Université de Paris X Nanterre
U.F.R. Segmi
Licence Economie-Gestion première année

Année 2007-2008

Cours de Mathématiques II
Chapitre 1. Algèbre linéaire

Table des matières

1 Espaces vectoriels

1.1 Espaces et sous-espaces vectoriels

Définition 1. *Un espace vectoriel sur* $\mathbb{R}$ *est un ensemble* E *contenant un élément appelé vecteur nul et noté* $\overrightarrow{0}$*, muni d'une addition interne notée* $+$ *et d'une multiplication externe notée* $\cdot$ *vérifiant :*

$$\forall x, y, z \in E\ , \quad (x+y)+z = x+(y+z)\ , \quad \textit{(associativité)} \tag{1}$$
$$\forall x, y \in E\ , \quad x+y = y+x\ , \quad \textit{(commutativité)} \tag{2}$$
$$\forall x \in E\ , x+0 = 0 \quad \textit{(élément neutre)} \tag{3}$$
$$\forall x \in E\ ;, \exists y \in E\ , \quad x+y = 0\ , \quad \textit{(symétrique)} \tag{4}$$
$$\forall x \in E\ , \quad 1 \cdot x = x\ , \tag{5}$$
$$\forall \lambda \in \mathbb{R}\ , \quad \forall x, y \in E\ , \quad \lambda \cdot (x+y) = \lambda \cdot x + \lambda \cdot y \tag{6}$$
$$\forall \lambda, \mu \in \mathbb{R}\ , \quad \forall x \in E\ , \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x \tag{7}$$
$$\forall \lambda, \mu \in \mathbb{R}\ , \quad \forall x \in E\ , \quad (\lambda \mu) \cdot x = \lambda \cdot (\mu \cdot x)\ . \tag{8}$$
$$\tag{9}$$

Les éléments de E *sont appelés des vecteurs, notés* u*,* v *... ou bien* $\overrightarrow{u}$*,* $\overrightarrow{v}$ *...Les éléments de* $\mathbb{R}$ *sont appelés des scalaires.*

Propriétés

$$\begin{aligned}
&\forall x \in E\ , \quad 0 \cdot x = \overrightarrow{0}\ , \\
&\forall \lambda \in \mathbb{R}\ , \quad \lambda \cdot \overrightarrow{0} = \overrightarrow{0}\ , \\
&\forall x \in E\ , \forall \lambda \in \mathbb{R}\ , \quad \lambda \cdot x = \overrightarrow{0} \Leftrightarrow \lambda = 0 \text{ ou } x = \overrightarrow{0}\ , \\
&\forall x, y \in E\ , \forall \lambda \in \mathbb{R}^*\ , \quad \lambda \cdot x = \lambda \cdot y \Leftrightarrow x = y\ , \\
&\forall x \in E\ , \forall \lambda, \mu \in \mathbb{R}^*\ , \quad \lambda \cdot x = \mu \cdot x \Leftrightarrow \lambda = \mu\ .
\end{aligned}$$

Exemple fondamental

Soit $n \geq 1$ et $\mathbb{R}^n$ l'ensemble des n-uplets de réels $\mathbf{x}$, représentés par une colonne à n lignes :

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\ .$$

L'addition et la multiplication externe sont définies coordonnée par coordonnée :

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}\ , \qquad \lambda \cdot \mathbf{x} = \begin{pmatrix} \lambda y_1 \\ x\vdots \\ \lambda y_n \end{pmatrix}\ .$$

Représentation géométrique Pour $n = 2$, on représente les vecteurs de $\mathbb{R}^2$ dans le plan.

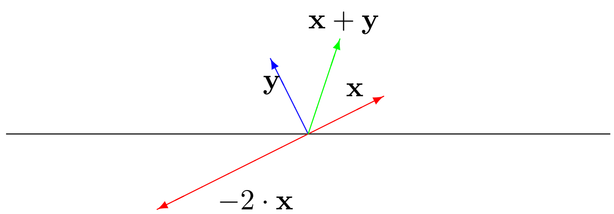

FIG. 1 – Les vecteurs $\mathbf{x} = (2, 1)$, $\mathbf{y} = (-1, 2)$, $-2 \cdot \mathbf{x}$ et $\mathbf{x} + \mathbf{y}$.

Définition 2 (Sous-espaces vectoriels)**.** *Une partie non vide F d'un espace vectoriel E est un sous espace vectoriel de E si*

$$\forall x, y \in F \ , \quad x + y \in F \ , \quad \textit{(stabilité pour l'addition)} \tag{10}$$

$$\forall x \in F \ , \forall \lambda \in \mathbb{R} \ , \quad \lambda \cdot y \in F \ , \quad \textit{(stabilité pour la multiplication par un scalaire).} \tag{11}$$

Remarque 3. Un sous-espace vectoriel n'est pas vide par définition, et contient donc nécessairement le vecteur nul, en raison de la stabilité par multiplication externe. Si $x \in F$, alors $\overrightarrow{0} = 0 \cdot x \in F$. Le singleton $\{\overrightarrow{0}\}$ est un sous-espace vectoriel de tout espace vectoriel E.

Remarque 4. Un sous-espace vectoriel est lui-même un espace vectoriel. On appelle sous-espace vectoriel propre d'un espace vectoriel E un sous-espace vectoriel de E qui n'est ni E lui-même, ni le singleton $\{\overrightarrow{0}\}$.

Exemple 5. La droite réelle $\mathbb{R}$ est un espace vectoriel qui n'admet que $\{\overrightarrow{0}\}$ et $\mathbb{R}$ lui même comme sous-espace vectoriel.

Exemple 6. Dans $\mathbb{R}^2$, la partie F définie par $F = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.

Exemple 7. Dans $\mathbb{R}^2$, la partie F définie par $F = \{(x, y) \in \mathbb{R}^2 \mid y = 2x + 1\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.

Exemple 8. Dans $\mathbb{R}^2$, la partie F définie par $F = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \ , y \geq 0\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.

Proposition 9. *Soit E un espace vectoriel et soit F et G deux sous-espaces vectoriels. L'intersection de F et G est un sous-espace vectoriel de E.*

Exemple 10. Soit dans $\mathbb{R}^3$ les sous-espaces F et G définis par

$$F = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 - x_2 = 0\} \ , \quad G = \{\mathbf{x} \in \mathbb{R}^3 \mid x_2 - 2x_3 = 0\} \ .$$

Alors

$$F \cap G = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 = x_2 = 2x_3\} \ .$$

Par exemple le triplet $(2, 2, 1)$ est dans $F \cap G$.

Remarque 11. Contrairement à l'intersection, la réunion de deux sous-espaces n'est pas un sous-espace, sauf si l'un contient l'autre.

1.2 Systèmes d'équations linéaires

La méthode du pivot de Gauss Une méthode pratique de résolution des systèmes linéaires est la méthode du pivot de Gauss, qui consiste à éliminer les inconnues une à une jusqu'à résoudre une équation à une inconnue. Cette méthode est itérative, et elle commence de façon identique pour 2 ou n inconnues.

1.2.1 Système de deux équations à deux inconnues

Considérons le système

$$\left\{ \begin{array}{ll} a_{1,1}x_1 + a_{1,2}x_2 & = b_1 \ , \\ a_{2,1}x_1 + a_{2,2}x_2 & = b_2 \ . \end{array} \right. \tag{12}$$

On suppose que les coefficients ne sont pas tous nuls. Quitte à faire une permutation on peut supposer $a_{1,1} \neq 0$. Retranchons alors membre à membre $a_{2,1}/a_{1,1}$ fois la première ligne à la seconde ; on obtient le système équivalent au précédent :

$$\left\{ \begin{array}{lll} a_{1,1}x_1 & +a_{1,2}x_2 & = b_1 \ , \\ & +\{a_{2,2} - a_{1,2}a_{2,1}/a_{1,1}\}x_2 & = b_2 - a_{2,1}b_1/a_{1,1} \ . \end{array} \right. \tag{13}$$

Si $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$, on obtient

$$x_2 = \frac{a_{1,1}b_2 - a_{2,1}b_1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \ .$$

En reportant dans la première équation de (13), on obtient

$$x_1 = \frac{a_{1,2}b_2 - a_{2,2}b_1}{a_{1,2}a_{2,1} - a_{1,1}a_{2,2}} \ .$$

On voit donc que dès que $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$, le système a une unique solution pour tout couple (b_1, b_2).

Si $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = 0$, le système (13) devient

$$\left\{ \begin{array}{rl} a_{1,1}x_1 + a_{1,2}x_2 & = b_1 \ , \\ 0 & = b_2 - a_{2,1}b_1/a_{1,1} \ . \end{array} \right.$$

Deux cas sont possibles :
- $a_{1,1}b_2 - a_{2,1}b_1 \neq 0$ et il n'y a aucune solution ;
- $a_{1,1}b_2 - a_{2,1}b_1 = 0$ et il y a une infinité de solutions ; ce sont les couples de la forme

  $$(\{b_1 -_{1,2} x_2\}/a_{1,1}, x_2) \ ,$$

  pour tout $x_2 \in \mathbb{R}$.

1.2.2 Système de trois équations à trois inconnues

Soit le système

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 & = b_1 \ , \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 & = b_2 \ , \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 & = b_3 \ . \end{cases} \tag{14}$$

Comme précédemment, on peut toujours supposer que $a_{1,1} \neq 0$. On va donc élminer la variable x_1 des deuxième et troisième équations en leur retranchant respectivement la première ligne multipliée par $a_{2,1}/a_{1,1}$ et $a_{3,1}/a_{1,1}$. On note les lignes L_1, L_2 et L_3 et on note ces opérations $L_2 - (a_{2,1}/a_{1,1})L_1$ et $L_3 - (a_{3,1}/a_{1,1})L_1$. Pour alléger les notations, on note

$$\begin{aligned} \tilde{a}_{i,j} &= a_{i,j} - a_{1,j}a_{i,1}/a_{1,1} \ , \quad 2 \leq i, j \leq 3 \ , \\ \tilde{b}_j &= b_j - (a_{i,1}/a_{1,1})b_1 \ , \quad 2 \leq j \leq 3 \ . \end{aligned}$$

On transcrit cette première étape de la façon suivante :

$$\begin{cases} a_{1,1}x_1 & +a_{1,2}x_2 & +a_{1,3}x_3 & = b_1 \ , & \\ & +\tilde{a}_{2,2}x_2 & +\tilde{a}_{2,3}x_3 & = \tilde{b}_2 \ , & \qquad L_2 \to L_2 - (a_{2,1}/a_{1,1})L_1 \\ & +\tilde{a}_{3,2}x_2 & +\tilde{a}_{3,3}x_3 & = \tilde{b}_3 \ , & \qquad L_3 \to L_3 - (a_{3,1}/a_{1,1})L_1 \ . \end{cases} \tag{15}$$

Le système (15) est équivalent au système d'origine, mais il est maintenant décomposé en un sous-système de deux équations à deux inconnues que l'on sait résoudre en appliquant la méthode précédente. Ce sous-système a soit aucune solution, soit une unique solution, soit une infinité de solutions. S'il n'a aucune solution, le système initial n'en a pas non plus. S'il admet des solutions, i.e. des couples (x_2, x_3) qui vérifient les deux dernières équations, chaque solution reportée dans la première équation permet de determiner un triplet (x_1, x_2, x_3) solution du système initial.

Exemple 12. Résoudre le système

$$\begin{cases} 2x_1 + 2x_2 + 4x_3 & = 8 \ , \\ 2x_1 - x_2 + 3x_3 & = 3 \ , \\ x_1 - x_2 + x_3 & = 0 \ . \end{cases} \tag{16}$$

On élimine x_1 des deuxièmes et troisièmes équations.

$$\begin{cases} 2x_1 + 2x_2 + 4x_3 & = 8 \ , & \\ -3x_2 - x_3 & = -5 \ , & \qquad L_2 \to L_2 - L_1 \\ -2x_2 - x_3 & = -4 \ , & \qquad L_3 \to L_3 - \frac{1}{2}L_1 \ . \end{cases}$$

On élimine x_2 de la troisième équation.

$$\begin{cases} 2x_1 + 2x_2 + 4x_3 & = 8 \ , & \\ -3x_2 - x_3 & = -5 \ , & \\ -\frac{1}{3}x_3 & = -\frac{2}{3} \ , & \qquad L_3 \to L_3 - \frac{2}{3}L_2 \ . \end{cases}$$

On peut résoudre maintenant variable par variable, en commençant par x_3 et en remontant.

$$x_3 = 2 \ , \quad x_2 = 1 \ , \quad x_1 = -1 \ .$$

L'unique solution du système (16) est donc le triplet $(-1, 1, 2)$. On vérifie en reportant cette solution dans le sytème original.

Exemple 13. Résoudre le système

$$\left\{ \begin{array}{ll} 2x_1 + x_2 - x_3 & = 1 \ , \\ 3x_1 - x_3 & = 1 \ , \\ -6x_1 + 2x_3 & = -2 \ . \end{array} \right. \tag{17}$$

On voit ici que les deux dernières équations sont proportionnelles, i.e. $L_3 = -2L_2$. On peut donc simplifier le système (17) en supprimant L_3, ce qui donne

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

La seconde équation donne $x_3 = 3x_1 - 1$, et en retranchant la seconde ligne à la première, on obtient $x_1 = x_2$. Le système (17) admet donc une infinité de solutions qui sont des triplets de la forme $(x, x, 3x - 1)$.

Exemple 14. Résoudre le système

$$\left\{ \begin{array}{ll} 2x_1 + 3x_2 + x_3 & = 1 \ , \\ 2x_1 + x_2 + x_3 & = 2 \ , \\ 2x_1 + 2x_2 + x_3 & = 3 \ . \end{array} \right. \tag{18}$$

On élimine x_1 des deuxième et troisième équations.

$$\left\{ \begin{array}{lll} 2x_1 + 3x_2 + x_3 & = 1 \ , & \\ -2x_2 & = 1 \ , & L_2 \to L_2 - L_1 \ . \\ -x_2 & = 2 \ , & L_3 \to L_3 - L_1 \ . \end{array} \right.$$

Les deux dernières équations sont incompatibles et le système (18) n'a aucune solution.

1.2.3 Cas général

Cette méthode se généralise itérativement à un système de taille quelconque, du type

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

où $k \leq n$ sont des entiers, $a_{i,j}$, $1 \leq i \leq k$, $1 \leq j \leq n$ et b_i, $1 \leq i \leq k$ sont des nombres réels non tous nuls et $x_1, \ldots, x_n$ sont les inconnues.

A chaque étape, on essaie d'éliminer une inconnue. Le coefficient non nul utilisé pour éliminer la variable x_i dans les lignes L_j, $j > i$, est appelé le pivot. La méthode est itérative et se poursuit tant qu'un pivot est disponible. Lorsqu'il n'y a plus de pivot, le système est résolu, avec aucune solution, une unique solution ou une infinité de solutions.

Lorsque le second membre est nul, le système est dit homogène, et il admet au moins le vecteur nul comme solution. L'ensemble des solutions est un sous-espace vectoriel de $\mathbb{R}^n$.

1.3 Combinaisons linéaires, familles génératrices

Définition 15 (Combinaison linéaire)**.** *Soit E un espace vectoriel et soit $\mathbf{x}_1, \ldots, \mathbf{x}_n$ des vecteurs de E. La combinaison linéaire des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ de coefficients $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ est le vecteur*

$$\lambda_1 \cdot \mathbf{x}_1 + \cdots + \lambda_n \cdot \mathbf{x}_n \ .$$

Exemple 16. Soit dans $\mathbb{R}^3$ les vecteurs

$$\mathbf{a} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \ , \quad \mathbf{b} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \ , \quad \mathbf{c} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \ .$$

Alors pour tous $x, y, z \in \mathbb{R}^3$, on a

$$x \cdot \mathbf{a} + y \cdot \mathbf{b} + z \cdot \mathbf{c} = \begin{pmatrix} 2x + 2y + z \\ 2x - y + 3z \\ x - y + z \end{pmatrix} \ .$$

Proposition 17. *Soit F un sous-espace vectoriel d'un espace vectoriel E. Toute combinaison linéaire de vecteurs de F est un vecteur de F.*

Définition 18 (Sous-espace engendré par une famille de vecteurs)**.** *Soient E un espace vectoriel. Le sous-espace vectoriel engendré par des vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de E, noté $\mathrm{Vect}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ est l'ensemble de toutes les combinaisons linéaires possibles de ces vecteurs :*

$$\mathrm{Vect}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \{\lambda_1 \cdot \mathbf{a}_1 + \cdots + \lambda_n \cdot \mathbf{a}_n \mid \lambda_1, \ldots, \lambda_n \in \mathbb{R}\} \ .$$

Le fait que le sous-espace engendré est bien un sous-espace vectoriel doit être prouvé.

Proposition 19. *Soient E un espace vectoriel et $\mathbf{a}_1, \ldots, \mathbf{a}_n$ des vecteurs de E. Alors $\mathrm{Vect}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ est un sous-espace vectoriel de E.*

Définition 20 (Famille génératrice)**.** *Soit E un espace vectoriel et F un sous-espace de E (éventuellement E lui même). On dit que les vecteurs $\mathbf{a}_1, \cdots, \mathbf{a}_n$ engendrent F, ou que la famille $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ est génératrice pour F si $F = \mathrm{vect}(\mathbf{a}_1, \cdots, \mathbf{a}_n)$.*

Exemple 21. Dans $\mathbb{R}^n$, pour $i = 1, \ldots, n$, notons $\mathbf{e}_i$ le vecteur défini par

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ i\text{-ème ligne} \\ \\ \\ \\ \end{matrix}$$

La famille $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ engendre $\mathbb{R}^n$. En effet, tout n-uplet $\mathbf{x}$ peut s'écrire

$$\mathbf{x} = x_1 \cdot \mathbf{e}_1 + \cdots x_n \mathbf{x}_n \ .$$

Remarque 22. Un sous-espace vectoriel possède une infinité de familles génératrices, qui peuvent être formée d'un nombre arbitrairement grand de vecteurs. Notamment, si $\mathcal{F}$ est une famille génératrice pour un sous-espace vectoriel F, alors toute famille de vecteurs de F contenant la famille $\mathcal{F}$ est encore génératrice pour F.

1.4 Familles libres, familles liées

Définition 23 (Famille liée)**.** *Soit E un espace vectoriel et $\mathbf{x}_1, \ldots, \mathbf{x}_k$ des vecteurs de E. On dit que la famille $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ est liée ou que les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_k$ sont linéairement dépendants si il existe une combinaison linéaire nulle de ces vecteurs dont les coefficients ne soient pas tous nuls. Soit*

$$\exists (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k \setminus \{0\} \ , \quad \lambda_1 \cdot \mathbf{x}_1 + \cdots + \lambda_k \cdot \mathbf{x}_k = \overrightarrow{0} \ .$$

Exemple 24. Dans $\mathbb{R}^2$, les vecteurs

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ , \quad \mathbf{x}_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

sont liés car $3 \cdot \mathbf{x}_1 - 2 \cdot \mathbf{x}_2 - \mathbf{x}_3 = \overrightarrow{0}$.

Proposition 25. *Une famille de vecteurs est automatiquement liée dès qu'elle vérifie au moins un des critères suivants :*

(i) *elle contient le vecteur nul ;*

(ii) *elle contient de fois le même vecteur ;*

(iii) *si dans $\mathbb{R}^n$ elle contient strictement plus que n vecteurs.*

Définition 26 (Famille libre)**.** *Une famille de vecteurs est dite libre si elle n'est pas liée. Les vecteurs composant la famille sont alors dits linéairement indépendants.*

Caractérisation Des vecteurs $\mathbf{x}_1, \cdots, \mathbf{x}_k$ d'un espace vectoriel E sont linéairement indépendants si et seulement si

$$\forall \lambda_1, \ldots, \lambda_n \in \mathbb{R}\ , \quad \lambda_1 \cdot \mathbf{x}_1 + \cdots, \lambda_n \cdot \mathbf{x}_n = \overrightarrow{0} \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_n = 0\ . \tag{19}$$

Exemple 27. Dans $\mathbb{R}^3$, les vecteurs

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \ , \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \ , \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

sont libres. En effet, soit a, b, c des réels. Si $a \cdot \mathbf{u}_1 + b \cdot \mathbf{u}_2 + c \cdot \mathbf{u}_3 = \overrightarrow{0}$, alors

$$\begin{cases} a + c = 0 \\ 2a + b - c = 0 \\ -b + c = 0 \end{cases} ,$$

d'où l'on déduit aisément que $a = b = c = 0$.

Remarque 28. En dehors des cas triviaux, prouver qu'une famille de vecteurs est libre ou liée revient à résoudre un système linéaire d'équations homogènes. Si l'unique solution est 0, la famille est liée. S'il existe des solutions non nulles, la famille est liée.

Exemple 29. Montrons que les vecteurs $\mathbf{e}_i$, $i = 1, \ldots, n$ de $\mathbb{R}^n$ définis dans l'exemple 21 sont linéairement indépendants. Soient $\lambda_1, \ldots, \lambda_n$ des réels. On a

$$\lambda_1 \cdot \mathbf{x}_1 + \cdots, \lambda_n \cdot \mathbf{x}_n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} ,$$

donc la propriété (19) est vérifiée.

1.5 Bases, dimension

Définition 30. *Soit E un espace vectoriel. Une famille de vecteurs de E est une base si et seulement si elle est à la fois libre et génératrice.*

Théorème 31. *Tout espace vectoriel non réduit au singleton $\{\overrightarrow{0}\}$ admet une base. Toutes les bases d'un espace vectoriel ont le même cardinal. Ce cardinal est appelé la dimension de l'espace vectoriel.*

Notation On note $\dim(E)$ la dimension de l'espace vectoriel E. Par convention, la dimension de l'espace vectoriel réduit au singleton $\{\overrightarrow{0}\}$ est 0.

Exemple 32. Les vecteurs définis dans l'exemple 21 forment une base de $\mathbb{R}^n$. L'espace vectoriel $\mathbb{R}^n$ est donc de dimension n.

Exemple 33. Un espace vectoriel de dimension 1 est appelé droite vectoriel. Tout vecteur non nul de l'espace forme une base.

Corollaire 34. *Soit E un espace de dimension n. Une famille libre a au plus n éléments. Une famille génératrice a au moins n éléments. Une famille libre à n élements est une base. Une famille génératrice à n éléments est une base.*

Ce résultat est important. Pour prouver qu'une famille de vecteurs d'un espace de dimension n est une base, il suffit de vérifier qu'elle a n éléments et qu'elle est libre, ce qui est généralement plus facile que de montrer qu'elle est génératrice.

Exemple 35. Soit dans $\mathbb{R}^3$ u_1, u_2 et u_3 les vecteurs définis dans l'exemple 27. Ils forment une famille libre à trois éléments, et donc une base de $\mathbb{R}^3$.

Théorème 36. *Soit E un espace vectoriel et soit $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une vase de E. Pour tout vecteur $\mathbf{x} \in E$, il existe un unique n-uplet de réels $\lambda_1, \ldots, \lambda_n$ tel que*

$$\mathbf{x} = \lambda_1 \cdot \mathbf{e}_1 + \cdots + \lambda_n \cdot \mathbf{e}_n \ .$$

Ce théorème signifie que tout vecteur d'un espace vectoriel s'écrit *de façon unique* comme combinaison linéaire des vecteurs d'une base. Ces coefficients sont appelés les *coordonnées* ou *composantes* du vecteur dans la base considérée. Etant donnés une base d'un espace vectoriel et un vecteur, déterminer les coordonnées du vecteur dans la base revient à résoudre un système d'équations linéaires dont on sait a priori qu'il admet une unique solution.

Exemple 37. Considérons à nouveau les vecteurs définis dans l'exemple 27. On a vu qu'ils forment une base, donc tout autre vecteur de $\mathbb{R}^3$ est une combinaison linéaire de ces trois vecteurs. Soit $\mathbf{x} \in \mathbb{R}^3$. Cherchons ses coordonnées a, b, c dans la base $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. Pour ceci, on doit résoudre le système

$$\left\{ \begin{array}{rl} a + c & = x_1 \ , \\ 2a + b - c & = x_2 \ , \\ -b + c & = x_3 \ . \end{array} \right.$$

Par la méthode du pivot, ou par une méthode élémentaire d'élimination, on obtient :

$$a = (x_2 + x_3)/2 \ , \quad b = x_1 - x_2/2 - 3x_3/2 \ , \quad c = x_1 - x_2/2 - x_3/2 \ .$$

Par exemple, si $\mathbf{x}$ est le triplet $(1, 2, 4)$, on a $\mathbf{x} = 3 \cdot \mathbf{u}_1 - 6 \cdot \mathbf{u}_2 - 2 \cdot \mathbf{u}_3$.

1.6 Caractérisation des sous-espaces vectoriels

Un sous-espace vectoriel étant en soi un espace vectoriel, il admet aussi des bases et une dimension.

Théorème 38. *Soit E un espace vectoriel, F un sous-espace de E. On a alors* $\dim(F) \leq \dim(E)$ *et l'égalité a lieu si et seulement si $F = E$. De plus si* $\dim(F) = k$, $\dim(E) = n$, *avec $0 < k < n$, et si $(\mathbf{e}_1, \cdots, \mathbf{e}_k)$ est une base de F, alors il existe des vecteurs $\mathbf{e}_{k+1}, \cdots, \mathbf{e}_n$ tels que $(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ soit une base de E.*

Remarque 39. La relation entre la dimension d'un sous-espace et l'espace E lui-même s'étend à deux sous-espaces. Si F et G sont deux sous-espaces tels que $F \subset G$, alors $\dim(F) \leq \dim(G)$ et l'égalité a lieu si et seulement si $F = G$. En particulier, si E est égal à $\mathbb{R}$, ou plus généralement si E est un espace vectoriel de dimension 1, ses seuls sous-espaces sont $\{\overrightarrow{0}\}$ et E lui-même.

Nous allons maintenant caractériser les sous-espaces vectoriels de $\mathbb{R}^2$ et $\mathbb{R}^3$.

1.6.1 Sous-espaces vectoriels de $\mathbb{R}^2$

L'espace $\mathbb{R}^2$ étant de dimension 2, ses seuls sous-espaces propres sont de dimension 1. Soit D un sous-espace de dimension 1 et $\mathbf{u}$ un vecteur non nul de D. Un vecteur $\mathbf{x}$ est donc colinéaire à $\mathbf{u}$: il existe donc un réel λ tel que $\mathbf{x} = \lambda \cdot \mathbf{u}$, i.e. $x_1 = \lambda u_1$ et $x_2 = \lambda u_2$. Le couple (x_1, x_2) vérifie donc

$$u_2 x_1 - u_1 x_2 = 0 \; . \tag{20}$$

Une droite vectorielle de $\mathbb{R}^2$ est donc caractérisée soit par un vecteur directeur, soit par une équation du type (20).

1.6.2 Sous-espaces vectoriels de $\mathbb{R}^3$

Les sous-espaces propres de $\mathbb{R}^3$ (ou d'un espace vectoriel de dimension 3) sont de dimension 1 ou 2.

Les droites vectorielles sont toujours caractérisées par un vecteur de base non-nul, mais maintenant par un système de deux équations linéaires. Soit par exemple la droite D engendrée par le vecteur $\mathbf{u}$. Un vecteur $\mathbf{x}$ de D est colinéaire à $\mathbf{u}$, donc il existe $\lambda \in \mathbb{R}$ tel que $\mathbf{x} = \lambda \cdot \mathbf{u}$, i.e. $x_1 = \lambda u_1$, $x_2 = \lambda u_2$ et $x_3 = \lambda u_3$. La droite D est donc l'ensemble des triplets (x_1, x_2, x_3) qui vérifient

$$\begin{cases} u_2 x_1 - u_1 x_2 & = 0 \; , \\ u_3 x_2 - u_2 x_3 & = 0 \; . \end{cases}$$

Exemple 40. La droite engendrée dans $\mathbb{R}^3$ par le triplet $(1, 2, 3)$ est caractérisée par le système d'équations

$$\begin{cases} 2x - y & = 0 \; , \\ 3y - 2z & = 0 \; . \end{cases}$$

Exemple 41. Le sous-espace vectoriel de $\mathbb{R}^3$ caractérisé par le système d'équations

$$\begin{cases} 2x - y &= 0 \ , \\ y - 2z &= 0 \ . \end{cases}$$

est une droite engendrée par le triplet $(1, 2, 1)$.

Les sous-espaces de dimension 2 (plans vectoriels) de $\mathbb{R}^3$ sont engendrés par deux vecteurs linéairement indépendants et caractérisés par une seule équation linéaire.

Exemple 42. Trouvons l'équation du plan vectoriel P engendré dans $\mathbb{R}^3$ par les vecteurs

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad , \quad \mathbf{v} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \ .$$

On vérifie tout d'abord que les vecteurs sont linéairement indépendants. Si $a \cdot \mathbf{u} + b\mathbf{v} = 0$, alors

$$a - b = 2a + b = 3a + b = 0$$

ce qui implique clairement que $a = b = 0$. Soit $\mathbf{x}$ un vecteur de P. Il existe des réels a et b tels que $\mathbf{x} = a \cdot \mathbf{u} + b\mathbf{v}$, donc

$$\begin{aligned} a - b &= x_1 \ , \\ 2a + b &= x_2 \ , \\ 3a + b &= x_3 \ . \end{aligned}$$

On obtient aisément $a = (x_1 + x_2)/3$ et $b = -x_1 - x_2 + x_3$, et en reportant par exemple dans la première équation, on trouve une équation caractérisant le plan P :

$$x_1 + 4x_2 - 3x_3 = 0 \ .$$

Proposition 43. *Soit E un espace vectoriel de dimension 3 et P et Q deux sous-espaces de E de dimension 2. Alors soit $P = Q$, soit $P \cap Q$ est un sous-espace de dimension 1, i.e. une droite vectorielle.*

Exemple 44. Soient P et Q les plans de $\mathbb{R}^3$ d'équations respectives $3x + 2y - 3z = 0$ et $-3x + y + 2z = 0$. Alors $P \cap Q$ est une droite vectorielle dont un vecteur de base est $(7, 3, 9)$.

2 Matrices

Définition 45. *Soit m et n deux entiers non nuls. Une matrice réelle A de taille $m \times n$ est un tableau de nombres réels à m lignes et n colonnes, dont les éléments sont notés $a_{i,j}$,*

$1 \leq i \leq m$, $1 \leq j \leq n$; *i est l'indice de ligne et j est l'indice de colonne. Une matrice A d'éléments* $a_{i,j}$ *est représentée sous la forme*

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

La longueur d'une ligne est égale au nombre de colonnes m *et la hauteur d'une colonne est égale au nombre de lignes* n*. Si* $m = 1$ *la matrice est appelée matrice ou vecteur ligne ; si* $n = 1$ *la matrice est appelée matrice ou vecteur colonne.*

La matrice dont tous les éléments sont nuls est appelée matrice nulle et notée 0 quelle que soit sa dimension.

Matrices carrées

Définition 46. *Une matrice* A *de taille* $n \times n$ *(ou simplement de taille* n*) est dite matrice carrée. Ses éléments dont l'indice de ligne est égal à l'indice de colonne sont appelés éléments diagonaux.*

Par convention, on identifiera une matrice carrée de taille 1 (a) et son unique élément a.

Définition 47 (Matrice diagonale)**.** *Une matrice carrée est dite diagonale si tous ses éléments non diagonaux sont nuls, i.e. si* $A = (a_{i,j})_{1 \leq i,j \leq n}$*, alors* A *est diagonale si et seulement si* $a_{i,j} = 0$ *si* $i \neq j$*.*

Par exemple, la matrice A est diagonale mais pas la matrice B.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 0 \end{pmatrix} \quad ,$$

Définition 48 (Matrice identité)**.** *La matrice identité de taille* n*, notée* I_n *est la matrice diagonale dont tous les éléments diagonaux sont égaux à 1.*

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad I_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

2.1 Opérations sur les matrices

Multiplication par un scalaire Soit $\lambda \in \mathbb{R}$ et A une matrice de taille $m \times n$ d'éléments $a_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$. La matrice λA est la matrice de taille $m \times n$ d'éléments $\lambda a_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$.

Par exemple,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} \quad , \quad 2A = \begin{pmatrix} 2 & 4 \\ 6 & -8 \\ 4 & 0 \end{pmatrix} \ ;$$

et pour toute matrice A, on a $0A = 0$.

Addition des matrices Soit A et B deux matrices de **même taille** $m \times n$, d'éléments respectifs $a_{i,j}$ et $b_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$. La somme des matrices A et B, notée $A+B$ est la matrice de **même taille** $m \times n$, dont les éléments sont $a_{i,j} + b_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$.

On peut évidemment additionner un nombre quelconque de matrices de **même taille** selon la même règle.

Soit par exemple

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -3 & 0 \\ 1 & -2 \\ -3 & 6 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 5 & -4 \end{pmatrix} \ ;$$

alors

$$A + B + C = \begin{pmatrix} 1-3+1 & 2+0+2 \\ 3+1+0 & -4-2-1 \\ 2-3+5 & 0+6-4 \end{pmatrix} = \begin{pmatrix} -1 & 4 \\ 1 & -2 \\ 0 & 2 \end{pmatrix} \ .$$

Remarquons aussi que additionner plusieurs fois une matrice avec elle même revient à la multiplier par un entier, ce qui est bien ce qu'on attend : pour toute matrice A :

$$\underbrace{A + \cdots + A}_{n \text{ fois}} = nA \ .$$

Multiplication des matrices Soit A et B deux matrices. Le produit AB est possible si le nombre de colonnes de A est égal au nombre de lignes de B. Le produit s'effectue alors "ligne par colonne". Plus précisément, soit A une matrice de taille $m \times n$ d'éléments $a_{i,k}$, $1 \leq i \leq m$, $1 \leq k \leq n$ et B une matrice de taille $n \times p$, d'éléments $b_{k,j}$, $1 \leq k \leq n$, $1 \leq j \leq p$. Le produit AB est une matrice de taille $m \times p$, d'éléments $c_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq p$ définis par

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

L'élément $c_{i,j}$ est le produit terme à terme de la ligne i de A et de la colonne j de B.

Soit par exemple

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -3 & 0 & 2 \\ 1 & -2 & 0 \\ -3 & 6 & 1 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \; .$$

On peut effectuer les produits BA, CB, AC et CA ; les produits AB et BC ne sont pas effectuables. Calculons AC en détaillant les opérations :

$$\begin{aligned} AC &= \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1\times 1 + 2\times 0 & 1\times 2 + 2\times(-1) & 1\times(-1) + 2\times 2 \\ 3\times 1 + (-4)\times 0 & 3\times 2 + (-4)\times(-1) & 3\times(-1) + (-4)\times 2 \\ 2\times 1 + 0\times 0 & 2\times 2 + 0\times(-1) & 2\times(-1) + 0\times 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 3 \\ 3 & 10 & -11 \\ 2 & 4 & -2 \end{pmatrix} \; , \end{aligned}$$

Calculons aussi CA.

$$CA = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 1 & 4 \end{pmatrix} \; .$$

On voit donc que le produit de deux matrices n'est pas toujours possible ; et lorsque l'on peut effectuer le produit dans les deux sens, les deux matrices obtenues ne sont pas toujours de la même taille, et même si elles le sont, elles ne sont pas toujours égales. On conclut que **le produit de matrices n'est pas commutatif**.

Exemple 49. Soit A et B deux matrices carrées définies par

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix} \; .$$

On peut effectuer les produits AB et BA, on obtient deux nouvelles matrices carrées de taille 2×2, mais elles ne sont pas égales. En effet, on calcule :

$$AB = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix} \quad , \quad BA = \begin{pmatrix} -3 & 3 \\ 0 & 3 \end{pmatrix} \; .$$

Propriétés du produit des matrices

(i) Le produit de matrices est associatif. Soit A, B, C trois matrices telles que les produits AB et $(AB)C$ soient possibles. Alors les produits BC et $A(BC)$ sont possibles et $(AB)C = A(BC)$. On note donc cette valeur commune ABC.

(ii) Le produit est distributif par rapport à l'addition. Soit A et B deux matrices de même taille et C telle que les produits AC et BC soient possibles ; alors $(A+B)C = AC+BC$. De même, si les produits CA et CB sont possibles, alors $C(A+B) = CA+CB$.

(iii) La matrice identité est l'élément neutre pour la multiplication. Soit A une matrice de taille $m \times n$. Alors

$$AI_n = A\ , \quad I_m A = A\ .$$

Produit de matrices et de vecteurs On peut faire le produit d'une matrice et d'un vecteur, ou de deux vecteurs pour obtenir, selon l'ordre des produits, un vecteur ou une matrice ou un scalaire.

(i) Si A est une matrice de taile $m \times n$ et x est un vecteur colonne à n éléments, Ax est un vecteur colonne à m éléments.

(ii) Si A est une matrice de taile $m \times n$ et y est un vecteur ligne à m éléments, yA est un vecteur ligne à n éléments.

(iii) Si x est un vecteur colonne à n éléments et y est un vecteur ligne à m éléments alors xy est une matrice de taille $n \times m$.

(iv) Si x est un vecteur colonne et y est un vecteur ligne et ont le même nombre d'éléments, alors yx est un scalaire.

(v) Si A est une matrice de taile $m \times n$, x est un vecteur colonne à n éléments, et y est un vecteur ligne à m éléments, alors yAx est un scalaire.

Exemple 50. Soit

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix}\ , \quad x = \begin{pmatrix} -3 \\ 2 \end{pmatrix}\ , \quad y = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}\ .$$

Alors

$$Ax = \begin{pmatrix} 1 \\ -17 \\ -6 \end{pmatrix}\ , \quad yA = \begin{pmatrix} 5 & -6 \end{pmatrix}\ , \quad xy = \begin{pmatrix} -3 & -6 & 3 \\ 2 & 4 & -2 \end{pmatrix}\ , \quad yAx = -27\ .$$

Le produit yx n'est pas possible.

Ecriture matricielle d'un système d'équations linéaires Considérons un système de m équations linéaires à n inconnues :

$$\left\{ \begin{array}{lcl} a_{1,1}x_1 & + \cdots + a_{1,n}x_n & = & b_1\ , \\ \vdots & & & \vdots \\ a_{m,1}x_1 & + \cdots + a_{m,n}x_n & = & b_m\ . \end{array} \right. \tag{21}$$

Posons

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} , \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} , \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} .$$

Le système (21) s'écrit matriciellement

$$Ax = b .$$

2.2 Matrices carrées inversibles

Définition 51. *Soit A une matrice carrée de taille n. Elle est dite inversible s'il existe une matrice B de même taille telle que*

$$AB = BA = I_n .$$

La matrice B est alors appelée inverse de A et notée A^{-1}.

De la définition, il ressort immédiatement que la matrice I_n est inversible, puisque $I_n I_n = I_n$. Si $n = 1$, une matrice de taille 1 est de la forme (a), et est inversible si et seulement si $a \neq 0$.

Remarque 52. Pour vérifier qu'une matrice B est l'inverse d'une matrice A, il suffit en fait de vérifier soit $AB = I_n$, soit $BA = I_n$. Chacune de ces deux conditions entraîne l'autre.

Une matrice carrée n'est pas nécessairement inversible. On a la condition nécessaire et suffisante suivante.

Théorème 53. *Soit A une matrice carrée de taille n. Elle est inversible si et seulement si pour tout $b \in \mathbb{R}^n$, le système $Ax = b$ admet une unique solution. La solution est alors $A^{-1}b$.*

Remarque 54. Une condition nécessaire et suffisante d'inversibilité est aussi que l'unique solution du système $Ax = 0$ est 0. On peut utiliser cette propriété pour montrer que A est non inversible en exhibant un vecteur x non nul tel que $Ax = 0$.

Exemple 55. Considérons les matrices suivantes

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix} , \quad B = \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix} .$$

A est inversible et l'on peut vérifier par exemple en résolvant le système $Ax = b$ par la méthode du pivot que son inverse est

$$A^{-1} = \frac{1}{9} \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix} .$$

La matrice B n'est pas inversible car

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

Proposition 56 (Inverse d'un produit)**.** *Soit A et B deux matrices inversibles de même taille. Alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

Démonstration. Posons $C = B^{-1}A^{-1}$. Il suffit de montrer que $ABC = I_n$. Par associativité du produit de matrice et du fait que I_n est l'élément neutre pour la multiplication des matrices, on a

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

□

Par récurrence, on étend cette propriété au produit de n matrices inversibles de même taille :

$$(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1} \ .$$

Méthode pratique d'inversion d'une matrice par la méthode du pivot La méthode du pivot donne une méthode pratique d'inversion d'une matrice dite méthode de la matrice témoin. Cette méthode consiste à transformer une matrice en la matrice identité par les opérations de la méthode du pivot. Simultanément, la matrice identité est transformée par les mêmes opérations. Si la matrice initiale A peut effectivement être transformée en l'identité, alors la matrice obtenue par transformation de l'identité est l'inverse de A. Si A ne peut être transformée en l'identité, c'est-à-dire qu'après un certain nombre d'opérations il n'est plus possible de choisir un pivot, c'est que la matrice A n'est pas inversible.

Exemple 57. Soit A la matrice

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

Appliquons la méthode. On annule tout d'abord les coefficients sous la diagonale de A, en reportant les opération sur la matrice identité.

$$L_2 \to L_2 + L_1 \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & -1 \\ 0 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ ;$$

$$L_3 \to L_3 - \frac{2}{3}L_2 \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & \frac{5}{3} \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{2}{3} & -\frac{2}{3} & 1 \end{pmatrix} \ .$$

On a déjà obtenu que la matrice est inversible. Sinon, un zéro serait apparu sur la diagonale. On annule maintenant les autres coefficients en remontant à partir de la dernière ligne.

$$L_2 \to L_2 + \frac{3}{5}L_3 \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{5}{3} \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{5} & \frac{3}{5} & \frac{3}{5} \\ -\frac{2}{3} & -\frac{2}{3} & 1 \end{pmatrix} \ ;$$

$$L_1 \to L_1 - \frac{1}{3}L_2 \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{5}{3} \end{pmatrix} \qquad \begin{pmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{3}{5} & \frac{3}{5} \\ -\frac{2}{3} & -\frac{2}{3} & 1 \end{pmatrix} \; ;$$

$$\begin{array}{c} L_2 \to L_2/3 \\ L_3 \to \frac{3}{5}L_3 \end{array} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ -\frac{2}{5} & -\frac{2}{5} & \frac{3}{5} \end{pmatrix} \; .$$

L'inverse de A est donc la matrice

$$\frac{1}{5}\begin{pmatrix} 4 & -1 & -1 \\ 1 & 1 & 1 \\ -2 & -2 & 3 \end{pmatrix} \; ,$$

ce que l'on vérifie aussitôt

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

Exercice 58 (Contrôle continu du 2 avril 2007). Soit les matrices

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 0 \end{pmatrix} \; , \quad B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 2 & 2 \end{pmatrix} \; .$$

a. *Calculer AB, $(AB)^2$ et en déduire $(AB)^{-1}$ sans calculs.*

$$AB = \begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix} \; , \quad (AB)^2 = I_2 \; .$$

On en déduit que la matrice AB est inversible et égale à son inverse. (On dit que c'est une matrice involutive).

b. *Expliquer pourquoi l'écriture $(AB)^{-1} = B^{-1}A^{-1}$ n'a pas de sens ici.*
Les matrices A et B ne sont pas carrées, donc ne sont pas inversibles.

c. *Calculer BA et montrer qu'elle n'est pas inversible.*

$$BA = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ -2 & 2 & -2 \end{pmatrix} \; .$$

On peut trouver un vecteur non nul $x \in \mathbb{R}^3$ tel que $BAx = 0$, donc BA n'est pas inversible. On a par exemple

$$BA\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 0 \; .$$

Remarque 59. Une autre façon de prouver qu'une matrice n'est pas inversible est de trouver une combinaison linéaire nulle des lignes ou des colonnes. Dans l'exemple ci-dessus, en notant L_1, L_2 et L_3 les lignes de BA, on voit aisément que $L_3 = 2L_2 - 2L_1$.

2.3 Transposition

Définition 60. *Soit A une matrice de taille $m \times n$. La transposée de A, notée ${}^tA$ est la matrice de taille $n \times m$ dont les lignes sont les colonnes de A. Si l'élément à l'intersection de la i-ème ligne et j-ème colonne de A est $a_{i,j}$, alors $a_{i,j}$ est l'élément à l'intersection de la j-ème ligne et de la i-ème colonne de ${}^tA$.*

Exemple 61.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \\ 2 & 0 \end{pmatrix} , \qquad {}^tA = \begin{pmatrix} 1 & 3 & 2 \\ 2 & -4 & 0 \end{pmatrix} .$$

Propriétés

(i) Pour toute matrice A, ${}^t({}^tA) = A$.

(ii) Si x est un vecteur colonne, ${}^tx$ est un vecteur ligne.

(iii) Si x est un vecteur ligne, ${}^tx$ est un vecteur colonne.

(iv) Si A et B sont deux matrices de même taille, alors ${}^t(A+B) = {}^tA + {}^tB$.

(v) Pour toutes matrices A et B telles que le produit AB soit possible, alors le produit ${}^tB\,{}^tA$ est possible et ${}^t(AB) = {}^tB\,{}^tA$.

(vi) Si A est inversible, alors ${}^tA$ est inversible et $({}^tA)^{-1} = {}^t(A^{-1})$.

Définition 62 (Matrice symétrique)**.** *Une matrice carrée est dite symétrique si elle est égale à sa transposée.*

Exemple 63. La matrice

$$\begin{pmatrix} 1 & 2 & a \\ b & 5 & c \\ 3 & b & -2 \end{pmatrix}$$

est symétrique si $a = 3$ et $b = c = 2$

Les matrices diagonales sont symétriques, notamment la matrice identité.

3 Produit scalaire

Définition 64. *Soit E un espace vectoriel. Un produit scalaire est une application de $E \times E$ dans $\mathbb{R}$, notée $< \cdot, \cdot >$ ayant les propriétés suivantes.*

(i) Symétrie : $\forall x, y \in E$, $< x, y >=< y, x >$.

(ii) Bilinéarité : $\forall u, v, x, y \in E$, $\forall \alpha, \beta, \lambda, \mu \in \mathbb{R}$,

$$< \alpha u + \beta v, \lambda x + \mu y >= \alpha\lambda < u, x > + \alpha\mu < u, y > + \beta\lambda < v, x > + \beta\mu < v, y > \ .$$

(iii) Positivité : $\forall x \in E$, $< x, x > \geq 0$ et $< x, x >= 0$ si et seulement si $x = 0$.

La norme associée au produit scalaire est définie pour tout $x \in E$ par

$$\|x\| = \sqrt{< x, x >} \, .$$

Exemple 65 (Produit scalaire euclidien). Sur l'espace vectoriel $\mathbb{R}^n$ dont les éléments sont des vecteurs colonnes à n éléments, on définit le produit scalaire euclidien et la norme euclidienne par

$$< x, y >= {}^t x \, y = \sum_{i=1}^n x_i y_i \ , \quad \|x\| = \sqrt{{}^t x \, x} = \sqrt{\sum_{i=1}^n x_i^2} \ .$$

Exemple 66. Soit $x, y \in E$ et $a, b \in \mathbb{R}$. Par définition de la norme et bilinéarité du produit scalaire, on a

$$\|ax + by\|^2 = a^2 \|x\|^2 + 2ab < x, y > + b^2 \|y\|^2 \ . \tag{22}$$

Proposition 67 (Inégalité de Cauchy-Schwarz)**.** *Pour tout x et y dans E, on a*

$$| < x, y > | \leq \|x\| \|y\| \ .$$

L'égalité n'a lieu que si x et y sont colinéaires.

Démonstration. Soit $x, y \in E$ et soit $\lambda \in \mathbb{R}$. Si $y = 0$ alors l'inégalité est forcément vérifiée. On peut donc supposer que $y \neq 0$. En appliquant la propriété (22), on a

$$\|x + \lambda y\|^2 = \|x\|^2 + 2\lambda < x, y > + \lambda^2 \|y\|^2 \ .$$

Ce trinôme du second degré en λ est toujours positif, donc son discriminant est négatif ou nul, i.e.

$$< x, y >^2 - \|x\|^2 \|y\|^2 \leq 0 \ .$$

Le cas d'égalité correspond à l'existence d'une racine double, c'est-à-dire à l'existence d'un λ_0 tel que $x + \lambda_0 y = 0$. $\square$

Proposition 68. *La norme associée à un produit scalaire sur un espace vectoriel E admet les propriétés suivantes :*

(i) Positivité : $\forall x \in E$, $\|x\| \geq 0$ et $\|x\| = 0$ si et seulement si $x = 0$.

(ii) Homogénéité : $\forall x \in E$, $\forall \lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.

(iii) Inégalité triangulaire : $\forall x, y \in E$,

$$\|x + y\| \leq \|x\| + \|y\| \ .$$

Démonstration. Les deux premières propriétés sont des conséquences immédiates de la définition. Montrons l'inégalité triangulaire. Par la propriété (22), on a

$$\|x+y\|^2 = \|x\|^2 + 2 < x, y > + \|y\|^2 \ .$$

En appliquant l'inégalité de Cauchy-Schwarz, on a donc

$$\|x+y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \ .$$

□

Définition 69 (Orthogonalité)**.** *Deux vecteurs x et y de E sont dits orthogonaux si leur produit scalaire est nul.*

Définition 70 (Cosinus)**.** *Soit $x, y \in E$. On appelle cosinus de x et y la fonction notée* $\cos(x, y)$ *définie par*

$$\cos(x,y) = \begin{cases} \frac{<x,y>}{\|x\|\|y\|} & \textit{si } x \neq 0 \textit{ et } y \neq 0 \ , \\ 1 & \textit{si } x = 0 \textit{ ou } y = 0 \ . \end{cases}$$

Proposition 71. *Pour tout $x, y \in E$ et $\lambda, \mu \in \mathbb{R}$, on a*

(i) Symétrie : $\cos(x, y) = \cos(y, x)$ *;*

(ii) Homogénéité : $\cos(\lambda x, \mu y) = \operatorname{sign}(\lambda\mu) \cos(x, y)$ *;*

(iii) $|\cos(x, y)| < 1$ *sauf si x et y sont colinéaires.*

Définition 72 (Angle non orienté)**.** *On appelle angle non orienté entre deux vecteurs x et y l'unique réel $\theta \in [0, \pi]$ tel que* $\cos(\theta) = \cos(x, y)$*.*

Deux vecteurs sont orthogonaux si et seulement si leur angle non orienté est $\pi/2$ et son colinéaires si et seulement si leur angle est 0 ou π.

Définition 73 (Sous-espaces orthogonaux)**.** *Soit F et G deux sous-espaces d'un espace vectoriel E muni d'un produit scalaire. On dit que F et G sont orthogonaux si tout vecteur de F est orthogonal à tout vecteur de G, soit*

$$\forall x \in F \ , \quad \forall y \in G \ , \quad < x, y >= 0 \ .$$

Proposition 74. *Soit F un sous-espace d'un espace vectoriel E muni d'un produit scalaire. Il existe un unique sous-espace G, orthogonal à F et tel que* $\dim F + \dim G = \dim E$*. Le sous-espace G est alors appelé l'orthogonal de F, et noté $F^{\perp}$.*

Exemple 75. Dans $\mathbb{R}^3$, muni du produit scalaire euclidien, l'orthogonal d'un plan (sous-espace de dimension 2) est une droite, et réciproquement. Soit P le plan engendré par les vecteurs

$$x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \ , \quad y = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \ .$$

Déterminons la droite D orthogonale à P. Tout vecteur z de D doit vérifier les deux conditions $< x, z >= 0$ et $< y, z >= 0$, soit, en notant z_1, z_2, z_3 les coordonnées de z :

$$\begin{aligned} z_1 - z_2 + 2z_3 &= 0 \ , \\ z_1 + z_2 - z_3 &= 0 \ . \end{aligned}$$

Ces équations sont les équations cartésiennes de la droite D. Un vecteur directeur de D a pour coordonnées $1, -3, -2$ et tous les vecteurs de D sont proportionnels à celui-ci.

4 Projection orthogonale

Dans cette section, on va traiter le problème de la projection orthogonale, tout d'abord sur une droite puis dans un cadre général.

4.1 Projection orthogonale sur une droite

Soit x un vecteur non nul de E et soit D la droite vectorielle engendrée par x, i.e. $D = \{\lambda x \mid \lambda \in \mathbb{R}\}$. Soit $y \in E$. On veut projeter le vecteur y orthogonalement sur la droite D. On note $\Pi_D(y)$ la projection orthogonale de y sur D. Elle est caractérisée par les deux propriétés suivantes :

(i) $\Pi_D(y) \in D$, i.e. $\Pi_D(y)$ est colinéaire à x

(ii) $y - \Pi_D(y)$ est orthogonal à x.

On en déduit la solution :

$$\Pi_D(y) = \frac{< x, y >}{\|x\|^2} x \ . \tag{23}$$

Démonstration. Il faut vérifier les deux conditions caractéristiques. La première l'est par construction. Pour vérifier l'orthogonalité de $y - \Pi_D(y)$ ainsi défini et x, on calcule le produit scalaire :

$$< y - \frac{< x, y >}{\|x\|^2} x, x >=< x, y > - \frac{< x, y >}{\|x\|^2} < x, x >= 0 \ ,$$

car par définition, $\|x\|^2 =< x, x >$. □

Exemple 76. Dans $\mathbb{R}^3$, on pose $x = {}^t(1\ , -1\ , 2)$ et $y = {}^t(1, 1, -1)$. La projection de y sur x est αx, avec

$$\alpha = \frac{< x, y >}{< x, x >} = \frac{{}^t xy}{{}^t xx} = \frac{-2}{6} = -\frac{1}{3} \ .$$

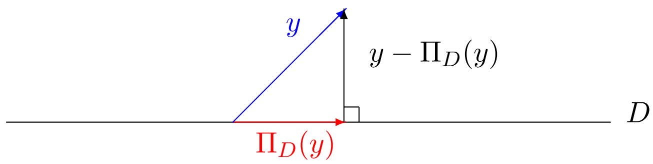

FIG. 2 – La projection orthogonale de y sur D

Remarque 77. Si x et y sont orthogonaux, alors la projection orthogonale de y sur la droite engendrée par x est le vecteur nul.

On va maintenant voir la projection orthogonale de y sur x comme le vecteur colinéaire à x le plus proche de y.

Proposition 78. *Soit x un vecteur non nul de E et soit $y \in E$. La projection orthogonale de y sur la droite D engendrée par x est l'unique vecteur colinéaire à x réalisant le minimum de $\|y - z\|$ pour z colinéaire à x, i.e.*

$$\Pi_D(y) = \arg\min_{z \in D} \|y - z\| \ .$$

Démonstration. Posons $\alpha =< x, y > / < x, x >$. Soit $z \in D$. En appliquant la propriété (22), on a

$$\begin{aligned} \|y - z\|^2 = \|y - \alpha x + \alpha x - z\|^2 &= \|y - \alpha x\|^2 - 2 < y - \alpha x, \alpha x - z > + \|\alpha x - z\|^2 \\ &= \|y - \alpha x\|^2 + \|\alpha x - z\|^2 \end{aligned}$$

car $y - \alpha x$ est orthogonal à $z - \alpha x$ (puisque z est aussi colinéaire à x). Le membre de droite est somme de deux carrés dont le premier $\|y - \alpha x\|$ est fixe. Cette somme donc est minimale lorsque le second est nul, c'est-à-dire lorsque $z = \alpha x$. □

Autre démonstration. On cherche le réel α qui minimise $\|y - \alpha x\|$, c'est-à-dire qui minimise $\|y - \alpha x\|^2$. Par la propriété (22), on a

$$\|y - \alpha x\|^2 = \|y\|^2 - 2\alpha < x, y > + \alpha^2 \|x\|^2 \ .$$

On note $f(\alpha)$ cette fonction. On a $f'(\alpha) = -2 < x, y > + 2\alpha \|x\|^2$. f' s'annule donc en $\alpha =< x, y > / < x, x >$ qui est donc un extremum. La dérivée seconde est positive, c'est donc un minimum. □

4.2 Cas général

Soit A une matrice de taille $m \times n$. Soit $A_1, \ldots, A_n$ ses colonnes. Ce sont des vecteurs de $\mathbb{R}^m$. Par exemple, si

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

ses colonnes sont

$$A_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \quad , \quad A_2 = \begin{pmatrix} 2 \\ -4 \\ 0 \end{pmatrix}$$

On cherche à déterminer la projection orthogonale d'un vecteur y sur le sous espace vectoriel de $\mathbb{R}^m$ engendré par $A_1, \cdots, A_n$. Cette projection, notée Π_A est caractérisée par les propriétés

(i) $\Pi_A(y)$ est une combinaison linéaire des colonnes de A ;

(ii) $y - \Pi_A(y)$ est orthogonal aux colonnes de A.

Théorème 79. *Soit A une matrice de taille $m \times n$ et soit $y \in \mathbb{R}^m$. Si la matrice ${}^tAA$ est inversible, alors la projection orthogonale de y sur le sous-espace de $\mathbb{R}^m$ engendré par les colonnes de A est*

$$A\,({}^tAA)^{-1}\,{}^tA\,y\ .$$

Remarque 80. La matrice ${}^tAA$ est de taille $m \times m$. Elle n'est pas toujours inversible. On admettra qu'une condition nécessaire et suffisante est que les n colonnes de A, qui sont de taille m, soient des vecteurs linéairement indépendants de $\mathbb{R}^m$. Une condition nécessaire (mais non suffisante) est donc que $m \geq n$. En effet, n vecteurs d'un espace vectoriel de dimension m ne peuvent être linéairement indépendants que si $m \geq n$.

Remarque 81. Vérifions que les produits de matrices ci-dessus sont licites :
- ${}^tAA$ est une matrice de taille $n \times n$, donc son inverse aussi ;
- $A\,({}^tAA)^{-1}$ est donc de taille $m \times n$;
- $A\,({}^tAA)^{-1}\,{}^tA$ est donc de taille $m \times m$, et on peut donc l'appliquer au vecteur y de $\mathbb{R}^m$.

Démonstration. Il faut prouver les deux propriétés caractéristique de la projection orthogonale.

(i) Le produit $({}^tAA)^{-1}\,{}^tA\,y$ est un vecteur ligne de taille n. Notons le $(b_1, \ldots, b_n)$. On peut alors écrire le vecteur $A({}^tAA)^{-1}\,{}^tA\,y$ de la façon suivante

$$A({}^tAA)^{-1}\,{}^tA\,y = \sum_{i=1}^{n} b_i A_i\ .$$

C'est-à-dire que la propriété caractéristique (i) est vérifiée.

(ii) Notons $w = A({}^tAA)^{-1}\,{}^tA\,y$. Soit u la matrice ligne de taille n dont les éléments sont les produits scalaires $< y - w, A_i >$. On veut montrer que chacun de ces produits scalaires est nul, c'est-à-dire que u est le vecteur nul. Or

$$u = {}^tA\ \left(y - A({}^tAA)^{-1}\,{}^tA\,y\right) = {}^tA\ y - ({}^tA\ A)\ ({}^tAA)^{-1}\,{}^tA\,y = {}^tA\ y - {}^tA\ y = 0\ .$$

□

4.3 Moindres carrés

Nous allons maintenant étudier le problème des moindres carrés dans le cas général. Soit comme précédemment A une matrice de taille $m \times n$ et $y \in \mathbb{R}^m$. On cherche la combinaison linéaire des n colonnes de A qui approche le mieux y au sens des moindres carrés, c'est-à-dire que l'on cherche $z \in \mathbb{R}^n$ tel que

$$z = \arg\min_{x \in \mathbb{R}^n} \|y - Ax\|$$

Théorème 82. *Si la matrice ${}^tAA$ est inversible, alors la solution du problème des moindres carrés est $({}^tAA)^{-1}\,{}^tA\,y$.*

Démonstration. Notons $z = ({}^tAA)^{-1}\,{}^tA\,y$. Alors $\Pi_A(y) = Az$. Soit $x \in \mathbb{R}^n$. Par définition de la projection orthogonale, $y - Az$ est orthogonal à Ax pour tout $x \in \mathbb{R}$, i.e. $< y - Az, Ax >= 0$ et donc bien évidemment $< y - Az, Az >= 0$. On a donc, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} \|y - Ax\|^2 &= \|y - Az - (Ax - Az)\|^2 \\ &= \|y - Az\|^2 - 2 < y - Az, Ax - Az > + \|(Ax - Az)\|^2 \\ &= \|y - Az\|^2 + \|(Ax - Az)\|^2 \geq \|y - Az\|^2 \,. \end{aligned}$$

Le minimum est donc bien atteint en z. □

Remarque 83. Si $m = n$ et si la matrice A est inversible, alors on a $({}^tAA)^{-1}\,{}^tA = A^{-1}$. Pour cette raison, la matrice $({}^tAA)^{-1}\,{}^tA$ est nommée lorsque $m > n$ et lorsqu'elle est définie, la pseudo-inverse de A.

Exemple 84. Considérons

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 3 & 1 \end{pmatrix} \quad , \qquad y = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \ .$$

On calcule

$$ {}^tAA = \begin{pmatrix} 14 & 2 \\ 2 & 3 \end{pmatrix} \ , \qquad ({}^tAA)^{-1} = \frac{1}{38}\begin{pmatrix} 3 & -2 \\ -2 & 14 \end{pmatrix} \ , \qquad {}^tAy = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \ ,$$
$$({}^tAA)^{-1}\,{}^tAy = \begin{pmatrix} 7/38 \\ 46/38 \end{pmatrix}$$

4.4 Systèmes surdéterminés d'équations linéaires

La méthode des moindres carrés peut-être considérée comme une résolution approchée d'un système d'équations linéaires lorsque le nombre d'équations est supérieur au nombre d'inconnues (surdétermination). En général, il n'y a alors pas de solution. On doit alors se contenter d'une solution approchée. La méthode des moindres carrés en fournit une.

Plus précisément si A est une matrice de taille $m \times n$ telle que ${}^tAA$ soit inversible, et b un vecteur de $\mathbb{R}^m$, on appelle solution au sens des moindres carrés de l'équation

$$Ax = b$$

l'unique vecteur z de $\mathbb{R}^n$ réalisant le minimum de $\|Az - b\|$, dont on vient de montrer qu'il est égal à $({}^tAA)^{-1}\,{}^tA\,b$.

Exemple 85 (Examen de juin 2007). *Résoudre au sens des moindres carrés le système suivant :*

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

Soient A la matrice et a le vecteur définis par

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

On cherche à résoudre au sens des moindres carrés l'équation $Au = b$ où u est un vecteur de $\mathbb{R}^3$. On sait que la solution est donnée par

$$u = ({}^tAA)^{-1}\,{}^tAa\ .$$

On pose $B = {}^tAA$ et $b = {}^tAa$. On cherche à résoudre le système $Bu = b$. Calculons tout d'abord B et b.

$$B = \begin{pmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & -2 \\ 2 & -2 & 4 \end{pmatrix}\ ,$$

$$b = \begin{pmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$$

On doit donc résoudre le système

$$\left\{\begin{array}{rcr} 4x + 2z &=& 1\ , \\ 4y - 2z &=& 3\ , \\ 2x - 2y + 4z &=& -1\ . \end{array}\right.$$

Par la méthode du pivot, on obtient aisément la solution $u = {}^t(1/4, 3/4, 0)$.

4.5 La matrice de projection orthogonale

Soit toujours A une matrice de taille $m \times n$ telle que ${}^tAA$ soit inversible. On a vu que la matrice de projection orthogonale sur le sous-espace vectoriel de $\mathbb{R}^m$ engendré par les n colonnes de A est la matrice P définie par

$$P = A\,({}^tAA)^{-1}\,{}^tA\,.$$

Propriétés de P

(i) P est une matrice carrée de taille $m \times m$

(ii) P est symétrique : ${}^tP = P$.

(iii) P est idempotente : $P^2 = P$.

(iv) Si y appartient au sous-espace vectoriel de $\mathbb{R}^m$ engendré par les n colonnes de A, alors $Py = y$.

(v) Si y est orthogonal au sous-espace vectoriel de $\mathbb{R}^m$ engendré par les n colonnes de A, alors $Py = 0$.

On peut aussi montrer que toute matrice symétrique idempotente est une matrice de projection orthogonale.

Exemple 86. Calculons la matrice de projection orthogonale correspondant à la matrice A de l'exemple 85. On a déjà calculé ${}^tAA$. On l'inverse par la méthode de la matrice témoin :

$$({}^tAA)^{-1} = \frac{1}{8}\begin{pmatrix} 3 & -1 & -2 \\ -1 & 3 & 2 \\ -2 & 2 & 4 \end{pmatrix}\,.$$

On obtient donc la matrice de projection orthogonale P.

$$\begin{aligned}
P = A({}^tAA)^{-1}{}^tA &= \frac{1}{8}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 3 & -1 & -2 \\ -1 & 3 & 2 \\ -2 & 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \\
&= \frac{1}{8}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 4 & -2 & 0 \\ -2 & 0 & 2 & 4 \\ 0 & -4 & 0 & 4 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}\,.
\end{aligned}$$

On vérifie que la matrice obtenue satisfait à $P^2 = P$.

Exemple 87 (Examen dee Juin 2007. Extrait).
Soit le plan $E = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z) = 0\}$.

(i) *Donner une base de E.*

Une base de E est par exemple $\{{}^t(1, -1, 0), {}^t(1, 0, -1)\}$

(ii) *Ecrire la matrice P de projection orthogonale sur E.*

Soit A la matrice de taille 3×2 dont les colonnes sont les deux vecteurs de la base de E ci-dessus. On sait que $P = A({}^tAA)^{-1}\,{}^tA$. Calculons tout d'abord ${}^tAA$ et son inverse.

$$ {}^tAA = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} , \qquad ({}^tAA)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} . $$

On obtient enfin

$$ \begin{aligned} P &= \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} . \end{aligned} $$

Il est important de remarquer que la matrice obtenue ne dépend pas de la base choisie. Un choix astucieux aurait consisté à prendre deux vecteurs de E orthogonaux. La matrice ${}^tAA$ aurait alors été diagonale, et donc d'inversion immédiate. Si l'on prend pour vecteurs de base ${}^t(1, 0, -1)$ et ${}^t(1, -2, 1)$, on obtient

$$ A = \begin{pmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 1 \end{pmatrix} , \quad {}^tAA = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix} , \quad ({}^tAA)^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/6 \end{pmatrix} , $$

$$ P = \begin{pmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/6 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 1 & -2 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} . $$

4.6 La droite des moindres carrés

En correspondance avec des points déterminés de la droite réelle $x_1, \ldots, x_n$ (que l'on peut supposer sans perte de généralité ordonnés) on fait des observations $y_1, \cdots, y_n$. Par exemple, on peut observer une certaine grandeur (prix, température, poids, etc.) en des instants régulièrement espacés $1, 2, \ldots, n$ (exprimés en seconde, ou heures, jours, semaines), ou irrégulièrement espacés. Les x_i peuvent être des tailles et les y_i des poids.

On cherche (à des fins de prédiction par exemple) à trouver une relation linéaire entre les x_i et les y_i, c'est-à-dire qu'on cherche à placer tous les points (x_i, y_i) sur une même droite. On cherche donc deux nombres réels a et b tels que

$$ \forall i = 1, \ldots, n \ , y_i = a + bx_i \ . $$

Bien entendu, il est improbable que tous ces points soient effectivement sur la même droite. On cherche donc à résoudre ce système d'équations de façon approchée, au sens des moindres carrés. On écrit ce système vectoriellement. En posant

$$A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \ , \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \ , \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \ , \quad \mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \ , \quad u = \begin{pmatrix} a \\ b \end{pmatrix} \ ,$$

on cherche à résoudre au sens des moindres carrés l'équation

$$y = a\mathbb{1} + bx \ ,$$

ou encore

$$Au = y \ ,$$

c'est-à-dire que l'on cherche le vecteur $u \in \mathbb{R}^2$ qui minimise

$$\|Au - y\|^2 \ .$$

La solution est donc donnée par $u = ({}^tAA)^{-1t}Ay$, si la matrice ${}^tAA$ est inversible.

Proposition 88. *Si les réels x_i ne sont pas tous égaux, alors la matrice ${}^tAA$ est inversible et la solution des moindres carrés est donnée par*

$$a = \frac{< x, x >< y, \mathbb{1} > - < x, y >< x, \mathbb{1} >}{n < x, x > -(< x, \mathbb{1} >)^2} \ ,$$
$$b = \frac{n < x, y > - < x, \mathbb{1} >< y, \mathbb{1} >}{n < x, x > -(< x, \mathbb{1} >)^2} \ .$$

Remarque 89. On peut exprimer plus explicitement ces quantités. Notons

$$\bar{x} = \frac{1}{n} < x, \mathbb{1} >= \frac{1}{n}\sum_{i=1}^{n} x_i \ , \quad \bar{y} = \frac{1}{n} < y, \mathbb{1} >= \frac{1}{n}\sum_{i=1}^{n} y_i \ ,$$
$$\overline{xy} = \frac{1}{n} < x, y >= \frac{1}{n}\sum_{i=1}^{n} x_i y_i \ , \quad \overline{x^2} = \frac{1}{n} < x, x >= \frac{1}{n}\sum_{i=1}^{n} x_i^2 \ .$$

On obtient

$$a = \frac{\overline{x^2}\bar{y} - \overline{xy}\bar{x}}{\overline{x^2} - \bar{x}^2} \ , \quad b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} \ . \tag{24}$$

On distinguera bien $\bar{x}^2$ et $\overline{x^2}$, c'est-à-dire le carré de la moyenne et la moyenne des carrés, qui sont jamais égaux sauf si tous les x_i sont égaux. Plus précisément, l'inégalité de Cauchy-Schwarz entraîne

$$\overline{x^2} \geq \bar{x}^2 \ ,$$

avec égalité si seulement si x est proportionnel au vecteur $\mathbb{1}$.

Remarque 90. Avec ces notations, on peut exprimer a simplement en fonction de b :

$$a = \bar{y} - b\bar{x} \ . \tag{25}$$

Exemple 91. Déterminer la droite la plus proche au sens des moindres carrés des points du plan $\mathbb{R}^2$ (0,0), (1,0) et (3,12).

On pose $x = (0, 1, 3)$ et $y = (0, 0, 12)$. On obtient

$$\bar{x} = 4/3 \ , \ \bar{y} = 4 \ , \ \overline{x^2} = 10/3 \ , \ \overline{xy} = 12 \ .$$

La pente de la droite de régression est donnée par la formule (24) :

$$b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{12 - 16/3}{10/3 - 16/9} = 30/7 \ .$$

Le coefficient a est donné par la formule (25) :

$$a = \bar{y} - b\bar{x} = 4 - (30/7) \times (4/3) = -12/7 \ .$$

5 Formes quadratiques

Définition 92. *Soit A une matrice carrée de taille n. On appelle forme quadratique associée à A l'application notée q_A, définie par*

$$q_A(x) = < Ax, x > = {}^t x \, {}^t Ax \ .$$

Soit $a_{i,j}$, $1 \leq i, j \leq n$ les éléments de A. On a alors

$$q_A(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i x_j \ . \tag{26}$$

Exemple 93. Soit

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

Soit $x \in \mathbb{R}^2$ de coordonnées x_1 et x_2. Alors

$$q_A(x) = x_1^2 + 3x_1x_2 + 3x_2^2 \ .$$

Considérons maintenant la matrice

$$B = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 3 \end{pmatrix} \ .$$

Alors $q_B(x) = x_1^2 + 3x_1x_2 + 3x_2^2 = q_A(x)$.

On voit sur cet exemple qu'il peut exister plusieurs matrices associées à la même forme quadratique. Cependant, il en existe une seule qui soit symétrique. Si l'on considère l'écriture générale (26), on peut la réécrire de la façon suivante

$$q_A(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_{i,j} + a_{j,i}}{2} x_i x_j \ .$$

Posons $b_{i,j} = (a_{i,j} + a_{j,i})/2$. La matrice B d'éléments $b_{i,j}$ est alors symétrique et l'on a $q_A = q_B$.

Ce calcul permet aussi de retrouver la matrice A ou la matrice symétrique associée à partir de l'écriture développée du type (26) d'une forme quadratique q_A.

Exemple 94. Soit la forme quadratique sur $\mathbb{R}^3$ définie par $q(x) = x_1^2 - x_2^2 + x_1x_2 + 3x_1x_3$. Elle est associée à la matrice symétrique B définie par

$$B = \begin{pmatrix} 1 & 1/2 & 3/2 \\ 1/2 & -1 & 0 \\ 3/2 & 0 & 0 \end{pmatrix} \ .$$

On peut aussi trouver une infinité de matrices non symétriques donnant la même forme quadratique, par exemple

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

On a bien $q = q_A = q_B$.

Définition 95 (Décomposition d'une forme quadratique)**.** *On appelle décomposition d'une forme quadratique q définie sur $\mathbb{R}^n$ toute écriture de la forme*

$$q(x) = \sum_{i=1}^{p} \alpha_i < b_i, x >^2$$

où $b_1, \ldots, b_p$ sont des vecteurs de $\mathbb{R}^n$ et les α_i sont des nombres réels non nuls. Si de plus la famille $(b_1, \ldots, b_p)$ est famille libre, la décomposition associée est appelée réduction.

Remarque 96. En posant $b_i' = \sqrt{|\alpha_i|} b_i$, on peut toujours supposer que les nombres α_i sont égaux soit à 1 soit à -1.

Exemple 97. Considérons la forme quadratique de l'exemple 93 : $q(x) = x_1^2 + 3x_1x_2 + 3x_2^2$. On a

$$\begin{aligned} q(x) &= (x_1 + 3x_2/2)^2 - 9x_2^2/4 + 3x_2^2 = (x_1 + 3x_2/2)^2 + 3x_2^2/4 \\ &= < b_1, x >^2 + < b_2, x >^2 \end{aligned}$$

avec

$$b_1 = \begin{pmatrix} 1 \\ 3/2 \end{pmatrix} , \quad b_2 = \begin{pmatrix} 0 \\ \sqrt{3}/2 \end{pmatrix} .$$

Les vecteurs b_1 et b_2 sont bien linéairement indépendants. On peut aussi écrire

$$\begin{aligned} q(x) &= 3(x_1/2 + x_2)^2 - 3x_1^2/4 + x_1^2 = 3(x_1/2 + x_2)^2 + 3x_1^2/4 \\ &=< b'_1, x >^2 + < b'_2, x >^2 \end{aligned}$$

avec

$$b'_1 = \begin{pmatrix} \sqrt{3}/2 \\ \sqrt{3} \end{pmatrix} , \quad b'_2 = \begin{pmatrix} \sqrt{3}/2 \\ 0 \end{pmatrix} .$$

Les vecteurs b'_1 et b'_2 sont bien linéairement indépendants. Enfin, on peut écrire

$$q(x) = \alpha < b_1, x >^2 + \alpha < b_2, x >^2 + \beta < b'_1, x >^2 + \beta < b'_2, x >^2 ,$$

où α et β sont n'importe quels réels non nuls tels que $\alpha + \beta = 1$. Mais la famille (b_1, b_2, b'_1, b'_2) est évidemment liée, puisque la dimension de l'espace est 2.

Remarque 98. On voit donc qu'on peut écrire une forme quadratique comme une somme de carrés de plusieurs façons, et qu'une décomposition n'est pas toujours une réduction.

Méthode de réduction de Gauss Cette méthode est une méthode pratique pour déterminer une réduction d'une forme quadratique en somme de carrés. C'est une méthode itérative qui consiste à isoler successivement les variables. On commence par rassembler tous les termes contenant la variable x_1 par exemple, puis on utilise la relation

$$x_1^2 + 2a_2x_1x_2 + 2a_3x_1x_3 + \cdots = (x_1 + a_2x_2 + a_3x_3 + \cdots)^2 - a_2^2x_2^2 - a_3x_3^2 - 2a_2a_3x_2x_3 - \cdots$$

pour isoler x_1 dans un facteur du type $< b_1, \mathbf{x} >^2$. On a alors

$$q(x_1, \ldots, x_n) =< x, b_1 >^2 + q_2(x_2, \ldots, x_n) ,$$

où q_2 est une forme quadratique qui ne dépend plus de x_1. On poursuit par la réduction de q_2. On peut formaliser cette méhode par une récurrence rigoureuse. Elle garantit que les vecteurs b_i obtenus (si l'on n'a pas fait d'erreur de calcul) sont linéairement indépendants.

Exemple 99. Déterminons une réduction de la forme définie dans l'exemple 94. On a

$$\begin{aligned} q(x) &= x_1^2 - x_2^2 + x_1x_2 + 3x_1x_3 \\ &= (x_1 + x_2/2 + 3x_3/2)^2 - x_2^2/4 - 9x_3^2/4 - 3x_2x_3/2 - x_2^2 \\ &= (x_1 + x_2/2 + 3x_3/2)^2 - 5x_2^2/4 - 3x_2x_3/2 - 9x_3^2/4 \\ &= (x_1 + x_2/2 + 3x_3/2)^2 - 5/4(x_2 + 3x_3/5)^2 + 9x_3^2/20 - 9x_3^2/4 \\ &= (x_1 + x_2/2 + 3x_3/2)^2 - 5/4(x_2 + 3x_3/5)^2 - 9x_3^2/5 \\ &=< x, b_1 >^2 - 5/4 < x, b_2 >^2 - 9/5 < x, b_3 >^2 , \end{aligned}$$

avec

$$b_1 = \begin{pmatrix} 1 \\ 1/2 \\ 3/2 \end{pmatrix} \ , \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 3/5 \end{pmatrix} \ , \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \ .$$

Signature et rang d'une forme quadratique La réduction d'une forme quadratique en somme de carrés n'est pas unique, mais le nombre de carrés et le signe des coefficients est toujours fixe.

Théorème 100. *Soit q une forme quadratique sur $\mathbb{R}^n$. Soit $\alpha_1, \ldots, \alpha_p$ des nombres réels et $(b_1, \ldots, b_p)$ une famille libre de vecteurs de $\mathbb{R}^n$ tels que*

$$q(x) = \sum_{i=1}^{p} \alpha_i < b_i, x >^2 \ .$$

Soit n_- et n_+ les nombres de coefficients α_i de signe négatifs et positifs, respectivement. Alors n_- et n_+ ne dépendent que de q et sont les mêmes pour toute réduction, et l'on a $n_- + n_+ \leq n$. Le couple (n_+, n_-) est appelé signature de la forme quadratique q, et l'entier $n_- + n_+$ est appelé le rang de q.

Si $n_- = 0$, alors pour tout $x \in \mathbb{R}^n$, $q(x) \geq 0$, et q est dite positive ; si de plus $n_+ = n$, alors pour tout $x \in \mathbb{R}^n \setminus \{0\}$, $q(x) > 0$ et q est dite définie positive.

Si $n_+ = 0$, alors pour tout $x \in \mathbb{R}^n$, $q(x) \leq 0$, et q est dite négative ; si de plus $n_- = n$, alors q est dite définie négative.

Exemple 101. Dans l'exemple 93, la signature est $(2, 0)$, le rang est 2. La signature de la forme définie dans l'exemple 94 a été réduite dans l'exemple 99. On a obtenu

$$q(x) =< x, b_1 >^2 - 5/4 < x, b_2 >^2 - 9/5 < x, b_3 >^2 ,$$

où les vecteurs b_1, b_2, b_3 sont linéairement indépendants, donc on obtient que la signature est $(1, 2)$ et le rang 3.

Exemple 102. Soit sur $\mathbb{R}^3$ la forme quadratique q définie par $q(x) = x_1 x_2$. On trouve sa signature en écrivant

$$q(x) = x_1 x_2 = \frac{1}{4}(x_1 + x_2)^2 - \frac{1}{4}(x_1 - x_2)^2 = \frac{1}{4} < x, b_1 >^2 - \frac{1}{4} < x, b_2 >^2 ,$$

avec

$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \ , \quad b_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \ .$$

Les vecteurs b_1, b_2 sont linéairement indépendants, donc on obtient que la signature est $(1, 1)$ et le rang 2.