alpha p(u)+\beta p(u').$$

Une projection p vérifie l'égalité $p^2 = p$. En effet, soit p la projection sur F parallèlement à G, tout vecteur u de E s'écrit de façon unique $u = v+w$ avec v élément de F et w élément de G. on a alors $p(u) = v$ et $p(v) = v$ car $v = v+0$ avec v élément de F et 0 élément de G. Ainsi

$$p^2(u) = p\,(p(u)) = p(v) = v = p(u).$$

Exemple 48 Nous avons vu que les sous-espaces vectoriels F'' et G'' de $\mathbb{R}^3$ suivants :

$$F'' = \{(x,y,z) \in \mathbb{R}^3 \mid x-y-z = 0\} \text{ et } G'' = \{(x,y,z) \in \mathbb{R}^3 \mid y = z = 0\}$$

sont supplémentaires dans $\mathbb{R}^3$. Soit p la projection sur F'' parallèlement à G''. D'après les calculs faits précédemment, on a $p(x,y,z) = (y+z,y,z)$.

Exemple 49 Nous avons vu que l'ensemble des fonctions paires $\mathcal{P}$ et l'ensemble des fonctions impaires I sont des sous espace vectoriels supplémentaires dans $A(\mathbb{R},\mathbb{R})$. Notons p la projection sur $\mathcal{P}$ parallèllement à I. Si f est un élément de $A(\mathbb{R},\mathbb{R})$, on a $p(f) = \alpha$ où

$$\begin{array}{rcl} \alpha : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x)+f(-x)}{2}. \end{array}$$

3.3.4 Applications linéaires et sous espaces vectoriels

Soient E et F deux ensembles et f une application de E dans F. Soit A un sous-ensemble de E. L'ensemble des images par f des éléments de A, appelé " image de A par f ", est notée $f(A)$. C'est un sous ensemble de F. On a

$$f(A) = \{y \in F \mid \exists x \in A, f(x) = y\}.$$

$f(E)$ s'appelle l'image de f et est noté Imf.

Dans toute la suite, E et F désigneront des K-espaces vectoriels.

Proposition 35 (Structure de l'image d'un sous espace vectoriel) *Soit f une application linéaire du K-espace vectoriel E dans le K-espace vectoriel F. Si A est un sous-espace vectoriel de E, alors $f(A)$ est un sous-espace vectoriel de F. En particulier* Imf *est un sous-espace vectoriel de F.*

Démonstration. Comme A est un sous-espace vectoriel de E, il contient l'élément 0_E, donc $f(0_E)$ (qui est égal à 0_F) appartient à $f(A)$. Donc $f(A)$ est non vide. Ensuite on montre que pour tout couple (y_1, y_2) d'éléments de $f(A)$ et pour tout scalaire α , l'élément $y_1 + \alpha y_2$ appartient à $f(A)$. En effet :

$$\begin{array}{l} y_1 \in f(A) \Longleftrightarrow \exists x_1 \in A, f(x_1) = y_1 \\ y_2 \in f(A) \Longleftrightarrow \exists x_2 \in A, f(x_2) = y_2 \end{array}$$

Comme f est linéaire, on a

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

Or $x_1 + \alpha x_2$ est un élément de A, car A est un sous-espace vectoriel de E, donc $y_1 + \alpha y_2$ est bien un élément de $f(A)$.

Définition 31 (Définition du noyau) *Soient E et F deux K-espaces vectoriels et f une application linéaire de E dans F. Le noyau de f, noté* Ker(f) *, est l'ensemble des éléments de E dont l'image est 0_F .*

$$\mathrm{Ker(f)} = \{\mathrm{x} \in \mathrm{E} \mid \mathrm{f(x)} = 0_{\mathrm{F}}\}$$

Proposition 36 *Soient E et F deux K-espaces vectoriels et f une application linéaire de E dans F. Le noyau de f est un sous-espace vectoriel de E.*

Démonstration. Ker(f) est non vide car il contient 0_E. Soient x_1 et x_2 deux éléments de Ker(f) et α un scalaire. Montrons que $x_1 + \alpha x_2$ est élément de Ker(f). On a, en utilisant la linéarité de f et le fait que x_1 et x_2 sont éléments de Ker(f) :

$$f(x_1 + \alpha x_2) = f(x_1) + \alpha f(x_2) = 0_F.$$

Exemple 50 *Soient E un K-espace vectoriel, et λ un élément de K. Considérons f_λ l'homothétie de rapport λ. On a :*

$$\begin{array}{rcl} f_\lambda : E & \to & E \\ x & \mapsto & \lambda x \end{array}$$

Si $\lambda = 0$, f_λ est l'application nulle de E, donc $\mathrm{Im(f_0)} = \{0_\mathrm{E}\}$ *et $Ker(f_0) = E$.*

Si $\lambda \neq 0$, f_λ est une bijection, alors $\mathrm{Im(f_\lambda)} = \mathrm{E}$ *, et le seul élément de E ayant pour image 0_E est 0_E, donc* $\mathrm{Ker(f_\lambda)} = \{0_\mathrm{E}\}$.

Exemple 51 Soient E un K-espace vectoriel, F et G deux sous-espaces vectoriels de E, supplémentaires, et p la projection sur F parallèlement à G. Si le vecteur u de E s'écrit d'une manière unique $u = u_F + u_G$ avec u_F élément de F et u_G élément de G, alors $p(u) = u_F$. Le noyau de p est l'ensemble des vecteurs u de E tels que $u_F = 0$, c'est donc G. Montrons que l'image de p est F. Il est immédiat que l'image de p est contenue dans F. Réciproquement tout élément de F est sa propre image, ce qui fournit l'inclusion $F \subset \mathrm{Im(p)}$

$$\mathrm{Ker(p) = G \quad et \quad Im(p) = F.}$$

Théorème 14 (Caractérisation des applications linéaires injectives) *Soient E et F deux K-espaces vectoriels et f une application linéaire de E dans F. L'application f est injective si et seulement si son noyau ne contient que le vecteur nul.*

Démonstration. Supposons que f soit injective et montrons que $\mathrm{Ker(f)} = \{0\}$. Soit x un élément de $\mathrm{Ker(f)}$. On a $f(x) = 0_F$. Or, comme f est linéaire, on a aussi $f(0_E) = 0_F$. De l'égalité $f(x) = f(0_E)$, on déduit $x = 0_E$ car f est injective. Donc $\mathrm{Ker(f)} = \{0_\mathrm{E}\}$.

Supposons maintenant que $\mathrm{Ker(f)} = \{0\}$. Soient x et y deux éléments de E tels que $f(x) = f(y)$. On a donc $f(x) - f(y) = 0_F$. Comme f est linéaire, on en déduit $f(x-y) = 0_F$, c'est à dire $x-y$ est élément de $\mathrm{Ker(f)}$. Donc $x-y = 0_E$, soit $x = y$. Ce qui achève la démonstration du théorème.

Remarque 37 Avec les notations du théorème, on a comme d'habitude, f surjective si et seulement si $\mathrm{Im}\, f = F$.

Exemple 52 Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}^3$ définie par :

$$\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R}^3 \\ (x,y) & \mapsto & (x,x,-x) \end{array}$$

On montre que f est linéaire. L'image de f est l'ensemble de tous les triplets $(x,x,-x)$ pour x parcourant $\mathbb{R}$. On a donc

$$\mathrm{Im(f)} = \mathrm{vect}\left((1,1,-1)\right).$$

Le noyau de f est l'ensemble des couples $(0,y)$ pour y parcourant $\mathbb{R}$. On a donc

$$\mathrm{Ker(f)} = \mathrm{vect}\left((0,1)\right).$$

f n'est ni surjective, ni injective.

Chapitre 4

Espaces vectoriels de type fini, bases

4.1 Espaces vectoriels de type fini

4.1.1 Ensemble fini de générateurs d'un espace vectoriel

Définition 32 *Soit E un espace vectoriel sur un corps K. Soit p un entier supérieur à 1 et p vecteurs de E, $v_1, \ldots, v_p$. Les vecteurs $v_1, \ldots, v_p$ engendrent E si tout élément de E est combinaison linéaire des vecteurs $v_1, \ldots, v_p$, ce qui peut s'écrire, avec le symbolisme mathématique :*

$$\forall x \in E,\ \exists(\alpha_1, \ldots, \alpha_p) \in K^p \mid x = \alpha_1 v_1 + \cdots + \alpha_p v_p$$

Les vecteurs $(v_1, \ldots, v_p)$ engendrent E si et seulement si $E = \text{vect}(v_1, \ldots, v_p)$.

Vocabulaire :
Si les vecteurs $v_1, \ldots, v_p$ engendrent E, ils constituent un ensemble fini de générateurs de E ou une famille de générateurs de E. Dans ce cas, l'ensemble $\{v_1, \ldots, v_p\}$ est appelée aussi partie génératrice de E.

Les termes "ensemble de générateurs" ou "famille de générateurs" sont les termes utilisés usuellement par la communauté mathématique, c'est pourquoi nous les avons indiqués ici. Ils n'ont cependant pas le même statut. En effet, quand on parle d'une famille d'éléments d'un ensemble, il peut y avoir des éléments égaux et l'ordre des éléments importe. En revanche, quand on parle d'ensemble, le mot désigne une liste d'objets distincts et l'ordre des éléments n'importe pas.

Proposition 37 *Soit E un K-espace vectoriel admettant une famille finie de générateurs $(v_1, \ldots, v_p)$. Alors, toute partie A de E contenant les vecteurs $v_1, \ldots, v_p$ est encore une partie génératrice de E.*

Démonstration. Ceci est tout à fait immédiat en reprenant la définition de sous-espace engendré par une partie et en utilisant le fait que les vecteurs $v_1, \ldots, v_p$ sont des éléments de A. □

Remarque 38 Si la famille de vecteurs $(v_1,\dots,v_p)$ engendrent E, une sous-famille de la famille peut ne pas engendrer E. Considérons par exemple $E=\mathbb{R}^3$. Soient les vecteurs $u=(1,0,0)$, $v=(0,1,0)$ et $w=(0,0,1)$. Ils engendrent E. En effet tout élément (x,y,z)de $\mathbb{R}^3$ peut s'écrire

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

En revanche, si l'on ne considère que la partie composée des éléments $(1,0,0)$ et $(0,1,0)$, elle n'engendre pas $\mathbb{R}^3$. Il suffit pour justifier cette affirmation de trouver un élément de $\mathbb{R}^3$ qui n'est pas combinaison linéaire des vecteurs u et v . Le vecteur w, par exemple, n'est pas une combinaison linéaire des vecteurs u et v sinon il existerait deux réels a et b tels que

$$w=(0,0,1)=a(1,0,0)+b(0,1,0).$$

entrainant, entre autre que $0=1$ (en regardant les troisièmes composantes).

Proposition 38 (Réduction d'une famille génératrice) *Si les vecteurs $v_1,\dots,v_p$ engendrent E et si l'un des vecteurs, par exemple v_p, est combinaison linéaire des autres, alors la partie $\{v_1,\dots,v_p\}-\{v_p\}=\{v_1,\dots,v_{p-1}\}$ engendre E.*

En effet, comme les vecteurs $v_1,\dots,v_p$ engendrent E, pour tout élément x de E, il existe des scalaires $(\lambda_1,\dots,\lambda_p)$ tels que

$$x=\lambda_1 v_1+\cdots+\lambda_p v_p.$$

Or l'hypothèse v_p est combinaison linéaire des vecteurs $(v_1,\dots,v_{p-1})$ se traduit par l'existence de scalaires $(\alpha_1,\dots,\alpha_{p-1})$ tels que .

$$v_p=\alpha_1 v_1+\cdots+\alpha_{p-1}v_{p-1}.$$

Alors, le vecteur x s'écrit :

$$x=\lambda_1 v_1+\cdots+\lambda_{p-1}v_{p-1}+\lambda_p\left(\alpha_1 v_1+\cdots+\alpha_{p-1}v_{p-1}\right).$$

soit

$$x=\left(\lambda_1+\lambda_p\alpha_1\right)v_1+\cdots+\left(\lambda_{p-1}+\lambda_p\alpha_{p-1}\right)v_{p-1}$$

ce qui prouve que x est combinaison linéaire des vecteurs $v_1,\dots,v_{p-1}$. Ceci achève la démonstration. Il est clair que si l'on remplace v_p par n'importe lequel des vecteurs v_i, la démonstration est la même. □

Remarque 39 Un K-espace vectoriel quelconque ne possède pas obligatoirement de système fini de générateurs. Par exemple l'espace vectoriel réel des fonctions polynômes sur $\mathbb{R}$.

Remarque 40 Soit E un K-espace vectoriel et F un sous-espace vectoriel de E. Si F admet une famille génératrice $(v_1,\dots,v_p)$, il résulte de la définition (appliquée à l'espace vectoriel F) que les vecteurs $v_1,\dots,v_p$ sont nécessairement éléments de F.

Exemple 53 Soit le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$ et les vecteurs $v=(1,0)$ et $w=(1,1)$. Les vecteurs v et w engendrent E. En effet, soit $u=(x,y)$ un élément quelconque de $\mathbb{R}^2$. Montrer que u est combinaison linéaire de v et w revient à démontrer l'existence de

deux réels α et β tels que $u = \alpha v + \beta w$. Il s'agit donc d'étudier l'existence de solutions du système :

$$\begin{cases} \alpha + \beta &= x \\ \beta &= y \end{cases}$$

Il a pour solution $\beta = y$ et $\alpha = x - y$ et ceci, quels que soient les réels x et y. Toujours dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$, il est facile de démontrer que $\{(1,0),(0,1)\}$ est aussi une partie génératrice de $\mathbb{R}^2$ $((x,y) = x(1,0) + y(0,1))$.

Ceci prouve qu'il peut exister plusieurs familles finies différentes, non incluses les unes dans les autres, engendrant le même espace vectoriel.

Exemple 54 Soit $P_n(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions polynômes de degré inférieur où égal à n. Soient les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ définies pour tout x de $\mathbb{R}$ par :

$$f_0(x) = 1,\ f_1(x) = x,\ \ldots, f_k(x) = x^k, \ldots, f_n(x) = x^n.$$

Les fonctions $(f_0, \ldots, f_n)$ constituent une famille génératrice de $P_n(\mathbb{R})$.

Exemple 55 Soit $E = \mathbb{R}$ considéré comme un $\mathbb{R}$-espace vectoriel. Soit a un élément non nul de $\mathbb{R}$. Alors $\{a\}$, où a est considéré comme un vecteur, est une partie génératrice de E. En effet, soit x un élément quelconque de $\mathbb{R}$. Il peut s'écrire

$$x = \frac{x}{a} a.$$

L'inverse de a, $\frac{1}{a}$, existe car a est non nul ; dans cette égalité, $\frac{x}{a}$ joue le rôle d'un scalaire et a celui d'un vecteur.

Définition 33 (Définition d'un espace vectoriel de type fini) *Un espace vectoriel est dit de type fini s'il admet une famille finie de générateurs.*

Exemple 56 Il résulte des trois exemples de la page précédente que $\mathbb{R}^2$, $P_n(\mathbb{R})$ et $\mathbb{R}$ sont des espaces vectoriels de type fini.

Il est clair, en particulier en considérant les espaces vectoriels $\mathbb{R}^2$ et $\mathbb{R}$, qu'il peut exister plusieurs familles finies différentes de générateurs d'un espace vectoriel de type fini. Cela a été vu dans le cadre du premier exemple, concernant $\mathbb{R}^2$. En ce qui concerne $\mathbb{R}$, tout élément non nul de $\mathbb{R}$ est un système générateur de $\mathbb{R}$.

De plus, si G est une famille finie de générateurs d'un espace vectoriel E, un élément peut avoir plusieurs décompositions sur cette famille de vecteurs. Par exemple, considérons $\mathbb{R}^2$ et les vecteurs $u = (1,0)$, $v = (0,1)$ et $w = (1,1)$. Il résulte de ce qui précède que $\{u,v,w\}$ est une partie génératrice de $\mathbb{R}^2$ (car elle contient $\{u,v\}$ qui est une partie génératrice d'après le premier exemple). Or, si a est un réel non nul quelconque, pour tout (x,y) de $\mathbb{R}^2$, il est possible d'écrire les deux décompositions distinctes suivantes :

$$(x,y) = xu + yv + 0w$$
$$(x,y) = (x-a)u + (y-a)v + aw.$$

4.1.2 Dépendance et indépendance linéaire

Cette définition est aussi importante que la définition d'une famille génératrice.

Définition 34 *Soit E un K-espace vectoriel. Une famille $(v_1, v_2, \ldots, v_n)$ de E est dite* linéairement indépendante *ou* libre *si toute combinaison linéaire nulle*

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

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

Par convention, on posera que l'ensemble vide est une famille libre.
On définit de même la notion de partie libre ou linéairement indépendante.

Remarque 41 *Toute famille $(v_1, v_2, \ldots, v_n)$ contenant deux vecteurs égaux est liée. En effet, si on suppose $v_i = v_j$ avec $i < j$, on a la relation de dépendance linéaire non triviale*

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_{j-1} - 1.v_j + 0v_{j+1} + \cdots + 0v_n.$$

Exemple 57 Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$, considérons la partie

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

On souhaite déterminer si elle est libre ou liée. On cherche $(\lambda_1, \lambda_2, \lambda_3)$ tel que

$$\begin{cases} \lambda_1 + 4\lambda_2 + 2\lambda_3 = 0 \\ 2\lambda_1 + 5\lambda_2 + \lambda_3 = 0 \\ 3\lambda_1 + 6\lambda_2 = 0 \end{cases}$$

Pour cela, il suffit d'effectuer la réduction de Gauss sur la matrice associée au système.

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

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

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

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

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

est donc liée. □

Exemple 58 Dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$, la famille $(1,i)$ est libre. En effet pour tous scalaires α et β, on a

$$\alpha 1+\beta i=0 \Longrightarrow \alpha=\beta=0.$$

Exemple 59 Dans le $\mathbb{R}$-espace vectoriel $A(\mathbb{R},\mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$, on considère la famille $(\cos,\sin)$. Montrons que c'est une famille libre. Supposons que l'on ait $\lambda\cos+\mu\sin=0$. Cela équivaut à

$$\forall x\in\mathbb{R},\ \ \lambda\cos x+\mu\sin x=0.$$

En particulier, pour $x=0$, cette égalité donne $\lambda=0$. Et pour $x=\dfrac{\pi}{2}$, elle donne $\mu=0$. Donc la famille $(\cos,\sin)$ est libre. En revanche la famille $(\cos^2,\sin^2,1)$ est liée car on a la relation de dépendance linéaire $\cos^2+\sin^2-1=0$.

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

Proposition 39 *i) La famille $\{v\}$ est linéairement indépendante si $v\neq 0$ et linéairement dépendante si $v=0$.*

ii) La famille $\{v_1,v_2\}$ est linéairement indépendante si et seulement si v_1 n'est pas un multiple de v_2 et v_2 n'est pas un multiple de v_1.

Démonstration. Le point i) est trivial. Pour le point ii), supposons d'abord la famille liée. Il existe donc λ_1,λ_2 non tous les deux nuls tels que $\lambda_1v_1+\lambda_2v_2=0$. Si c'est λ_1 qui n'est pas nul, on peut diviser par λ_1, ce qui donne $v_1=-\frac{\lambda_2}{\lambda_1}v_2$ et v_1 est un multiple de v_2. Si c'est λ_2 qui n'est pas nul, alors de même v_2 est un multiple de v_1. On vient de montrer que si la famille est liée, alors v_1 est un multiple de v_2 *ou* v_2 est un multiple de v_1, ce qui est la négation logique de l'assertion "v_1 n'est pas un multiple de v_2 *et* v_2 n'est pas un multiple de v_1 ".

Réciproquement, si v_1 est un multiple de v_2, alors il existe un scalaire μ tel que $v_1=\mu v_2$, soit $1v_1+(-\mu)v_2=0$ ce qui est une relation de dépendance linéaire entre v_1 et v_2 puisque $1\neq 0$. De même, v_2 est un multiple de v_1, alors la famille est liée, d'où la réciproque. □

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

Théorème 15 *Soit E un K-espace vectoriel. Une famille $\mathcal{F}=(v_1,v_2,\ldots,v_n)$ de $n\geq 2$ vecteurs de E est linéairement dépendante si et seulement si au moins un des vecteurs de $\mathcal{F}$ est combinaison linéaire des autres vecteurs de $\mathcal{F}$.*

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

$$\lambda_1v_1+\lambda_2v_2+\cdots+\lambda_nv_n=0,$$

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

$$\lambda_kv_k=-\lambda_1v_1-\lambda_2v_2-\cdots-\lambda_nv_n,$$

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

$$v_k=-\frac{\lambda_1}{\lambda_k}v_1-\frac{\lambda_2}{\lambda_k}v_2-\cdots-\frac{\lambda_n}{\lambda_k}v_n,$$

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

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

$$v_k = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n,$$

où v_k ne figure pas au second membre. Passant v_k au second membre, il vient

$$0 = \mu_1 v_1 + \mu_2 v_2 + \cdots - v_k + \cdots + \mu_n v_n,$$

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

Proposition 40 *a) Toute partie contenant une partie liée est liée.*
b) Toute partie contenue dans une partie libre est libre.

Démonstration. a) Soient $A = \{v_1, \ldots, v_p\}$ une partie liée et $A' = \{v_1, \ldots, v_p, \ldots, v_n\}$ une partie contenant A. Il existe donc des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_p$ non tous nuls, tels que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = 0.$$

Si l'on pose $\lambda_{p+1} = \lambda_{p+2} = \cdots = \lambda_n = 0$, on peut écrire l'égalité :

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p + \lambda_{p+1} v_{p+1} + \cdots + \lambda_n v_n = 0.$$

Il existe donc une combinaison linéaire nulle, des vecteurs $v_1, v_2, \ldots, v_p, \ldots, v_n$, à coefficients non tous nuls (l'un des λ_i avec i compris entre 1 et p est non nul) et par conséquent A' est une partie liée.

b) C'est la contraposée de a). □

Corollaire 16 *Toute partie contenant le vecteur 0_E est liée.*

Proposition 41 (Adjonction d'un vecteur à une partie libre) *Soient E un K-espace vectoriel et $\{v_1, v_2, \ldots, v_n\}$ une partie libre de E. Si u est un vecteur de E tel que $\{v_1, v_2, \ldots, v_n, u\}$ soit une partie liée de E, alors le vecteur u est combinaison linéaire des vecteurs $v_1, v_2, \ldots, v_n$.*

Démonstration. Les vecteurs $v_1, v_2, \ldots, v_n, u$ sont linéairement dépendants. Il existe donc des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta$ non tous nuls tels que

$$(1) \qquad \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n + \beta u = 0.$$

Le coefficient β peut-il être nul ? Si β est nul, l'égalité (1) devient :

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.$$

avec au moins un des coefficients non nul ; ceci est impossible car contraire à l'hypothèse $\{v_1, v_2, \ldots, v_n\}$ partie libre. Donc β est non nul, il est donc inversible dans K et on peut déduire de l'égalité (1) l'égalité suivante :

$$u = -\frac{\alpha_1}{\beta} v_1 - \frac{\alpha_2}{\beta} v_2 + \cdots - \frac{\alpha_n}{\beta} v_n.$$

Ce qui signifie que u est combinaison linéaire des vecteurs $v_1, v_2, \ldots, v_n$. □

Remarque 42 Il est intéressant de décortiquer cet énoncé : l'hypothèse $\{v_1, v_2, \ldots, v_n, u\}$ " est une partie liée " implique immédiatement que l'un des vecteurs est combinaison linéaire des autres, et l'hypothèse $\{v_1, v_2, \ldots, v_n\}$ "est une partie libre " permet de conclure que c'est le vecteur que l'on a rajouté u qui est combinaison linéaire des autres.

Proposition 42 *Soient E et F deux K-espaces vectoriels et $f : E \to F$ une application linéaire. Soient $v_1, v_2, \ldots, v_p$ p vecteurs de E.*

a) Si la famille $(f(v_1), f(v_2), \ldots, f(v_p))$ est libre alors la famille $(v_1, v_2, \ldots, v_p)$ est libre.

b) On suppose f injective. Si la famille $(v_1, v_2, \ldots v_p)$ est libre,alors la famille $(f(v_1), f(v_2), \ldots, f(v_p))$ est libre.

Démonstration. a) On suppose

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = 0.$$

On applique f à cette égalité, il vient (en utilisant la linéarité de f) :

$$\lambda_1 f(v_1) + \lambda_2 f(v_2) + \cdots + \lambda_p f(v_p) = 0$$

d'où l'on déduit $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$ car la famille $(f(v_1), f(v_2), \ldots, f(v_p))$ est libre.

b) On suppose

$$\lambda_1 f(v_1) + \lambda_2 f(v_2) + \cdots + \lambda_p f(v_p) = 0.$$

En utilisant la linéarité de f, cette égalité s'écrit

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p) = 0.$$

Comme f est injective, on en déduit

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = 0$$

ce qui implique, $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$ car la famille $(v_1, v_2, \ldots v_p)$ est libre. □

4.1.3 Notion de bases dans un espace vectoriel de type fini

La notion de base généralise la notion de repères. Dans $\mathbb{R}^2$, un repère est donné par un couple de vecteurs non colinéaires. Dans $\mathbb{R}^3$, un repère est donné par un triplet de vecteurs non coplanaires. Dans un repère, un vecteur se décompose suivant les vecteurs de bases. Il en sera de même pour les bases d'un espace vectoriel.

Définition 35 *Soit E un espace vectoriel sur un corps K. Une base finie de E est un n-uplet d'éléments de E , $(v_1, \ldots, v_n)$, où n est un entier supérieur ou égal à 1, vérifiant les deux conditions suivantes :*

(1) La partie $\{v_1, v_2, \ldots, v_n\}$ est une partie génératrice de E.

(2) La partie $\{v_1, v_2, \ldots, v_n\}$ est une partie libre de E.

Remarque 43 Il existe une notion de base infinie. Mais cela sort du cadre de ce cours où ne sera traitée que la notion de base finie.

Remarque 44 Il faut observer que la définition donnée introduit un ordre sur les vecteurs d'une base puisque une base est d'abord un n-uplet. Soit $\mathcal{B} = (v_1, \ldots, v_n)$ une base de E. Il est clair que si l'on change l'ordre des vecteurs, c'est-à-dire si l'on considère $(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$ où σ est une bijection de $\{1, \ldots, n\}$ dans $\{1, \ldots, n\}$, les deux conditions (1) et (2) sont évidemment satisfaites puisque

$$\{v_1, \ldots, v_n\} = \{v_{\sigma(1)}, \ldots, v_{\sigma(n)}\}.$$

Alors $(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$ est une base de E mais elle est différente de $\mathcal{B}$ si σ est différente de l'identité, puisque les deux n-uplets $(v_1, \ldots, v_n)$ et $(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$ sont différents. L'importance de l'ordre sera visible lorsque on étudiera la notion de matrice associée à une application linéaire.

Théorème 17 *Soient E un K-espace vectoriel et $v_1, \ldots, v_n$ n vecteurs de E. Les conditions suivantes sont équivalentes :*

(i) Le n-uplet $(v_1, \ldots, v_n)$ est une base de E.

(ii) Tout vecteur de E s'écrit de manière unique comme combinaison linéaire des vecteurs $v_1, \ldots, v_n$; c'est-à-dire que pour tout vecteur v de E, il existe un n-uplet unique $(\alpha_1, \ldots, \alpha_n)$ de K^n tel que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Démonstration. Supposons que $(v_1, \ldots, v_n)$ soit une base de E. Alors, comme $(v_1, \ldots, v_n)$ est une famille génératrice, tout vecteur v de E s'écrit comme combinaison linéaire de $v_1, \ldots, v_n$. Montrons que cette écriture est unique. Supposons qu'il existe deux écritures de v comme combinaison linéaire de $v_1, \ldots, v_n$ à savoir

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \\ v &= \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \end{aligned}$$

et montrons que pour tout i dans $[1, n]$, $\alpha_i = \beta_i$. On a

$$(\alpha_1 - \beta_1) v_1 + (\alpha_2 - \beta_2) v_2 + \cdots + (\alpha_n - \beta_n) v_n = 0.$$

Comme la famille $v_1, \ldots, v_n$ est libre, on en déduit que, pour tout i dans $[1, n]$, $\alpha_i - \beta_i = 0$.

Supposons maintenant que tout vecteur v de E s'écrive de façon unique comme combinaison linéaire de $v_1, \ldots, v_n$. Alors la famille $(v_1, \ldots, v_n)$ est génératrice. Montrons qu'elle est libre. Supposons que l'on ait une relation de dépendance linéaire

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0_E.$$

0_E s'écrit aussi

$$0_E = 0 v_1 + 0 v_2 + \cdots + 0 v_n,$$

Par unicité de l'écriture de 0_E comme combinaison linéaire de $v_1, \ldots, v_n$, on a : pour tout i dans $[1, n]$, $\lambda_i = 0$. Donc la famille $v_1, \ldots, v_n$ est libre. □

Vocabulaire : Si v s'écrit $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$, les scalaires $(\alpha_1, \ldots, \alpha_n)$ s'appellent les coordonnées de v dans la base $(v_1, v_2, \ldots, v_n)$. On utilisera souvent la matrice colonne $[v]_{(v_1, v_2, \ldots, v_n)} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ des coordonnées de v dans la base $(v_1, v_2, \ldots, v_n)$.

Proposition 43 *L'application*

$$\begin{array}{rcl} \chi : K^n & \to & E \\ (\alpha_1, \alpha_2, \ldots, \alpha_n) & \mapsto & \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \end{array}$$

est un isomorphisme.

Démonstration. Si $(\alpha_1, \ldots, \alpha_n)$ et $(\beta_1, \ldots, \beta_n)$ sont deux n-uplets et si λ est un scalaire, on a

$$\begin{array}{l} \chi(\alpha_1, \ldots, \alpha_n) = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \\ \chi(\beta_1, \ldots, \beta_n) = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \end{array}$$

Donc

$$(\lambda\alpha_1 + \beta_1)v_1 + \ldots (\lambda\alpha_n + \beta_n) = \lambda\chi(\alpha_1, \ldots, \alpha_n) + \chi(\beta_1, \ldots, \beta_n).$$

Or le membre de gauche n'est autre que

$$\chi(\lambda\alpha_1 + \beta_1, \ldots, \lambda\alpha_n + \beta_n).$$

χ est donc bien linéaire. De plus χ est bijective car tout élément de E s'écrit de façon unique comme combinaison linéaire des vecteurs $v_1, \ldots, v_n$. □

Donnons maintenant des exemples. La plupart d'entre eux ont été déjà cités dans les paragraphes concernant les notions de partie génératrice finie ou de famille libre.

Exemple 60 Soient les vecteurs $e_1 = (1,0)$ et $e_2 = (0,1)$. Alors (e_1, e_2) est une base de $\mathbb{R}^2$ appelée base canonique de $\mathbb{R}^2$.

Exemple 61 Soient les vecteurs $e_1 = (1,0,0,\ldots,0)$, $e_2 = (0,1,0,\ldots,0), \ldots, e_n = (0,0,\ldots,0,1)$. Alors $(e_1, \ldots, e_n)$ est une base de $\mathbb{R}^n$ appelée la base canonique de $\mathbb{R}^n$. Un élément $x = (x_1, \ldots, x_n)$ de $\mathbb{R}^n$ s'écrit dans la base canonique sous la forme :

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Ce qui signifie que la i-ème composante de x est égale à la i-ème coordonnée de x dans la base canonique. Le n-uplet des coordonnées de x dans la base canonique est égal à x ; cela justifie la dénomination de "base canonique".

Plus généralement, $(e_1, \ldots, e_n)$ est une base du K-espace vectoriel K^n appelée base canonique.

Exemple 62 On considère le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$. Soient les vecteurs $\varepsilon_1 = (1,0)$ et $\varepsilon_2 = (1,1)$. Alors $(\varepsilon_1, \varepsilon_2)$ est une base de $\mathbb{R}^2$. En effet, nous avons vu que tout élément (x,y) de $\mathbb{R}^2$ s'écrivait de façon unique comme combinaison linéaire de ε_1 et de ε_2, à savoir

$$(x,y) = (x-y)\varepsilon_1 + y\varepsilon_2.$$

Les coordonnées de (x,y) dans la base $(\varepsilon_1, \varepsilon_2)$ sont $(x-y, y)$.

Cet exemple prouve qu'il peut y avoir des bases différentes sur un même espace vectoriel. Cependant, une remarque peut être faite : les deux bases ont chacune deux éléments.

Exemple 63 Considérons $P_n(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel réel des fonctions polynômes de degré inférieur ou égal à n. Étant donné un entier naturel compris entre 0 et n, on définit f_k comme suit :

$$\forall x \in \mathbb{R}, f_k(x) = x^k.$$

Alors $(f_0, \ldots, f_k)$ est une base de $P_n(\mathbb{R})$ appelée souvent base canonique de $P_n(\mathbb{R})$. On sait qu'une fonction polynôme sur $\mathbb{R}$ de degré inférieur ou égal à n est une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ telle qu'il $n+1$ éléments $a_0, a_1, \ldots a_n$ tels que

$$\forall x \in \mathbb{R},\ \ f(x) = a_0 + a_1 x + \cdots + a_n x^n.$$

on a alors $f = a_0 f_0 + a_1 f_1 + \cdots + +a_n f_n$, ce qui prouve que la famille $(f_k)_{k\in[1,n]}$ engendre $P_n(\mathbb{R})$. Il est démontré dans les exercices que cette famille est libre.

Soit $f : x \mapsto f(x) = x^2 - x + 2$. Alors f s'écrit $f = 2f_0 - f_1 + f_2$. Les coordonnées de f dans la base (f_0, f_1, f_2) sont $(2, -1, 1)$. Ses coordonnées dans la base (f_2, f_1, f_0) sont $(1, -1, 2)$. Cela illustre bien l'importance de l'ordre introduit sur les vecteurs d'une base.

Exemple 64 $(1, i)$ est une base de $\mathbb{C}$ considéré comme un $\mathbb{R}$-espace vectoriel. En effet tout nombre complexe s'écrit de façon unique sous la forme $a + ib$.

Exemple 65 Considérons K comme un K-espace vectoriel. Tout élément non nul de K est une base de E. En effet, si x est un élément non nul de K, tout élément y de K peut s'écrire $y = \frac{y}{x}x$. Dans cette égalité $\frac{y}{x}$ joue le rôle d'un scalaire. Ceci prouve que $\{x\}$ engendre K considéré comme un K-espace vectoriel. Comme x est un élément non nul de K, $\{x\}$ est une partie libre d'où le résultat.

Si K est un corps infini (par exemple $\mathbb{R}$ ou $\mathbb{C}$), cet exemple prouve qu'il peut y avoir une infinité de bases sur un même espace vectoriel.

Exemple 66 Soit r un entier compris entre 1 et n et s un entier compris entre 1 et p. On désigne par $E_{r,s}$ la matrice à n lignes et p colonnes dont tous les éléments sont nuls sauf celui de la r-ième ligne et de la s-ième colonne qui est égal à 1. Les matrices $E_{r,s}$ forment une base de $M_{n,p}(K)$.

En effet, soit $A = (a_{i,j})$un élément de $M_{n,p}(K)$. On peut écrire

$$A = \sum_{(i,j)\in[1,n]\times[1,p]} a_{i,j} E_{i,j}.$$

Ce qui prouve que les matrices $E_{r,s}$ engendrent $M_{n,p}(K)$. De plus, il est évident que

$$\sum_{(i,j)\in[1,n]\times[1,p]} \lambda_{i,j} E_{i,j} \Longrightarrow \forall (i,j) \in [1,n] \times [1,p],\ \ \lambda_{i,j} = 0.$$

Nous allons maintenant donner deux caractérisations des bases.

Proposition 44 (Première caractérisation d'une base) *Soit E un espace vectoriel. Un n-uplet $(v_1, \ldots v_n)$ est une base de E si et seulement si l'ensemble $\{v_1, \ldots v_n\}$ est une partie libre maximale.*

Que signifie la phrase " est une partie libre maximale" ? Cela signifie que la partie vérifie les deux propriétés suivantes :

(a) La partie $\{v_1, \ldots, v_n\}$ est libre.

(b) Quel que soit le vecteur w de $E - \{v_1, \ldots, v_n\}$, la partie $\{v_1, \ldots v_n, w\}$ n'est pas libre.

C'est une partie maximale au sens de l'inclusion, c'est à dire que toute partie finie contenant strictement $\{v_1, \ldots v_n\}$ n'est pas libre.

Démonstration. Il s'agit en fait de démontrer que si $\{v_1, \ldots v_n\}$ est une partie libre, la propriété (b) ci-dessus et la propriété " engendre E " sont équivalentes.

Soit donc $\{v_1, \ldots v_n\}$ une partie libre de vecteurs de E satisfaisant à la propriété (b) ci-dessus. Alors, pour tout vecteur w de $E - \{v_1, \ldots, v_n\}$, la partie $\{v_1, \ldots v_n, w\}$ n'est pas libre, donc est liée. Or il a été vu dans le paragraphe "Dépendance et Indépendance linéaire", que sous ces hypothèses, le vecteur w est combinaison linéaire des vecteurs $v_1, \ldots v_n$. Ceci prouve que tout élément de E est combinaison linéaire des vecteurs $\{v_1, \ldots v_n\}$ qui forment donc une partie génératrice de E.

Réciproquement, supposons que $v_1, \ldots v_n$ soit une famille libre engendrant E. Alors, tout vecteur w de $E - \{v_1, \ldots, v_n\}$ est une combinaison linéaire des vecteurs $\{v_1, \ldots v_n\}$, ce qui prouve que la partie $\{v_1, \ldots v_n, w\}$ est liée, ce qui équivaut à la propriété (b). □

Proposition 45 (Deuxième caractérisation d'une base) *Soit E un espace vectoriel. Un n-uplet $(v_1, \ldots v_n)$ est une base de E si et seulement si l'ensemble $\{v_1, \ldots v_n\}$ est une partie génératrice minimale de E.*

Que signifie la phrase " est une partie génératrice minimale" ? Cela signifie que la partie vérifie les deux propriétés suivantes :

(a) Les vecteurs $v_1, \ldots, v_n$ engendrent E.

(b) Si on enlève un vecteur de la partie $\{v_1, \ldots, v_n\}$, la partie obtenue n'est plus génératrice.

C'est une partie minimale au sens de l'inclusion, c'est-à-dire que toute partie finie contenue strictement dans $\{v_1, \ldots, v_n\}$ n'est pas une partie génératrice de E.

Démonstration. De même que précédemment, il s'agit de montrer que si $\{v_1, \ldots v_n\}$ est une partie génératrice de E, il y a équivalence entre la propriété (b) ci-dessus et la propriété "les vecteurs sont linéairement indépendants". Cette preuve comporte deux étapes.

Soit donc $\{v_1, \ldots v_n\}$ une partie génératrice de E, satisfaisant à la propriété (b) ci-dessus. Supposons que les vecteurs $v_1, \ldots v_n$ ne soient pas linéairement indépendants. Ils sont donc linéairement dépendants ce qui signifie que l'un des vecteurs est combinaison linéaire des autres. Autrement dit, il existe un entier i compris entre 1 et n tel que $v_i = \sum_{j=1, j \neq i}^{n} \alpha_j v_j$. Compte tenu de ce qui a été vu dans le paragraphe "Ensemble fini de générateurs d'un espace vectoriel", cela implique que $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ est une partie génératrice de E. Or ceci est contraire à l'hypothèse (b), donc les vecteurs sont linéairement indépendants.

Réciproquement, soit $G = \{v_1, \ldots v_n\}$ une partie libre engendrant E. Considérons la partie obtenue en supprimant un vecteur de cette partie, soit v_i. La partie $G' =$

$\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ ne peut engendrer E car le vecteur v_i ne peut être combinaison linéaire des éléments de G' puisque les vecteurs $v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n$ sont linéairement indépendants. La propriété (b) est donc satisfaite. □

Voyons maintenant des théorèmes d' existence d'une base finie.

Théorème 18 (Théorème d'existence de parties libres et génératrices) *Soit E un K-espace vectoriel de type fini, G une partie génératrice finie de E et L une partie libre incluse dans G. Alors, il existe une partie B vérifiant les trois propriétés suivantes :*
- $L \subset B \subset G$
- *B est libre*
- *B engendre E.*

La démonstration que nous donnons de ce théorème est un algorithme.

Démonstration. • ou bien L est une partie génératrice de E et c'est fini puisque c'est une partie génératrice et libre,

• ou bien L n'est pas une partie génératrice et il existe au moins un élément g_1 de G qui n'est pas combinaison linéaire des éléments de L. Alors la partie $L_1 = L \cup \{g_1\}$ vérifie les propriétés suivantes :

$$L_1 \text{ libre}$$
$$L \subsetneq L_1 \subset E$$

On recommence le même raisonnement à partir de L_1.

• ou bien L_1 est une partie génératrice de E et c'est fini (partie génératrice et libre),

• ou bien L_1 n'est pas une partie génératrice de E et il existe au moins un élément g_2 de G qui n'est pas combinaison linéaire des éléments de L_1. Alors la partie $L_2 = L_1 \cup \{g_2\}$ vérifie les propriétés suivantes :

$$L_2 \text{ libre}$$
$$L \subsetneq L_1 \subsetneq L_2 \subset E$$

L'algorithme consiste donc à construire une suite, strictement croissante pour l'inclusion, de parties libres contenues dans G, où, si L_{r-1} n'engendre pas E, L_r est obtenue à partir de L_{r-1} en lui ajoutant un vecteur g_r de G tel que $L_{r-1} \cup \{g_r\}$ soit libre. Comme la partie G est finie, le processus s'arrête et il existe un entier s tel que L_s engendre E. Alors L_s sera une partie finie, libre et génératrice, et sera donc une base de E. □

Corollaire 19 (Théorème d'existence d'une base) *Tout espace vectoriel de type fini (c'est-à-dire admettant une famille finie de générateurs), non réduit à $\{0\}$, admet une base.*

Démonstration. Soit $G = \{g_1, g_2, \ldots, g_n\}$ une partie génératrice non vide de E. Comme $E \neq \{0\}$, il existe i compris entre 1 et n tel que $g_i \neq 0$. On applique le théorème précédent à $L = \{g_i\}$ et G. □

Exemple 67 Soit $P(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions polynômes réelles et E le sous-espace de $P(\mathbb{R})$ engendré par les éléments f_1, f_2, f_3, f_4, f_5 définies par :

$$\forall x \in \mathbb{R}, f_1(x) = 1, f_2(x) = x, f_3(x) = x+1, f_4(x) = 1+x^3, f_5 = x - x^3.$$

Comme f_1 est non nulle, $L = \{f_1\}$ est libre. Considérons f_2. Comme les éléments f_1 et f_2 sont linéairement indépendants, $\{f_1, f_2\}$ est une partie libre. Considérons f_3 : ce

vecteur est combinaison linéaire des vecteurs f_1 et f_2 car $f_3 = f_1 + f_2$ donc $\{f_1, f_2, f_3\}$ est liée. Considérons alors f_4. Un calcul rapide prouve que les vecteurs f_1, f_2 et f_4 sont linéairement indépendants. Alors $\{f_1, f_2, f_4\}$ est une partie libre. Il ne reste que le vecteur f_5 à considérer. Il s'agit, pour pouvoir conclure, d'étudier la linéaire indépendance des vecteurs f_1, f_2, f_4, f_5. Or un calcul rapide montre l'égalité

$$f_1 + f_2 - f_4 - f_5 = 0.$$

ce qui prouve que la famille $\{f_1, f_2, f_4, f_5\}$ est liée. Donc avec les notations de l'algorithme $s = 2$ et $L_2 = \{f_1, f_2, f_4\}$ est une base de E.

Une conséquence extrêmement importante de ce qui précède est le théorème suivant :

Théorème 20 (Théorème de la "base incomplète") *Soit E un K-espace vectoriel de type fini, non réduit à $\{0\}$. Soit G une partie génératrice finie de E et L une partie libre de E. Alors il existe une partie G' de G telle que, en notant $\{v_1, v_2, \ldots, v_n\}$ la partie $L \cup G'$, $(v_1, v_2, \ldots, v_n)$ soit une base de E.*

Démonstration. Pour justifier ce théorème fondamental, il suffit d'utiliser la propriété précédente en partant de la partie $G_1 = L \cup G$ et de la partie libre L incluse dans $L \cup G$. □

L'algorithme du pivot de Gauss fournit une méthode pour extraire une base d'une famille génératrice comme le montre l'exemple suivant.

Exemple 68 Dans le $\mathbb{R}$-espace vectoriel $M_2(\mathbb{R})$ des matrice carrées d'ordre 2 à coefficients dans $\mathbb{R}$, on considère les quatre éléments suivants : $e_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $e_4 = \begin{pmatrix} 3 & 3 \\ 4 & 2 \end{pmatrix}$. Soit $E = \operatorname{vect}(e_1, e_2, e_3, e_4)$ le sous espace vectoriel de $M_2(\mathbb{R})$ engendré par e_1, e_2, e_3, e_4.

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \operatorname{vect}(e_1, e_2, e_3, e_4)$$

$$\iff \exists (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4 \mid \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 + \lambda_4 e_4$$

$$\iff \exists (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4 \mid \begin{cases} a = \lambda_1 + \lambda_2 + \quad + 3\lambda_4 \\ b = \lambda_1 + \quad + \lambda_3 + 4\lambda_4 \\ c = \quad + \lambda_3 + 3\lambda_4 \\ d = \quad + \lambda_2 + \quad + 2\lambda_4 \end{cases}$$

Donc $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ appartient à $\operatorname{vect}(e_1, e_2, e_3, e_4)$ si et seulement si le système ci-dessus a au moins une solution. Résolvons ce système par la méthode du pivot de Gauss :

$$\begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 1 & 0 & 1 & 4 & b \\ 0 & 0 & 1 & 3 & c \\ 0 & 1 & 0 & 2 & d \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 0 & -1 & 1 & 1 & b-a \\ 0 & 0 & 1 & 3 & c \\ 0 & 1 & 0 & 2 & d \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 0 & -1 & 1 & 1 & b-a \\ 0 & 0 & 1 & 3 & c \\ 0 & 0 & 1 & 3 & d-a+b \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 0 & -1 & 1 & 1 & b-a \\ 0 & 0 & 1 & 3 & c \\ 0 & 0 & 0 & 0 & d-a+b-c \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 0 & -1 & 0 & -2 & b-a-c \\ 0 & 0 & 1 & 3 & c \\ 0 & 0 & 0 & & d-a+b-c \end{pmatrix} \sim$$

$$\begin{pmatrix} 1 & 1 & 0 & 3 & a \\ 0 & 1 & 0 & 2 & a+c-b \\ 0 & 0 & 1 & 3 & c \\ 0 & 0 & 0 & 0 & d-a+b-c \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 1 & b-c \\ 0 & 1 & 0 & 2 & a+c-b \\ 0 & 0 & 1 & 3 & c \\ 0 & 0 & 0 & & d-a+b-c \end{pmatrix}.$$

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \operatorname{vect}(e_1, e_2, e_3, e_4) \Longleftrightarrow d-a+b-c=0.$$

Donc

$$\operatorname{vect}(e_1, e_2, e_3, e_4) = \left\{ \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M_2(\mathbb{R}) \mid d-a+b-c=0 \right\}.$$

Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ un élément de $\operatorname{vect}(e_1, e_2, e_3, e_4)$. La résolution du système précédent nous donne toutes les façons d'écrire A comme combinaison linéaire de (e_1, e_2, e_3, e_4). On a

$$A = (b-c-\lambda_4)e_1 + (a-b+c-2\lambda_4)e_2 + (c-3\lambda_4)e_3 + \lambda_4 e_4.$$

Pour chaque valeur de λ_4, on obtient une écriture de A comme combinaison linéaire de (e_1, e_2, e_3, e_4). Il y a donc une infinité de façons d'écrire A comme combinaison linéaire de (e_1, e_2, e_3, e_4).

Si on retire e_4, ce qui correspond à prendre λ_4 égal à 0, on obtient

$$A = (b-c)e_1 + (a-b+c)e_2 + ce_3.$$

Ceci nous donne toutes les façons d'écrire A comme combinaison linéaire de (e_1, e_2, e_3). La matrice A s'écrit donc de façon unique comme combinaison linéaire de (e_1, e_2, e_3). La famille (e_1, e_2, e_3) est donc une base de $\operatorname{vect}(e_1, e_2, e_3, e_4)$.

Théorème 21 (Base d'une somme directe) *Soit E un K-espace vectoriel. Soient F et G deux sous-espaces vectoriels de E tels que $E = F \oplus G$. On suppose que F et G sont de type fini. Soient $B_F = (a_1, a_2, \ldots, a_r)$ une base de F et $(b_1, b_2, \ldots, b_s)$une base de G. Alors E est de type fini et $(a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s)$ est une base de E.*

Démonstration. Comme E est somme directe de F et de G, tout élément x de E s'écrit (de manière unique) comme somme d'un élément y de F et d'un élément z de G, soit $x = y+z$. D'après la définition de la notion de base, il existe des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_r$ tels que :

$$y = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r$$

et des scalaires $\beta_1, \beta_2, \ldots, \beta_s$ tels que :

$$z = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_s b_s.$$

Alors l'égalité $x = y+z$ implique l'égalité :

$$x = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_r a_r + \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_s.$$

Ce qui prouve que $a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s$ est une famille génératrice de E. L'espace vectoriel E est donc de type fini puisque il existe une famille génératrice finie de E.

Il reste à montrer que $a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s$ est une famille libre. Soient donc des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_r, \mu_1, \mu_2, \ldots, \mu_s$ tels que :

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r + \mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_s b_s = 0.$$

Cette égalité peut encore s'écrire :

$$(\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r) + (\mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_s b_s) = 0.$$

Comme $\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r$ est élément de F et $\mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_s b_s$ est élément de G, on a d'après la définition de la somme directe :

$$\begin{array}{l} \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r = 0 \\ \mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_s b_s = 0 \end{array}$$

Or les vecteurs $a_1, a_2, \ldots, a_r$ sont libres, de même que les vecteurs $b_1, b_2, \ldots, b_s$ donc

$$\begin{array}{l} \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r = 0 \Longrightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_r = 0 \\ \mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_r b_r = 0 \Longrightarrow \mu_1 = \mu_2 = \cdots = \mu_r = 0 \end{array}$$

Ce qui achève la démonstration. □

Remarque 45 *Par abus de langage, on pourra dire que* $(a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s)$ *est la "réunion" des bases considérées.*

Cette propriété se généralise au cas où E est somme directe d'un nombre fini p de sous-espaces vectoriels de type fini.

Théorème 22 (Base d'une somme directe) *Soit E un espace vectoriel sur un corps K. Soient $F_1, F_2, \ldots, F_p$ p sous-espaces vectoriels de E tels que $E = F_1 \oplus F_2 \oplus \cdots \oplus F_p$. On suppose que $F_1, F_2, \ldots, F_p$ sont de type fini. Soient $B_{F_i} = (a_1^i, a_2^i, \ldots, a_{r_i}^i)$ une base de F_i. Alors E est de type fini et $(a_1^1, \ldots, a_{r_1}^1, a_1^2, \ldots, a_{r_2}^2, \ldots, a_1^p, \ldots, a_{r_p}^p)$ est une base de E.*

4.1.4 Dimension d'un espace vectoriel de type fini

Tout espace vectoriel de type fini, non réduit à $\{0\}$ possède des bases finies. L'objet de cette section est de prouver que le nombre d'éléments d'une base d'un espace vectoriel est un invariant de cet espace vectoriel, ce qui permet de définir la notion de dimension.

Pour démontrer que toutes les bases d'un espace vectoriel de type fini ont le même nombre d'éléments, il faut tout d'abord comparer le nombre d'éléments d'une partie libre et d'une partie génératrice de cet espace vectoriel.

Lemme 1 *Soit E un espace vectoriel de type fini, non réduit à 0_E, engendré par une partie G de E ayant n éléments $G = \{g_1, g_2, \ldots, g_n\}$. Alors toute partie $F = \{u_1, u_2, \ldots u_n, u_{n+1}\}$ de $n+1$ éléments est liée.*

Une autre façon d'exprimer le résultat de ce lemme est : "toute partie libre de E a un nombre d'éléments inférieur à celui d'une partie génératrice de E".

Démonstration. La preuve de ce lemme est hors programme. Elle peut se faire en raisonnant par récurrence sur l'entier n.

On démontre par récurrence que pour tout la propriété suivante est vraie : "Dans un espace vectoriel engendré par n vecteurs, toute partie ayant $n+1$ éléments est liée".

On vérifie que la propriété est vraie pour $n = 1$. Soit E un espace vectoriel engendré par un vecteur noté g_1, et $\{v_1, v_2\}$ une partie de E ayant deux éléments. Les vecteurs

v_1 et v_2 peuvent s'écrire comme combinaisons linéaires du vecteur g_1, autrement dit, il existe des scalaires α_1, α_2 tels que $v_1 = \alpha_1 g_1$ et $v_2 = \alpha_2 g_1$. Ce qui donne la relation : $\alpha_2 v_1 - \alpha_1 v_2 = 0_E$ On suppose v_2 non nul (sinon il est évident que $\{v_1, v_2\}$ est liée), le scalaire α_2 est donc non nul. On a trouvé une combinaison linéaire nulle des vecteurs v_1, v_2, avec des coefficients non tous nuls. Donc la famille $\{v_1, v_2\}$ est liée.

On démontre maintenant que si la propriété est vraie au rang $n-1$ ($n \geq 1$), alors elle vraie pour l'entier n. Soit E un espace vectoriel engendré par n vecteurs notés $g_1, g_2, \ldots, g_n$, et $\{v_1, v_2, \ldots, v_n, v_{n+1}\}$ une partie de E ayant $n+1$ éléments. Tout vecteur v_j, pour $j = 1, 2, \ldots, n+1$, est combinaison linéaire de $g_1, g_2, \ldots, g_n$, il existe donc des scalaires $\alpha_1^j, \alpha_2^j, \ldots, \alpha_n^j$ tels que :

$$v_j = \alpha_1^j g_1 + \alpha_2^j g_2 + \cdots + \alpha_n^j g_n.$$

Remarque :

On est contraint d'utiliser ici deux indices i, j pour les scalaires (Attention ! j n'est pas un exposant) car deux informations sont nécessaires : l'indice j indique qu'il s'agit de la décomposition du vecteur v_j, et i indique à quel vecteur de la partie génératrice est associé ce coefficient.

En particulier, pour $j = n+1$, le vecteur v_{n+1} s'écrit :

$$v_{n+1} = \alpha_1^{n+1} g_1 + \alpha_2^{n+1} g_2 + \cdots + \alpha_n^{n+1} g_n.$$

Si v_{n+1} est nul, c'est terminé, la partie est liée ; sinon, v_{n+1} est non nul, et au moins un des coefficients α_j^{n+1} est non nul. On suppose, pour alléger l'écriture, que α_n^{n+1} est non nul (sinon il suffit de changer l'ordre des vecteurs). On construit une nouvelle famille de n vecteurs de E de telle sorte que ces vecteurs soient combinaisons linéaires de $g_1, g_2, \ldots, g_{n-1}$, c'est-à-dire appartiennent au sous-espace engendré par $(g_1, g_2, \ldots, g_{n-1})$. Pour $j = 1, 2, \ldots, n$, on définit w_j par :

$$w_j = \alpha_n^{n+1} v_j - \alpha_n^j v_{n+1} = \sum_{k=1}^{n} (\alpha_n^{n+1} \alpha_k^j - \alpha_n^j \alpha_k^{n+1}) g_k.$$

Le coefficient de g_n est nul. Donc w_j combinaison linéaire de $g_1, g_2, \ldots, g_{n-1}$. On a n vecteurs qui appartiennent à un espace vectoriel engendré par $n-1$ vecteurs ; on peut appliquer l'hypothèse de récurrence : la famille $\{w_1, w_2, \ldots, w_n\}$ est liée. Par conséquent il existe des scalaires non tous nuls λ_1, λ_2, $\ldots$, λ_n tels que :

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n = 0$$

En remplaçant les w_j par leur expression en fonction des vecteurs v_i, on obtient :

$$\alpha_n^{n+1} \lambda_1 v_1 + \alpha_n^{n+1} \lambda_2 v_2 + \cdots + \alpha_n^{n+1} \lambda_n v_n - (\lambda_1 \alpha_n^1 + \cdots + \lambda_n \alpha_n^n) v_{n+1} = 0_E$$

Le coefficient α_n^{n+1} a été supposé non nul et au moins un des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_n$ est non nul, on a donc une combinaison linéaire nulle des vecteurs $v_1, v_2, \ldots, v_n, v_{n+1}$ avec des coefficients qui ne sont pas tous nuls ; ceci prouve que ces vecteurs sont linéairement dépendants. La démonstration par récurrence est ainsi achevée. □

Théorème 23 (Définition de la dimension) *Dans un espace vectoriel de type fini E, non réduit à* $\{0_E\}$*, toutes les bases ont le même nombre d'éléments. Ce nombre entier, taille commune de toutes les bases de E, est appelé dimension de E sur K, et noté* $\dim_K E$ *(ou seulement* $\dim E$ *s'il n'y a pas ambiguïté sur le corps K)*

Démonstration. L'espace vectoriel E étant de type fini et non réduit à $\{0\}$, il existe des bases de E ayant un nombre fini d'éléments. Soient $B = \{u_1, u_2, \ldots, u_n\}$ et $B' = \{v_1, v_2, \ldots, v_p\}$ deux bases de E. Si n était distinct de p l'un de ces deux entiers serait strictement supérieur à l'autre, par exemple $n > p$. Alors, d'après le lemme précédent, $\{v_1, v_2, \ldots, v_p\}$ étant génératrice de E, la partie $\{u_1, u_2, \ldots, u_n\}$ serait liée, ce qui contredit l'hypothèse que B est une base de E. □

Méthodologie :

Pour déterminer la dimension d'un espace vectoriel de type fini (différent de $\{0\}$), il suffit de trouver une partie de E à la fois libre et génératrice de E, le cardinal (nombre d'éléments) de cette partie donne la dimension de E.

Convention :

L'espace vectoriel $\{0\}$ ne possède pas de base, la définition de la dimension ne peut donc pas s'appliquer ici ; on convient de lui attribuer pour dimension 0.

Vocabulaire

Par analogie avec la géométrie :

Un espace vectoriel de dimension 1 est appelé droite vectorielle.

Un espace vectoriel de dimension 2 est appelé plan vectoriel.

Dans un espace vectoriel de dimension n, un sous-espace de dimension $n-1$ est appelé hyperplan.

Remarque 46 Dans la littérature mathématique, on rencontre souvent l'expression "espace vectoriel de dimension finie" au lieu de "espace vectoriel de type fini".

Exemple 69 La base canonique de $\mathbb{R}^2$ est $((1,0),(0,1))$. Toutes les bases de $\mathbb{R}^2$ ont donc deux éléments. La dimension de $\mathbb{R}^2$ est donc 2.

Plus généralement, la base canonique de K^n, muni de sa structure d'espace vectoriel sur K, est $(e_1, e_2, \ldots, e_n)$ où, pour $i = 1, 2, \ldots, n$, e_i est le n-uplet dont toutes les composantes sont nulles sauf la ième qui vaut 1. Toutes les bases de K^n ont donc n éléments : la dimension de K^n sur K est donc égale à n. En particulier $\mathbb{R}$ muni de sa structure de $\mathbb{R}$-espace vectoriel, et $\mathbb{C}$ muni de sa structure de $\mathbb{C}$-espace vectoriel ont pour dimension 1.

Exemple 70 La dimension du $\mathbb{R}$-espace vectoriel $P_n(\mathbb{R})$ (espace des fonctions polynômes à coefficients réels de degré inférieur ou égal à n) est égale à $n+1$. La base canonique de $P_n(\mathbb{R})$ est $(f_0, f_1, \ldots, f_n)$ où, pour $i = 0, 1, \ldots, n$, f_i est l'application définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, f_i(x) = x^i.$$

La dimension de $P_n(\mathbb{R})$ est $n+1$.

Exemple 71 La dimension du $\mathbb{R}$-espace vectoriel $\mathbb{C}$ est égale à 2. Si l'ensemble des nombres complexes $\mathbb{C}$ est muni de sa structure d'espace vectoriel sur $\mathbb{R}$ (la loi externe est la multiplication par un scalaire réel), la base canonique de $\mathbb{C}$ est $(1, i)$. La dimension de $\mathbb{C}$ sur $\mathbb{R}$ ($\dim_{\mathbb{R}}\mathbb{C}$) est donc égale à 2.

Exemple 72 Le K-espace vectoriel $M_{n,p}(K)$ est de dimension $n \times p$ puisque la famille $(E_{r,s})_{(r,s)\in[1,n]\times[1,p]}$ en constitue une base.

4.1.5 Propriétés d'un espace vectoriel de dimension n ($n > 0$)

Lorsqu'un espace vectoriel est de type fini, le fait de connaître sa dimension est une information très riche ; les propriétés et théorèmes suivants montrent comment exploiter cette information.

Proposition 46 *Soit E un K-espace vectoriel de dimension n non nulle, alors :*

1. Toute partie libre de E a au plus n éléments.

2. Toute partie génératrice de E a au moins n éléments.

Démonstration. L'espace vectoriel E étant de dimension n (non nulle), il existe une partie de E ayant n éléments qui détermine une base de E, c'est-à-dire à la fois libre et génératrice.

(1) E étant engendré par une partie ayant n éléments, toute partie libre de E a au maximum n éléments, sinon elle est liée (conséquence du lemme).

(2) Il existe une partie libre de E ayant n éléments, par conséquent toute partie génératrice de E a au minimum n éléments (conséquence du lemme).

Théorème 24 *Soient E un K-espace vectoriel de dimension n non nulle, et $u_1, u_2, \ldots, u_n$ n vecteurs de E :*

1. Si $\{u_1, u_2, \ldots, u_n\}$ est une partie libre alors $(u_1, u_2, \ldots, u_n)$ est une base de E.

2. Si $\{u_1, u_2, \ldots, u_n\}$ est une partie génératrice de E alors $(u_1, u_2, \ldots, u_n)$ est une base de E.

Autrement dit, lorsque le nombre de vecteurs considéré est exactement égal à la dimension de l'espace vectoriel, l'une des deux conditions : générateurs, ou linéairement indépendants suffit pour que ces vecteurs déterminent une base de E.

Démonstration. C'est une conséquence immédiate du théorème précédent. Si la dimension de E est égale à n toute partie libre ayant exactement n éléments est une partie libre maximale donc détermine une base de E.

De même, une partie génératrice de E ayant n éléments est une partie génératrice minimale, donc détermine une base de E. □

4.1.6 Sous espaces vectoriels de type fini

Tout sous-espace vectoriel d'un K-espace vectoriel étant lui même un K-espace vectoriel, la question est de savoir s'il est de type fini ou s'il ne l'est pas.

Par exemple l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ contient des sous-espaces vectoriels de type fini comme l'ensemble des fonctions polynômes réelles de degré inférieur ou égal à n (n étant un nombre entier donné), mais il contient aussi des sous-espaces vectoriels qui ne sont pas de type fini comme l'ensemble de toutes les fonctions polynômes réelles. Cependant, la réponse est plus précise lorsque l'espace vectoriel considéré est lui-même de type fini.

Théorème 25 *Soit E un K-espace vectoriel de type fini. Alors tout sous-espace vectoriel F de E est de type fini, et sa dimension est inférieure ou égale à celle de E ; la dimension de F est égale à celle de E si et seulement si le sous-espace F est égal à l'espace E tout entier.*

Démonstration. La démonstration est triviale dans le cas où le sous espace F est réduit à $\{0\}$.

Soit donc F un sous-espace vectoriel de E et soit n la dimension de E ; n est un entier strictement positif puisque E, qui contient F, n'est pas réduit à $\{0\}$.

Soit v un élément non nul de F : $\{v\}$ est une partie libre de F, donc F contient des parties libres. Toute partie libre d'éléments de F étant une partie libre d'éléments de E (voir la définition des parties libres), comme E est de dimension n, toutes les parties libres de F ont au plus n éléments.

On considère l'ensemble A des entiers k tels qu'il existe une partie libre de F ayant k éléments :

$$A = \{k \in \mathbb{N}, \exists \{f_1, f_2, \ldots, f_k\} \subset F \text{ et} \{f_1, f_2, \ldots, f_k\} \text{ partie libre de F}\}$$

Cet ensemble A, non vide ($1 \in A$), est un sous-ensemble borné de $\mathbb{N}$ (puisque tout élément de A est compris entre 1 et n) donc il admet un maximum. Soit p ce maximum et soit $\{v_1, v_2, \ldots, v_p\}$ une partie libre de F ayant p éléments ; cette partie libre est donc une partie libre maximale de F.

D'après la propriété des parties libres maximales d'un espace vectoriel, la partie $\{v_1, v_2, \ldots, v_p\}$ est une partie génératrice de F et donc détermine une base de F.

On a ainsi démontré simultanément que

• F est de type fini (puisque $\{v_1, v_2, \ldots, v_p\}$ est une partie génératrice de F),

• dimF = p, donc dimF $\leq$ dimE (puisque toute partie libre de F a au plus n éléments).

De plus, lorsque $p = n$, le p-uplet $(v_1, v_2, \ldots, v_p)$, qui est une base de F, est aussi une base de E (car $\{v_1, v_2, \ldots, v_p\}$ est alors une partie libre de E ayant exactement n éléments), c'est donc une base de E. Tout élément de E s'écrit comme une combinaison linéaire de $v_1, v_2, \ldots, v_p$, donc appartient à F, d'où $E = F$.

Exemple 73 Si E est un K-espace vectoriel de dimension 2, les sous-espaces vectoriels de E sont

• soit de dimension 0, c'est alors le sous-espace $\{0\}$,

• soit de dimension 1, ce sont tous les sous-espaces engendrés par les vecteurs non nuls u de E,

• soit de dimension 2, c'est alors l'espace E tout entier.

Vocabulaire Dans un K-espace vectoriel E de dimension n ($n \geq 2$), tout sous-espace vectoriel de E de dimension 1 est appelé droite vectorielle de E, tout sous-espace vectoriel de E de dimension 2 est appelé plan vectoriel de E, et tout sous-espace vectoriel de E de dimension $n-1$ est appelé hyperplan de E : il y a identité entre les notions d'hyperplan et de plan vectoriel lorsque $n = 3$, et entre les notions d'hyperplan et de droite vectorielle lorsque $n = 2$.

Dans le cas plus général où E est un K-espace vectoriel quelconque et F et G deux sous-espaces vectoriels de type fini de E, la comparaison des dimensions de F et de G ne donne pas d'information sur F et G, les dimensions peuvent être égales sans que ces sous-espaces soient égaux, mais si F est contenu dans G, il peut être considéré comme un sous-espace vectoriel de G et le théorème précédent permet de déduire le corollaire suivant :

Corollaire 26 *Soient F et G deux sous espaces vectoriels de type fini de E, tels que F soit contenu dans G. Alors la dimension de F est inférieure ou égale à la dimension de G et les dimensions de F et de G sont égales si et seulement F et G sont égaux.*

Exemple 74 Deux droites F et G d'un K-espace vectoriel E sont soit égales, soit d'intersection réduite au vecteur nul.

Exemple 75 Soient F et G les sous-espaces vectoriels de $\mathbb{R}^3$ suivants :

$$F := \{(x,y,z) \in \mathbb{R}^3 \mid 2x-3y+z=0\}$$
$$G := \text{vect}(u,v) \text{ où } u=(1,1,1) \text{ et } v=(2,1,-1)$$

On veut montrer que $F = G$. On remarque que les vecteurs u et v ne sont pas colinéaires, donc G est de dimension 2, et de plus ils appartiennent à F, donc G est contenu dans F. Pour trouver la dimension de F, on pourrait déterminer une base de F, on montrerait alors que la dimension de F est 2. Mais il est plus judicieux ici de remarquer que F est contenu strictement dans $\mathbb{R}^3$ (par exemple : le vecteur $(1,0,0)$ de $\mathbb{R}^3$ n'est pas dans F), donc la dimension de F est strictement inférieure à 3 ; mais puisque F contient G, la dimension de F est supérieure ou égale à 2 , donc la dimension de F ne peut être que 2.

On a donc démontré que G est contenu dans F et que F et G ont la même dimension ; ceci entraine que F est égal à G.

Étudions maintenant la somme de deux sous espaces de type fini.

Proposition 47 *Soient F et G deux sous-espaces vectoriels de type fini d'un K-espace vectoriel E.*

1- Si $f_1, f_2, \ldots, f_p$ est une famille génératrice de F, et $g_1, g_2, \ldots, g_q$ une famille génératrice de G, alors la famille $f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q$ est une famille génératrice de $F+G$.

2- Si, de plus, $(f_1, f_2, \ldots, f_p)$ est une base de F, et $(g_1, g_2, \ldots, g_q)$ une base de G, alors le $p+q$-uplet $(f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q)$ est une base de $F+G$ si et seulement si la somme de F et de G est directe.

Démonstration. 1. Tout élément u de $F+G$ est la somme d'un élément v de F et d'un élément w de G :

$$u = v + w.$$

Or v est une combinaison linéaire de $f_1, f_2, \ldots, f_p$ et w est une combinaison linéaire de $g_1, g_2, \ldots, g_q$ donc u est une combinaison linéaire de $f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q$ donc $\{f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q\}$ est bien une partie génératrice de $F+G$. On en déduit que $F+G$ est de type fini et on a

$$\dim(\mathrm{F}+\mathrm{G}) \leq \mathrm{p}+\mathrm{q} = \dim\mathrm{F} + \dim\mathrm{G}.$$

2. Compte tenu du 1, il suffit de montrer que, lorsque $\{f_1, f_2, \ldots, f_p\}$ est une partie libre de F, et $\{g_1, g_2, \ldots, g_q\}$ une partie libre de G, alors $\{f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q\}$ est une partie libre de $F+G$ si et seulement si $F+G = F \oplus G$.

Supposons que $\{f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q\}$ soit une partie libre de $F+G$ et montrons que $F+G = F \oplus G$. Écrivons 0_E comme la somme d'un élément de F et d'un élément de G : $0_E = u+v$ avec $u \in F$ et $v \in G$ et montrons que $u = v = 0_E$. Comme u est dans F, il s'écrit comme combinaison linéaire de $\{f_1, f_2, \ldots, f_p\}$:

$$\exists(\lambda_1, \lambda_2, \ldots, \lambda_p) \in K^p, \;\; u = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_p f_p.$$

Comme v est dans G, il s'écrit comme combinaison linéaire de $\{g_1, g_2, \ldots, g_q\}$:

$$\exists(\mu_1, \mu_2, \ldots, \mu_p) \in K^q, \;\; v = \mu_1 g_1 + \mu_2 g_2 + \cdots + \mu_q g_q.$$

Comme $u+v=0_E$, on a

$$\lambda_1 f_1+\lambda_2 f_2+\cdots+\lambda_p f_p+\mu_1 g_1+\mu_2 g_2+\cdots+\mu_p g_p=0_E.$$

Comme $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ est une partie libre de $F+G$, on a $\lambda_1=\lambda_2=\cdots=\lambda_p=\mu_1=\mu_2=\cdots=\mu_q=0$. Donc $u=v=0_E$.

Réciproquement, on suppose que $F+G=F\oplus G$ et on veut montrer que $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ est une partie libre de $F+G$. On suppose

$$\lambda_1 f_1+\lambda_2 f_2+\cdots+\lambda_p f_p+\mu_1 g_1+\mu_2 g_2+\cdots+\mu_p g_p=0_E.$$

Or $\lambda_1 f_1+\lambda_2 f_2+\cdots+\lambda_p f_p$ est un élément de F et $\mu_1 g_1+\mu_2 g_2+\cdots+\mu_p g_p$ est un élément de G. Comme la somme de F et de G est directe, on en déduit

$$\begin{cases} \lambda_1 f_1+\lambda_2 f_2+\cdots+\lambda_p f_p=0_E \\ \mu_1 g_1+\mu_2 g_2+\cdots+\mu_p g_p=0_E \end{cases}$$

Comme $\{f_1,f_2,\ldots,f_p\}$ est une partie libre de F, tous les λ_i sont nuls. De même $\{g_1,g_2,\ldots,g_q\}$ étant une partie libre de G donc tous les μ_i sont nuls. On a montré ainsi que $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ est une partie libre de $F+G$. □

De la proposition précédente résulte le théorème suivant :

Théorème 27 *Soient F et G deux sous-espaces vectoriels de type fini d'un K-espace vectoriel E. Alors la somme $F+G$ est un sous-espace vectoriel de type fini de E et sa dimension est inférieure ou égale à la somme des dimensions de F et de G.*

La dimension de $F+G$ est égale à la somme des dimensions de F et de G si et seulement si la somme est directe.

Démonstration. Soient $\{f_1,f_2,\ldots,f_p\}$ et $\{g_1,g_2,\ldots,g_q\}$ des bases de F et de G. D'après la proposition précédente $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ est une partie génératrice de $F+G$, donc $F+G$ est de type fini et sa dimension est inférieure ou égale au nombre d'éléments de la famille $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ donc inférieure ou égale à $p+q=\mathrm{dimF}+\mathrm{dimG}$.

De plus,

•Si la somme $F+G$ est directe, alors $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ est une base de $F+G$. Donc la dimension de $F+G$ est égale à $p+q=\mathrm{dimF}+\mathrm{dimG}$.

• Réciproquement, si la dimension de $F+G$ est égale à $p+q=\mathrm{dimF}+\mathrm{dimG}$ Alors la famille génératrice $\{f_1,f_2,\ldots,f_p,g_1,g_2,\ldots,g_q\}$ qui a $p+q$ éléments est une famille génératrice de $F+G$ ayant $\dim(\mathrm{F}+\mathrm{G})$ éléments, elle détermine donc une base de $F+G$. Ce qui entraine que la somme $F+G$ est directe. □

Remarque 47 Les assertions du théorème reste vraies pour la somme d'un nombre quelconque de sous espaces vectoriels d'un espace vectoriel de type fini.

De la proposition précédente résulte aussi le théorème fondamental suivant :

Théorème 28 *Tout sous-espace vectoriel d'un K-espace vectoriel de type fini admet un supplémentaire.*

Démonstration. Soient E un K-espace vectoriel de type fini, $B = (e_1, \ldots, e_n)$ une base de E, F un sous-espace vectoriel de E, $B_F = (f_1, f_2, \ldots, f_p)$ une base de F. La famille $e_1, \ldots, e_n$ est donc une famille génératrice de E et la partie $\{f_1, f_2, \ldots, f_p\}$ une partie libre de E. On peut donc appliquer le théorème de la base incomplète. D'après ce théorème, il existe une partie C de $\{e_1, \ldots, e_n\}$ telle que $\{f_1, f_2, \ldots, f_p\} \cup C$ détermine une base B' de E. Soit G le sous-espace vectoriel de E engendré par C, alors B' est à la fois une base de E et une base de $F + G$. Ceci prouve bien que la somme $F + G$ est directe et que $E = F + G = F \oplus G$. □

Exemple 76 Soit $F = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - 3y + z = 0\}$. On a vu dans l'exemple précédent que $F = \text{vect}(u, v)$ où $u = (1, 1, 1)$ et $v = (2, 1, -1)$. Le sous-espace F est de dimension 2. Pour construire un supplémentaire de F, donc trouver G tel que $F \oplus G = \mathbb{R}^3$, il suffit de remarquer que ce supplémentaire est forcément de dimension 1 (voir le théorème sur la dimension d'une somme directe), donc qu'une base de G n'a qu'un élément qu'on note w, et que (u, v, w) doit être une base de $\mathbb{R}^3$. N'importe quel élément de $\mathbb{R}^3$, n'appartenant pas à F (donc forcément non nul) convient : En effet soit w n'appartenant pas à F, alors $F \cap \mathbb{R}w = \{0_E\}$. donc la somme $F + \mathbb{R}w$ est directe, sa dimension est alors égale à 3, donc $F \oplus \mathbb{R}w = E$. Le sous-espace $\mathbb{R}w$ est bien un supplémentaire de F.

On peut choisir pour w n'importe quel triplet ne vérifiant pas l'égalité $2x - 3y + z = 0$. Il existe une infinité de tels triplets non colinéaires, ce qui prouve l'existence d'une infinité de supplémentaires de F.

4.1.7 Rang d'une famille finie de vecteurs

Soit E un K-espace vectoriel et $(v_1, \ldots, v_p)$ une famille finie de vecteurs de E. Le sous-espace vectoriel engendré par $(v_1, \ldots, v_p)$ est de type fini (puisqu'il admet trivialement une famille finie de générateurs). On peut donc donner la définition suivante :

Définition 36 (Définition du rang d'une famille finie de vecteurs) *Soit E un K-espace vectoriel et une famille finie de vecteurs de E, $(v_1, \ldots, v_p)$. Le rang de la famille $(v_1, \ldots, v_p)$ (on dit aussi rang des vecteurs $v_1, \ldots, v_p$) est la dimension du sous-espace vectoriel de E engendré par les vecteurs $v_1, \ldots, v_p$.*

Notation Le rang de la famille $(v_1, \ldots, v_p)$ est noté $\text{rg}(v_1, \ldots, v_p)$.

Proposition 48 *Soit E un K-espace vectoriel et $(v_1, \ldots, v_p)$ une famille de p vecteurs non tous nuls de E. Alors :*

1) Les inégalités suivantes sont satisfaites :

$$0 < rg(v_1, \ldots, v_p) \leq p.$$

2) Le rang de $(v_1, \ldots, v_p)$ est le nombre maximum d'éléments d'une famille libre extraite de $(v_1, \ldots, v_p)$. Donc $\text{rg}(v_1, \ldots, v_p) = r$ si et seulement il existe une famille libre de r vecteurs extraite de $(v_1, \ldots, v_p)$ et si toute famille de q vecteurs, avec $q > r$, extraite de $(v_1, \ldots, v_p)$, est liée. En particulier $\text{rg}(v_1, \ldots, v_p) = p$ si et seulement les vecteurs $(v_1, \ldots, v_p)$ sont linéairement indépendants.

3) Si $\text{rg}(v_1, \ldots, v_p) = r$, toute partie libre de r éléments extraite de $(v_1, \ldots, v_p)$ détermine une base du sous-espace vectoriel engendré par les vecteurs $v_1, \ldots, v_p$

4) Soient F un K-espace vectoriel et $f : E \to F$ une application linéaire. On a l'inégalité

$$\mathrm{rg}(f(e_1), f(e_2), \ldots, f(e_p)) \leq \mathrm{rg}(e_1, e_2, \ldots, e_p).$$

Si, de plus, f est injective, le rang des vecteurs $v_1, \ldots, v_p$ est égal au rang des vecteurs $f(v_1), \ldots, f(v_p)$.

Démonstration. les points 1,2 et 3 sont immédiats. Le quatrième point découle de la proposition 42.

Remarque 48 Le cas où tous les vecteurs sont nuls est immédiat : il est clair que l'on a l'équivalence suivante :

$$\mathrm{rg}(v_1, \ldots, v_p) = 0 \Longleftrightarrow v_1 = v_2 = \cdots = v_p = 0.$$

Remarque 49 Si E est un espace vectoriel de type fini, il est évident d'après les propriétés des sous-espaces vectoriels de type fini, que le rang d'une famille finie de vecteurs de E est inférieur ou égal à la dimension de E. Il est égal à la dimension de E si et seulement si la famille engendre E.

Proposition 49 *L'espace vectoriel engendré par une famille de vecteurs n'est pas modifié par les trois opérations élémentaires suivantes sur les vecteurs :*

- *On échange deux vecteurs.*
- *On multiplie un vecteur de la famille par un scalaire non nul.*
- *On rajoute à l'un des vecteurs une combinaison linéaire des autres vecteurs (substitution).*

En particulier le rang d'une famille de vecteurs n'est pas modifié par les trois opérations élémentaires précédentes sur les vecteurs.

Remarque 50 Cet énoncé suppose évidemment que l'on considère une famille d'au moins deux vecteurs. Cela ne pose aucun problème dans la mesure où la détermination du rang d'une famille de vecteurs ne comportant qu'un vecteur est immédiate : ou bien ce vecteur est nul et le rang est égal à 0 ou bien ce vecteur est non nul et le rang est égal à 1.

Démonstration. Le premier point de la proposition est évident. Soit F le sous-espace vectoriel de E engendré par les vecteurs $v_1, \ldots, v_p$. Démontrons le deuxième point. Pour simplifier l'exposition, on suppose que c'est le premier vecteur que l'on multiplie par un scalaire non nul λ. Soit $(\alpha_1, \ldots, \alpha_p)$ p scalaires. L'égalité

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \frac{\alpha_1}{\lambda}\lambda v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

montre qu'une combinaison linéaire des vecteurs $(v_1, \ldots, v_p)$ est une combinaison linéaire des vecteurs $(\lambda v_1, \ldots, v_p)$ et vice versa. Le deuxième point en découle. Montrons maintenant le troisième point. Comme précédemment on suppose que c'est au vecteur v_1 que l'on rajoute une combinaison linéaire des autres. Il est clair que cela ne nuit pas à la généralité de la démonstration. On pose $v_1' = v_1 + \sum_{i=2}^{p} \alpha_i v_i$. La définition de v_1' implique que v_1' appartient au sous-espace vectoriel engendré par les vecteurs $(v_1, \ldots, v_p)$. On a donc l'inclusion

$$\mathrm{vect}(v_1', \ldots, v_p) \subset \mathrm{vect}(v_1, \ldots, v_p).$$

La définition de v'_1 implique aussi l'égalité

$$v_1 = v'_1 - \sum_{i=2}^{p} \alpha_i v_i$$

et donc, comme précédemment, l'inclusion de $\mathrm{vect}(v_1, \ldots, v_p)$ dans $\mathrm{vect}(v'_1, \ldots, v_p)$ D'où l' égalité :

$$\mathrm{vect}(v_1, \ldots, v_p) = \mathrm{vect}(v'_1, \ldots, v_p).$$

□

Dans la plupart des exemples, la recherche du rang concernera des vecteurs appartenant à un espace vectoriel E de type fini, vecteurs que l'on exprimera donc dans une base de E. On se place donc désormais dans cette situation. On s'appuiera alors sur la propriété suivante :

Proposition 50 *Soit E un K-espace vectoriel de type fini et n sa dimension. Soient $(e_1, \ldots, e_n)$ une base de E et $w_1, \ldots, w_s$ des vecteurs de E, dont les coordonnées dans la base $(e_1, \ldots, e_n)$ se présentent de la manière suivante :*

$$\begin{array}{lclclclclcl}
w_1 & = & a_{i_1,1}e_{i_1} & +\cdots+ & a_{i_2,1}e_{i_2} & +\cdots+ & a_{i_s,1}e_{i_s} & +\cdots+ & a_{n,1}e_n \\
w_2 & = & & & a_{i_2,2}e_{i_2} & +\cdots+ & a_{i_s,2}e_{i_s} & +\cdots+ & a_{n,2}e_n \\
\vdots & = & & & & \ddots & & & \\
w_s & = & & & & & a_{i_s,s}e_{i_s} & +\cdots+ & a_{n,s}e_n
\end{array}$$

avec, $i_1 < i_2 < \cdots < i_s$ et, pour tous les entiers j compris entre 1 et s, $a_{i_j,j}$ non nul. Alors les vecteurs $(w_1, \ldots, w_s)$ sont linéairement indépendants.

Démonstration. On démontre par récurrence sur s que s vecteurs $w_1, \ldots, w_s$ satisfaisant les hypothèses de la proposition sont linéairement indépendants.

Si $s = 1$, on a vecteur le w_1 non nul. Il constitue donc une famille libre.

On suppose $s \geq 2$ et l'assertion vraie pour $s-1$. Soit une combinaison linéaire nulle des vecteurs $w_1, \ldots, w_s$:

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_s w_s = 0.$$

La coordonnée de $\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_s w_s = 0$ sur le vecteur e_{i_1} est $\lambda_1 a_{i_1,1}$. Comme $a_{i_1,1}$ est non nul, $\lambda_1 = 0$. On est donc ramené à la combinaison linéaire $\lambda_2 w_2 + \cdots + \lambda_s w_s = 0$ où les $s-1$ vecteurs $w_2, \ldots, w_s$ vérifient les hypothèses de la proposition. □

Une conséquence immédiate de cette proposition est :

Corollaire 29 *Si des vecteurs $w_1, \ldots, w_s$ vérifient les hypothèses de la proposition précédente alors leur rang est exactement égal à s.*

Proposition 51 *Soit E un K-espace vectoriel de dimension finie et soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Soient $v_1, \ldots, v_p$ p vecteurs. Notons $[v_i]_{\mathcal{B}}$ la matrice colonne des coordonnées de v_i dans la base $\mathcal{B}$. Le rang des vecteurs $v_1, \ldots, v_p$ dans E est le rang des vecteurs $[v_1]_{\mathcal{B}}, \ldots, [v_p]_{\mathcal{B}}$ dans $M_{n,1}(K)$.*

Démonstration. Cela découle du fait que l'application

$$\begin{array}{rcl} E & \to & M_{n,1}(K) \\ v & \mapsto & [v]_{\mathcal{B}} \end{array}$$

est un isomorphisme.

Définition 37 *On définit le rang d'une matrice comme étant le rang de ses vecteurs colonnes.*

Remarque 51 Notons M la matrice dont la j-ième colonne est $[v_j]_{\mathcal{B}}$. Le rang des vecteurs $v_1, \ldots, v_p$ est égal au rang de M.

D'après la proposition 49, on ne change pas le rang d'une matrice par des opérations élémentaires sur les colonnes.

Définition 38 *On dit qu'une matrice est échelonnée par rapport aux colonnes si sa transposée est échelonnée (par rapport aux lignes).*

Soit A une matrice. L'algorithme de Gauss nous dit que, par des opérations élémentaires sur les lignes, on peut transformer A^T en une matrice échelonnée (par rapport aux lignes) R. Par transposition, on en déduit que, par des opérations élémentaires sur les colonnes, on peut transformer A en la matrices R^T qui est échelonée par rapport aux colonnes. Le rang d'une matrice échelonnée par rapport aux colonnes est facile à calculer (voir proposition 50).

Exemple 77 Calculons le rang de la famille des 5 vecteurs suivants de $\mathbb{R}^4$.

$$\begin{aligned} v_1 &= (1,1,1,1) \\ v_2 &= (-1,2,0,1) \\ v_3 &= (3,2,-1,-3) \\ v_4 &= (3,5,0,-1) \\ v_5 &= (3,8,1,1) \end{aligned}$$

On est ramené à calculer le rang de la matrice

$$\begin{pmatrix} 1 & -1 & 3 & 3 & 3 \\ 1 & 2 & 2 & 5 & 8 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & 1 & -3 & -1 & 1 \end{pmatrix}$$

Pour les opérations élémentaires sur les colonnes, on utilise les mêmes notations que pour les opérations sur les lignes.

En faisant les opérations $C_2 \longleftarrow C_2 + C_1, C_3 \longleftarrow C_3 - 3C_1$, $C_4 \longleftarrow C_4 - 3C_1, C_5 \longleftarrow C_5 - 3C_1$, on obtient

$$\begin{pmatrix} 1 & -1 & 3 & 3 & 3 \\ 1 & 2 & 2 & 5 & 8 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & 1 & -3 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & -1 & 2 & 5 \\ 1 & 1 & -4 & -3 & -2 \\ 1 & 2 & -6 & -4 & -2 \end{pmatrix}$$

On échange C_2 et C_3 pour éviter d'introduire des fractions. On obtient :

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

En faisant les opérations $C_3 \longleftarrow C_3 + 3C_2$, $C_4 \longleftarrow C_4 + 2C_2$ et $C_5 \longleftarrow C_5 + 5C_2$, on obtient

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & 2 & 5 \\ 1 & -4 & 1 & -3 & -2 \\ 1 & -6 & 2 & -4 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & -11 & -22 \\ 1 & -6 & -16 & -16 & -32 \end{pmatrix}.$$

Enfin, en faisant les opérations $C_4 \longleftarrow C_4 - C_3$ et $C_5 \longleftarrow C_5 - 2C_3$, on obtient

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & -11 & -22 \\ 1 & -6 & -16 & -16 & -32 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -4 & -11 & 0 & 0 \\ 1 & -6 & -16 & 0 & 0 \end{pmatrix}.$$

On en déduit que le rang des vecteurs v_1, v_2, v_3, v_4, v_5 est 3.

Remarque 52 *En fait, nous avons même démontré que*

$$\text{vect}(v_1, v_2, v_3, v_4, v_5) = \text{vect}\left(v_1, (0,-1,-4,-6), (0,0,-11,-16)\right).$$

Exemple 78 *Considérons les trois vecteurs suivants dans* $\mathbb{R}^5$ *:* $\nu_1 = (1,2,1,2,0)$, $\nu_2 = (1,0,1,4,4)$ *et* $\nu_3 = (1,1,1,0,0)$. *Montrons que la famille* (ν_1, ν_2, ν_3) *est libre dans* $\mathbb{R}^5$. *Pour cela calculons le rang de cette famille de vecteurs ou, ce qui revient au même, celui de la matrice suivante*

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

Par des opérations élémentaires sur les colonnes, on obtient :

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

La famille (ν_1, ν_2, ν_3) *est de rang 3, elle est donc libre dans* $\mathbb{R}^5$.

Exemple 79 *Considérons les quatre vecteurs suivants dans* $\mathbb{R}^3$ *:* $\nu_1 = (1,2,3)$, $\nu_2 = (2,0,6)$, $\nu_3 = (3,2,1)$ *et* $\nu_4 = (-1,2,2)$. *Montrons que la famille* $(\nu_1, \nu_2, \nu_3, \nu_4)$ *engendre* $\mathbb{R}^3$. *Pour cela calculons le rang de cette famille de vecteurs ou, ce qui revient au même, celui de la matrice suivante*

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

Par des opérations élémentaires sur les colonnes, on obtient :

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 2 & 2 \\ 3 & 6 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & -4 & 4 \\ 3 & 0 & -8 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & 0 & -8 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & 0 & -8 & 0 \end{pmatrix}.$$

La famille $(\nu_1, \nu_2, \nu_3, \nu_4)$ *étant de rang 3, ce qui veut exactement dire que Vect* $(\nu_1, \nu_2, \nu_3, \nu_4)$ *est un sous espace vectoriel de dimension 3 de* $\mathbb{R}^3$. *On a donc Vect* $(\nu_1, \nu_2, \nu_3, \nu_4) = \mathbb{R}^3$. *Autrement dit la famille* $(\nu_1, \nu_2, \nu_3, \nu_4)$ *engendre* $\mathbb{R}^3$.

4.2 Applications linéaires en dimension finie

L'étude des propriétés des applications linéaires définies sur un espace vectoriel de type fini conduit à des résultats très riches et très utilisés.

4.2.1 Construction et caractérisation

Théorème 30 (Construction d'une application linéaire) *Soient E et F deux espaces vectoriels sur un même corps K. On suppose que l'espace vectoriel E est de type fini. Soit n ($n \geq 1$) sa dimension. Alors, si $(e_1, \ldots, e_n)$ est une base de E, pour tout n-uplet $(a_1, \ldots, a_n)$ d'éléments de F, il existe une et une seule application linéaire f de E dans F telle que :*

$$\forall i \in [1,n], \;\; f(e_i) = a_i.$$

Remarque 53 Le théorème ne fait aucune hypothèse sur la dimension de l'espace vectoriel F, espace vectoriel d'arrivée de f.

Une application linéaire d'un espace vectoriel de type fini dans un espace vectoriel quelconque est entièrement déterminée par les images des vecteurs d'une base de l'espace vectoriel de départ.

Méthodologie de la preuve du théorème :

La conclusion du théorème comporte deux points : l'existence et l'unicité d'une application linéaire satisfaisant à certaines propriétés. La démonstration va donc comporter deux parties :

• une première qui consistera à prouver que si une telle application existe, elle est unique ;

• la deuxième qui consistera à montrer l'existence d'une telle application linéaire par sa construction explicite.

Il peut paraître curieux de commencer par l'unicité, mais la plupart du temps, dans une situation de ce type, c'est ce qui est fait. Cela permet en effet de déterminer, si elle existe, la seule application qui peut convenir. Nous avons déjà rencontré cette méthode dans l'exemple 38

Démonstration.

Commençons par démontrer l'unicité. Supposons qu'il existe une application linéaire $f : E \to F$ telle que

$$\forall i \in [1,n], \;\; f(e_i) = a_i.$$

Soit x un élément de E. Il existe des scalaires $x_1, x_2, \ldots, x_n$ uniques tels que $x = \sum_{i=1}^{n} x_i e_i$. Comme f est linéaire, on a

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

Donc, si elle existe, f est unique

Démontrons maintenant l'existence de f. Nous avons montré que la seule solution possible au problème posé est l'application

$$\begin{array}{rcl} L : E & \to & F \\ x & \mapsto & \sum_{i=1}^{n} x_i a_i \;\; \text{si } x = \sum_{i=1}^{n} x_i e_i \end{array}$$

Pour achever la démonstration il s'agit de vérifier que cette application est linéaire et qu'elle vérifie la condition imposée, à savoir : pour tout i dans $[1,n]$, $L(e_i) = a_i$.

Soient $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$ deux éléments de E, et α un scalaire. Alors, tous calculs faits, cela donne

$$\alpha x + y = \sum_{i=1}^{n} (\alpha x_i + y_i) e_i$$

et les scalaires $\alpha x_i + y_i$ sont les coordonnées de $\alpha x + y$ dans la base $(e_1, \ldots, e_n)$. Donc d'après la définition de L, on a

$$L(\alpha x + y) = \sum_{i=1}^{n} (\alpha x_i + y_i) a_i = \alpha \left(\sum_{i=1}^{n} x_i a_i \right) + \left(\sum_{i=1}^{n} y_i a_i \right) = \alpha L(x) + \beta L(y).$$

L'application L est donc linéaire.

Pour justifier qu'elle vérifie la condition imposée, il suffit de remarquer que la i-ème composante de e_i sur la base $(e_1, \ldots, e_n)$ est égale à 1 et que toutes les autres sont nulles. Alors

$$L(e_i) = 1a_i + \sum_{j=1, j \neq i}^{n} 0 a_j = a_i.$$

Ce qui termine la preuve du théorème. □

4.2.2 Rang d'une application linéaire

Proposition 52 *Soient E et F deux espaces vectoriels sur un même corps K et f une application linéaire de E dans F. On suppose l'espace vectoriel E de type fini. Alors, l'image de f est un espace vectoriel de type fini. Plus précisément, si n est la dimension de E et $(e_1, \ldots, e_n)$ une base de E, alors $(f(e_1), \ldots, f(e_n))$ est une famille génératrice de* $\operatorname{Im} f$.

Démonstration. Il suffit de démontrer que tout élément de $\operatorname{Im} f$ est combinaison linéaire des vecteurs $f(e_1), \ldots, f(e_n)$.

Soit y un élément quelconque de $\operatorname{Im} f$. Il existe donc un élément x de E tel que $y = f(x)$. Comme $(e_1, \ldots, e_n)$ est une base de E, il existe des scalaires $(x_1, \ldots, x_n)$ tels que $x = \sum_{i=1}^{n} x_i e_i$. En utilisant la linéarité de f, on en déduit $f(x) = \sum_{i=1}^{n} x_i f(e_i)$, ce qui achève la démonstration. □

Définition 39 (Définition du rang d'une application linéaire) *Soient E et F deux espaces vectoriels sur un même corps K et f une application linéaire de E dans F. On suppose l'espace vectoriel E de type fini. La dimension de l'espace vectoriel* $\operatorname{Im} f$ *est appelée rang de f et notée* $\operatorname{rg}(f)$.

Remarque 54 D'après la deuxième partie de la proposition précédente, la dimension de $\operatorname{Im}(f)$ est le rang du système de vecteurs $f(e_1), \ldots, f(e_n)$, ce qui explique à postériori la dénomination rang de l'application linéaire f. Il en résulte que le rang d'une application linéaire est inférieur ou égal à la dimension de l'espace vectoriel de départ. Pour déterminer une base et la dimension de l'image d'une application linéaire dont l'espace de départ est de type fini, on détermine les vecteurs $f(e_1), \ldots, f(e_n)$ et on peut utiliser les techniques de détermination du rang d'une famille finie de vecteurs.

Proposition 53 *Soient E et F deux espaces vectoriels et f un isomorphisme de E dans F. Si E (respectivement F) est de type fini, alors F (respectivement E) est de type fini et on a* dimE = dimF

Démonstration. Si E est de dimension finie, alors comme $F = \mathrm{Im}(f)$, F est engendré par l'image d'une base de E, on a donc dimF $\leq$ dimE. De même $f^{-1} : F \to E$ est un isomorphisme, donc $f^{-1}(F) = E$ et, d'après le résultat précédent, on a dimE $\leq$ dimF.

Si c'est F qui est de dimension finie, on fait le même raisonnement avec f^{-1}. □

Proposition 54 *Soient E et F deux K-espaces vectoriels de dimension finie et $f : E \to F$ une application linéaire. On a*

$$rg(f) \leq Inf\,(dim(E), dim(F))\,.$$

Démonstration. Si $(e_1, e_2, \ldots, e_n)$ $(n = dimE)$ est une base de E, alors $(f(e_1), f(e_2), \ldots, f(e_n))$ est une famille génératrice de $Im(f)$. Comme $Im(f)$ a une famille génératrice à n éléments, sa dimension est inférieure ou égale à n.
Comme $Im(f)$ est un sous-espace vectoriel de F, sa dimension est inférieure ou égal à la dimension de F.

4.2.3 Théorème du rang

De tous ces résultats, on va déduire le théorème dit " Théorème du rang " qui est un résultat tout à fait fondamental dans la théorie des applications linéaires en dimension finie. On se place toujours dans la même situation : E et F sont deux espaces vectoriels sur un même corps K. L'espace vectoriel E est supposé de type fini et f est une application linéaire de E dans F.

• Il résulte des propriétés générales des applications linéaires que le noyau et l'image d'une application linéaire sont des sous-espaces vectoriels respectivement de l'espace de départ et de l'espace d'arrivée.

• Il résulte des propriétés des sous espaces d'un espace de type fini que le noyau d'une application linéaire d'un espace vectoriel de type fini dans un espace quelconque est de type fini.

• Nous savons que l'image d'une application linéaire d'un espace vectoriel de type fini dans un espace quelconque est de type fini.

L'objet du théorème du rang est de donner une relation entre la dimension du noyau et la dimension de l'image de f.

Théorème 31 (Théorème du rang) *Soient E et F deux espaces vectoriels sur un même corps K, E de type fini. Soit f une application linéaire de E dans F. Alors*

$$\mathrm{dimE} = \mathrm{dimKerf} + \mathrm{dimImf}.$$

Dans la pratique, il suffit donc de déterminer la dimension du noyau ou celle de l'image d'une application linéaire pour avoir les deux dimensions.

Démonstration. Si f est injective, en désignant par $(e_1, \ldots, e_n)$ une base de E, nous avons vu que la famille à n éléments $(f(e_1), \ldots, f(e_n))$ est une famille libre de F donc une famille libre de $\mathrm{Im}(f)$. De plus, $\{f(e_1), \ldots, f(e_n)\}$ est une partie génératrice de

$\mathrm{Im}(f)$. Donc $(f(e_1),\dots,f(e_n))$ est une base de $\mathrm{Im}(f)$. La dimension de $\mathrm{Im}\,f$ est donc égale à n qui est la dimension de E et le théorème du rang est vrai.

Si f n'est pas injective, le noyau de f est un sous espace de E de dimension p avec $1 \leq p \leq n$. Soit $(\varepsilon_1,\dots,\varepsilon_p)$ une base de Kerf. D'après le théorème de la base incomplète, il existe $n-p$ vecteurs $\varepsilon_{p+1},\dots,\varepsilon_n$ de E tels que $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)$ soit une base de E. Alors $\mathrm{Im}\,f$ est engendré par les vecteurs $f(\varepsilon_1),f(\varepsilon_2),\dots,f(\varepsilon_n)$. Mais, comme pour tout i compris entre 1 et p on a $f(\varepsilon_i)=0$, $\mathrm{Im}\,f$ est engendrée par les vecteurs $f(\varepsilon_{p+1}),\dots,f(\varepsilon_n)$. Montrons que ces vecteurs sont linéairement indépendants. Soient $\alpha_{p+1},\dots,\alpha_n$ des scalaires tels que

$$\alpha_{p+1}f(\varepsilon_{p+1})+\cdots+\alpha_n f(\varepsilon_n)=0.$$

Puisque f est linéaire cette égalité équivaut à l'égalité

$$f\left(\alpha_{p+1}\varepsilon_{p+1}+\cdots+\alpha_n\varepsilon_n\right)=0$$

qui prouve que le vecteur $\alpha_{p+1}\varepsilon_{p+1}+\cdots+\alpha_n\varepsilon_n$ appartient au noyau de f. Il existe donc des scalaires $\lambda_1,\dots,\lambda_p$ tels que

$$\alpha_{p+1}\varepsilon_{p+1}+\cdots+\alpha_n\varepsilon_n=\lambda_1\varepsilon_1+\cdots+\lambda_p\varepsilon_p.$$

Comme $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)$ est une base de E, les vecteurs $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$ sont linéairement indépendants et par conséquent :

$$\forall i\in[1,p],\ \ \lambda_i=0\ \text{ et }\ \forall i\in[p+1,n],\ \ \alpha_i=0.$$

Les vecteurs $f(\varepsilon_{p+1}),\dots,f(\varepsilon_n)$ définissent donc une base de $\mathrm{Im}\,f$. Le sous espace vectoriel $\mathrm{Im}\,f$ est donc de dimension $n-p$. Ce qui achève la démonstration. □

On remarquera le rôle essentiel joué par le théorème de la base incomplète dans cette démonstration.

4.2.4 Application linéaire entre deux espaces de même dimension

Théorème 32 *Soient E et F deux espaces vectoriels de type fini sur un même corps K. On suppose qu'ils ont même dimension. Soit f une application linéaire de E dans F. Alors f est injective si et seulement si elle est surjective et donc si et seulement si elle est bijective.*

Autrement dit, dans le cas d'une application linéaire entre deux espaces de même dimension, il suffit pour démontrer qu'elle est bijective, de démontrer l'une des deux propriétés injectivité ou surjectivité.

Démonstration. Cela est immédiat à partir du théorème du rang. En effet la propriété f injective équivaut, d'après le théorème du rang, à $\mathrm{rg}(f)=\mathrm{dimE}$. D'après l'hypothèse sur l'égalité des dimensions de E et de F, Ceci équivaut à la $\mathrm{rg}(f)=\mathrm{dimF}$. Donc $\mathrm{Im}\,f=F$ et f est surjective. On démontre de manière analogue que si f est surjective alors f est injective. Cela achève la démonstration. □

Chapitre 5

Applications linéaires et matrices

Les résultats qui sont développés ici, décrivant un lien entre la notion de matrice et celle d'application linéaire, sont fondamentaux.

5.1 Matrice associée à une application linéaire

Soient E et F deux espaces vectoriels de type fini sur un même corps K. Soit p la dimension de E et $(e_1,\dots,e_p)$ une base de E. Soit n la dimension de F et $(f_1,\dots,f_n)$ une base de F. Soit ϕ une application linéaire de E dans F.

L'étude des propriétés des applications linéaires entre deux espaces de type fini permet d'affirmer que :

- l'application linéaire est déterminée de façon unique par l'image d'une base de E, donc par les vecteurs $\phi(e_1),\phi(e_2),\dots\phi(e_p)$.

Si j est un entier compris entre 1 et p, $\phi(e_j)$ est un vecteur de F et s'écrit de manière unique comme combinaison linéaire des vecteurs de la base $B_F=(f_1,f_2,\dots,f_n)$ de F.

Il existe n scalaires uniques $a_{1,j},a_{2,j},\dots,a_{n,j}$ tels que

$$\phi(e_j)=a_{1,j}f_1+a_{2,j}f_2+\cdots+a_{n,j}f_n.$$

Donc, l'application linéaire est entièrement déterminée par les coefficients $(a_{i,j})_{(i,j)\in[1,n]\times[1,p]}$. Il est donc naturel d'introduire la définition suivante :

Définition 40 *On appelle matrice associée à l'application linéaire ϕ par rapport aux bases B_E et B_F la matrice à n lignes et p colonnes dont la j-ième colonne est constituée par les coordonnées du vecteur $\phi(e_j)$ dans la base $B_F=(f_1,f_2,\dots,f_n)$ à savoir*

$$\begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}.$$

Notation : la matrice associée à l'application linéaire ϕ par rapport aux bases B_E et B_F sera notée $[\phi]_{B_E}^{B_F}$

la notation (E,B_E) signifie que l'on considère l'espace vectoriel E muni de la base B_E .

Remarque 55 *Le type de la matrice associée à l'application linéaire ϕ par rapport aux bases B_E et B_F dépend uniquement de la dimension de E et de celle de F. En effet cette matrice a un nombre de lignes égal à la dimension de l'espace d'arrivée de ϕ et un nombre de colonnes égal à la dimension de l'espace de départ de ϕ.*

2) Des bases étant choisies respectivement dans E et F, il y a unicité de la matrice associée à ϕ. Mais, la matrice trouvée dépend entièrement de ce choix de bases.

Exemple 80 Soit L l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par

$$\begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R}^2 \\ (x_1,x_2,x_3) & \mapsto & (x_1+x_2,x_1+x_3) \end{array}$$

Soient (e_1,e_2,e_3) la base canonique de $\mathbb{R}^3$ et (f_1,f_2) la base canonique de $\mathbb{R}^2$. Déterminons la matrice associée à L dans les bases (e_1,e_2,e_3) et (f_1,f_2).

On a

$$\phi(e_1) = (1,1) = f_1 + f_2.$$

La première colonne de la matrice $[\phi]_{(e_1,e_2,e_3)}^{(f_1,f_2)}$ est donc $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. De même, on a

$$\phi(e_2) = (1,0) = f_1.$$

La deuxième colonne de la matrice $[\phi]_{(e_1,e_2,e_3)}^{(f_1,f_2)}$ est donc $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Enfin on a

$$\phi(e_3) = (0,1) = f_2.$$

La troisième colonne de la matrice $[\phi]_{(e_1,e_2,e_3)}^{(f_1,f_2)}$ est donc $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Il en résulte que

$$[\phi]_{B_E}^{B_F} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

En revanche, la matrice de ϕ dans les bases (e_1,e_2,e_3) et (f_2,f_1) est

$$[\phi]_{(e_1,e_2,e_3}^{(f_2,f_1)} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Sur cet exemple, on voit bien la nécessité de définir une base d'un espace de dimension n comme un n-uplet et non pas comme une partie.

On va maintenant changer la base de l'espace de départ et conserver celle de l'espace d'arrivée. Soient les vecteurs $\varepsilon_1 = (1,1,0)$, $\varepsilon_2 = (1,0,1)$ et $\varepsilon_3 = (0,1,1)$ de $\mathbb{R}^3$. On montre facilement que ces vecteurs déterminent une base de $\mathbb{R}^3$. On considère alors les bases $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ et (f_1,f_2) de $\mathbb{R}^3$ et $\mathbb{R}^2$ respectivement. Alors $L(\varepsilon_1) = 2f_1 + f_2$, $L(\varepsilon_2) = f_1 + f_2$, $L(\varepsilon_3) = f_1 + f_2$ et on a

$$[L]_{(\varepsilon_1,\varepsilon_2,\varepsilon_3)}^{(f_1,f_2)} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Cet exemple illustre bien le fait que la matrice dépend du choix des bases.

5.1.1 Cas d'un endomorphisme d'un espace vectoriel de dimension n

Quelles que soient les bases choisies, la matrice associée à un endomorphisme est une matrice carrée d'ordre n.

Il y a deux grandes catégories de choix de bases dans cette situation :

• Ou bien on prend la même base sur E espace de départ et E espace d'arrivée (ce qui n'avait pas de sens dans le cas général d'une application linéaire entre deux espaces différents). Dans ce cas, la matrice associée à l'endomorphisme en choisissant B comme base, à la fois sur E espace de départ et E espace d'arrivée, est notée $[\phi]_B$.

• Ou bien on prend des bases distinctes.

Matrice associée à l'application identique

Soit donc E un espace vectoriel de dimension égale à n. L'endomorphisme considéré est l'application identique de E, notée Id_E. Soient B et B' deux bases distinctes de E.

• Première situation : On se place dans le schema suivant

$$\begin{array}{rcl} Id_E : (E,B) & \to & (E,B) \\ x & \mapsto & Id_E(x) = x \end{array}$$

Il est facile de voir que $[\phi]_B = I_n$. Bien noter que ce résultat ne dépend de la base B choisie sur E.

• Deuxième situation : on se place dans le schéma suivant :

$$\begin{array}{rcl} (E,B) & \to & (E,B') \\ x & \mapsto & Id_E(x) = x \end{array}$$

où B et B' sont deux bases différentes de E. Si $B = (e_1,\ldots,e_n)$ et $B' = (e'_1,\ldots,e'_n)$, on a $Id_E(e_j) = e_j = \sum_{i=1}^{n} a_{i,j}e'_j$ et $[Id_E]_B^{B'}$ est la matrice dont la j-ième colonne est formée de e_j par rapport à $B' = (e'_1,e'_2,\ldots,e'_n)$ soit $\begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$

Définition 41 *Cette matrice est appelée matrice de passage de la base B' à la base B.*

Elle joue un rôle fondamental lorsque l'on cherche une relation entre les matrices associées à une même application linéaire avec des choix de bases différents.

5.2 Propriété relative à la structure d'espace vectoriel de $L(E,F)$.

5.2.1 Matrices associées à la somme de deux applications linéaires et au produit d'une application linéaire par un scalaire

On sait que la somme de deux applications linéaires d'un K-espace E dans un K-espace vectoriel F est encore une application linéaire de E dans F. Il en est de même pour le produit d'une application linéaire par un scalaire.

Dans le contexte où nous sommes, si les espaces vectoriels considérés sont de type fini, la question qui se pose immédiatement est la suivante :

Quelle est la matrice associée à une somme d'applications linéaires ou au produit d'une application linéaire par un scalaire ?

Proposition 55 *Soient E et F deux espaces vectoriels de type fini sur un même corps K. Soient B_E et B_F des bases de E et F respectivement. Soient h et g deux applications linéaires de E dans F et un scalaire quelconque. Alors on a :*

$$[h+g]_{B_E}^{F_F} = [h]_{B_E}^{B_F} + [g]_{B_E}^{B_F}$$

$$[\alpha h]_{B_E}^{B_F} = \alpha [h]_{B_E}^{B_F}.$$

La matrice associée à la somme de deux applications linéaires est la somme des matrices à condition de considérer respectivement toujours la même base sur l'espace de départ et la même sur l'espace d'arrivée.

Démonstration. Soient p la dimension de E et n la dimension de F, $B_E = (e_1, \ldots, e_p)$ et $B_F = (f_1, f_2, \ldots, f_n)$. Soient $A = (a_{i,j}) = [h]_{B_E}^{B_F}$ et $B = (b_{i,j}) = [g]_{B_E}^{B_F}$.

Cela signifie que, pour tout j compris entre 1 et p on a les égalités :

$$h(e_j) = \sum_{i=1}^{n} a_{i,j} f_i \text{ et } g(e_j) = \sum_{i=1}^{n} b_{i,j} f_i$$

d'où l'on déduit immédiatement l'égalité $(h+g)(e_j) = \sum_{i=1}^{n} (a_{i,j} + b_{i,j}) f_i$.

Cela prouve, d'après la définition de la matrice associée à une application linéaire que le terme général de la matrice associée à $f+g$ par rapport aux bases B_E et B_F est $a_{i,j} + b_{i,j}$ qui est le terme général de la matrice $A+B$. La démonstration de la deuxième formule est tout à fait semblable. □

Théorème 33 (Théorème d'isomorphisme entre et $L(E,F)$ et $M_{\mathrm{dimF,dimE}}(K)$) *Soient E et F deux espaces vectoriels de type fini. Soit p la dimension de E et n celle de F. Les espaces vectoriels $L(E,F)$ et $M_{n,p}(K)$ sont isomorphes.*

Pour prouver ce théorème, on va construire effectivement un isomorphisme entre ces deux espaces vectoriels.

Démonstration. La preuve repose sur le fait qu'une application linéaire définie sur un espace de type fini est déterminée de façon unique par les images des vecteurs d'une base.

Soit $B_E = (e_1, e_2, \ldots, e_p)$ une base de E et $B_F = (f_1, f_2, \ldots, f_n)$ une base de F D'après la proposition précédente, l'application

$$\begin{array}{rcl} L : L(E,F) & \to & M_{\mathrm{dimF,dimE}}(K) \\ h & \mapsto & [h]_{B_E}^{B_F} \end{array}$$

est linéaire.

Il reste à démontrer que L est une bijection. Montrons que L est injective. Soit h dans $L(E,F)$ telle que $[h]_{B_E}^{B_F}$ est nulle. Alors tous les coefficients de $[h]_{B_E}^{B_F}$ sont nuls. On en déduit que, pour tout j dans $[1,p]$, $h(e_j) = 0$. Donc h est nulle.

Montrons que L est surjective. Soit $M = (m_{i,j})$ un élément de $M_{n,p}$. Il existe une unique application linéaire $\phi : E \to F$ telle que $\phi(e_j) = \sum_{i=1}^{n} m_{i,j} f_i$. On a $L(\phi) = M$. $\square$

Remarque 56 *L'isomorphisme L qui vient d'être construit dépend des bases choisies sur E et F.*

La première conséquence de ce théorème est un résultat sur la dimension de $L(E,F)$.

Théorème 34 (Dimension de $L(E,F)$) *Soient E et F deux espaces vectoriels de type fini, dont les dimensions sont respectivement égale à p et n. Alors $L(E,F)$ est de type fini et sa dimension est égale à $n \times p$.*

Démonstration. Cela découle du fait que $L(E,F)$ est isomorphe à $M_{n,p}(K)$ qui est de dimension $n \times p$.

5.3 Produit de matrices et composition d'applications linéaires

Théorème 35 (Matrice associée à la composée de deux applications linéaires) *Soient E, F et G trois espaces vectoriels de type fini sur un même corps K, de dimension respectivement égale à p,n et q. Soient une $B_E = (e_1, \ldots, e_p)$ une base de E, $B_F = (f_1, f_2, \ldots, f_n)$ une base de F et une base de $B_G = (g_1, g_2, \ldots, g_q)$ une base de G. Soient ϕ une application linéaire de E dans F et ψ une application linéaire de F dans G. Alors on a :*

$$[\psi \circ \phi]_{B_E}^{B_G} = [\psi]_{B_F}^{B_G} [\phi]_{B_E}^{B_F}.$$

Autrement dit, à condition de bien choisir les bases, la matrice associée à la composition de deux applications linéaires est le produit des matrices associées à chacune d'elle, dans le même ordre.

Démonstration. Un peu lourde quant aux notations, cette preuve est pourtant simple quant aux idées car elle est uniquement basée sur la définition de la matrice associée à une application linéaire par rapport à des bases choisies, et sur la définition du produit de deux matrices.

La première étape consiste à fixer les notations

Soit $A = [\phi]_{B_E}^{B_F} = (a_{i,j})_{(i,j)\in[1,n]\times[1,p]}$ et $B = [\psi]_{B_F}^{B_G} = (b_{r,s})_{(r,s)\in[1,q]\times[1,n]}$. Les coefficients $a_{i,j}$ et $b_{r,s}$ sont caractérisés par les égalités suivantes :

$$\forall j \in [1,p], \phi(e_j) = \sum_{i=1}^{n} a_{i,j} f_i \quad (1)$$

$$\forall s \in [1,n], \psi(f_s) = \sum_{r=1}^{q} b_{r,s} g_r \quad (2)$$

La deuxième étape consiste à chercher la matrice associée à $\psi \circ \phi$ par rapport aux bases B_E et B_G.

Exprimons, pour tout j dans $[1,p]$, $\psi \circ \phi(e_j)$ dans la base B_G car le terme de la r-ième ligne, j-ième colonne de la matrice cherchée est la coordonnée de $\psi \circ \phi(e_j)$ sur le vecteur g_r.

Les relations (1), impliquent :

$$\psi \circ \phi(e_j) = \psi\left(\sum_{i=1}^{n} a_{i,j} f_i\right) = \sum_{i=1}^{n} a_{i,j} \psi(f_i).$$

D'où, en utilisant les relations (2), on a pour tout $j \in [1, p]$

$$\psi \circ \phi(e_j) = \sum_{i=1}^{n} a_{i,j} \left(\sum_{r=1}^{q} b_{r,i} g_r\right).$$

Les propriétés des lois d'un espace vectoriel, permettent d'écrire :

$$\psi \circ \phi(e_j) = \sum_{r=1}^{q} \left(\sum_{i=1}^{n} a_{i,j} b_{r,i}\right) g_r.$$

Le coefficient de $(\psi \circ \phi)(e_j)$ sur le vecteur g_r de la base B_G est donc égal à $\sum_{i=1}^{n} a_{i,j} b_{r,i} = \sum_{i=1}^{n} b_{r,i} a_{i,j}$ (puisque les coefficients sont dans un corps commutatif).

On reconnaît là exactement le terme général de la matrice BA, ce qui achève la démonstration. □

Exemple 81 Soient les espaces vectoriels $E = \mathbb{R}^2$, $F = \mathbb{R}^3$ et $G = \mathbb{R}$. Soient $B_E = (e_1, e_2)$, $B_F = (f_1, f_2, f_3)$ et $B_G = (g_1)$ leurs bases canoniques respectives. Soit ϕ l'application linéaire de E dans F telle que $[\phi]_{B_E}^{B_F} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 7 \end{pmatrix}$ et ψ l'application linéaire de F dans G telle que $[\psi]_{B_F}^{B_G} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$.

On se propose de déterminer $\psi \circ \phi$, application linéaire de $E = \mathbb{R}^2$ dans $\mathbb{R}$. Pour cela, il suffit de déterminer sa matrice par rapport aux bases canoniques $B_E = (e_1, e_2)$ et $B_G = (g_1)$.

D'après le théorème précédent, on a $[\psi \circ \phi]_{B_E}^{B_G} = [\psi]_{B_F}^{B_G} [\phi]_{B_E}^{B_F}$ Donc

$$[\psi \circ \phi]_{B_E}^{B_G} = [\psi]_{B_F}^{B_G} [\phi]_{B_E}^{B_F} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Cela signifie que $\psi \circ \phi(e_1) = g_1$ et $\psi \circ \phi(e_2) = 0$. D'où l'expression de $\psi \circ \phi$:

$$\forall (x, y) \in \mathbb{R}^2, \ (\psi \circ \phi)(x, y) = x.$$

Cet exemple met bien en évidence le gain, en termes de quantité de calculs, réalisé en passant par l'intermédiaire des matrices.

Dans le cas particulier de la puissance d'un endomorphisme de E, nous obtenons :

Corollaire 36 *Soit E un espace vectoriel de type fini et B_E une base de E. Soit f une application linéaire de E dans E. Alors :*

$$\forall n \in \mathbb{N}, \ [f^n]_{B_E} = ([f]_{B_E})^n.$$

Démonstration. La démonstration est une récurrence immédiate sur n.

Théorème 37 (Caractérisation de la matrice d'un isomorphisme) *Soient E et F deux K-espaces vectoriels de type fini, de même dimension. Une condition nécessaire et suffisante pour qu'une application ϕ linéaire de E dans F soit un isomorphisme est que la matrice associée à ϕ par rapport à des bases B_E et B_F quelconques de E et F respectivement, soit inversible.*

De plus, si ϕ est un isomorphisme de E dans F, et si $A = [\phi]_{B_E}^{B_F}$, la matrice de ϕ^{-1} par rapport aux bases B_F et B_E est égale à A^{-1}, inverse de la matrice A. Cela s'écrit :

$$\left([\phi]_{B_E}^{B_F}\right)^{-1} = [\phi^{-1}]_{B_F}^{B_E}.$$

L'hypothèse E et F ont la même dimension joue un rôle fondamental tout au long de la démonstration.

Démonstration. Soit donc deux espaces vectoriels E et F de dimension n et un isomorphisme ϕ de E dans F. La théorie des applications linéaires permet de dire que ϕ^{-1} est une application linéaire de F dans E. Cela peut se traduire par les égalités :

$$\phi \circ \phi^{-1} = Id_F \text{ et } \phi^{-1} \circ \phi = Id_E.$$

On a

$$Id_E : (E, B_E) \xrightarrow{\phi} (F, B_F) \xrightarrow{\phi^{-1}} (E, B_E)$$
$$Id_F : (F, B_F) \xrightarrow{\phi^{-1}} (E, B_E) \xrightarrow{\phi} (F, B_F)$$

d'où $[\phi^{-1} \circ \phi]_{B_E} = [Id_E]_{B_E} = I_n$ et $[\phi \circ \phi^{-1}]_{B_F} = [Id_F]_{B_F} = I_n$. Compte tenu du théorème général que nous venons d'obtenir pour la matrice associée à la composée d'applications linéaires, cela donne les égalités matricielles suivantes :

$$[\phi^{-1}]_{B_F}^{B_E}[\phi]_{B_E}^{B_F} = I_n \text{ et } [\phi]_{B_E}^{B_F}[\phi^{-1}]_{B_F}^{B_E} = I_n.$$

Cela prouve que la matrice $[\phi]_{B_E}^{B_F}$ est inversible et que son inverse est la matrice $[\phi^{-1}]_{B_F}^{B_E}$.

Démontrons maintenant la réciproque : Soient E et F deux espaces vectoriels de même dimension égale à n, B_E et B_F des bases de E et F respectivement. Soit ϕ une application linéaire de E dans F dont la matrice $A = [\phi]_{B_E}^{B_F}$ par rapport aux bases B_E et B_F est inversible. Soit ψ l'application linéaire de F dans E dont la matrice par rapport aux bases B_F et B_E est A^{-1}. Alors l'égalité $AA^{-1} = A^{-1}A = I_n$ peut s'écrire $[\phi]_{B_E}^{B_F}[\psi]_{B_F}^{B_E} = [\psi]_{B_F}^{B_E}[\phi]_{B_E}^{B_F} = I_n$.

Or, on sait que la matrice associée à l'application identique d'un espace vectoriel de type fini sur lui-même est la matrice unité à condition de prendre la même base sur l'espace de départ et d'arrivée. Donc la matrice I_n peut être considérée soit comme la matrice de l'identité de E par rapport à la base B_E, soit comme la matrice de l'identité de F par rapport à la base B_F. Compte tenu du théorème sur la matrice d'une composition d'applications linéaires, cela implique

$$[\psi \circ \phi]_{B_E} = [Id_E]_{B_E} \text{ et } [\phi \circ \psi]_{B_F} = [Id_F]_{B_F}$$

Or, des bases étant choisies, l'application qui à une application linéaire associe sa matrice par rapport à ces bases, est une bijection donc est en particulier injective.

Donc, des égalités précédentes, on déduit $\psi \circ \phi = Id_E$ et $\phi \circ \psi = Id_F$. Cela prouve que ϕ est inversible et que son application réciproque est égale à ψ.

Comme on a l'égalité $A^{-1} = [\psi]_{B_F}^{B_E}$, on obtient

$$\left([\phi]_{B_E}^{B_F}\right)^{-1} = [\phi^{-1}]_{B_F}^{B_E}.$$

□

Dans le cas particulier d'un endomorphisme.

Corollaire 38 *Soit E un K-espace vectoriel de type fini.*

Une condition nécessaire et suffisante pour qu'une application linéaire f de E dans E soit un automorphisme est que la matrice associée à f dans une base quelconque de E soit inversible.

De plus, si f est un automorphisme de E et si $A = [f]_{B_E}$, la matrice de f^{-1} dans la base B_E est égale à inverse de la matrice A. Cela s'écrit :

$$([f]_{B_E})^{-1} = [f^{-1}]_{B_E}.$$

5.4 Traduction matricielle de l'action d'une application linéaire sur un vecteur

Soient E et F deux espaces vectoriels de type fini sur un même corps K et ϕ une application linéaire de E dans F. Le but de ce paragraphe est de traduire l'égalité vectorielle $y = \phi(x)$ par une égalité matricielle et d'étudier des applications de ce résultat.

Notation : Soit E un espace vectoriel de dimension finie et soit $B_E = (e_1, e_2, \ldots, e_p)$ une base de E. Soit x un élément de E. Il existe un p-uplet unique d'éléments de K, $(x_1, x_2, \ldots, x_p)$ tel que

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

La matrice colonne des coordonnées de x est noté $[x]_{B_E}$. On a donc $[x]_{B_E} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$

Proposition 56 *Soit E un espace vectoriel de dimension finie et B_E une base de E. Soit F un espace vectoriel de dimension finie et B_F une base de F. Soit $\phi : E \to F$ une application linéaire et $[\phi]_{B_E}^{B_F}$ la matrice de ϕ dans les bases B_E et B_F. On a*

$$[\phi(x)]_{B_F} = [\phi]_{B_E}^{B_F} [x]_{B_E}.$$

Démonstration. On pose $B_E = (e_1, \ldots, e_p)$, $B_F = (f_1, f_2, \ldots, f_n)$, $[\phi]_{B_E}^{B_F} = A = (a_{i,j})$ et $[x]_{B_E} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$.

$$\phi(x) = \phi\left(\sum_{i=1}^{p} x_i e_i\right) = \sum_{i=1}^{p} x_i \phi(e_i) = \sum_{i=1}^{p} x_i \left(\sum_{k=1}^{n} a_{k,i} f_k\right).$$

En utilisant la commutativité de K, on a

$$\phi(x) = \left(\sum_{i=1}^{p} a_{1,i}x_i\right) f_1 + \cdots + \left(\sum_{i=1}^{p} a_{n,i}x_i\right) f_n.$$

La matrice colonne des coordonnées de $\phi(x)$ dans la base $(f_1, f_2, \ldots, f_n)$ est $\begin{pmatrix} \sum_{i=1}^{p} a_{1,i}x_i \\ \sum_{i=1}^{p} a_{2,i}x_i \\ \vdots \\ \sum_{i=1}^{p} a_{n,i}x_i \end{pmatrix}$.

Or la matrice $[\phi(x)]_{B_F} = \begin{pmatrix} \sum_{i=1}^{p} a_{1,i}x_i \\ \sum_{i=1}^{p} a_{2,i}x_i \\ \vdots \\ \sum_{i=1}^{p} a_{n,i}x_i \end{pmatrix}$ n'est autre que $A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$. □

Exemple 82 Soit E une K-espace vectoriel de dimension 3 et $B_E = (e_1, e_2, e_3)$ une base de E. Les éléments de E sont donc des combinaisons linéaires de e_1, e_2 et e_3. Soit f l'endomorphisme de E dont la matrice dans la base B_E est égale

$$[f]_{B_E} = A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

On se propose de déterminer le noyau de f. On a

$$\begin{aligned} x_1e_1 + x_2e_2 + x_3e_3 \in \mathrm{Ker(f)} &\iff f(x_1e_1 + x_2e_2 + x_3e_3) = 0 \\ &\iff [f(x_1e_1 + x_2e_2 + x_3e_3)]_{B_E} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff [f]_{B_E}[x_1e_1 + x_2e_2 + x_3e_3]_{B_E} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ &\iff \begin{cases} x_1 + 2x_2 + x_3 = 0 \\ 2x_1 + 3x_2 + 2x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \end{aligned}$$

On résoud ce système par la méthode du pivot de Gauss. On trouve

$$\mathrm{Kerf} = \{\mathrm{x_1e_1 + x_2e_2 + x_3e_3 \in E \mid x_2 = 0,\ x_1 + x_3 = 0}\} = 0 = \{\mathrm{te_1 - te_3 \mid t \in \mathbb{R}}\}$$

5.5 Formule de changement de bases

5.5.1 Matrice de passage d'une base à une autre

Soit E un espace vectoriel de type fini, n sa dimension. On sait que toutes les bases de E ont n éléments.

Définition 42 *Soit B une base de E. Soit B' une autre base définie par la donnée des coordonnées de ses vecteurs dans la base B.*

On appelle matrice de passage de la base B à la base B' et on note $P_{BB'}$ la matrice carrée d'ordre n dont la j-ième colonne est formée des coordonnées du j-ième vecteur de la base B', par rapport à la base B.

Attention à l'ordre des bases dans cette définition.

Exemple 83 Soit l'espace vectoriel réel $\mathbb{R}^2$. On considère la base canonique et la base $B' = (\varepsilon_1, \varepsilon_2)$ avec $\varepsilon_1 = e_1 + e_2$ et $\varepsilon_2 = e_2$. La matrice de passage de la base B à la base B' est la matrice $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ dont la première colonne est donnée par les coordonnées du vecteur ε_1 sur la base (e_1, e_2) et la deuxième par les coordonnées ε_2 de sur la base (e_1, e_2).

On va interpréter cette matrice comme la matrice associée à l'application identique de E par rapport à des bases bien choisies.

Proposition 57 (Interprétation d'une matrice de passage entre deux bases) *Soient B et B' deux bases de E. La matrice de passage de B à B' est égale à la matrice de l'application identique de E, avec B' comme base de l'espace de départ et B comme base de l'espace d'arrivée.*

Remarque 57 Cette interprétation est un outil fondamental pour ce qui suit. Elle permet d'obtenir les résultats de façon très élégante et avec un minimum de calculs.

Démonstration. On pose $B = (e_1, e_2, \ldots, e_n)$ et $B' = (e'_1, e'_2, \ldots, e'_n)$. On considère

$$\begin{array}{rcl} Id_E : (E, B') & \to & (E, B) \\ x & \mapsto & Id_E(x) = x \end{array}$$

$Id_E(e'_j) = e'_j = \sum_{i=1}^{n} a_{i,j} e_i$ et $[Id]_{B'}^{B}$ est la matrice dont le j-ième colonne est formée des coordonnées de e'_j par rapport à B, soit $\begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$. Cette colonne est la j-ième colonne de $P_{BB'}$. □

Proposition 58 (Inverse d'une matrice de passage) *La matrice de passage d'une base B à une base B' est inversible et son inverse est égale à la matrice de passage de la base B' à la base B.*

Démonstration. On a $P_{B,B'} = [Id_E]_{B'}^{B}$. Donc, d'après le théorème caractérisant la matrice d'un isomorphisme, $P_{B,B'}^{-1} = \left([Id_E]_{B'}^{B}\right)^{-1} = [Id_E^{-1}]_{B}^{B'}$. Or $Id_E^{-1} = Id_E$. Donc $P_{B,B'}^{-1} = [Id_E]_{B}^{B'} = P_{B'B}$. □

Proposition 59 *Soient B, B' et B'' trois bases d'un espace vectoriel de type fini E. Alors*

$$P_{BB''} = P_{BB'}P_{B'B''}.$$

Démonstration. $Id_E : (E,B'') \to (E,B)$ se factorise de la façon suivante :

$$(E,B'') \xrightarrow{Id_E} (E,B') \xrightarrow{Id_E} (E,B).$$

Cette factorisation permet d'écrire l'égalité suivante $[Id_E]^B_{B''} = [Id_E]^B_{B'}[Id_E]^{B'}_{B''}$. Soit $P_{BB''} = P_{BB'}P_{B'B''}$. □

Nous allons maintenant interpréter une matrice carrée inversible comme la matrice de passage d'une base à une autre base.

Proposition 60 *Soit E un K-espace vectoriel de dimension n et B_E une base de E. Soit M une matrice carrée d'ordre n inversible. Il existe une unique base B'_E de E telle que M soit la matrice de passage de B_E à B'_E.*

Démonstration. On pose $B_E = (e_1, e_2, \ldots, e_n)$. soit e'_j le vecteur de E dont la matrice colonne des coordonnées dans la base B_E est la j-ième colonne de M. Le rang de la matrice M est le rang de la famille $e'_1, e'_2, \ldots, e'_n$. La famille $e'_1, e'_2, \ldots, e'_n$ étant de rang n, elle est linéairement indépendante. Comme E est de dimension n, $B'_E = (e'_1, e'_2, \ldots, e'_n)$ est une base de E. Il est évident que M est la matrice de passage de la base B_E à la base B'_E. □

Nous allons maintenant étudier l'effet d'un changement de bases sur les coordonées d'un vecteur.

Proposition 61 *Soit E un K-espace vectoriel de dimension n. Soit $B = (e_1, e_2, \ldots, e_n)$ et $B' = (e'_1, e'_2, \ldots, e'_n)$ deux bases de E. Soit $P_{BB'}$ la matrice de passage de B à B'. Soit x un vecteur de E. Si $x = \sum_{i=1}^{n} x_i e_i$, la matrice colonne des coordonnées de x dans la base B est $[x]_B = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. De même notons $[x]_{B'}$ la matrice colonne des coordonnées de x dans la base B'. On a la relation*

$$[x]_B = P_{BB'}[x]_{B'}.$$

Démonstration. $P_{BB'}$ est la matrice de $Id_E : (E,B') \to (E,B)$. On a $[Id_E(x)]_B = [Id_E]^B_{B'}[x]_{B'}$, soit $[x]_B = P_{BB'}[x]_{B'}$. □

5.5.2 Formule de changement de bases

Théorème 39 (Formule de changement de base) *Soient E et F deux K-espaces vectoriels de type fini, B_E et B'_E deux bases de E, et B_F et B'_F deux bases de F. Soit ϕ une application linéaire de E dans F.*

Alors, la matrice associée à ϕ par rapport aux bases B_E et B_F, et la matrice associée à ϕ par rapport aux bases B'_E et B'_F sont liées par la formule :

$$[\phi]^{B'_F}_{B'_E} = P_{B'_F B_F}[\phi]^{B_F}_{B_E}P_{B_E B'_E} = \left(P_{B_F B'_F}\right)^{-1}[\phi]^{B_F}_{B_E}P_{B_E B'_E}.$$

Démonstration. L'application $\phi : (E, B'_E) \to (F, B'_F)$ se factorise de la façon suivante

$$(E, B'_E) \xrightarrow{Id_E} (E, B_E) \xrightarrow{\phi} (F, B_F) \xrightarrow{Id_F} (F, B'_F)$$

On a donc l'égalité de matrices suivante :

$$\begin{aligned} [\phi]_{B'_E}^{B'_F} &= [Id_F]_{B_F}^{B'_F} [\phi]_{B_E}^{B_F} [Id_E]_{B'_E}^{B_E} \\ &= P_{B'_F B_F} [\phi]_{B_E}^{B_F} P_{B_E B'_E}. \end{aligned}$$

Dans le cas particulier d'un endomorphisme, on obtient le théorème suivant :

Corollaire 40 (Formule de changement de bases pour les endomorphismes) *Soit E un K-espace vectoriel de dimension n. Soient B_E et B'_E deux bases de E et $P = P_{B_E B'_E}$ la matrice de passage de B_E à B'_E. Soit ϕ un endomorphisme de E. Notons $[\phi]_{B_E}$ (respectivement $[\phi]_{B'_E}$) la matrice de ϕ dans la base B_E (respectivement B'_E). On a*

$$[\phi]_{B'_E} = P^{-1} [\phi]_{B_E} P.$$

Remarque 58 Nous avons vu qu'une matrice carrée inversible pouvait être interprétée comme une matrice de passage. La formule de changement de bases pour les endomorphismes montre alors que deux matrices semblables représentent le même endomorphisme dans des bases différentes.

5.6 Rang d'une matrice et applications linéaires

Rappelons le définition du rang d'une application linéaire.

Définition 43 *Soient E et F deux K-espaces vectoriels de dimension finie et f une application linéaire de E dans F. La dimension de l'espace vectoriel* $\operatorname{Im} f$ *est appelé rang de f et est noté $rg(f)$.*

Proposition 62 *Soient E et F deux K-espaces vectoriels de dimension finie et f une application linéaire de E dans F.Soit B_E une base de E et B_F une base de F. Le rang de f est le rang de sa matrice dans les bases B_E et B_F.*

Démonstration. Soient $B_E = (e_1, e_2, \ldots, e_p)$ et $B_F = (f_1, f_2, \ldots, f_n)$. Comme la famille $f(e_1), f(e_2), \ldots, f(e_p)$ engendre $\operatorname{Im} f$, le rang de f est le rang de la famille $f(e_1), f(e_2), \ldots, f(e_p)$. D'après les résultats que nous avons vu dans le chapitre 4, le rang de la famille $f(e_1), f(e_2), \ldots, f(e_p)$ est le rang de la matrice de f dans les bases B_E et B_F. □

Théorème 41 (Matrice inversible et rang) *Une matrice carrée d'ordre n est inversible si et seulement si elle est de rang n.*

Démonstration. Soit A une matrice carrée d'ordre n. Soit f l'endomorphisme de K^n dont la matrice dans la base canonique est A. On a les équivalences suivantes :

$$\begin{aligned} A \text{ de rang } n &\iff f \text{ de rang } n \\ &\iff f \text{ surjective} \\ &\iff f \text{ bijective} \\ &\iff A \text{ inversible.} \end{aligned}$$

Nous avons utilisé le fait qu'un endomorphisme d'un espace vectoriel de dimension finie est bijectif si et seulement si il est surjectif et le théorème sur la caractérisation de la matrice d'un isomorphisme. □

Chapitre 6

Déterminants

Ce chapitre ne concerne que *les matrices carrées.* On travaillera systématiquement dans $M_n(K)$.

6.1 Théorie des déterminants

6.1.1 Définition et premières propriétés

Théorème 42 (Théorème d'existence et d'unicité du déterminant) *Il existe une unique application de $M_n(K)$ dans K, appelée* déterminant, *telle que*

i) Le déterminant est linéaire par rapport à chaque vecteur-colonne, les autres étant fixés.

ii) Si une matrice A a deux vecteurs colonnes égaux, alors son déterminant est nul.

iii) Le déterminant de la matrice identité I_n vaut 1.

On a plusieurs notations pour les déterminants :

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

On note a_i la $i$ème colonne de A et

$$\det A = \begin{vmatrix} a_1 & a_2 & \cdots & a_n \end{vmatrix}.$$

Avec cette notation, la propriété i) s'écrit : pour tout λ dans K,

$$\begin{aligned} \begin{vmatrix} a_1 & a_2 & \cdots a_{i-1} & a_i + \lambda a_i' & a_{i+1} & \cdots & a_n \end{vmatrix} &= \begin{vmatrix} a_1 & a_2 & \cdots a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{vmatrix} \\ &+ \lambda \begin{vmatrix} a_1 & a_2 & \cdots a_{i-1} & a_i' & a_{i+1} & \cdots & a_n \end{vmatrix}. \end{aligned}$$

Une application de $M_n(K)$ dans K qui satisfait la propriété i) est appelée *forme multilinéaire.* Si elle satisfait ii), on dit qu'elle est *alternée.* Le déterminant est donc la seule forme multilinéaire alternée qui vaut 1 sur la matrice I_n. Les autres formes multilinéaires alternées sont les multiples scalaires du déterminant. On verra plus loin comment on peut calculer effectivement les déterminants.

Donnons maintenant quelques propriétés importantes du déterminant.

Proposition 63 *Soit A une matrice $n \times n$ et A' la matrice obtenue en échangeant deux colonnes distinctes de A. Alors on a* $\det A' = -\det A$.

Démonstration. Soit $A = \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix}$. On va échanger les colonnes i et j, ce qui donne la matrice $A' = \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix}$, où le vecteur a_j se retrouve en colonne i et le vecteur a_i en colonne j (on pris ici $i < j$, sans perte de généralité).

Introduisons alors une troisième matrice

$$\tilde{A} = \begin{pmatrix} a_1 & \cdots & a_i + a_j & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix}.$$

Cette matrice a deux colonnes distinctes égales, donc d'après ii)

$$\det \tilde{A} = 0.$$

D'un autre côté, nous pouvons développer ce déterminant en utilisant la propriété i) de multilinéarité, c'est-à-dire linéarité par rapport à chaque colonne. Ceci donne

$$\begin{aligned} \det \tilde{A} &= \det \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix} + \det \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_j + a_i & \cdots & a_n \end{pmatrix} \\ &= \det \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix} + \det \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_i & \cdots & a_n \end{pmatrix} \\ &\quad + \det \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_j & \cdots & a_n \end{pmatrix} + \det \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix} \\ &= \det A + 0 + 0 + \det A', \end{aligned}$$

encore grâce à i) pour les deux déterminants du milieu. □

Proposition 64 *Soit A une matrice $n \times n$ et A' la matrice obtenue en ajoutant à une colonne de A une combinaison linéaire des autres colonnes de A. Alors on a* $\det A' = \det A$.

Démonstration. Soit $A = \begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix}$ et donnons nous des scalaires λ_j, $j = 1, \ldots, n$, $j \neq i$. On pose

$$A' = \begin{pmatrix} a_1 & \cdots & a_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda_j a_j & \cdots & a_n \end{pmatrix}.$$

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

$$\det A' = \det A + \sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda_j \det \begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_n \end{pmatrix}.$$

Or chacun des déterminants apparaissant sous le signe de sommation est nul, puisqu'il concerne une matrice dont les colonnes i et j sont égales. □

Corollaire 43 *Si une colonne de A est combinaison linéaire des autres colonnes alors* $\det A = 0$.

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

6.1.2 Déterminants de matrices particulières

Calculons le déterminant de matrices triangulaires en utilisant les propriétés du déterminant.

Proposition 65 *Si A est une matrice triangulaire supérieure ou inférieure, alors on a*

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

Autrement dit, pour une matrice triangulaire, le déterminant est égal au produit des termes diagonaux.

Démonstration. On traite le cas des matrices triangulaires supérieures, le cas des matrices triangulaires inférieures est identique. Soit donc

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Par linéarité par rapport à la première colonne, on a

$$\det A = a_{11}\begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

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

$$\det A = a_{11}\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

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

$$\det A = a_{11}a_{22}\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix},$$

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

$$\det A = a_{11}a_{22}a_{33}\cdots a_{nn}\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = a_{11}a_{22}a_{33}\cdots a_{nn}\det I_n,$$

d'où le résultat par iii). □

Remarque 59 La façon de procéder doit rappeler l'algorithme de Gauss. C'est en fait le même argument mais avec des substitutions de colonnes.

Notons que le résultat s'applique en particulier aux *matrices diagonales*, lesquelles sont à la fois triangulaires supérieures et triangulaires inférieures.

Corollaire 44 *Soit E une matrice élémentaire de la méthode de Gauss.*
i) Si E est la matrice d'une substitution de lignes, alors $\det E = 1$.
ii) Si E est la matrice d'un échange de lignes, alors $\det E = -1$.
iii) Si E est la matrice d'une multiplication d'une ligne par $\lambda \neq 0$*, alors* $\det E = \lambda$.
Dans tous les cas, ce déterminant est non nul.

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

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

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

Remarque 60 Les matrices élémentaires de la méthode de Gauss sont soit triangulaire (substitution), soit symétriques c'est à dire égales à leur transposée (échange de lignes et homothétie). Par conséquent, $\det(E_i) = \det(E_i^T)$.

6.1.3 Démonstration du théorème d'existence et d'unicité

Démonstration. Pour démontrer l'existence d'un objet mathématique satisfaisant aux conditions du théorème - définition, on donne une formule qui, de plus, permet de calculer effectivement le déterminant d'une matrice, et on vérifie que les propriétés caractéristiques des déterminants sont satisfaites.

Notation : Soit une matrice carrée d'ordre n. Il est évident que si l'on supprime une ligne et une colonne dans M, la matrice obtenue est à $n-1$ lignes et colonnes. On note la matrice obtenue en supprimant la i-ème ligne et la j-ième colonne $M_{i,j}$. Le théorème d'existence peut s'énoncer de la façon suivante :

Théorème 45 (Existence du déterminant) *Les formules suivantes :*
- *Si a est un élément quelconque de K,* $\det(a) = a$
- *Si* $M = (m_{i,j})$ *est une matrice carrée d'ordre n.*

$$\det(M) = (-1)^{i+1} m_{i,1} \det(M_{i,1}) + (-1)^{i+2} m_{i,2} \det(M_{i,2}) + \cdots + (-1)^{i+n} m_{i,n} \det(M_{i,n})$$

définissent par récurrence, pour tout entier n supérieur ou égal à 1, une application de $M_n(K)$ *dans K qui satisfait aux propriétés caractérisant les déterminants.*

Démonstration du théorème d'existence du déterminant

La démonstration du théorème d'existence du déterminant est hors programme. Elle se fait par récurrence sur l'ordre des matrices.

Dans le cas $n = 1$. Il est évident que toutes les propriétés souhaitées sont satisfaites. Supposons maintenant que l'application $\det : M_{n-1}(K) \to K$ soit définie et satisfasse les propriétés (1), (2) et (3). Pour faciliter l'exposition, la preuve va être faite pour $i = 1$. Soit $M = (m_{i,j})$ notée aussi $M = (C_1, \ldots, C_n)$ où C_j est la j-ième colonne de M.

• Propriété (1). Il s'agit de vérifier que l'application

$$M \mapsto \det(M) = (-1)^2 m_{1,1}\det(M_{1,1}) + (-1)^3 m_{1,2}\det(M_{1,2}) + \cdots + (-1)^{1+n} m_{1,n}\det(M_{1,n})$$

est linéaire par rapport à chaque colonne. Soit j un entier compris entre 1 et n. Montrons que l'application

$$C_j \mapsto \det(M) = (-1)^2 m_{1,1}\det(M_{1,1}) + (-1)^3 m_{1,2}\det(M_{1,2}) + \cdots + (-1)^{1+n} m_{1,n}\det(M_{1,n})$$

est linéaire. Si $C_j = D_j + \beta D'_j$, cette décomposition modifie le coefficient $m_{1,j}$ ($m_{1,j} = d_{1,j} + \beta d'_{1,j}$) et les matrices $M_{1,k}$ avec $k \neq j$ puisque dans ces matrices la j-ième colonne reste). Cela donne

$$\begin{aligned}
\det(M) &= (-1)^{1+j} m_{1,j}\det(M_{1,j}) + \sum_{1\leq k\leq n, k\neq j} (-1)^{1+k} m_{1,k}\det(M_{1,k}) \\
&= (-1)^{1+j}(d_{1,j} + \beta d'_{1,j})\det(M_{1,j}) \\
&+ \sum_{1\leq k\leq n, k\neq j} (-1)^{1+k} m_{1,k}\det\left((\hat{C}_1^1, \ldots, \hat{D}_j^1 + \beta \hat{D}_j'^1, \ldots, \hat{C}_n^1)_k\right)
\end{aligned}$$

où $\hat{C}_r^1$ désigne la r-ième colonne de M à laquelle on a supprimé la première ligne et $(\hat{C}_1^1, \ldots, \hat{C}_j^1, \ldots \hat{C}_n^1)_k$ la matrice déduite de $(C_1, C_2, \ldots, C_n)$ en supprimant la k-ième colonne et la première ligne. Par conséquent les matrices $(\hat{C}_1^1, \ldots, \hat{D}_j^1 + \beta \hat{D}_j'^1, \ldots, \hat{C}_n^1)_k$ qui interviennent dans la somme précédente possèdent $n-1$ lignes et $n-1$ colonnes et on peut donc leur appliquer l'hypothèse de récurrence. D'où

$$\begin{aligned}
\det(M) &= \left[(-1)^{1+j} d_{1,j}\det(M_{1,j}) + \sum_{1\leq k\leq n, k\neq j} (-1)^{1+k} m_{1,k}\det\left((\hat{C}_1^1, \ldots, \hat{D}_j^1, \ldots, \hat{C}_n^1\right)_k\right] \\
&+ \beta\left[(-1)^{1+j} d'_{1,j}\det(M_{1,j}) + \sum_{1\leq k\leq n, k\neq j} (-1)^{1+k} m_{1,k}\det\left((\hat{C}_1^1, \ldots, \hat{D}_j'^1, \ldots, \hat{C}_n^1\right)_k\right] \\
&= \det(C_1, \ldots, D_j, \ldots, C_n) + \beta\det(C_1, \ldots, D'_j, \ldots, C_n)
\end{aligned}$$

Ce qui achève la démonstration de la propriété (1).

• Propriété (2)

Si la matrice M a deux colonnes égales, par exemple C_r et C_s avec r et s distincts, il est clair que les colonnes obtenues en supprimant la première ligne, $\hat{C}_s$ et $\hat{C}_r$ sont encore égales. Donc toutes les matrices d'ordre $n-1$, $(\hat{C}_1, \ldots, \hat{C}_r, \ldots, \hat{C}_s, \ldots, \hat{C}_n)_k$ avec k différent de r et s, ont deux colonnes égales. D'où, par hypothèse de récurrence,

$$\forall k \in \{1, 2, \ldots, n\} - \{r, s\},\ \ \det(\hat{C}_1, \ldots, \hat{C}_r, \ldots, \hat{C}_s, \ldots, \hat{C}_n)_k = 0.$$

Donc

$$\begin{aligned}
\det(M) &= \sum_{1\leq k\leq n} (-1)^{1+k} m_{1k}\det\left((\hat{C}_1, \ldots, \hat{C}_j, , \ldots, \hat{C}_n)_k\right) \\
&= (-1)^{1+s} m_{1,s}\det\left((\hat{C}_1, \ldots, \hat{C}_r, , \ldots, \hat{C}_n)_s\right) + (-1)^{1+r} m_{1,r}\det\left((\hat{C}_1, \ldots, \hat{C}_s, , \ldots, \hat{C}_n)_r\right)
\end{aligned}$$

où $\hat{C}_s^1$ et $\hat{C}_r^1$ sont égales. Supposons, par exemple, s supérieur ou égale à r. Il faut faire $s-r-1$ échanges de colonnes pour amener la s-ième colonne à la r-ième place. Compte tenu de l'hypothèse de récurrence, on en déduit

$$\det\left((\hat{C}_1, \ldots, \hat{C}_s, , \ldots, \hat{C}_n)_r\right) = (-1)^{s-r-1}\det\left((\hat{C}_1, \ldots, \hat{C}_r, , \ldots, \hat{C}_n)_s\right).$$

d'où

$$\begin{aligned}\det(M) &= (-1)^{1+s} m_{1,s} \det\left((\hat{C}_1, \ldots, \hat{C}_r, , \ldots, \hat{C}_n)_s\right) \\ &+ (-1)^{1+r}(-1)^{s-r-1} m_{1,r} \det\left((\hat{C}_1, \ldots, \hat{C}_r, , \ldots, \hat{C}_n)_s\right).\end{aligned}$$

Puisque $m_{1,r} = m_{1,s}$ (les colonnes de rang r et de rang s sont égales), on en déduit $\det(M) = 0$.

• Propriété (3) Si l'on considère la matrice unité I_n, ses coefficients $m_{i,j}$ sont tels que :

$$\begin{aligned} i = j &\Longrightarrow m_{i,j} = 1 \\ i \neq j &\Longrightarrow m_{i,j} = 0. \end{aligned}$$

Donc $\det(I_n) = (-1)^{1+1} \det(I_n)_{1,1}$. Or, la matrice obtenue à partir de la matrice unité en supprimant la première ligne et la première colonne est la matrice unité d'ordre $n-1$. En lui appliquant l'hypothèse de récurrence, on a $\det(I_n)_{1,1} = 1$. On en déduit $\det(I_n) = 1$. Ce qui achève la preuve du théorème d'existence du déterminant.

Nous admettrons l'unicité du déterminant. □

Remarque 61 La définition donnée ci-dessus suppose le choix d'un indice i de ligne et peut paraître arbitraire. Alors se pose naturellement la question : que se passe-t-il si l'on prend une autre valeur pour i ? L'unicité du déterminant d'une matrice permet de répondre : quelle que soit la ligne choisie, le résultat est le même.

6.1.4 Propriétés du déterminant

Théorème 46 *On a*

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

Démonstration. Cette démonstration est hors programme. Commençons par faire la remarque préliminaire suivante :

Remarque préliminaire :

Si M est une matrice carrée et E une matrice élémentaire, la matrice EM^T est la matrice obtenue à partir de M^T en faisant l'opération élémentaire correspondante sur les lignes. Donc, par transposition, multiplier une matrice à droite par la transposée d'une matrice élémentaire effectue l'opération élémentaire correspondante sur les colonnes de M. Dans le cas où E est la matrice d'une substitution de lignes, on ne modifie pas le déterminant puisqu'on ajoute à une colonne un multiple d'une autre colonne. Dans la cas où E est la matrice d'un échange de lignes, on multiplie le déterminant par -1 puisqu'on échange deux colonnes. Dans le cas où E est la matrice d'une multiplication d'une ligne par λ, on multiplie le déterminant par λ par linéarité par rapport à la colonne multipliée par λ. Dans tous les cas, on a :

$$\det(ME^T) = \det(M)\det(E).$$

Passons maintenant à la démonstration du théorème. Supposons d'abord que B soit inversible. L'algorithme de Gauss appliquée à la matrice B^T fournit des matrices élémentaires E_j telles que

$$(E_p E_{p-1} \cdots E_2 E_1) B^T = I_n.$$

On a alors :

$$B(E_1^T E_2^T \cdots E_p^T) = I_n^T = I_n$$

D'après la remarque préliminaire appliquée p fois, on a

$$\det B(E_1^T E_2^T \cdots E_p^T) = \det B \det(E_1)\det(E_2)\ldots\det(E_p) = 1$$

On en déduit

$$\det B = \frac{1}{\det(E_p)\cdots\det(E_2)\det(E_1)}.$$

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

$$\det(C(E_1^T E_2^T \cdots E_p^T)) = \det(E_p)\cdots\det(E_2)\det(E_1)\det C.$$

Or

$$C(E_1^T E_2^T \cdots E_p^T) = A(B(E_1^T E_2^T \cdots E_p^T)) = A,$$

d'où le résultat dans ce cas.

Si B n'est pas inversible, $\operatorname{rg} B < n$, il existe donc une relation de dépendance linéaire entre les colonnes de B (ce qui revient à dire qu'il existe une matrice colonne x telle que $Bx = 0$) et donc $\det B = 0$. Or $Bx = 0 \Longrightarrow ABx = 0$. On voit que AB n'est pas inversible non plus, d'où $\det(AB) = 0 = \det A \det B$ également dans ce cas. □

Un des usages des déterminants est de caractériser les matrices inversibles.

Corollaire 47 *Une matrice carrée d'ordre n, A, est inversible si et seulement si son déterminant est non nul. De plus si A est inversible, alors* $\det(A^{-1}) = \frac{1}{\det A}$.

Démonstration. Si A n'est pas inversible, alors elle est de rang strictement inférieur à n. Il existe donc une relation de dépendance linéaire entre ses colonnes, c'est à dire qu'au moins l'une de ses colonnes est combinaison linéaire des autres. On en déduit $\det(A) = 0$. Si A est inversible, il existe une matrice A^{-1} telle que $AA^{-1} = A^{-1}A = I_n$, donc $\det(A)\det(A^{-1}) = \det I_n = 1$. On en déduit que $\det(A)$ est non nul et $\det(A^{-1}) = \frac{1}{\det A}$. □

Corollaire 48 *Deux matrices semblables ont même déterminant.*

Démonstration. Soit $A' = P^{-1}AP$ avec $P \in GL_n(\mathbb{R})$. Par multiplicativité du déterminant, on en déduit que

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

puisque $\det P^{-1} = 1/\det P$. □

Corollaire 49 *On a* $\det(A^T) = \det A$.

Démonstration. Cette démonstration est hors programme. Soit $A \in M_n(K)$. Par l'algorithme de Gauss, on a une factorisation $A = M^{-1}U$ avec U échelonnée réduite, donc en particulier triangulaire supérieure et $M^{-1} = E_1^{-1}E_2^{-1}\cdots E_{p-1}^{-1}E_p^{-1}$. Par conséquent, en transposant on a aussi $A^T = U^T(M^{-1})^T$ avec $(M^{-1})^T = (E_p^{-1})^T(E_{p-1}^{-1})^T\cdots(E_2^{-1})^T(E_1^{-1})^T$.

Utilisant la multiplicativité du déterminant, on en déduit

$$\begin{cases} \det A &= \dfrac{\det U}{\det E_1 \det E_2 \cdots \det E_p}, \\ \det A^T &= \dfrac{\det U^T}{\det E_1^T \det E_2^T \cdots \det E_p^T}. \end{cases}$$

Or U est triangulaire supérieure, son déterminant est le produit de ses termes diagonaux. Par conséquent, U^T est triangulaire inférieure et son déterminant est le produit de ces mêmes termes diagonaux, c'est-à-dire $\det U = \det U^T$.

De même, les matrices E_i sont soit triangulaires (substitution), soit symétriques c'est-à-dire égales à leur transposée (échange de lignes et homothétie). Par conséquent, $\det E_i = \det E_i^T$ aussi, d'où le résultat. □

Remarque 62 Tout ce que l'on a dit des déterminants à propos des colonnes est donc vrai pour les lignes. Ainsi, le déterminant est multilinéaire par rapport aux lignes, si une matrice a deux lignes égales, son déterminant est nul, on ne modifie pas un déterminant en ajoutant à une ligne une combinaison linéaire des autres lignes, etc. □

Définition 44 *Soit A une matrice $n \times n$ et A_{ij} la matrice $(n-1) \times (n-1)$ obtenue en effaçant la ligne i et la colonne j de A. On appelle* mineur de A relatif à a_{ij} *le déterminant $\Delta_{ij} = \det A_{ij}$. On appelle* cofacteur de A relatif à a_{ij} *le nombre $C_{ij} = (-1)^{i+j}\Delta_{ij}$.*

Théorème 50 (développement suivant une ligne ou une colonne) *On a les formules suivantes :*

$$\begin{aligned}
\forall i, \quad \det A &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\Delta_{ij} = \sum_{j=1}^{n} a_{ij}C_{ij} \\
&\qquad \textit{(développement par rapport à la ligne } i\textit{)}, \\
\forall j, \quad \det A &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\Delta_{ij} = \sum_{i=1}^{n} a_{ij}C_{ij} \\
&\qquad \textit{(développement par rapport à la colonne } j\textit{)}.
\end{aligned}$$

Démonstration. Nous avons déja démontré la formule de développement suivant une ligne lors de la démonstration du théorème d'existence et d'unicité du déterminant. Comme $\det A = \det A^T$, on en déduit la formule de développement par rapport à une colonne.

Remarque 63 Le développement par rapport à une ligne permet de ramener le calcul d'un déterminant $n \times n$ à celui de n déterminants $(n-1) \times (n-1)$. Par récurrence descendante, on se ramène ainsi au calcul de $n!$ déterminants 1×1. Il faut remarquer que le nombre $n!$ croît extrêmement vite avec n. Ainsi, pour une modeste matrice 25×25, on a $25! \approx 1{,}5 \times 10^{25}$.

Exemple 84 On déduit du développement par rapport à la première ligne des expressions explicites pour les déterminants 2×2 et 3×3.

Il faut d'abord remarquer qu'un déterminant 1×1 est de la forme $\det\big(a\big) = a$. C'est en effet visiblement la seule forme multilinéaire alternée qui vaut 1 sur la matrice $\big(1\big)$.

Considérons maintenant un déterminant 2×2.

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

En effet, $A_{11} = \big(a_{22}\big)$ et $A_{12} = \big(a_{21}\big)$, d'où $C_{11} = a_{22}$ et $C_{12} = -a_{21}$. Cette formule de développement de déterminant est la seule formule explicite à connaître par cœur.

Le cas des déterminants 3×3 est déjà beaucoup plus compliqué.

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

Cette expression, qu'il est inutile de chercher à retenir, contient $6 = 3!$ produits de trois coefficients de A (un par colonne) affectés des signes $+$ ou $-$ suivant la nature d'une certaine permutation associée au produit en question. Pour un déterminant 4×4, on aurait $24 = 4!$ produits de quatre coefficients de A, et pour un déterminant $n \times n$, on aurait $n!$ produits de n coefficients de A affectés de signes $+$ ou $-$.

Mentionnons la *règle de Sarrus*, une astuce mnémotechnique qui permet de retrouver les déterminants 3×3 (et seulement ceux-là, cette règle ne se généralise pas à d'autres dimensions). On écrit la matrice en tableau et on lui ajoute en bas ses deux premières lignes. On obtient ainsi un tableau 5×3

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

Les produits de trois termes affectés du signe $+$ apparaissent dans les trois diagonales descendantes du haut à gauche vers le bas à droite, tandis que les produits de trois termes affectés du signe $-$ apparaissent dans les trois diagonales montantes du bas à gauche vers le haut à droite.

En résumé, le développement par rapport à une ligne ou une colonne n'est utile pour calculer explicitement un déterminant que si la matrice dont on part a des propriétés particulières, par exemple beaucoup de zéros, ou s'y ramène par des opérations qui ne modifient pas le déterminant. □

Terminons cette section par deux avertissements. D'une part

$$\boxed{\det(A+B) \neq \det A + \det B.}$$

L'exemple suivant

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

le montre amplement. Le déterminant *n'est pas* linéaire. En fait, il n'y a pas de formule simple pour exprimer le déterminant d'une somme de matrices (il y a des formules relativement compliquées). D'autre part

$$\boxed{\det(\lambda A) \neq \lambda \det A.}$$

Ici il y a une formule simple. En effet

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

puisque λI_n est diagonale et son déterminant est le produit de ses termes diagonaux, soit ici λ^n.

6.1.5 Interprétation géométrique des déterminants

On a une interprétation géométrique de $\mathbb{R}^n$ pour $n = 1,2,3$. On va voir qu'en dimension 2, les déterminants sont liés aux questions de surface et en dimension 3 aux questions de volume.

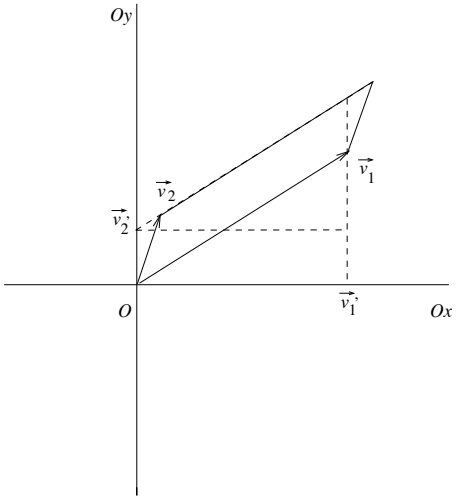
En dimension 2, deux vecteurs v_1, v_2 déterminent un parallélogramme, alors qu'en dimension 3, trois vecteurs v_1, v_2, v_3 déterminent un parallélépipède.

On prendra comme unité de surface la surface du carré unité dont les côtés sont les vecteurs de la base canonique, et comme unité de volume, le volume du cube unité construit de la même façon en dimension 3.

Proposition 66 *La surface du parallélogramme est donnée par* $|\det\begin{pmatrix} v_1 & v_2 \end{pmatrix}|$. *Le volume du parallélépipède est donné par* $|\det\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}|$.

Démonstration. Traitons le cas $n = 2$. Le résultat est vrai si $\begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$. En effet, dans ce cas on a affaire à un rectangle de côtés $|a|$ et $|d|$, donc de surface $|ad|$, alors que le déterminant de la matrice vaut ad.

Supposons que $\{v_1, v_2\}$ est une famille libre. Notons $\begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Si $a_{11} \neq 0$, alors $v_2' = v_2 - \frac{a_{12}}{a_{11}} v_1$ est un multiple de e_2, c'est-à-dire que sa première composante est nulle. L'opération ne change ni le déterminant, ni la surface du parallélogramme. Comme la famille de départ était libre, $v_2' \neq 0$ et ce vecteur a une deuxième composante a_{22}' non nulle. On pose alors $v_1' = v_1 - \frac{a_{21}}{a_{22}'} v_2'$, ce qui produit un vecteur multiple de e_1. L'opération ne change ni le déterminant ni la surface des parallélogrammes. On est donc ramené au premier cas d'un rectangle aux côtés parallèle aux axes, pour lequel le résultat est déjà acquis.



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

Si $a_{11} = 0$, alors $a_{12} \neq 0$ puisque la famille est libre, et on échange les rôles de v_1 et v_2.

Enfin, si la famille est liée, alors le déterminant vaut 0. Dans ce cas, le parallélogramme est réduit à un segment et est donc de surface nulle.

Le cas tridimensionnel se traite de façon analogue. □

6.1.6 Déterminant d'un endomorphisme

Théorème 51 (définition du déterminant d'un endomorphisme) *Soit E un K espace vectoriel de dimension finie et f un endomorphisme de E. Toutes les matrices associées à f par rapport à des bases différentes ont le même déterminant appelé déterminant de f et noté* $\det(f)$.

Démonstration. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. Soient $[f]_{\mathcal{B}}$ et $[f]'_{\mathcal{B}}$ les matrices de f respectivement par rapport à la base $\mathcal{B}$ et à la base $\mathcal{B}'$. Soit $P_{\mathcal{B},\mathcal{B}'}$ la matrice de passage de $\mathcal{B}$ à la base $\mathcal{B}'$. On a

$$[f]_{\mathcal{B}'} = P^{-1}_{\mathcal{B},\mathcal{B}'}[f]_{\mathcal{B}}P_{\mathcal{B},\mathcal{B}'}.$$

Les matrices $[f]_{\mathcal{B}}$ et $[f]_{\mathcal{B}'}$ étant semblables, elles ont le même déterminant.

Théorème 52 (Propriétés du déterminant d'un endomorphisme) *Soit E un K espace vectoriel de dimension finie.*

1. $\det(Id_E) = 1$.

2. Soient f et g deux endomorphismes de E. On a $\det(f \circ g) = \det(f)\det(g)$.

3. Un endomorphisme f de E est inversible si et seulement si $\det(f) \neq 0$.

Démonstration. 1. Soit $\mathcal{B}$ une base de E. On a

$$\det(Id_E) = \det[Id_E]_{\mathcal{B}} = \det I_n = 1.$$

2. $f \circ g$ a pour matrice $[f \circ g]_{\mathcal{B}} = [f]_{\mathcal{B}}[g]_{\mathcal{B}}$. On a donc :

$$\det(f \circ g) = \det[f \circ g]_{\mathcal{B}} = \det([f]_{\mathcal{B}}[g]_{\mathcal{B}}) = \det([f]_{\mathcal{B}})\det([g]_{\mathcal{B}}) = \det(f)\det(g).$$

3. On a les équivalences suivantes :

$$\begin{aligned} f \text{ est inversible} &\iff [f]_{\mathcal{B}} \text{ est inversible} \\ &\iff \det([f]_{\mathcal{B}}) \neq 0 \\ &\iff \det(f) \neq 0. \end{aligned}$$

6.2 Applications des déterminants

6.2.1 Expression de l'inverse d'une matrice à l'aide du déterminant

Définition 45 *On introduit la* matrice des cofacteurs de A, $\mathrm{cof}A = (C_{ij})$ *avec* $C_{ij} = (-1)^{i+j}\det A_{ij}$.

C'est aussi une matrice $n \times n$. Noter la disposition en échiquier des signes $(-1)^{i+j}$, commençant par un $+$ au coin supérieur gauche

$$\begin{pmatrix} + & - & + & \cdots & (-1)^{1+n} \\ - & + & - & \cdots & (-1)^{2+n} \\ + & - & + & \cdots & (-1)^{3+n} \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}.$$

Théorème 53 *On a pour toute matrice A*

$$A(\mathrm{cof}A)^T = (\mathrm{cof}A)^T A = (\det A)I_n.$$

En particulier, si A est inversible, alors

$$A^{-1} = \frac{1}{\det A}(\mathrm{cof}A)^T.$$

Démonstration. Posons $B = A(\mathrm{cof}A)^T$. Par la formule générale du produit matriciel,

$$(B)_{ij} = \sum_{k=1}^{n} a_{ik}(\mathrm{cof}A)^T_{kj} = \sum_{k=1}^{n} a_{ik}C_{jk}.$$

Si $i = j$, on reconnaît le développement du déterminant de A par rapport à la ligne i, donc

$$\forall i, \quad (B)_{ii} = \det A.$$

Si $i \neq j$, on reconnaît le développement par rapport à la ligne j du déterminant de la matrice $\tilde{A}$ dans laquelle la ligne j a été remplacée par la ligne i (ce qui ne change pas les cofacteurs considérés). Donc

$$\forall i \neq j, \quad (B)_{ij} = \det\tilde{A} = 0.$$

On procède de même pour $(\mathrm{cof}A)^T A$ avec les développements par rapport aux colonnes.

Dans le cas où A est inversible, alors $\det A \neq 0$ et il suffit de diviser par $\det A$ pour obtenir la formule pour A^{-1}. □

Remarque 64 Sauf pour $n = 2$, ou pour des matrices très particulières, *ce n'est pas la bonne façon de calculer explicitement l'inverse d'une matrice.* Si on a vraiment besoin de l'inverse d'une matrice, alors on a tout intérêt à utiliser la méthode de Gauss dès que $n \geq 3$. Par contre, la formule précédente est intéressante pour la théorie. □

6.2.2 Application des déterminants à l'indépendance linéaire de vecteurs

Soit E un K-espace vectoriel de dimension n. Soit $\mathcal{B}$ une base de E. Soient $v_1, \ldots, v_n$ n vecteurs de E. Soit M la matrice de $M_n(K)$ dont la j-ième colonne est formée des coordonnées du vecteur v_j par rapport à la base $\mathcal{B}$. On appelle déterminant des vecteurs $v_1, \ldots, v_n$ et on note $\det_{\mathcal{B}}(v_1, \ldots, v_n)$ le déterminant de la matrice M.

Remarque 65 Compte tenu des propriétés du déterminant d'une matrice, on a les propriétés suivantes :

• L'application $(v_1, \ldots, v_n) \mapsto \det_{\mathcal{B}}(v_1, \ldots, v_n)$ est linéaire par rapport à chaque variable.

• Si deux des vecteurs de la famille sont égaux, $\det_{\mathcal{B}}(v_1, \ldots, v_n)$=0.

Théorème 54 *Soit E un K espace vectoriel de dimension n. Soient n vecteurs de E et $\mathcal{B}$ une base de E. Ces vecteurs sont linéairement indépendants si et seulement si* $\det_{\mathcal{B}}(v_1, \ldots, v_n) \neq 0$.

Démonstration. les vecteurs $v_1,\dots,v_n$ sont linéairement indépendants si et seulement si la famille $(v_1,\dots,v_n)$ est de rang n. Soit M la matrice de $M_n(K)$ dont la j-ième colonne est formée des coordonnées du vecteur v_j par rapport à la base $\mathcal{B}$. Le rang de la famille $(v_1,\dots,v_n)$ est égal le rang de la matrice M et M est de rang n si et seulement si elle est inversible.

Définition 46 *Soit n et q deux entiers naturels. Soit $A=(a_{i,j})$ une matrice à n lignes et q colonnes à coefficients dans K. Soit p un entier inférieur à n et à q. On appelle mineur d'ordre p le déterminant d'une matrice carrée d'ordre p obtenue à partir de A en supprimant $n-p$ lignes et $q-p$ colonnes.*

Théorème 55 (Caractérisation de l'indépendance linéaire de p vecteurs) *Soit E un K -espace vectoriel de dimension n et $(e_1,\dots,e_n)$ une base de E. Soit p un entier inférieur ou égal à n et $v_1,\dots,v_p$ p vecteurs de E. Pour tout j compris entre 1 et n, on pose $v_j=\sum_{i=1}^{n} a_{i,j}e_i$. Alors les p vecteurs $(v_1,\dots,v_p)$ sont linéairement indépendants si et seulement si il existe un mineur d'ordre p non nul extrait de la matrice $(a_{i,j})$ de $M_{n,p}(K)$.*

Démonstration. Supposons $v_1,\dots,v_p$ linéairement indépendants.

Si $p=n$, le résultat est une conséquence immédiate de la propriété précédente.

Si $p<n$, on peut appliquer le théorème de la base incomplète à la partie libre $\{v_1,\dots,v_p\}$. Il existe donc des vecteurs $e_{k_{p+1}},\dots,e_{k_n}$ tels que $(v_1,\dots,v_p,e_{k_{p+1}},\dots,e_{k_n})$ soit une base de E. Soit la base de E obtenue en renumérotant les éléments de la base $(e_1,\dots,e_n)$ de la manière suivante. Si $\{1,\dots,n\}=\{k_1,\dots,k_p\}\cup\{k_{p+1},\dots,k_n\}$. On pose

$$\forall i\in[1,n],\ \ \varepsilon_i=e_{k_i}.$$

Alors, la matrice carrée d'ordre n des composantes des vecteurs $(v_1,\dots,v_p,e_{k_{p+1}},\dots,e_{k_n})$ par rapport à la base $(\varepsilon_1,\dots,\varepsilon_n)$ est

$$N=\begin{pmatrix} a_{k_1,1} & \dots & a_{k_p,p} & 0 & \dots & 0\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ a_{k_p,1} & \dots & a_{k_p,p} & 0 & \dots & 0\\ a_{k_{p+1},1} & \dots & a_{k_{p+1},p} & 1 & \dots & 0\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ a_{k_n,1} & \dots & a_{k_n,p} & 0 & \dots & 1 \end{pmatrix}.$$

Son déterminant est non nul puisque les vecteurs $(v_1,\dots,v_p,e_{k_{p+1}},\dots,e_{k_n})$ forment une base de E. On calcule le déterminant de N en développant par rapport à la dernière colonne autant de fois que nécessaire. On voit que

$$\det(N)=\begin{vmatrix} a_{k_1,1} & \dots & a_{k_p,p}\\ \vdots & \ddots & \vdots\\ a_{k_p,1} & \dots & a_{k_p,p} \end{vmatrix}.$$

Le mineur $\begin{vmatrix} a_{k_1,1} & \dots & a_{k_p,p}\\ \vdots & \ddots & \vdots\\ a_{k_p,1} & \dots & a_{k_p,p} \end{vmatrix}$ est donc non nul.

Montrons maintenant la réciproque. On pose

$$N = \begin{pmatrix} a_{k_1,1} & \dots & a_{k_p,p} \\ \vdots & \ddots & \vdots \\ a_{k_p,1} & \dots & a_{k_p,p} \end{pmatrix}$$

et on suppose $\det(N) \neq 0$. On considère la relation

$$\lambda_1 v_1 + \cdots + \lambda_p v_p = 0$$

En exprimant chaque v_i dans la base $(e_1, \dots, e_n)$, on voit aisément que cette relation équivaut au système suivant

$$\begin{cases} a_{1,1}\lambda_1 & + & \dots & + & a_{1,p}\lambda_p & = & 0 \\ a_{2,1}\lambda_1 & + & \dots & + & a_{2,p}\lambda_p & = & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1}\lambda_1 & + & \dots & + & a_{n,p}\lambda_p & = & 0 \end{cases}$$

Ce qui implique

$$\begin{cases} a_{k_1,1}\lambda_1 & + & \dots & + & a_{k_1,p}\lambda_p & = & 0 \\ a_{k_2,1}\lambda_1 & + & \dots & + & a_{k_2,p}\lambda_p & = & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k_1,1}\lambda_1 & + & \dots & + & a_{k_p,p}\lambda_p & = & 0 \end{cases}$$

On a donc

$$C \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = 0.$$

Comme C est inversible, cela implique $\lambda_1 = \cdots = \lambda_p = 0$. Ce qu'il fallait démontrer.

6.2.3 Application à la détermination du rang d'une matrice

Rappelons la définition du rang d'une matrice.

Définition 47 *Le rang d'une matrice est le nombre maximum de vecteurs colonnes linéairement indépendants.*

Notation : le rang de la matrice A est noté $rg(A)$.

Rappelons que l'on ne change pas le rang d'une matrice $A = (C_1, \dots, C_n)$ par les opérations élémentaires suivantes sur les colonnes :
- Permutation de deux colonnes.
- On multiplie une colonne par un scalaire non nul.
- On ajoute à une colonne un multiple d'une autre colonne.

Théorème 56 *Le rang d'une matrice est le plus grand entier r tel qu'il existe une matrice carrée d'ordre r extraite de M de déterminant non nul.*

Démonstration. Cela découle de la caractérisation de l'indépendance de r colonnes à l'aide des mineurs.

Exemple 85 *Soit α un paramètre réel. Calculons le rang de la matrice A à coefficients réels suivante :*

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 1 & \alpha & 1 \end{pmatrix}$$

On obtient les mineurs d'ordre 3 de A en supprimant une colonnes. Calculons le mineur d'ordre 3 obtenu en supprimant la première colonne $\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & \alpha & 1 \end{vmatrix}$. On a

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

Si α est différent de 2, le rang de la matrice A est 3. Si $\alpha = 2$, on vérifie que les 4 mineurs d'ordre 3 de A sont nuls. Donc A est de rang inférieur ou égal à 2. Or $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}$ est un mineur d'ordre 2 non nul. Donc si α est égal à 2, le rang de A est 2.

Proposition 67 *Soit A une matrice à n lignes et p colonnes. Le rang de A est égal au rang de sa transposée.*

Démonstration. Cette démonstration est hors programme.

Les mineurs de A^T sont obtenus à partir des mineurs de A par transposition. Comme les déterminants d'une matrice et de sa transposée sont égaux, la proposition découle de la caractérisation du rang d'une matrice à l'aide des mineurs. □

Corollaire 57 *On ne change pas le rang d'une matrice par les opérations élémentaires suivantes sur les lignes :*

- *Permutation de deux lignes.*
- *On multiplie une ligne par un scalaire non nul.*
- *On ajoute à une ligne un multiple d'une autre ligne.*

Démonstration. Faire une opération élémentaire sur les lignes de A, c'est faire une opération élémentaire sur les colonnes de A^T. Or, nous savons qu'une opération élémentaire sur les colonnes de A^T ne change pas le rang de A^T. Comme le rang de A est égal au rang de A^T, le corollaire en découle.

Remarque 66 *L'algorithme de Gauss permet de transformer, par une succession d' opérations sur les lignes une matrice A en une matrice échelonnée R dont le rang est facile à calculer. D'après le corollaire précédent, on a $rg(A) = rg(R)$.*

Exemple 86 *Calculons le rang de la matrice*

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

Nous avons vu dans l'exemple 2 du chapitre 1 que, par des opérations élémentaires sur les lignes, la matrice A se transformait en la matrice

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

dont le rang est 3. Le rang de la matrice A est donc 3.

Chapitre 7

Diagonalisation

De nombreux problèmes, dans des domaines variés et pas seulement mathématiques, nécessitent pour leur résolution de savoir calculer des puissances de matrices. Dans le chapitre 1, nous avons vu comment calculer les puissances d'une matrice semblable à une matrice diagonale. Dans ce chapitre, nous nous intéresserons au problème suivant : Si E est un espace vectoriel de type fini, de dimension supérieure ou égale à 1 et f un endomorphisme de E, il s'agit de déterminer s'il existe une base de E telle que la matrice de f par rapport à cette base soit diagonale

Dans ce chapitre, nous supposerons que K est $\mathbb{R}$ ou $\mathbb{C}$. Tant que cela sera possible, nous traiterons simultanément le cas d'un espace vectoriel réel et celui d'un espace vectoriel complexe. Cependant la nature du corps de base joue un rôle important dans cette théorie.

7.1 Endomorphisme diagonalisable, valeur propre, vecteur propre

Définition 48 (Endomorphisme diagonalisable) *Soit E un K-espace vectoriel de dimension finie et f un endomorphisme de E. On dit que f est diagonalisable s'il existe une base de E telle que la matrice de f par rapport à cette base soit diagonale.*

Plus précisément, si n est la dimension de E, un endomorphisme f de E est diagonalisable si et seulement s'il existe une base $(e_1,\dots,e_n)$ de E et des éléments $\lambda_1,\dots,\lambda_n$ de K, tels que la matrice associée à f dans la base $(e_1,e_2,\dots,e_n)$ soit la matrice diagonale $\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$

Remarque sur la notation : les scalaires $\lambda_1,\dots,\lambda_n$ ne sont pas nécessairement distincts.

Compte tenu de la définition de la matrice d'un endomorphisme par rapport à une base cela signifie que :

$$\forall i \in [1,n],\ \ f(e_i) = \lambda_i e_i.$$

Les scalaires λ et les vecteurs v, liés par une relation de la forme $f(v) = \lambda v$,

jouent donc manifestement un rôle important dans cette théorie. Cela nous conduit à la définition des notions de valeur propre et de vecteur propre.

Définition 49 (Définition d'un vecteur propre) *Soit E un K-espace vectoriel de dimension finie et f un endomorphisme de E.*

Un vecteur v de E est appelé vecteur propre de f s'il vérifie les deux conditions :

- *v est non nul,*
- *il existe un élément λ du corps des scalaires K tel que $f(v) = \lambda v$.*

Définition 50 (Définition d'une valeur propre) *Soit E un espace vectoriel de type fini sur K et f un endomorphisme de E. Un élément λ du corps des scalaires K est appelé valeur propre de f s'il existe un vecteur v, non nul, tel que $f(v) = \lambda v$.*

Remarque 67 Attention : un vecteur propre est non nul.

Vocabulaire : Soit v un vecteur non nul et λ un élément de K tels que $f(v) = \lambda v$. On dit alors que v est un vecteur propre associé à la valeur propre λ. Les deux notions de valeur propre et de vecteur propre sont donc étroitement liées.

Exemple 87 Soit E un espace vectoriel réel de dimension 2 et (e_1, e_2) une base de E. On considère l'endomorphisme f de E défini par $f(e_1) = e_2$, $f(e_2) = e_2$. Il est immédiat que 1 est une valeur propre puisqu'il existe un vecteur non nul, à savoir e_2, tel que $f(e_2) = e_2$. Le vecteur e_2 est un vecteur propre associée à la valeur propre 1. Dans cet exemple, il y a une valeur propre visible mais l'existence d'autres valeurs propres n'a pas été étudiée.

Exemple 88 Soit f un endomorphisme non injectif d'un K-espace vectoriel E. Cela signifie que son noyau n'est pas réduit au vecteur nul, autrement dit qu'il existe un vecteur v non nul tel que $f(v) = 0$. Ceci équivaut à dire que le scalaire 0_K est une valeur propre pour f.

7.1.1 Caractérisation des valeurs propres

Soit f un endomorphisme d'un K-espace vectoriel E de dimension finie égale à n ($n \geq 1$).

Un scalaire λ est une valeur propre de f si et seulement si :

$$\begin{array}{l}
\exists v \in E - \{0\},\ f(v) = \lambda v \Longleftrightarrow \\
\exists v \in E - \{0\},\ f(v) - \lambda v = 0 \Longleftrightarrow \\
\exists v \in E - \{0\},\ (f - \lambda Id_E)\, v = \Longleftrightarrow \\
\mathrm{Ker(f - \lambda Id_E)} \neq \{0\} \Longleftrightarrow \\
\mathrm{rg}(f - \lambda Id_E) < n \Longleftrightarrow \\
\det(f - \lambda Id_E) = 0
\end{array}$$

Nous avons donc démontré la propriété suivante :

Proposition 68 (Caractérisation d'une valeur propre) *Un élément du corps de base K de l'espace vectoriel est une valeur propre de f si et seulement si* $\det(f - \lambda Id_E) = 0$.

Cette propriété donne donc un procédé pratique pour déterminer les valeurs propres d'un endomorphisme.

Exemple 89 Soit f l'endomorphisme de $\mathbb{R}^2$ défini par $f(e_1) = 2e_1 + e_2$, $f(e_2) = e_1 + 2e_2$ où (e_1, e_2) désigne la base canonique de $\mathbb{R}^2$.

Pour déterminer ses valeurs propres il faut, d'après la caractérisation précédente, chercher les éléments λ de $\mathbb{R}$, tels que $\det(f - \lambda Id_{\mathbb{R}^2}) = 0$. Pour cela il est naturel d'écrire la matrice A associée à f dans la base canonique et de calculer $\det(A - xI_2)$ qui est égal à $\det(f - xId_{\mathbb{R}^2})$.

On a $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ et par conséquent

$$\det(A - xI_2) = \begin{vmatrix} 2-x & 1 \\ 1 & 2-x \end{vmatrix} = (x-1)(x-3)$$

Les réels 1 et 3 sont donc les valeurs propres de f.

Exemple 90 Soit g l'endomorphisme de $\mathbb{R}^2$ défini par $g(e_1) = e_1 + e_2$, $g(e_2) = -e_1 + e_2$ où (e_1, e_2) désigne la base canonique de $\mathbb{R}^2$. De même que précédemment, on écrit la matrice B associée à g dans la base canonique et on calcule $\det(B - xI_2)$. On a $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ et par conséquent $\det(B - xI_2) = \begin{vmatrix} 1-x & -1 \\ 1 & 1-x \end{vmatrix} = x^2 - 2x + 2$. Or il n'y a pas de réels x tels que $x^2 - 2x + 2$ soit nul (le discriminant du trinôme est strictement négatif). Donc l'endomorphisme g n'admet pas de valeurs propre.

Exemple 91 Soit h l'endomorphisme de $\mathbb{R}^3$ défini par $h(e_1) = e_1$, $h(e_2) = e_3$ et $h(e_3) = -e_2$ où (e_1, e_2, e_3) désigne la base canonique de $\mathbb{R}^3$.

De même que précédemment, on écrit la matrice C associée à h dans la base canonique et on calcule $\det(C - xI_3)$.

On a $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ et par conséquent

$$\det(C - xI_3) = \begin{vmatrix} 1-x & 0 & 0 \\ 0 & -x & -1 \\ 0 & 1 & -x \end{vmatrix} = -(x-1)(x^2+1)$$

La seule valeur réelle de x annulant $\det(C - xI_3)$ est $x = 1$. Donc h a une seule valeur propre qui est 1.

7.1.2 Fonction polynôme caractéristique

Soit x un élément du corps de base K. On a vu apparaitre naturellement l'expression $\det(f - xId_E)$. Pour calculer le déterminant de l'endomorphisme de E,$f - xId_E$, il est nécessaire d'introduire la matrice associée à f par rapport à une base de E. Soit donc $\mathcal{B}$ une base de E et A la matrice associée à f par rapport à cette base. Alors la matrice associée à $f - xId_E$ est $A - xId_E$ et par conséquent $\det(f - xId_E) = \det(A - xI_n)$. Si $A = (a_{i,j})_{(i,j)\in[1,n]}$, on a :

$$\det(f - xId_E) = \det(A - xI_n) = \begin{vmatrix} a_{1,1} - x & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - x & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n,1} & \dots & \dots & a_{n,n} - x \end{vmatrix}$$

L'expression explicite de ce déterminant prouve que c'est une expression polynômiale en x, de degré n, dont le coefficient du terme de plus haut degré est égal à $(-1)^n$. Cette remarque nous amène à la définition suivante :

Définition 51 (Fonction polynôme caractéristique) *Soit f un endomorphisme d'un K-espace vectoriel E de dimension n, entier supérieur ou égal à 1 (K est égal à $\mathbb{R}$ ou $\mathbb{C}$). On appelle fonction polynôme caractéristique et on note $P_{car,f}$ la fonction polynômiale définie par*

$$\forall x \in K,\ P_{car,f}(x) = \det(f - xId_E).$$

Remarque 68 Plus précisément, on définit le polynôme caractéristique de f. De façon un peu informelle, un *polynôme à une indéterminée* à coefficients dans K est une expression de la forme

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

avec $a_i \in K$. Si $a_n \neq 0$, alors l'entier n est le *degré* de P. Par convention, le degré du polynôme nul est $-\infty$. La lettre X désigne l'*indéterminée*. On peut lui donner un sens mathématique précis, L'ensemble de tous les polynômes à une indéterminée à coefficients dans K est noté $K[X]$. En fait, un polynôme à coefficients dans K définit une application polynomiale de K dans K par

$$x \mapsto P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n.$$

Si $K = \mathbb{R}$ ou $\mathbb{C}$, on peut sans danger confondre le polynôme et la fonction polynomiale associée.

Si K est un corps commutatif quelconque, on ne peut plus confondre un polynôme et la fonction polynomiale qui lui est associée. Soit f un endomorphisme d'un K-espace vectoriel de dimension finie E. Si A est la matrice de f dans une base $\mathcal{B}$, on définit le polynôme caractéristique $P_{car,f}$ de f comme étant le déterminant à coefficients dans $K[X]$

$$\det(f - XId_E) = \det(A - XI_n) = \begin{vmatrix} a_{1,1} - X & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n,1} & \dots & \dots & a_{n,n} - X \end{vmatrix}.$$

Il faut s'assurer au préalable que si A' est la matrice de f dans une autre base $\mathcal{B}'$, on a $\det(A - XId_E) = \det(A' - XId_E)$.

Définition 52 *On dit qu'une fonction polynôme de K est scindée si elle est produit de fonctions polynômes de K de degré 1.*

Exemple 92 la fonction polynomiale $x^2 + 1$ n'est pas scindée sur $\mathbb{R}$ mais elle l'est sur $\mathbb{C}$ puisque l'on peut écrire $x^2 + 1 = (x + i)(x - i)$.

Proposition 69 (Condition nécessaire de diagonalisabilité) *Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. Si f est diagonalisable, alors sa fonction polynôme caractéristique est scindée.*

Démonstration. Si f est diagonalisable, il existe une base de vecteurs propres pour f, $(v_1,\dots,v_n)$. Si v_i est vecteur propre pour la valeur propre λ_i, la matrice de f dans la base $(v_1,\dots,v_n)$ est

$$D=\begin{pmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & 0 & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \lambda_n\end{pmatrix}.$$

On a $P_{car,f}(x)=\det(D-xI_n)=(\lambda_1-x)(\lambda_2-x)\dots(\lambda_n-x)$.

7.1.3 Caractérisation des valeurs propres d'un endomorphisme à l'aide du polynôme caractéristique

Si P est une fonction polynomiale à coefficients dans K et si λ est un élément de K, on dit que λ est une racine de P si $P(\lambda)=0$. Le théorème suivant est une conséquence immédiate de ce qui a été vu précédemment.

Théorème 58 (Valeurs propres et polynôme caractéristique) *Soit f un endomorphisme d'un K- espace vectoriel E de type fini. Un élément λ de K est valeur propre de f si et seulement si il est racine de la fonction polynôme caractéristique de f.*

Remarque 69 L'existence et le nombre de valeurs propres d'un endomorphisme dépendent essentiellement du corps de base de l'espace vectoriel. Si l'on considère, par exemple, l'endomorphisme h de $\mathbb{R}^3$ défini par $h(e_1)=e_1$, $h(e_2)=e_3$ et $h(e_3)=-e_2$ où $(e_1,\dots,e_3)$ désigne la base canonique de $\mathbb{R}^3$. Sa fonction polynôme caractéristique est $P_{car,h}(x)=(1-x)(x^2+1)$. Elle n'a qu'une seule racine réelle, qui est donc la seule valeur propre de h. Mais si on considère l'endomorphisme h' de $\mathbb{C}^3$ défini par $h'(e_1)=e_1$, $h'(e_2)=e_3$ et $h'(e_3)=-e_2$ où (e_1,e_2,e_3) où (e_1,e_2,e_3) désigne la base canonique de $\mathbb{C}^3$. Sa fonction polynôme caractéristique est $P_{car,h'}(x)=(1-x)(x^2+1)$. C'est une fonction de $\mathbb{C}$ dans $\mathbb{C}$ et elle a trois racines qui sont $1,i$ et $-i$.

Un endomorphisme d'un espace vectoriel complexe (c'est-à-dire dont le corps de base est $\mathbb{C}$) admet toujours des valeurs propres (puisque une fonction polynôme à coefficients dans $\mathbb{C}$ a toujours des racines d'après le théorème de D'Alembert-Gauss) alors qu'un endomorphisme d'un espace vectoriel réel peut ne pas avoir de valeurs propres (l'endomorphisme g des exemples précédents n'a pas de valeur propre).

7.1.4 Sous-espace propre associé à une valeur propre

Une fois déterminées les valeurs propres d'un endomorphisme, s'il y en a, on peut rechercher les vecteurs propres associés. Cela revient à résoudre l'équation linéaire $f(v)=\lambda v$, c'est-à-dire à déterminer $\mathrm{Ker}(\mathrm{f}-\lambda\mathrm{Id_E})$.

Définition 53 (Sous-espace propre associé à une valeur propre) *Soit f un endomorphisme d'un K-espace vectoriel E de type fini et λ une valeur propre de f. On appelle sous-espace propre associé à la valeur propre λ le noyau de $f-\lambda Id_E$*

Notation : Le sous espace propre associé à la valeur propre λ sera noté E_λ.

Il résulte donc de la définition que le sous-espace propre associé à une valeur propre λ est un sous-espace vectoriel dont les éléments sont le vecteur nul et les vecteurs propres associés à λ.

Compte tenu de cette définition on a les équivalences :

$$\lambda \text{ valeur propre} \Longleftrightarrow E_\lambda \neq \{0_E\} \Longleftrightarrow \dim E_\lambda \geq 1.$$

7.2 Version matricielle

7.2.1 Notion de matrice diagonalisable, de valeur propre d'une matrice, de vecteur propre d'une matrice

Définition 54 (Définitions et premières propriétés) *Soit M une matrice carrée d'ordre n à coefficients dans K.*

(1) On dit que M est diagonalisable si elle est semblable à une matrice diagonale c'est-à-dire s'il existe deux matrices D et P de $M_n(K)$ telles que D soit diagonale, P inversible et $M = PDP^{-1}$.

(2) Une matrice colonne V appartenant à $M_{n,1}(K)$ est un vecteur propre de M si :

$$V \neq 0 \text{ et } \exists \lambda \in K, MV = \lambda V.$$

(3) Un élément λ de K est une valeur propre de M s'il existe V, non nul, appartenant à $M_{n,1}$ tel que $MV = \lambda V$.

(4) Un élément λ de K est une valeur propre de M si et seulement si $\det(M - \lambda I_n) = 0$.

(5) On appelle fonction polynôme caractéristique de M la fonction polynôme $x \mapsto \det(M - xI_n)$. On la note $P_{car,M}$.

(6) Un élément λ de K est une valeur propre de M si et seulement c'est une racine de la fonction polynôme caractéristique de M.

(7) Le sous-espace propre associé à la valeur propre λ est égal à l'ensemble des V appartenant à $M_{n,1}(K)$ tels que $(M - \lambda I_n)V = 0$, autrement dit l'ensemble des matrices colonnes $V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ telles que $(M - \lambda I_n)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

A la matrice M, on associe l'endomorphisme L_M de $M_{n,1}(K)$ défini par

$$\forall V \in M_{n,1}(K),\ L_M(V) = MV.$$

La matrice de L_M dans la base canonique $\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\right)$ est M. Ainsi : un élément V de $M_{n,1}(K)$ est un vecteur propre de M si et seulement si c'est un vecteur propre de L_M, un scalaire λ est valeur propre de M si et seulement si c'est une valeur propre de L_M, le polynôme caractéristique de M est le polynôme caractéristique de L_M et le sous espace propre de M associé à la valeur propre λ est le sous espace propre de L_M associé à la valeur propre λ.

La proposition suivante résulte immédiatement de la définition de la fonction polynôme caractéristique d'une matrice et des propriétés des déterminants.

Proposition 70 (polynôme caractéristique et matrices semblables) *Soient M et N deux matrices semblables de $M_n(K)$ alors $P_{car,M} = P_{car,N}$.*

Démonstration. Soient M et N deux matrices semblables. Il existe donc une matrice inversible Q de $M_n(K)$ telle que $M = QNQ^{-1}$. Alors $M - xI_n = Q(N - xI_n)Q^{-1}$. Les matrices $M - xI_n$ et $N - xI_n$ sont donc semblables et par conséquent ont le même déterminant. □

7.2.2 Relation entre endomorphisme diagonalisable et matrice diagonalisable

En partant d'un endomorphisme...

Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. Soit B une base de E et M la matrice associée à f dans B. Dire que f est diagonalisable équivaut à dire qu'il existe une base B' de E dans laquelle la matrice de f est une matrice diagonale D. On sait que deux matrices sont associées à un même endomorphisme par rapport à des bases différentes si et seulement si elles sont semblables. Donc f est diagonalisable si et seulement si M est semblable à une matrice diagonale, c'est-à-dire si et seulement si M est diagonalisable.

En partant d'une matrice carrée...

Soit M une matrice carrée appartenant à $M_n(K)$, on peut considérer l'unique endomorphisme L_M de $M_{n,1}(K)$ qui lui est associé dans la base canonique. On a

$$\begin{array}{rcl} L_M : M_{n,1}(K) & \to & M_{n,1}(K) \\ X & \mapsto & MX. \end{array}$$

Alors d'après les définitions et les propriétés rappelées ci-dessus, M est diagonalisable si et seulement si L_M est diagonalisable.

Soit E un K-espace vectoriel de dimension finie. Il est immédiat que les définitions sont telles que si M est la matrice associée à un endomorphisme f de E par rapport à une base B, alors :

• les valeurs propres de l'un sont les valeurs propres de l'autre (puisque $\det(f - xId_E) = \det(M - xI_n)$),

• la matrice colonne V associée à un vecteur propre v de f dans la base B est un vecteur propre de M et réciproquement le vecteur v de E qui admet comme composantes dans la base B les coefficients d'une matrice colonne V vecteur propre de M est un vecteur propre de f (puisque $f(v) = \lambda v \Longleftrightarrow MV = \lambda V$).

Par conséquent, les sous-espaces propres associés à λ valeur propre respectivement de f et M sont isomorphes, par l'isomorphisme qui associe à un vecteur la matrice colonne de ses composantes dans la base B.

Il y a donc coïncidence parfaite entre ces notions relatives à une matrice ou à un endomorphisme. Cela justifie la similitude de vocabulaire.

Tous les théorèmes qui suivent, dont la finalité est de trouver des conditions pour qu'un endomorphisme ou une matrice soit diagonalisable, sont donc communs. Les démonstrations théoriques sont faites la plupart du temps dans le cadre vectoriel, c'est-à-dire pour les endomorphismes car elles s'y expriment plus simplement.

En revanche, pour faire les calculs explicites, on utilise le calcul matriciel.

Attention : une matrice à coefficients dans $\mathbb{R}$ peut aussi être considérée comme une matrice à coefficients dans $\mathbb{C}$. Il faut bien préciser, pour les matrices, le corps dans

lequel on se place. Nous verrons qu'il existe des exemples de matrices à coefficients réels, non diagonalisables dans $\mathbb{R}$ et diagonalisables dans $\mathbb{C}$.

Ce type de problème ne se pose pas pour un endomorphisme car le corps de base de l'espace vectoriel est fixé au départ.

7.3 Propriétés des sous-espaces propres

Théorème 59 (Somme directe de sous-espaces propres) *Soit f un endomorphisme d'un espace vectoriel E de type fini qui admet au moins deux valeurs propres distinctes. Soit p un entier supérieur ou égal à 2, $\lambda_1,\dots,\lambda_p$ des valeurs propres distinctes de f et les sous-espaces propres associés $E_{\lambda_1},E_{\lambda_2},\dots,E_{\lambda_p}$. Alors la somme $E_{\lambda_1}+E_{\lambda_2}+\cdots+E_{\lambda_p}$ est directe, ce qui est noté*

$$E_{\lambda_1}+E_{\lambda_2}+\cdots+E_{\lambda_p}=E_{\lambda_1}\oplus E_{\lambda_2}\oplus\cdots\oplus E_{\lambda_p}.$$

Démonstration. Faisons une démonstration par récurrence. Si $p=2$, soient λ_1,λ_2 deux valeurs propres distinctes de f. On va montrer que $E_{\lambda_1}+E_{\lambda_2}=E_{\lambda_1}\oplus E_{\lambda_2}$. Pour cela,il suffit de prouver que si v_1 et v_2 sont deux vecteurs respectivement de E_{λ_1} et E_{λ_2}, alors $v_1+v_2=0\Longrightarrow v_1=v_2=0$. Or

$$v_1+v_2=0\Longrightarrow f(v_1)+f(v_2)=f(0)=0.$$

Donc les vecteurs v_1 et v_2 vérifient les deux égalités

$$\begin{cases} v_1 & + & v_2 & = & 0\\ \lambda_1 v_1 & + & \lambda_2 v_2 & = & 0\end{cases}$$

Ces égalités impliquent l'égalité $(\lambda_2-\lambda_1)v_2=0$.(C'est la relation obtenue en prenant la deuxième ligne moins λ_1 fois la première). Or, comme d'après l'hypothèse, λ_1 et λ_2 sont deux valeurs propres distinctes, il en résulte immédiatement $v_2=0$. D'où, d'après la première relation, $v_1=0$.

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

Multiplions cette relation par λ_k. Il vient

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

Appliquons également à cette égalité l'endomorphisme f. Il vient

$$f(x_1)+f(x_2)+\cdots+f(x_k)=\lambda_1 x_1+\lambda_2 x_2+\cdots+\lambda_k x_k=0.$$

Soustrayons les deux égalités obtenues membre à membre. On obtient

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

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

Compte tenu du vocabulaire introduit et des résultats préliminaires qui viennent d'être démontrés, la définition d'endomorphisme diagonalisable peut être traduite de la façon suivante :

Théorème 60 *Soit f un endomorphisme d'un K-espace vectoriel E de type fini. On suppose que f (respectivement M) possède des valeurs propres, soit $\lambda_1, \ldots, \lambda_r$. Alors les conditions suivantes sont équivalentes :*

1. L'endomorphisme f est diagonalisable.

2. Il existe une base de E formée de vecteurs propres de f.

3. L'espace vectoriel E est somme directe des sous-espaces propres, c'est-à-dire $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.

La réunion de bases des sous-espaces propres forme alors une base de E constituée de vecteurs propres de f.

Remarque 70 La troisième condition équivaut à : $\dim(\mathrm{E}) = \sum_{\mathrm{i}=1}^{\mathrm{r}} \dim \mathrm{E}_{\lambda_\mathrm{i}}$.

Le point (iii) de ce théorème conduit immédiatement à une propriété utile dans la pratique.

Proposition 71 (Condition suffisante de diagonalisation) *Soit f un endomorphisme d'un K-espace vectoriel E de dimension n (ou M une matrice carrée d'ordre n à coefficients dans K). Si le polynôme caractéristique de f (respectivement de M) admet n racines distinctes, alors f (respectivement M) est diagonalisable.*

Démonstration. En effet, dans ce cas l'endomorphisme f a n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$. Les sous-espaces propres E_{λ_i} en somme directe. Pour chaque espace propre, on a $\dim \mathrm{E}_{\lambda_\mathrm{i}} \geq 1$. Donc

$$n \leq \sum_{i=1}^{n} \dim \mathrm{E}_{\lambda_\mathrm{i}} = \dim\Big(\bigoplus_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}_{\lambda_\mathrm{i}}\Big) \leq \dim \mathrm{E} = \mathrm{n},$$

puisque $\bigoplus_{i=1}^{n} E_{\lambda_i}$ est un sous espace vectoriel de E. On en déduit que

$$\dim\Big(\bigoplus_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}_{\lambda_\mathrm{i}}\Big) = \dim \mathrm{E},$$

et donc que

$$\bigoplus_{i=1}^{n} E_{\lambda_i} = E$$

d'où la diagonalisabilité de f. □

Ce corollaire n'est qu'une *condition suffisante*. De nombreuses matrices avec des valeurs propres multiples sont aussi diagonalisables.

Remarque 71 Si un endomorphisme est diagonalisable, la décomposition de l'espace E en somme directe de sous-espaces propres permet de mieux comprendre cet endomorphisme. En effet, on a déjà noté que sa restriction à un sous-espace propre est l'application linéaire la plus simple qui soit, une homothétie. □

Remarque 72 ATTENTION, même sur $\mathbb{C}$ il existe des matrices qui ne sont pas diagonalisables. Par exemple

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

est triangulaire supérieure. Sa fonction polynôme caractéristique est $P_{car,A}(x) = x^2$. Elle admet donc une seule valeur propre à savoir 0. Un rapide calcul montre que le sous espace propre associé à la valeur propre 0 est de dimension 1. Donc $\mathbb{C}^2$ ne peut pas être somme directe des espaces propres, et A n'est pas diagonalisable

On peut aussi le voir directement. Si A est diagonalisable, alors

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

contradiction. Donc A n'est pas diagonalisable. □

Exemple 93 Considérons la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. On calcule sa fonction polynôme caractéristique

$$P_{car,A}(x) = \begin{vmatrix} 1-x & 1 & 1 \\ 1 & 1-x & 1 \\ 1 & 1 & 1-x \end{vmatrix} = (3-x)x^2.$$

Le sous espace propre E_0 associée à la valeur propre 0 est le noyau. Déterminons-le :

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

$$E_0 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}.$$

Une base de E_0 est $((1,0,-1),(0,1,-1))$.

Déterminons le sous espace propre associé à la valeur propre 3. On a

$$(x_1, x_2, x_3) \in E_3 \Longleftrightarrow \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ce qui équivaut au système

$$\begin{cases} -2x_1 + x_2 + x_3 &= 0 \\ x_1 - 2x_2 + x_3 &= 0 \\ x_1 + x_2 - 2x_3 &= 0 \end{cases}$$

On résoud ce système par la méthode du pivot de Gauss et on obtient une base de E_3 formée du vecteur $(1,1,1)$. On a donc

$$A = P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} P^{-1} \text{ avec } P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Exemple 94 Considérons $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Sa fonction polynôme caractéristique est $x^2 + 1$. Elle n'a donc pas de valeur propre sur $\mathbb{R}$. Elle n'est donc pas diagonalisable sur $\mathbb{R}$. En revanche, elle a deux valeurs propres distinctes sur $\mathbb{C}$, à savoir i et $-i$. Elle est donc diagonalisable sur $\mathbb{C}$.

7.4 Application au calcul des puissances d'une matrice diagonalisable

Soit A une matrice diagonalisable. Il existe une matrice diagonale D et une matrice inversible P telle que $A = PDP^{-1}$. Nous avons vu au chapitre 1 qu'on avait alors $A^k = PD^kP^{-1}$ et que D^k était très facile à calculer. Etudions un exemple.

Exemple 95 Soit $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Sa fonction polynôme caractéristique est égal à $P_{car,A}(x) = (x-1)(x-3)$. Il a deux valeurs propres distinctes à savoir 1 et 3. Comme on a une matrice carrée d'ordre 2, elle est diagonalisable. Tous calculs faits, on a

$$A = P\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}P^{-1} \text{ avec } P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Donc

$$A^k = P\begin{pmatrix} 1 & 0 \\ 0 & 3^k \end{pmatrix}P^{-1} = \frac{1}{2}\begin{pmatrix} 1+3^k & -1+3^k \\ -1+3^k & 1+3^k \end{pmatrix}.$$