ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Algèbre Linéaire

Bachelor 1ère année

2009 - 2010

Génie Civil

&

Sciences et Ingénierie de l'Environnement

Support du cours de **Dr. Lara Thomas**

Polycopié initial élaboré par

Prof. Eva Bayer Fluckiger
Dr. Philippe Chabloz

Septembre 2009

Table des matières

Chapitre 1

Systèmes d'équations linéaires et matrices

L'algèbre linéaire est un outil essentiel pour toutes les branches des mathématiques appliquées, en particulier lorsqu'il s'agit de modéliser puis résoudre numériquement des problèmes issus de divers domaines : des sciences physiques ou mécaniques, des sciences du vivant, de la chimie, de l'économie, des sciences de l'ingénieur,...

Par exemple, la physique abonde de relations linéaires : les lois fondamentales du mouvement sont presque toutes linéaires, ou se déduisent de lois linéaires. Les systèmes électriques sont fondamentalement décrits par des lois linéaires ($V = RI$, etc.) C'est pourquoi, le présent cours commence avec une étude des équations linéaires et de leur résolution.

1.1 Introduction aux systèmes d'équations linéaires

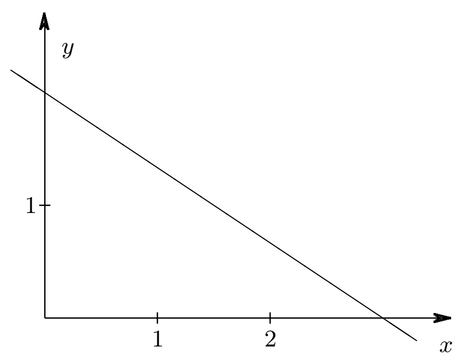
L'équation d'une droite dans le plan xy s'écrit

$$a_1x + a_2y = b$$

où a_1, a_2 et b sont des paramétres réels. Cette équation s'appelle équation linéaire dans les variables (ou inconnues) x et y.

Exemple 1.1.

$$2x + 3y = 6$$



Exemple 1.2. Les équations suivantes ne sont pas des équations linéaires :

$$2x + y^2 = 1$$

$$y = \sin(x)$$

$$x = \sqrt{y}$$

Définition 1.3. De maniére générale, on appelle *équation linéaire dans les variables* $x_1, ..., x_n$ toute relation de la forme

$$a_1x_1 + \cdots + a_nx_n = b \tag{1.1}$$

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

Il importe d'insister ici que ces équations linéaires sont *implicites*, c'est-à-dire qu'elles décrivent des relations entre les variables, mais ne donnent pas directement les valeurs que peuvent prendre les variables.

Résoudre une équation signifie donc la rendre *explicite*, c'est-à-dire rendre plus apparentes les valeurs que les variables peuvent prendre.

Une solution de l'équation linéaire (1.1) est un n-uple $s_1, \ldots, s_n$ de valeurs des variables $x_1, \ldots, x_n$ qui satisfont à l'équation (1.1). Autrement dit

$$a_1s_1 + \cdots + a_ns_n = b.$$

Par la suite, nous étudierons l'ensemble des solutions d'une équation linéaire.

Exemple 1.4. Trouvons l'ensemble des solutions de l'équation

$$x_1 - 4x_2 + 13x_3 = 5.$$

Nous donnons des valeurs arbitraires s et t à x_2 et x_3 respectivement et résolvons l'équation par rapport à x_1 :

$$x_2 = s, \qquad x_3 = t \qquad \text{et} \qquad x_1 = 4s - 13t + 5.$$

L'ensemble des solutions est alors

$$x_1 = 4s - 13t + 5, \qquad x_2 = s, \qquad x_3 = t$$

où s et t sont des nombres réels quelconques.

Définition 1.5. Un ensemble fini d'équations linéaires dans les variables $x_1, \ldots, x_n$ s'appelle *un système d'équations linéaires.* Tout $n-$uplet de nombres $s_1, \ldots, s_n$ satisfaisant chacune des équations s'appelle solution du système d'équations linéaires.

Exemple 1.6. Le système

$$\left\{\begin{array}{rcrcrcl} x_1 & - & 3x_2 & + & x_3 & = & 1 \\ -2x_1 & + & 4x_2 & - & 3x_3 & = & 9 \end{array}\right.$$

admet comme solution

$$x_1 = -18\,, \qquad x_2 = -6\,, \qquad x_3 = 1\,.$$

Par contre

$$x_1 = 7 \qquad x_2 = 2 \qquad x_3 = 0$$

ne satisfait que la première équation. Ce n'est donc pas une solution du système.

Définition 1.7. Un système d'équations est dit *incompatible* ou *inconsistant* s'il n'admet pas de solutions.

Exemple 1.8. Le système linéaire

$$\left\{\begin{array}{rcl} x_1 + x_2 & = & 1 \\ 2x_1 + 2x_2 & = & 1 \end{array}\right.$$

est clairement incompatible.

Considérons le système

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{1.2}$$

avec $a_{11} \cdot a_{12} \neq 0$ et $a_{21} \cdot a_{22} \neq 0$.

Ces deux équations représentent deux droites d_1 et d_2 dans le plan x_1x_2 et une solution du système est un point (s_1, s_2) qui est sur les deux droites. Trois cas se présentent alors :

(1) Les droites d_1 et d_2 se coupent en un seul point. Dans ce cas, illustré par la figure 1.1, le système (1.2) a une seule solution.

(2) Les droites d_1 et d_2 sont parallèles. Alors le système est incompatible et n'a pas de solution. La figure 1.2 illustre cette situation.

(3) Les droites d_1 et d_2 sont confondues et, dans ce cas, le système a une infinité de solutions.

Nous verrons plus loin que ces trois cas de figures (aucune solution, une seule solution, une infinité de solutions) sont les seuls cas qui peuvent se présenter pour n'importe quel système d'équations linéaires.

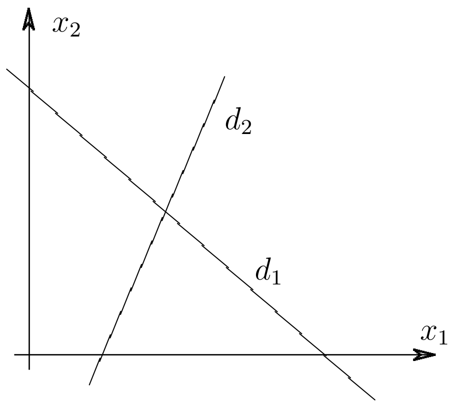


FIG. 1.1 – Droites se coupant en un seul point

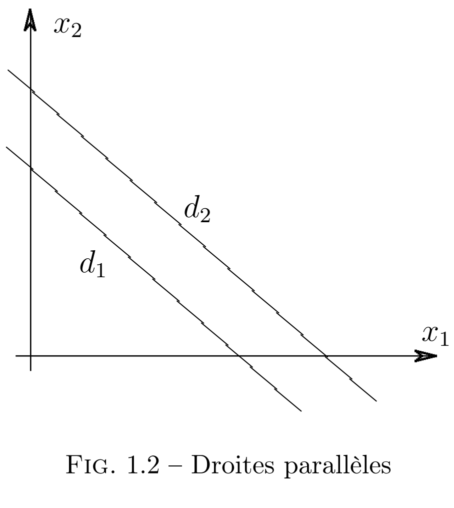


FIG. 1.2 – Droites parallèles

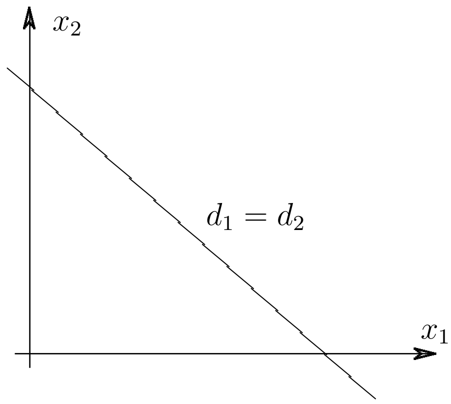


FIG. 1.3 – Droites confondues

1.2 Systèmes linéaires et matrices

Considérons un système quelconque de m équations à n inconnues,

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m. \end{array}\right. \tag{1.3}$$

où le nombre réel a_{ij} est le coefficient de la j-ème inconnue dans la i-ème équation.

Définition 1.9 (Matrice augmentée)**.** Nous obtenons la matrice augmentée associée au système en «oubliant» les variables x_i et les signes «+» et «=». La matrice augmentée aasociée au système (1.3) est alors

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

Exemple 1.10. Considérons le système linéaire

$$\left\{\begin{array}{rcl} x_1 + x_2 + 7x_3 & = & -1 \\ 2x_1 - x_2 + 5x_3 & = & -5 \\ -x_1 - 3x_2 - 9x_3 & = & -5. \end{array}\right.$$

Sa matrice augmentée est

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

La méthode de base pour résoudre un système d'équations linéaires est de remplacer le système par un autre, plus simple, ayant le méme ensemble de solutions. Ceci se fait par une succession d'opérations, appelées opérations élémentaires :

(1) multiplier une équation par une constante non nulle ;

(2) permuter deux équations ;

(3) ajouter un multiple d'une équation à une autre équation.

Les opérations (1), (2) et (3) ne changent pas l'ensemble des solutions. Elles correspondent à des opérations élémentaires sur les lignes de la matrice augmentée. Ces opérations sont les suivantes :

(1) multiplier une ligne par une constante non nulle ;

(2) permuter deux lignes ; $(*)$

(3) ajouter un multiple d'une ligne à une autre ligne.

Exemple 1.11. Utilisons ces opérations élémentaires pour résoudre le système suivant.

$$\left\{\begin{array}{rcr} x+y+7z & = & -1 \\ 2x-y+5z & = & -5 \\ -x-3y-9z & = & -5 \end{array}\right.$$

Nous calculons la matrice augmentée associée au système :

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 2 & -1 & 5 & -5 \\ -1 & -3 & -9 & -5 \end{pmatrix}$$

puis faisons les opérations élémentaires nécessaires sur le système et sur la matrice augmentée.

(3) $\ell_2 \longrightarrow \ell_2 - 2\ell_1$

$$\left\{\begin{array}{rcr} x+y+7z & = & -1 \\ -3y-9z & = & -3 \\ -x-3y-9z & = & -5 \end{array}\right.$$

(3) $\ell_2 \longrightarrow \ell_2 - 2\ell_1$

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & -3 & -9 & -3 \\ -1 & -3 & -9 & -5 \end{pmatrix}$$

Nous remarquons que les opérations élémentaires peuvent être faites uniquement sur la matrice augmentée pour revenir à la fin au système d'équations. C'est ce que nous faisons dans la suite.

(3) $\ell_3 \longrightarrow \ell_3 + \ell_1$

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & -3 & -9 & -3 \\ 0 & -2 & -2 & -6 \end{pmatrix}$$

(1) $\ell_2 \longrightarrow -\frac{1}{3}\,\ell_2$

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & -2 & -2 & -6 \end{pmatrix}$$

(3) $\ell_3 \longrightarrow \ell_3 + 2\ell_2$

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 4 & -4 \end{pmatrix}$$

(1) $\ell_3 \longrightarrow \frac{1}{4}\ell_3$

$$\begin{pmatrix} 1 & 1 & 7 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

(3) $\ell_1 \longrightarrow \ell_1 - 7\ell_3$

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

(3) $\ell_2 \longrightarrow \ell_2 - 3\ell_3$

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

(3) $\ell_1 \longrightarrow \ell_1 - \ell_2$

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

Cette matrice augmentée correspond au système

$$\left\{ \begin{array}{rrrcr} x & & & = & 2 \\ & y & & = & 4 \\ & & z & = & -1. \end{array} \right.$$

On obtient ainsi l'unique solution su système : $x = 2$, $y = 4$ et $z = -1$.

Cet exemple est généralisé dans le paragraphe suivant.

1.3 Elimination Gaussienne

Il s'agit d'une méthode qui permet de trouver l'ensemble des solutions de n'importe quel système d'équations linéaires. La méthode consiste à mettre la matrice augmentée du système sous une forme simple, dite forme échelonnée (réduite) par une série d'opérations élémentaires (1), (2), (3) de $(*)$.

Commençons par poser la définition suivante :

Définition 1.12 (matrice échelonnée)**.** Une matrice est appelée *matrice échelonnée* si elle a les propriétés suivantes :

(i) Dans toute ligne non nulle, le premier élément non nul vaut 1. Il est appelé le 1 directeur.

(ii) Les lignes dont tous les éléments sont nuls sont regroupées en bas de la matrice.

(iii) Dans deux lignes successives (contiguës) ayant des éléments non nuls, le 1 directeur de la ligne inférieure se trouve à droite du 1 directeur de la ligne supérieure.

Exemple 1.13. La matrice

$$\begin{pmatrix} 0 & 1 & -3 & -6 & 0 \\ 1 & 3 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 7 \end{pmatrix}$$

satisfait la propriété (i) mais pas la propriété (iii), alors que la matrice suivante

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

ne satisfait pas la propriété (i). En revanche, la matrice

$$\begin{pmatrix} 0 & 1 & 7 & -6 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

satisfait (i), (ii) et (iii) : elle est donc sous forme échelonnée. Finalement, la matrice

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

satisfait (i), (ii) mais pas (iii).

On peut raffiner un peu la définition précédente en posant :

Définition 1.14 (matrice échelonnée réduite)**.** Si, en plus des propriétés (i)-(iii) ci-dessus, la matrice satisfait la propriété (iv) ci-dessous, on parle de matrice *échelonnée réduite* :

(iv) Toute colonne contenant un 1 directeur a des zéros partout ailleurs.

Exemple 1.15. La matrice

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

est échelonnée réduite, alors que la matrice

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

est sous forme échelonnée non réduite (é cause de la 3-ème colonne).

On peut transformer n'importe quelle matrice en une matrice échelonnée (réduite) en utilisant l'algorithme de Gauss.

1.3.1 Algorithme d'élimination de Gauss

Cet algorithme permet de transformer n'importe quelle matrice sous sa forme échelonnée réduite à l'aide des opérations élémentaires (1)-(3) de $(*)$. Voici la marche à suivre illustrée par un exemple.

(1) Identifier la colonne se trouvant le plus à gauche contenant au moins un élément non nul.
Exemple :

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

$\uparrow 2^{\text{ème}}$ colonne

(2) Permuter, s'il le faut, la première ligne avec une autre, pour que l'élément en haut de la colonne identifiée en (1) devienne non nul.
Exemple (suite) :

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

$\ell_1 \longleftrightarrow \ell_2$

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

(3) Si l'élément se trouvant en haut de la dite colonne vaut a, multiplier la première ligne par $\frac{1}{a}$ pour y faire apparaître le 1 directeur.
Exemple (suite) :

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

$\ell_1 \longrightarrow \frac{1}{3}\ell_1$

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

(4) Ajouter des multiples adéquats de la première ligne aux lignes en-dessous pour annuler les éléments en dessous du 1 directeur.
Exemple (suite) :

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

$\ell_3 \longrightarrow \ell_3 - 3\ell_1$

$$\begin{pmatrix} 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

(5) Couvrir la première ligne de la matrice, et aller à (1)
Exemple (suite) :

$$\begin{pmatrix} 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

(4) $\ell_2 \longrightarrow \ell_2 - \ell_1$

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

(3) $\ell_2 \longrightarrow -\frac{1}{2}\ell_2$

$$\begin{pmatrix} 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(6) La matrice entiére est échelonnée.
Exemple (suite) : On remet la première ligne en place

$$\begin{pmatrix} 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(7) Pour la mettre sous la forme échelonnée réduite, il faut ajouter à une ligne des multiples adéquats des lignes situées au-dessous d'elle en allant du bas vers le haut.
Exemple (suite) :

$$\begin{pmatrix} 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$\ell_2 \longrightarrow \ell_2 - 3\ell_3$

$$\begin{pmatrix} 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$\ell_1 \longrightarrow \ell_1 - \frac{1}{3}\ell_3$

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

Les deux exemples ci-dessous illustrent encore l'algorithme. L'exemple 1.16 illustre le point (7) à partir d'une matrice qui est déjà sous forme échelonnée mais pas réduite. Dans l'exemple 1.17, on effectue l'algorithme dans son entier.

Exemple 1.16.

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

$\ell_2 \longrightarrow \ell_2 - 2\ell_3$

$$\begin{pmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

$\ell_1 \longrightarrow \ell_1 - 4\ell_2$

$$\begin{pmatrix} 1 & 0 & 0 & -29 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

Exemple 1.17.

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

(1) $\uparrow$ $1^{\text{ère}}$ colonne

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

(2) $\ell_1 \longleftrightarrow \ell_3$

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

$\uparrow$ $2^{\text{ème}}$ colonne

(4) $\ell_2 \longrightarrow \ell_2 - \ell_1$

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

On remet en place la première ligne pour obtenir

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

La matrice est maintenant sous forme échelonnée. Il reste à la mettre sous la forme échelonnée réduite.

(7) $\ell_2 \longrightarrow \ell_2 + 2\ell_3$

$$\begin{pmatrix} 1 & -3 & 2 & 0 & -1 \\ 0 & 1 & 0 & 10 & -5 \\ 0 & 0 & 1 & 5 & -4 \end{pmatrix}$$

(7) $\ell_1 \longrightarrow \ell_1 - 2\ell_3$

$$\begin{pmatrix} 1 & -3 & 0 & -10 & 7 \\ 0 & 1 & 0 & 10 & -5 \\ 0 & 0 & 1 & 5 & -4 \end{pmatrix}$$

(7) $\ell_1 \longrightarrow \ell_1 + 3\ell_2$

$$\begin{pmatrix} 1 & 0 & 0 & 20 & -8 \\ 0 & 1 & 0 & 10 & -5 \\ 0 & 0 & 1 & 5 & -4 \end{pmatrix}$$

La matrice est sous forme échelonnée réduite. Un système dont la matrice augmentée est sous forme échelonnée réduite est très simple à résoudre comme nous allons le voir ci-aprés.

1.3.2 Méthode de résolution d'un système d'équations linéaires

Après avoir mis la matrice augmentée du système sous forme échelonnée réduite, on procède selon les deux étapes suivantes.

(1) Donner une valeur arbitraire à chaque variable dont la colonne ne contient pas de 1 directeur. Ces variables sont les variables libres.

(2) Résoudre chaque équation en exprimant la variable correspondant au 1 directeur, appelée variable directrice, en fonction des autres variables.

Exemple 1.18. La matrice

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

est échelonnée réduite.
Elle correspond au système

$$\begin{cases} x & & & = -2 \\ & y & & = 4 \\ & & z & = 1. \end{cases}$$

Toutes les variables sont des variables directrices.
La solution est donc

$$x = -2, \qquad y = 4, \qquad z = 1.$$

Exemple 1.19. La matrice

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

est échelonnée réduite. Elle correspond au système

$$\begin{cases} 0x_1 + x_2 + 3x_3 + 0x_4 & = -1 \\ x_4 & = 5. \end{cases}$$

Les variables directrices sont x_2 et x_4 car les colonnes 2 et 4 de la matrice contiennent un 1 directeur, alors que x_1 et x_3 sont les variables libres.

Posons

$$x_1 = s, \qquad x_3 = t.$$

On obtient

$$x_2 = -1 - 3t, \qquad x_4 = 5$$

et l'ensemble des solutions du système est :

$$x_1 = s, \qquad x_2 = -1 - 3t, \qquad x_3 = t, \qquad x_4 = 5, \qquad \text{pour tout } t, s \in \mathbb{R}.$$

1.4 Systèmes homogènes d'équations linéaires

Un système homogène est un système dont les termes constants sont tous nuls. Il est de la forme

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\ \cdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & 0. \end{array}\right.$$

Tout système homogène d'équations linéaires est consistant, car il a au moins la solution dite triviale $x_1 = x_2 = \cdots = x_n = 0$. Un système homogène d'équations linéaires a ou bien une seule solution (la solution triviale), ou bien une infinité de solutions. En effet, supposons que le système admette la solution

$$x_1 = t_1, \ldots, x_n = t_n$$

avec au moins l'un des $t_i \neq 0$. Alors, pour un nombre réel k quelconque,

$$x_1 = kt_1, \ldots, x_n = kt_n$$

est aussi solution du système.

Théorème 1.20. *Tout système homogène d'équations linéaires dont le nombre d'inconnues est plus grand que le nombre d'équations a une infinité de solutions.*

Démonstration. Soit m le nombre de colonnes et n le nombre d'équations. La matrice augmentée du système a alors $m+1$ colonnes et n lignes. Il s'ensuit que sa forme échelonnée réduite doit comporter au moins une colonne sans 1 directeur. Supposons que ce soit la j-ème avec $1 \leq j \leq m$. Cette colonne correspond à une variable libre $x_j = s$ et il y a donc une infinité de solutions puisque le système est compatible. $\square$

Exemple 1.21. Considérons le système homogène

$$\left\{\begin{array}{rcrcrcrcrcl} 3x_1 & + & 3x_2 & - & 2x_3 & & & - & x_5 & = & 0 \\ -x_1 & - & x_2 & + & x_3 & + & 3x_4 & + & x_5 & = & 0 \\ 2x_1 & + & 2x_2 & - & x_3 & + & 2x_4 & + & 2x_5 & = & 0 \\ & & & & x_3 & + & 8x_4 & + & 4x_5 & = & 0 \end{array}\right.$$

Sa matrice augmentée est

$$\begin{pmatrix} 3 & 3 & -2 & 0 & -1 & 0 \\ -1 & -1 & 1 & 3 & 1 & 0 \\ 2 & 2 & -1 & 2 & 2 & 0 \\ 0 & 0 & 1 & 8 & 4 & 0 \end{pmatrix}$$

et la forme échelonnée réduite est

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 13 & 0 \\ 0 & 0 & 1 & 0 & 20 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Cette matrice correspond donc au système

$$\left\{\begin{array}{rcrrrcrcl} x_1 & + & x_2 & & & + & 13x_5 & = & 0 \\ & & & x_3 & & + & 20x_5 & = & 0 \\ & & & & x_4 & - & 2x_5 & = & 0 \end{array}\right.$$

Les variables directrices sont x_1, x_3 et x_4 alors que les variables libres sont x_2 et x_5. Posons alors

$$\begin{array}{rcl} x_2 & = & s \\ x_5 & = & t\,. \end{array}$$

On obtient

$$\begin{aligned} x_1 &= -s - 13t \\ x_3 &= -20t \\ x_4 &= 2t\,. \end{aligned}$$

L'ensemble des solutions est donc

$$x_1 = -s - 13t, \qquad x_2 = s, \qquad x_3 = -20t, \qquad x_4 = 2t\,, \qquad x_5 = t,$$

qui est bien infini.

Chapitre 2

Eléments du calcul matriciel

2.1 Quelques définitions et opérations

Définition 2.1 (Matrice)**.** Une *matrice (réelle)* A est un tableau rectangulaire de nombres (réels). Elle est dite de *taille* $m \times n$ si le tableau posséde m lignes et n colonnes. Les nombres du tableau sont appelés les éléments de A. L'élément situé à la i-ème ligne et à la j-ème colonne est noté a_{ij}. La matrice A est également notée

$$A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$$

ou plus simplement

$$A = (a_{ij})_{ij}$$

si le nombre de lignes et de colonnes est connu par ailleurs.

Exemple 2.2.

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

est une matrice 2×3 avec, par exemple, $a_{11} = 1$ et $a_{23} = 7$.

Si $n = m$, la matrice est dite *carrée.*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

matrice carrée $n \times n$

Dans le cas d'une matrice carrée, les éléments $a_{11}, a_{22}, \dots a_{nn}$ sont appelés *les éléments diagonaux.*

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Deux matrices sont égales lorsqu'elles ont la même taille et que les éléments correspondants sont égaux.

Définition 2.3 (Somme de deux matrices)**.** On peut définir la somme de deux matrices si elles sont de même taille. Soient A et$B$ deux matrices de taille $m \times n$. On définit leur somme $C = A + B$, de taille $m \times n$, par

$$c_{ij} = a_{ij} + b_{ij}.$$

En d'autres termes, on somme composante par composante.

Exemple 2.4.

$$A = \begin{pmatrix} 3 & -2 \\ 1 & 7 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 5 \\ 2 & -1 \end{pmatrix}, \qquad A + B = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}$$

$$A = \begin{pmatrix} 4 & 1 \\ -3 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 \\ 8 \end{pmatrix}, \qquad A + B \text{ n'est pas définie.}$$

La matrice (de taille $m \times n$) dont tous les éléments sont des zéros est appelée la *matrice nulle* et notée 0_{nm} ou plus simplement 0. C'est *l'élément neutre* pour l'addition, c'est-à-dire que $A + 0 = A$.

Définition 2.5 (Produit d'une matrice par un scalaire)**.** Le produit d'une matrice A par un scalaire k est formé en multipliant chaque élément de A par k. Il est noté kA.

Exemple 2.6. Soient

$$A = \begin{pmatrix} 3 & -2 \\ 0 & 6 \end{pmatrix} \qquad \text{et} \qquad k = -3.$$

$$\text{Alors} \qquad kA = \begin{pmatrix} -9 & 6 \\ 0 & -18 \end{pmatrix}.$$

La matrice $(-1)A$ est notée $-A$ et la différence $A - B$ est définie par $A + (-B)$.

Exemple 2.7.

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

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

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

2.2 Le produit matriciel

Le produit AB de deux matrices A et B est défini seulement si le nombre de colonnes de A est égal au nombre de lignes de B :

Définition 2.8 (Produit de deux matrices)**.** Soit $A = (a_{ij})$ une matrice $m \times n$ et $B = (b_{ij})$ une matrice $n \times p$. Alors le produit $C = AB$ est une matrice $m \times p$ dont les éléments c_{ij} sont définis par

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^{n} a_{ik}bkj.$$

$$\begin{array}{cc} AB = C : & \begin{pmatrix} b_{11} & \dots & b_{1r} \\ b_{21} & & \vdots \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mr} \end{pmatrix} \\ \begin{pmatrix} a_{11} & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nm} \end{pmatrix} & \begin{pmatrix} c_{11} & \dots & c_{1r} \\ c_{21} & & \vdots \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nr} \end{pmatrix} \end{array}$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2m}b_{m1}$$

Exemple 2.9.

$$\underset{1\times 3}{(2 \quad 3 \quad -1)}\underset{3\times 1}{\begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix}} = \underset{1\times 1}{(8-6-3) = (-1)}$$

Exemple 2.10.

$$\underset{3\times 1}{\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}} \underset{1\times 4}{\begin{pmatrix} -1 & 2 & 3 & 0 \end{pmatrix}} = \underset{3\times 4}{\begin{pmatrix} -3 & 6 & 9 & 0 \\ 2 & -4 & -6 & 0 \\ -1 & 2 & 3 & 0 \end{pmatrix}}$$

2.2.1 Matrice identité

Définition 2.11. La matrice carrée $n \times n$

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

s'appelle *la matrice identité*. Ses éléments diagonaux sont égaux à 1 et tous ses autres éléments sont égaux à 0.

Dans le calcul matriciel, la matrice identité joue un réle analogue à celui du nombre 1 dans l'arithmétique des scalaires. C'est l'élément neutre pour la multiplication. En d'autres termes, si A une matrice $m \times n$, alors

$$I_m A = A \qquad \text{et} \qquad AI_n = A.$$

2.3 Règles du calcul matriciel

Sous l' hypothèse que les tailles des matrices soient compatibles avec les opérations indiquées, on a les règles suivantes :

(a) Commutativité de la somme : $A + B = B + A$

(b) Associativité de la somme : $A + (B + C) = (A + B) + C$.

(c) Associativité du produit : $A(BC) = (AB)C$

(d) Distributivité du produit par rapport à la somme : $A(B + C) = AB + AC$
$(B + C)A = BA + CA$

(e) $A + 0 = A$

(f) $AI = IA = A$

(g) $A \cdot 0 = 0 \cdot A = 0$

ATTENTION !
Le produit des matrices n'est **pas nécessairement commutatif**. On peut avoir

$$AB \neq BA.$$

Exemple 2.12.

$$A = \begin{pmatrix} 1 & 0 \\ -2 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}$$

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

ATTENTION ! Il peut arriver que le produit de deux matrices non nulles soit nul. En d'autres termes, on peut avoir $A, B \neq 0$ et $AB = 0$.

Exemple 2.13.

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ce qui précède implique, par distributivité, que l'on peut avoir $AB = AC$ et $B \neq C$.

Exemple 2.14.

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & -1 \\ 5 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 5 \\ 5 & 4 \end{pmatrix}$$
$$AB = AC = \begin{pmatrix} -5 & -4 \\ 15 & 12 \end{pmatrix}.$$

2.4 Ecriture matricielle des systèmes d'équations linéaires

Le système linéaire

$$\left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ & \cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \right.$$

peut s'écrire sous forme matricielle :

$$\underbrace{\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{B}.$$

On appelle A la matrice des coefficients du système. Le vecteur x est une solution du système si et seulement si

$$Ax = B.$$

Théorème 2.15. *Un système d'équations linéaires n'a soit aucune solution, soit une seule solution, soit une infinité de solutions.*

Démonstration. Soit $Ax = B$ la représentation matricielle du système. On est nécessairement dans l'un des cas ci-dessous :

(a) le système est incompatible (aucune solution) ;

(b) le système a une seule solution ;

(c) le système a plusieurs solutions.

Pour démontrer le théorème , il suffit alors de montrer que dans le cas (c) il y a une infinité de solutions.

Soient x_1 et x_2 des solutions distinctes du système. Alors $Ax_1 = B$ et $Ax_2 = B$. Donc $Ax_1 - Ax_2 = 0$ et $A(x_1 - x_2) = 0$.

Posons

$$x_0 = x_1 - x_2\,.$$

On a $x_0 \neq 0$, car $x_1 \neq x_2$ et l'expression

$$x_1 + kx_0$$

est une solution du système pour tout nombre réel k. En effet, $A(x_1 + kx_0) = Ax_1 + kAx_0 = B + 0$. $\square$

Théorème 2.16. *Supposons que le système*

$$\left\{\begin{array}{rcl} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n & = & c_1 \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n & = & c_2 \\ \cdots & & \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n & = & c_m \end{array}\right.$$

détermine les variables $y_1, \dots, y_n$ en fonction de constantes $c_1, \dots, c_m$, et que le système

$$\left\{\begin{array}{rcl} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1p}x_p & = & y_1 \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2p}x_p & = & y_2 \\ \cdots & & \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{np}x_p & = & y_n \end{array}\right.$$

exprime les variables $x_1, \dots, x_p$ en fonction des variables $y_1, \dots, y_n$.

Écrivons ces systèmes sous la forme compacte

$$Ay = c, \qquad Bx = y.$$

Alors le système déterminant les variables $x_1, \dots, x_p$ en fonction des constantes $c_1, \dots, c_m$ est donné par

$$(AB)x = c.$$

Quelques cas particuliers

Dans le cas particulier ù $n = m = 2$ (2 équations à 2 inconnues) le système linéaire correspond à l'intersection de deux droites dans le plan. Nous avons vu, dans le chapitre 1, que trois cas pouvaient se présenter : les droites sont soit parallèles, soit sécantes, soit confondues et ces trois cas correspondent aux trois cas du théorème ci-dessus.

Si le système est homogène, les deux droites passent par le point $(0, 0)$ et ne peuvent donc être parallèles. Le cas sans solution est donc exclu.

Dans le cas ù l'on a 2 équations ($m = 2$) à 3 inconnues ($n = 3$), ceci correspond à l'intersection de deux plans dans l'espace. Trois cas se présentent alors :

- les plans sont parallèles et il n'y a alors aucune solution au système ;
- les plans sont confondus et il y a une infinité de solutions au système ;
- les plans se coupent en une droite et il y a une infinité de solutions ;

Du point de vue du nombre de solutions, nous constatons qu'il n' y a que deux possibilités, à savoir aucune solution ou une infinité de solutions. Mais les deux derniers cas ci-dessus sont néanmoins très différents géométriquement et il semblerait que dans le second cas (plans confondus), l'infinité de solutions soit plus grande que dans le troisième cas. Les chapitres suivants nous permettront de rendre rigoureuse cette impression.

2.5 L'inversion des matrices

Définition 2.17 (Matrice inverse)**.** Soit A une matrice carrée de taille $n \times n$. S'il existe une matrice carrée B de taille $n \times n$ telle que

$$AB = I \qquad \text{et} \qquad BA = I,$$

on dit que A est *inversible* et on appelle B un *inverse de A*.

(on verra plus tard qu'il suffit de vérifier une seule des conditions $AB = I$, $BA = I$)

Exemple 2.18. La matrice

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

est inversible et un de ses inverses est

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

En effet, on a

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Exemple 2.19. La matrice

$$A = \begin{pmatrix} 3 & 0 \\ 5 & 0 \end{pmatrix}$$

n'est pas inversible. En effet, soit

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

une matrice quelconque. Alors le produit

$$BA = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} * & 0 \\ * & 0 \end{pmatrix}$$

ne peut jamais être égal à la matrice identité.

Théorème 2.20. *Si B et C sont des inverses de A, alors*

$$B = C.$$

Démonstration. On a $I = AC = BA$ du fait que B et C sont des inverses de A ; donc

$$B = BI = B(AC) = (BA)C = IC = C. \qquad \square$$

Si A est une matrice inversible, son inverse est noté A^{-1}. On a donc

$$AA^{-1} = I \qquad \text{et} \qquad A^{-1}A = I.$$

2.5.1 Matrices 2×2

Considérons les matrices 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

On vérifie que

$$AB = BA = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Donc A est inversible si $ad - bc \neq 0$, et on a alors

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

2.5.2 Puissances d'une matrice

Soit A une matrice $n \times n$. On définit

$$A^m = \underbrace{AA \cdots A}_{\text{m facteurs}}$$

Si A est inversible, on définit

$$A^{-m} = \left(A^{-1}\right)^m = \underbrace{A^{-1}A^{-1} \cdots A^{-1}}_{\text{m facteurs}}$$

Théorème 2.21. *Soit A une matrice inversible. Alors* *(a) A^{-1} est inversible et $(A^{-1})^{-1} = A$;*

(b) A^m est inversible et

$$\begin{aligned}(A^m)^{-1} &= \underbrace{A^{-1}A^{-1}\cdots A^{-1}}_{m \text{ facteurs}} \\ &= (A^{-1})^m = A^{-m};\end{aligned}$$

(c) kA est inversible si $k \neq 0$ et $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Théorème 2.22. *Soient A et B deux matrices $n \times n$ inversibles. Alors*

(a) AB est inversible et

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

Démonstration. Il suffit de montrer

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

Cela suit de

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(AA^{-1})B = B^{-1}IB = B^{-1}B = I, \\ \text{et } (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I. \square\end{aligned}$$

De façon analogue, on montre que si $A_1, \ldots, A_m$ sont inversibles, alors

$$(A_1A_2\ldots A_m)^{-1} = A_m^{-1}A_{m-1}^{-1}\ldots A_1^{-1}.$$

Exemple 2.23.

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$$

$$B = \begin{pmatrix} -9 & -4 \\ 2 & 1 \end{pmatrix} \qquad B^{-1} = \begin{pmatrix} -1 & -4 \\ 2 & 9 \end{pmatrix}$$

$$AB = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} -9 & -4 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -16 & -7 \\ -39 & -17 \end{pmatrix}$$

$$B^{-1}A^{-1} = \begin{pmatrix} -1 & -4 \\ 2 & 9 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} 17 & -7 \\ -39 & 16 \end{pmatrix}$$

On a alors bien

$$(AB)(B^{-1}A^{-1}) = \begin{pmatrix} -16 & -7 \\ -39 & -17 \end{pmatrix} \cdot \begin{pmatrix} 17 & -7 \\ -39 & 16 \end{pmatrix} = \begin{pmatrix} -272+273 & 0 \\ 0 & 273-272 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

2.6 Les matrices élémentaires

Définition 2.24 (Matrice élémentaire)**.** On dit qu'une matrice E est *élémentaire* si elle peut être obtenue par une seule opération élémentaire sur les lignes de la matrice identité (voir §1.2 ($*$) pour la définition des opérations élémentaires).

Il existe donc trois types de matrices élémentaires correspondant aux trois opérations élémentaires.

(1) La matrice $E_i(c)$ est la matrice élémentaire obtenue en multipliant par c la i-ème ligne de I_n, ù c est un nombre réel non nul.

Exemple 2.25.

$$E_2(5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(2) La matrice E_{ij} est la matrice élémentaire obtenue en permutant les i-ème et j-ème lignes de I_n.

Exemple 2.26.

$$E_{24} = E_{42} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(3) La matrice $E_{ij}(c)$ est la matrice élémentaire obtenue en ajoutant c fois la j-ème ligne de I_n à la i-ème ligne.

Exemple 2.27.

$$E_{21}(-5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

L'opération élémentaire «permuter les lignes i et j» correspond à multiplier une matrice *sur la gauche* par la matrice élémentaire E_{ij} ; et de même pour toutes autres opérations élémentaires. C'est ce qu'indique le théorème suivant :

Théorème 2.28. *Si la matrice élémentaire E est le résultat d'une opération élémentaire effectuée sur I_m, alors pour toute matrice A de taille $m \times n$ le produit matriciel EA est égal à la matrice obtenue en effectuant la même opération élémentaire sur A.*

Ainsi, multiplier une matrice A sur la gauche par E_{ij} revient à échanger les lignes i et j de A ; multiplier A sur la gauche par $E_i(c)$ revient à multiplier la ligne i de A par c ; et multiplier A sur la gauche par $E_{ij}(c)$ revient à ajouter c fois la $i$ème ligne à la $j$éme.

Exemples :

(1)

$$E_1(2) \cdot A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 2a_{11} & 2a_{12} & 2a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

(2)

$$E_{23} \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \\ a_{21} & a_{22} \end{pmatrix}.$$

(3)

$$E_{21}(9) \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 9 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = \begin{pmatrix} a_{11} \\ 9a_{11} + a_{21} \\ a_{31} \end{pmatrix}.$$

Les opérations élémentaires sur les lignes sont réversibles. Ceci entraîne l'inversibilité des matrices élémentaires.

Théorème 2.29. *Toute matrice élémentaire est inversible. En particulier, on a :*

$$\begin{aligned} [E_{ij}(c)]^{-1} &= E_{ij}(-c) \\ E_{ij}^{-1} &= E_{ij} \\ [E_i(c)]^{-1} &= E_i\left(\frac{1}{c}\right). \end{aligned}$$

Exemple 2.30. On a

$$E_{21}(-4) = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix} \quad \text{et} \quad E_{21}(4) = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}$$

et

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

Définition 2.31. On dit que deux matrices sont *équivalentes par lignes* si l'une peut être obtenue à partir de l'autre par une suite d'opérations élémentaires sur les lignes.

Théorème 2.32. *Pour toute matrice A de taille $n \times n$, les affirmations suivantes sont équivalentes :*

(a) A est inversible.

(b) Le système $AX = B$ a une et une seule solution pour toute matrice B de taille $n \times 1$. Cette solution est donnée par $X = A^{-1}B$.

(c) $AX = 0$ n'a que la solution triviale $X = 0$.

(d) A est équivalente par lignes à I_n.

(e) A est un produit de matrices élémentaires.

Démonstration. (a) $\Rightarrow$ (b) Si A est inversible, on a les équivalences suivantes :

$$AX = B \iff A^{-1}AX = A^{-1}B \iff X = A^{-1}B$$

ce qui prouve (b).

(b) $\Rightarrow$ (c) C'est évident car (c) est un cas particulier de (b) avec $B = 0$.

(c) $\Rightarrow$ (d) Par hypothèse, le système $AX = 0$ est équivalent au système

$$\left\{\begin{array}{cccccc} X_1 & & & & = & 0 \\ & X_2 & & & = & 0 \\ \vdots & & & & \vdots & \\ & & & X_n & = & 0 \end{array}\right. .$$

La matrice associée à ce dernier système est la matrice identité. La matrice A est donc équivalente par lignes à I_n et ceci prouve le point (d).

(d) $\Rightarrow$ (e) On sait, par hypothèse, qu'une succession d'opérations élémentaires sur A conduit à la matrice I_n. Par le théorème 2.28, ceci signifie qu'il existe des matrices élémentaires $E_1, \ldots E_r$ telles que

$$E_r \cdot E_{r-1} \cdots E_1 \cdot A = I_n.$$

Comme une matrice élémentaire est inversible, ceci implique que

$$A = E_1^{-1} E_2^{-1} \cdots E_r^{-1}.$$

Mais l'inverse d'une matrice élémentaire est encore une matrice élémentaire et l'on a le résultat cherché.

(e) $\Rightarrow$ (a) Ceci découle du fait que toute matrice élémentaire est inversible et que le produit de matrices inversibles est encore inversible. □

2.7 Calcul de l'inverse d'une matrice

Le théorème précédent donne une méthode pour déterminer l'inverse d'une matrice inversible. La méthode consiste à faire les opérations élémentaires mettant la matrice A sous la forme échelonnée réduite, qui est I_n. On fait ensuite les mémes opérations élémentaires sur la matrice I_n. On aboutit alors à A^{-1}. En pratique, on fait les deux opérations en même temps selon la procédure suivante :

Former la matrice $(A : I)$ et effectuer sur cette matrice augmentée les opérations élémentaires mettant A sous la forme échelonnée réduite. On obtient alors la matrice $(I : A^{-1})$.

Exemple 2.33. Calculons l'inverse de

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

$$(A \quad : \quad I) = \left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 4 & 0 & -1 & : & 0 & 1 & 0 \\ -1 & 2 & 2 & : & 0 & 0 & 1 \end{array}\right)$$

$$\ell_2 := \ell_2 - 4\ell_1$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & -8 & -5 & : & -4 & 1 & 0 \\ -1 & 2 & 2 & : & 0 & 0 & 1 \end{array}\right)$$

$$\ell_3 := \ell_3 + \ell_1$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & -8 & -5 & : & -4 & 1 & 0 \\ 0 & 4 & 3 & : & 1 & 0 & 1 \end{array}\right)$$

$$\ell_2 := -\frac{1}{8}\ell_2$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 5/8 & : & 1/2 & -1/8 & 0 \\ 0 & 4 & 3 & : & 1 & 0 & 1 \end{array}\right)$$

$$\ell_3 := \ell_3 - 4\ell_2$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 5/8 & : & 1/2 & -1/8 & 0 \\ 0 & 0 & 1/2 & : & -1 & 1/2 & 1 \end{array}\right)$$

$$\ell_3 := 2\ell_3$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 5/8 & : & 1/2 & -1/8 & 0 \\ 0 & 0 & 1 & : & -2 & 1 & 2 \end{array}\right)$$

$$\ell_2 := \ell_2 - \frac{5}{8}\ell_3$$

$$\left(\begin{array}{rrrcrrr} 1 & 2 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & : & -2 & 1 & 2 \end{array}\right)$$

$$\ell_1 := \ell_1 - 2\ell_2 - \ell_3$$

$$\left(\begin{array}{rrrcrrr} 1 & 0 & 0 & : & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & : & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & : & -2 & 1 & 2 \end{array}\right)$$

$$A^{-1} = \frac{1}{4}\left(\begin{array}{rrr} -2 & 2 & 2 \\ 7 & -3 & -5 \\ -8 & 4 & 8 \end{array}\right)$$

2.8 Matrices triangulaires

Définition 2.34. Soit A une matrice de taille $n \times n$. On dit que A est *triangulaire inférieure* si ses éléments au dessus de la diagonale sont nuls, autrement dit si

$$i < j \Longrightarrow a_{ij} = 0.$$

Une matrice triangulaire inférieure a donc la forme suivante :

$$\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

On dit que A est *triangulaire supérieure* si ses éléments en dessous de la diagonale sont nuls, autrement dit si

$$i > j \Longrightarrow a_{ij} = 0.$$

Une matrice triangulaire supérieure a donc la forme suivante :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & \dots & a_{2n} \\ \vdots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & a_{nn} \end{pmatrix}$$

Exemple 2.35. Matrices triangulaires inférieures :

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

Exemple 2.36. Matrices triangulaires supérieures :

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

Définition 2.37. Une matrice qui est triangulaire inférieure et triangulaire supérieure est dite *diagonale*.

Exemple 2.38. Exemples de matrices diagonales :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

Théorème 2.39. *Une matrice A de taille $n \times n$, triangulaire, est inversible si et seulement si ses éléments diagonaux sont tous non nuls.*

Démonstration. Supposons que A est triangulaire supérieure.

Si les éléments de la diagonale sont tous non nuls, alors, en multipliant chaque ligne i par l'inverse de l'élément diagonal a_{ii}, on obtient la forme échelonnée de A. Elle ne contient que des 1 sur la diagonale. De ce fait, la forme échelonnée réduite de A sera la matrice identité. Le théorème 2.32 permet de conclure que A est inversible.

Inversement, supposons qu'au moins l'un des éléments diagonaux soit nul et notons a_{mm} le premier élément nul de la diagonale. En multipliant les lignes 1 à $m-1$ par l'inverse de leur élément diagonal, on obtient une matrice de la forme

$$\begin{pmatrix} 1 & * & \cdots & & & \cdots & * \\ 0 & \ddots & * & \cdots & & \cdots & * \\ 0 & 0 & 1 & * & & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & a_{ll} & \cdots & * \\ \vdots & \vdots & \vdots & \cdots & 0 & \ddots & \vdots \\ 0 & \cdots & & & \cdots & 0 & a_{nn} \end{pmatrix}$$

où $l = m+1$.

Il est alors clair que la colonne m de la forme échelonnée ne contiendra pas de 1 directeur. La forme échelonnée réduite de A ne peut donc pas être I_n et par le théorème 2.32, A n'est pas inversible.

Dans le cas d'une matrice triangulaire inférieure, on utilise la transposition qui fait l'objet de la section suivante et on obtient une matrice triangulaire supérieure. On applique alors la démonstration ci-dessus. □

2.9 La transposition

Soit A la matrice de taille $m \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Définition 2.40. On appelle matrice transposée de A, la matrice A^T de taille $n \times m$ définie par :

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

Autrement dit, la i-ème colonne de A^T est la i-ème ligne de A, ou encore

$$a^T_{ij} = a_{ji}.$$

Exemple 2.41.

$$\begin{aligned} (1 \quad -2 \quad 5)^T &= \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} \\ \begin{pmatrix} 0 & 3 \\ 1 & -5 \\ -1 & 2 \end{pmatrix}^T &= \begin{pmatrix} 0 & 1 & -1 \\ 3 & -5 & 2 \end{pmatrix} \\ \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}^T &= \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}. \\ (4)^T &= (4). \end{aligned}$$

Théorème 2.42. *L'opération de transposition obéit aux règles suivantes :*

(a) $(A+B)^T = A^T + B^T$

(b) $(kA)^T = kA^T$

(c) $(AB)^T = B^T A^T$

(d) $(A^T)^T = A$.

(e) Si A est inversible, alors A^T l'est aussi et on a $(A^T)^{-1} = (A^{-1})^T$ qui sera notée A^{-T}.

2.10 La trace

Soit A la matrice $n \times n$

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

Définition 2.43. On appelle *trace de* A,et on note $\mathrm{trace}(A)$, le nombre obtenu en additionnant les éléments diagonaux de A. Autrement dit,

$$\mathrm{trace}(A) = a_{11} + \dots + a_{nn}.$$

Exemple 2.44. Soient

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix} \quad et \quad B = \begin{pmatrix} 1 & 1 & 2 \\ 5 & 2 & 8 \\ 11 & 0 & -10 \end{pmatrix}$$

Alors $\mathrm{trace}(A) = 2 + 5 = 7$ et $\mathrm{trace}(B) = 1 + 2$ -10 =-7.

Théorème 2.45. *Soient A et B deux matrices $n \times n$. Alors*

(a) $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$ *;*

(b) $\mathrm{trace}(\lambda A) = \lambda\ \mathrm{trace}(A)$ *pour tout* $\lambda \in \mathbb{R}$ *;*

(c) $\mathrm{trace}(A^T) = \mathrm{trace}(A)$ *;*

(d) $\mathrm{trace}(AB) = \mathrm{trace}(BA)$*.*

Démonstration. (a) Pour tout $1 \le i \le n$, $(A + B)_{ii} = A_{ii} + B_{ii}$. Ainsi, on a bien $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$.

(b) On a $\mathrm{trace}(\lambda A) = \lambda A_{11} + \dots + \lambda A_{nn} = \lambda(A_{11} + \dots + A_{nn})$.

(c) Etant donné que la transposition ne change pas les éléments diagonaux, la trace de A est égale à la trace de A^T.

(d) On a

$$AB_{ii} = A_{i1}B_{1i} + A_{i2}B_{2i} + \dots + A_{in}B_{ni}.$$

Ainsi,

$$\begin{array}{rcllll} \mathrm{trace}(AB) & = & A_{11}B_{11} & +A_{12}B_{21} & +\dots & +A_{1n}B_{n1} \\ & + & A_{21}B_{12} & +A_{22}B_{22} & +\dots & +A_{2n}B_{n2} \\ & & \vdots & & & \\ & + & A_{n1}B_{1n} & +A_{n2}B_{2n} & +\dots & +A_{nn}B_{nn} \end{array}$$

On peut réarranger les termes pour obtenir

$$\begin{array}{clllll} & A_{11}B_{11} & +A_{21}B_{12} & +\dots & +A_{n1}B_{1n} \\ + & A_{12}B_{21} & +A_{22}B_{22} & +\dots & +A_{n2}B_{2n} \\ & \vdots & & & \\ + & A_{1n}B_{n1} & +A_{2n}B_{n2} & +\dots & +A_{nn}B_{nn} \end{array}$$

En utilisant la commutativité de la multiplication dans $\mathbb{R}$, la premiére ligne devient

$$B_{11}A_{11} + B_{12}A_{21} + \dots + B_{1n}A_{n1}$$

qui vaut BA_{11}. En faisant de même avec les autres lignes, on voit finalement que

$$\mathrm{trace}(AB) = BA_{11} + \dots + BA_{nn} = \mathrm{trace}(BA).$$

□

2.11 Matrices symétriques

Définition 2.46. Une matrice A de taille $n \times n$ est dite *symétrique* si elle est égale à sa transposée, c'est-à-dire si

$$A = A^T$$

ou encore si $a_{ij} = a_{ji}$ pour tout $i, j = 1, \ldots, n$.

Exemple 2.47. Les matrices

$$\begin{pmatrix} -1 & 0 & 5 \\ 0 & 2 & -1 \\ 5 & -1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix}, \qquad I_n \qquad \text{et} \qquad 0_{n,n}$$

sont symétriques.

Théorème 2.48. *Pour une matrice B quelconque, les matrices BB^T et B^TB sont symétriques.*

Démonstration. Par le théorème 2.42, on a

$$\begin{array}{rcccl} (BB^T)^T & = & (B^T)^T B^T & = & BB^T \\ (B^TB)^T & = & B^T(B^T)^T & = & B^TB. \end{array}$$

□

2.12 Matrices antisymétriques

Définition 2.49. Une matrice A de taille $n \times n$ est dite *antisymétrique* si

$$A^T = -A$$

c'est-à-dire si $a_{ij} = -a_{ji}$ pour tout $i, j = 1, \ldots, n$.

Exemple 2.50.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 4 & 2 \\ -4 & 0 & -5 \\ -2 & 5 & 0 \end{pmatrix}.$$

Remarquons que les éléments diagonaux d'une matrice antisymétrique sont toujours tous nuls.

Théorème 2.51. *Toute matrice A de taille $n \times n$ est la somme d'une matrice symétrique B et d'une matrice antisymétrique C.*

Démonstration. Posons $B = (A + A^T)/2$ et $C = (A - A^T)/2$. On a alors $A = B + C$; et B est symétrique, car $B^T = (A^T + (A^T)^T)/2 = B$; et C est antisymétrique, car $C^T = (A^T - (A^T)^T)/2 = -C$. □

Exemple 2.52. Soit

$$A = \begin{pmatrix} 2 & 10 \\ 8 & -3 \end{pmatrix}$$

$$A = \underbrace{\begin{pmatrix} 2 & 9 \\ 9 & -3 \end{pmatrix}}_{\text{symétrique}} + \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\text{antisymétrique}}.$$

Chapitre 3

Le déterminant

3.1 Permutations et déterminants

Nous allons construire dans ce chapitre une fonction — appelée *le déterminant* — qui associe un nombre réel à chaque matrice carrée et qui permettra de caractériser facilement les matrices inversibles puisque ce sont celles dont le déterminant est non nul.

Exemple 3.1. Soit

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

On a vu que si $ad - bc \neq 0$, alors A est inversible. On a alors

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

On définit alors le déterminant de A comme étant

$$\det(A) = ad - bc.$$

On va maintenant généraliser cette notion à des matrices carrées de taille $n \times n$.

Définition 3.2. On appelle *permutation* de l'ensemble d'entiers $\{1, \ldots, n\}$ un arrangement de ceux-ci sans omissions ni répétitions. Autrement dit, une permutation de $\{1, \ldots, n\}$ est une bijection de l'ensemble $\{1, \ldots, n\}$ sur lui-même.

Une permutation quelconque σ de $\{1, \ldots, n\}$ sera notée

$$\sigma = (j_1, j_2, \ldots, j_n)$$

où $j_1 = \sigma(1)$, $j_2 = \sigma(2)$,..., $j_n = \sigma(n)$.
L'ensemble de toutes les permutations de n éléments sera noté $\mathcal{S}_n$.

Exemple 3.3. Il y a deux permutations de l'ensemble $\{1, 2\}$:

$$\sigma_1 = (1, 2) \qquad \text{et} \qquad \sigma_2 = (2, 1)$$

$(1, 2)$ est l'identité car $\sigma_1(1) = 1$ et $\sigma_1(2) = 2$.

Exemple 3.4. Il y a 6 permutations de l'ensemble $\{1, 2, 3\}$:

$$(1, 2, 3) \quad (1, 3, 2) \quad (2, 1, 3) \quad (3, 2, 1) \quad (2, 3, 1) \quad (3, 1, 2).$$

Plus généralement, l'ensemble $\{1, \ldots, n\}$ a

$$n! = 1 \cdot 2 \cdots n$$

permutations. En effet, il y a n possibilités pour le premier nombre, $n - 1$ possibilités pour le deuxième et ainsi de suite ce qui nous donne $n(n-1)(n-2) \cdots 2 \cdot 1$ différents arrangements possibles des nombres $1, 2, \ldots, n$.

Définition 3.5. Dans une permutation on a une *inversion* si un nombre plus grand précède un nombre plus petit.

De manière plus précise, le nombre d'inversions d'une permutation

$$(j_1, j_2, \ldots, j_n)$$

est la somme du
- nombre de successeurs de j_1 plus petits que j_1, plus
- le nombre de successeurs de j_2 plus petits que j_2, plus
- ...
- le nombre de successeurs de j_{n-1} plus petits que j_{n-1}.

Exemple 3.6. La permutation

$$(4, 2, 5, 3, 1) \qquad \text{contient 7 inversions.}$$

En effet,il y a 3 successeurs plus petits que 4, 1 successeur plus petit que 2, 2 successeurs plus petits que 5, 1 successeur plus petit que 3 et pas de successeur plus petit que 1. En additionnant ces nombres, on obtient bien 7.

Exemple 3.7. La permutation

$$(6, 1, 3, 4, 5, 2)$$

contient $5 + 0 + 1 + 1 + 1 = 8$ inversions.

Définition 3.8. Une permutation ayant un nombre pair d'inversions est appelée *permutation paire*, sinon elle est appelée *permutation impaire*. On définit la *signature* de la permutation σ comme suit :

$$\operatorname{sign}(\sigma) = \begin{cases} 1 & \text{si } \sigma \text{ est paire} \\ -1 & \text{si } \sigma \text{ est impaire.} \end{cases}$$

Exemple 3.9. Classification des permutations de $\{1, 2, 3\}$:

Permutation	Nbre d'inversions	Parité
$(1, 2, 3)$	0	paire
$(1, 3, 2)$	1	impaire
$(2, 1, 3)$	1	impaire
$(2, 3, 1)$	2	paire
$(3, 1, 2)$	2	paire
$(3, 2, 1)$	3	impaire

Lemme 3.10. *Soit $n \geq 1$, $i, j \in \{1, \ldots, n\}$ avec $i < j$ et $\sigma \in \mathcal{S}_n$.*
Posons $\sigma' = (\sigma(1), \ldots, \sigma(i-1), \sigma(j), \sigma(i+1), \ldots, \sigma(j-1), \sigma(i), \sigma(j+1), \ldots, \sigma(n))$. Alors

$$\operatorname{sign}(\sigma) = -\operatorname{sign}(\sigma').$$

«Démonstration» : Nous illustrons la méthode de la démonstration par un cas particulier. Considérons les deux permutations de $\mathcal{S}_8$ suivantes :

$$\sigma = (1, 2, 5, 7, 6, 3, 4, 8) \qquad \text{et} \qquad \sigma' = (1, 3, 5, 7, 6, 2, 4, 8).$$

Pour calculer leur signature, il faut calculer le nombre d'inversions de σ et de σ'. On voit que 1, 4 et 8 ont le même nombre de successeurs plus petits dans σ et σ'.
Pour passer de σ à σ', on permute 2 et 3. Dans σ, 3 n'est pas un successeur de 2 plus petit, alors que dans σ', 2 est un successeur de 3 plus petit.
Dans σ, 5 n'est pas un successeur de 2 plus petit, mais 3 est un successeur de 5 plus petit,alors que dans σ', 5 n'est pas un successeur de 3 plus petit, mais 2 est un successeur de 5 plus petit. En répétant le même raisonnement avec 7 et 6, on remarque que le nombre de successeurs de 5, 7 et 6 plus petits est le même que cela soit dans σ ou dans σ'. Globalement, on voit donc que σ' a une et une seule inversion de plus que σ. Ainsi, leurs signatures sont opposées.

□

Définition 3.11. Soit

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

une matrice carrée de taille $n \times n$. Un *produit élémentaire de* A est le produit de n éléments de A, choisis de façon qu'aucun couple d'entre eux ne provienne de la méme ligne ou de la même colonne. Autrement dit, tous les éléments du produit sont dans des lignes et des colonnes différentes.

Exemple 3.12. Les produits élémentaires de la matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

sont $a_{11}a_{22}$ et $a_{12}a_{21}$. Les produits élémentaires de

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

sont :

$$\begin{array}{ccc} a_{11}a_{22}a_{33} & a_{11}a_{32}a_{23} & a_{21}a_{12}a_{33} \\ a_{21}a_{32}a_{13} & a_{31}a_{12}a_{23} & a_{31}a_{22}a_{13} \end{array}$$

Plus généralement, à partir d'une matrice de taille $n \times n$, on peut former $n\,!$ produits élémentaires. En effet, on constate qu'un produit élémentaire de A n'est rien d'autre qu'un produit $a_{1j_1}a_{2j_2}\dots a_{nj_n}$ où $(j_1, j_2, \dots, j_n)$ est un élément de $\mathcal{S}_n$.

Définition 3.13. Un *produit élémentaire signé* d'une matrice A est un produit

$$\text{sign}(\sigma) \cdot a_{1j_1} \cdot a_{2j_2} \dots a_{nj_n}$$

où $\sigma = (j_1, j_2, \dots, j_n)$ est une permutation à n éléments.

Exemple 3.14. Les produits élémentaires signés de

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

sont $a_{11}a_{22}$ (la permutation (1, 2) est paire) et $-a_{21}a_{12}$ (la permutation (2,1) est impaire).

Les produits élémentaires signés de

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

sont $a_{11}a_{22}a_{33}$, $-a_{11}a_{23}a_{32}$, $-a_{12}a_{21}a_{33}$, $a_{12}a_{23}a_{31}$, $a_{13}a_{21}a_{32}$ et $-a_{13}a_{22}a_{31}$.

Définition 3.15. Le *déterminant* d'une matrice A est le nombre obtenu en effectuant la somme de tous les produits élémentaires signés de A. Il est noté $\det(A)$. Autrement dit,

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathcal{S}_n} \text{sign}(\sigma) \quad a_{1i_1} \dots a_{ni_n} \,, \\ \text{où} \quad \sigma &= (i_1, \dots, i_n). \end{aligned}$$

Exemple 3.16.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} \,.$$

Exemple 3.17.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{aligned} \det(A) &= \operatorname{sign}((1,2,3))a_{11}a_{22}a_{33} + \operatorname{sign}((2,3,1))a_{12}a_{23}a_{31} + \operatorname{sign}((3,1,2))a_{13}a_{21}a_{32} \\ &+ \operatorname{sign}((3,2,1))a_{13}a_{22}a_{31} + \operatorname{sign}((2,1,3))a_{12}a_{21}a_{33} + \operatorname{sign}((1,3,2))a_{11}a_{23}a_{32} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{aligned}$$

Théorème 3.18. *Si A est une matrice ayant une ligne formée de zéros, alors* $\det(A) = 0$.

Démonstration. Par définition, le déterminant est la somme des produits élémentaires signés de A. Mais chacun de ces produits élémentaires contient un élément nul provenant de la ligne de zéros de A. Donc $\det(A) = 0$. □

Théorème 3.19. *Le déterminant d'une matrice A triangulaire (inférieure ou supérieure) est égal au produit $a_{11}a_{22}a_{33}\ldots a_{nn}$ des éléments diagonaux.*

Démonstration. Le seul produit élémentaire signé non nul est $a_{11}a_{22}\ldots a_{nn}$. La permutation correspondante est $(1, 2, \ldots, n)$ qui contient 0 inversions et qui est donc une permutation paire. On a donc bien

$$\det(A) = a_{11}a_{22}a_{33}\ldots a_{nn}.$$

□

Examinons maintenant ce qui se passe si deux lignes de la matrice sont égales.

Exemple 3.20. Soit

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

on a

$$\det(A) = abf - abf + ace - ace + bcd - bcd = 0$$

et l'on remarque que tous les produits élémentaires apparaissent deux fois avec des signes opposés (cf. lemme 3.10).

Ceci nous amène au théorème suivant :

Théorème 3.21. *Soit A une matrice avec deux lignes égales. Alors*

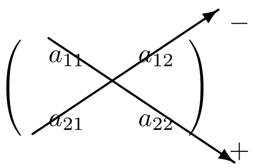
$$\det(A) = 0.$$

Démonstration. Dans le déterminant d'une telle matrice, tous les produits élémentaires apparaissent deux fois, avec des signes opposés. Donc $\det(A) = 0$. □

3.1.1 Méthode pour calculer des déterminants de matrices de taille 2×2 et 3×3

Nous décrivons ici la règle de Sarus pour calculer des déterminants 2×2 et 3×3.

Matrice 2×2



$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Matrice 3×3

On recopie les colonnes 1 et 2 à la suite de la colonne 3 et on calcule comme suit :

$$\begin{array}{ccccc} & & - & - & - \\ \left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right. & & & \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array} \\ & & + & + & + \end{array}$$

$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$.

Exemple 3.22. Calculer

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

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

donc $\det = -1$.

ATTENTION : Cette méthode ne s'applique pas pour les matrices de dimensions supérieures é 3.

3.2 Déterminants et opérations élémentaires

Nous allons voir que la réduction d'une matrice à la forme échelonnée nous fournit une méthode efficace pour calculer son déterminant.

Théorème 3.23. *Soit A une matrice de taille $n \times n$, et soit E une matrice élémentaire ($E = E_i(k), E_{ij}$ ou $E_{ij}(k)$). Alors*

(1) $\det(E_i(k)) = k$

(2) $\det(E_{ij}) = -1$

(3) $\det(E_{ij}(k)) = 1$

(4) $\det(EA) = \det(E)\det(A)$

Démonstration. (1) Soit $k \in \mathbb{R}$, $k \neq 0$. Rappelons que

$$E_i(k) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & 1 & & & & \vdots \\ \vdots & & & k & & & \vdots \\ \vdots & & & & 1 & & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} \det(E_i(k)) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) \quad a_{1j_1} \dots a_{nj_n}, \\ \text{où} \quad \sigma &= (j_1, \dots, j_n). \end{aligned}$$

Comme il n'y a qu'un seul élément non nul dans chaque ligne et dans chaque colonne, le seul produit élémentaire non nul est

$$1 \cdots 1\, k\, 1 \cdots 1 = k.$$

De plus, la permutation $(1, 2, \ldots, n)$ n'a pas d'inversion. Sa signature est donc 1. Ainsi

$$\det(E_i(k)) = k.$$

(2) Sans perte de généralité, on peut supposer que $i < j$. On a

$$E_{ij} = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 0 &&& 1 &&& \\ &&& 1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& 1 &&& \\ && 1 &&& 0 &&& \\ &&&&&& 1 && \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{pmatrix}$$

Comme avant, il y a un seul produit produit élémentaire non nul, qui vaut 1. Le déterminant sera donc ± 1. Il reste à déterminer la signature de la permutation

$$(1, 2, \ldots, (i-1), j, (i+1), (i+2), \ldots, (j-1), i, (j+1), \ldots, n).$$

j a $(j-1)$ successeurs plus petits. Les nombres compris entre $(i+1)$ et $(j-1)$ ont chacun un succeseur plus petit. Or, il y a $(j-1)-1$ nombres entre i et j. Ainsi, le nombre d'inversions de la permutation est

$$(j-i) + (j-i) - 1 = 2j - 2i - 1.$$

C'est un nombre impair. La signature est donc -1 et le déterminant de E_{ij} est -1.

(3) En écrivant la matrice $E_{ij}(k)$, on voit que le seul produit élémentaire non nul est 1 et que la signature de la permutation à étudier est celle de $(1, 2, \ldots, n)$. C'est une permutation paire, ce qui implique que $\det(E_{ij}(k)) = 1$.

(4) Pour montrer que $\det(EA) = \det(A)\det(E)$, nous allons considérer trois cas, $E = E_i(k)$, $E = E_{ij}$ et $E = E_{ij}(k)$.
Premier cas $E = E_i(k)$, $k \neq 0$ et $A = (a_{ij})$.

$$EA = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \ldots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Le déterminant est la somme des produits élémentaires signés. Chaque produit élémentaire a exactement un élément de chaque ligne,en particulier un élément de la i-ème ligne. Ainsi, dans chaque terme de la somme, on peut mettre k en évidence. Finalement,

$$\begin{aligned} \det(EA) &= k \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)} \\ &= \det(E)\det(A). \end{aligned}$$

Deuxième cas $E = E_{ij}$. On peut supposer que $i < j$. Posons $B = E_{ij}A$. On a

$$E_{ij}A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Les produits élémentaires de B et de A sont les mêmes. Comme $\det(E_{ij}) = -1$, pour montrer que $\det(E_{ij}A) = \det(E_{ij})\det(A)$ il faut montrer que les produits élémentaires signés de A et de B sont opposés. Par définition du déterminant,

$$\begin{aligned} \det(B) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)\, b_{1\sigma(1)} \dots b_{i\sigma(i)} \dots b_{j\sigma(j)} \dots b_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} \dots a_{j\sigma(i)} \dots a_{i\sigma(j)} \dots a_{n\sigma(n)} \end{aligned}$$

La deuxième égalité vient du fait que la i-ème ligne de B est la j-ème ligne de A (et réciproquement). Posons

$$\sigma' = (\sigma(1), \dots, \sigma(i-1), \sigma(j), \sigma(i+1), \dots, \sigma(j-1), \sigma(i), \sigma(j+1), \dots, \sigma(n)).$$

σ' est la composition de σ avec la permutation qui échange i et j. Par le lemme 3.10, $\operatorname{sign}(\sigma') = -\operatorname{sign}(\sigma)$. Ainsi

$$\begin{aligned} \det(B) &= \sum_{\sigma' \in \mathcal{S}_n} (-\operatorname{sign}(\sigma'))\, a_{1\sigma'(1)} \dots a_{n\sigma'(n)} \\ &= -\sum_{\sigma' \in \mathcal{S}_n} \operatorname{sign}(\sigma')\, a_{1\sigma'(1)} \dots a_{n\sigma'(n)} \\ &= -\det(A). \end{aligned}$$

Troisième cas $E = E_{ij}(k)$. On peut supposer que $i < j$. Posons $C = E_{ij}(k)A$. On a

$$C = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + ka_{j1} & \dots & a_{in} + ka_{jn} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Alors,

$$\begin{aligned} \det(C) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)\, b_{1\sigma(1)} \dots b_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} \dots a_{(i-1)\sigma(i-1)}(a_{i\sigma(i)} + ka_{j\sigma(i)})a_{(i+1)\sigma(i+1)} \dots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} \dots a_{n\sigma(n)} \\ &+ \underbrace{k \sum_{\sigma \in \mathcal{S}_n} \operatorname{sign}(a_{1\sigma(1)} \dots a_{(i-1)\sigma(i-1)})a_{j\sigma(i)}a_{(i+1)\sigma(i+1)} \dots a_{n\sigma(n)}}_{=\alpha} \\ &= \det(A) + k\,\alpha \end{aligned}$$

Comme $\det(E_{ij}(k)) = 1$, pour montrer que $\det(C) = \det(A)\det(E_{ij}(k))$, il suffit de montrer que $\alpha = 0$. Mais,

$$\alpha = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{(i-1)1} & \dots & a_{(i-1)n} \\ a_{j1} & \dots & a_{jn} \\ a_{(i+1)1} & \dots & a_{(i+1)n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

et la i-ème ligne de cette matrice est $a_{j1} \dots a_{jn}$. Elle a donc deux lignes identiques (la i-ème et la j-ème), ce qui implique que $\alpha = 0$.

□

Ce théorème nous permet de calculer le déterminant d'une matrice de façon relativement simple, en utilisant l'algorithme de Gauss pour réduire la matrice à la forme échelonnée (qui est triangulaire) et en utilisant les théorèmes 3.23 et 3.19. En effet, si

$$A = E_1 \dots E_r D$$

où $D = (d_{ij})$ est une matrice échelonnée (triangulaire). Alors

$$\det(A) = \det(E_1)\dots\det(E_r)\det(D) = \det(E_1)\dots\det(E_r)d_{11}d_{22}\dots d_{nn}.$$

Exemple 3.24. Calculer $\det(A)$, où

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix} \\ \det(A) &= \det \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix} \\ &= (-1)\det \begin{pmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{pmatrix} \\ &= (-3)\det \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{pmatrix} \\ &= (-3)\det \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{pmatrix} \\ &= (-3)\det \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{pmatrix} \\ &= (-3)(-55)\det \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \\ &= (-3)(-55) = 165. \end{aligned}$$

Exemple 3.25. Soit

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

Alors, en soustrayant deux fois la ligne 1 à la ligne 2, on obtient

$$\begin{aligned}\det(A) &= \det\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 0 & 0 & 0 \\ 3 & 9 & 1 & 5 \\ 1 & 1 & 4 & 8 \end{pmatrix} \\ &= 0.\end{aligned}$$

Notation : Pour toute matrice carrée A, on note

$$|A| = \det(A).$$

Théorème 3.26. *Soit A une matrice carrée. Alors A est inversible si et seulement si* $\det(A) \neq 0$. *On a, dans ce cas,*

$$\det(A^{-1}) := \frac{1}{\det(A)}.$$

Démonstration. Supposons d'abord que A est inversible. On peut alors écrire A comme produit de matrices élémentaires, $A = E_1 \cdots E_r$. En appliquant successivement le théorème 3.23, on a

$$\det(A) = \det(E_1 \cdots E_r) = \det(E_1)\det(E_2 \cdots E_r) = \cdots = \det(E_1)\cdots\det(E_r).$$

Comme le déterminant d'une matrice élémentaire n'est jamais nul, on en déduit que le déterminant de A n'est pas nul.

Ensuite, $A^{-1} = E_r^{-1} \cdots E_1^{-1}$, et on vérifie aisément que $\det(E^{-1}) = \det(E)^{-1}$ pour toute matrice élémentaire E. On a donc

$$\det(A^{-1}) = \det(E_r)^{-1} \cdots \det(E_1)^{-1},$$

et donc $\det(A^{-1}) = \det(A)^{-1}$.

Réciproquement, supposons que $\det(A) \neq 0$. Nous montrerons qu'alors A est équivalente par lignes à I, ce qui implique, par le théorème 2.32, que A est inversible.

Soit R la forme échelonnée réduite de A. On peut donc trouver des matrices élémentaires $E_1, \cdots, E_k$ telles que

$$\begin{aligned} E_k \cdots E_1 A &= R, \qquad \text{ou encore} \\ A &= E_1^{-1} \cdots E_k^{-1} R\,. \end{aligned}$$

On en déduit que

$$|A| = |E_1^{-1}| \cdots |E_k^{-1}|\ |R|\,.$$

Mais par hypothèse $|A| \neq 0$. Donc $|R| \neq 0$.

On en déduit que chaque ligne de R contient un 1 directeur. Donc $R = I$. □

Le théorème suivant est essentiel et nous affirme que le déterminant est multiplicatif :

Théorème 3.27. *Soient A et B deux matrices de taille $n \times n$. Alors*

$$\det(AB) = \det(A) \cdot \det(B)\,.$$

Démonstration. Si A et B sont les deux inversibles, on les écrit comme produit de matrices élémentaires : $A = E_1 \dots E_r$ et $B = F_1 \dots F_s$. On a alors

$$\begin{aligned}\det(AB) &= \det(E_1 \cdots E_r F_1 \cdots F_s) = \det(E_1)\det(E_2 \cdots E_r F_1 \dots F_s) \\ &= \cdots = \det(E_1)\cdots\det(E_r)\det(F_1 \cdots F_s) \\ &= \det(E_1 \cdots E_r)\det(F_1 \cdots F_s) = \det(A)\det(B).\end{aligned}$$

Si A ou B n'est pas inversible, alors AB n'est pas inversible non plus ; et $\det(AB) = 0$. □

Le déterminant est invariant par transposition :

Théorème 3.28. *Soit A une matrice de taille $n \times n$. Alors*

$$\det(A) = \det(A^T).$$

Autrement dit, le déterminant d'une matrice est égal à celui de sa transposée.

Démonstration. La démonstration se fait comme ci-dessus : supposons d'abord que A est inversible. On peut alors l'écrire comme produit de matrices élémentaires, $A = E_1 \cdots E_r$. On a alors

$$A^T = E_r^T \cdots E_1^T$$

et

$$|A^T| = |E_r^T| \cdots |E_1^T| = |E_r| \cdots |E_1| = |A|.$$

D'autre part, si A n'est pas inversible, alors A^T n'est pas inversible non plus, et $|A| = |A^T| = 0$. □

Comme la transposition transforme une ligne en une colonne (et réciproquement), ce thèorème nous permet de formuler le principe suivant :

Principe 3.29. *Pour toute propriìé des déterminants où il est question des lignes de la matrice, on a une propriété analogue concernant les colonnes de la matrice.*

Résumé des résultats sur le déterminant
A matrice carrée de taille $n \times n$

- $\det(A) = \sum \operatorname{sign}(\sigma) a_{1,j_1} \cdots a_{n,j_n}$ où $\sigma = (j_1, \ldots, j_n)$.
- Si A a une ligne nulle, alors $\det(A) = 0$.
- Si A a deux lignes égales, alors $\det(A) = 0$.
- Si A est une matrice triangulaire (inférieure ou supérieure) alors $\det(A)$ est le produit de ses éléments diagonaux.

$$A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \quad ou \quad \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ * & & a_{nn} \end{pmatrix}$$

$$\det(A) = a_{11} \cdots a_{nn}\,.$$

- $\det(AB) = \det(A)\ \det(B)$
- $\det(A^T) = \det(A)$
- A est inversible si et seulement si

$$\det(A) \neq 0\,.$$

- Si A est inversible, alors

$$\det(A) = \frac{1}{\det(A^{-1})}$$

Matrices élémentaires :

- $\det(E_i(k)) = k$
- $\det(E_{ij}) = -1$
- $\det(E_{ij}(k)) = 1$.

3.3 Les cofacteurs et la règle de Cramer

Définition 3.30 (Mineur et cofacteur)**.** Soit $A = (a_{ij})$ une matrice de taille $n \times n$. On appelle *mineur de l'élément a_{ij}* le déterminant de la sous-matrice obtenue en éliminant la i-ème ligne et la j-ème colonne de la matrice A. On le note M_{ij}. On appelle *cofacteur de a_{ij}* la quantité

$$C_{ij} = (-1)^{i+j} \cdot M_{ij}.$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{ij} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

$$M_{ij} = \det \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-11} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+11} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}$$

Exemple 3.31. Soit

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

Calculons $M_{11}, C_{11}, M_{32}, C_{32}$:

$$\begin{aligned} M_{11} &= \begin{vmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1 \\ C_{11} &= (-1)^{1+1} M_{11} = 1\,. \end{aligned}$$

$$\begin{aligned} M_{32} &= \begin{vmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 4 & 1 \end{vmatrix} = -11 \\ C_{32} &= (-1)^{3+2}(-11) = 11\,. \end{aligned}$$

Pour déterminer si $C_{ij} = M_{ij}$ ou $C_{ij} = -M_{ij}$, on peut utiliser le schéma suivant :

$$A = \begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

$$C_{11} = M_{11}, \quad C_{12} = -M_{12}, \quad C_{21} = -M_{21}, \ \text{etc...}$$

3.3.1 Calcul du déterminant par la méthode des cofacteurs

Soit A une matrice de taille 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Nous savons que

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} \\ &\quad + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\ &\quad - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \,. \end{aligned}$$

On peut le réécrire de la manière suivante :

$$\begin{aligned} \det(A) &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) \\ &\quad + a_{12}(a_{23}a_{31} - a_{21}a_{33}) \\ &\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \,. \end{aligned}$$

Les termes entre parenthèses sont les cofacteurs des éléments a_{11}, a_{12}, a_{13}.
Donc

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \,.$$

Exemple 3.32.

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

$$\begin{aligned} \det(A) &= 1C_{11} + 2C_{12} + 3C_{13} \\ &= 1\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} + 2\cdot(-1)\begin{vmatrix} 4 & 1 \\ 0 & 1 \end{vmatrix} + 3\begin{vmatrix} 4 & 2 \\ 0 & 1 \end{vmatrix} \\ &= 1 - 8 + 12 \\ &= 5 \,. \end{aligned}$$

De manière analogue, on obtient

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \\ &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \,. \end{aligned}$$

Nous avons vu que les propriétés du déterminant relatives aux lignes conduisent à des propriétés analogues relatives aux colonnes. On a donc aussi :

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} \\ &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} \,. \end{aligned}$$

Ces expressions sont appelées les développements du déterminant en cofacteurs par rapport aux lignes, respectivement aux colonnes, de la matrice A.

Pour une matrice A de taille $n \times n$, on a les développements en cofacteurs analogues. Nous résumons ceci dans le théorème suivant :

Théorème 3.33 (Déterminant par la méthode des cofacteurs)**.** *Développement par rapport à la i-ème ligne :*

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

Développement par rapport à la j-ème colonne :

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

3.3.2 Calcul de l'inverse par la méthode des cofacteurs

Reprenons la formule du déterminant développé selon la i-ème ligne :

$$a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}\,.$$

Remplaçons les éléments a_{ij} par ceux d'une autre ligne, disons la k-ème, avec $k \neq i$. Nous obtenons

$$a_{k1}C_{i1} + a_{k2}C_{i2} + \cdots + a_{kn}C_{in}.$$

Cette expression est égale à zéro. Autrement dit,

Théorème 3.34.

$$a_{k1}C_{i1} + a_{k2}C_{i2} + \cdots + a_{kn}C_{in} = 0 \quad si \quad k \neq i$$

«Démonstration» : Nous allons le vérifier dans le cas particulier où $n = 3$. La démonstration est analogue dans le cas général. Soit

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Remplaçons la 3-ème ligne par la 1-ère. On obtient

$$A' = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{pmatrix}$$

On a $\det(A') = 0$, car A' a deux lignes égales. Calculons $\det(A')$ par développement en cofacteurs par rapport à la 3ème ligne. On a :

$$\det(A') = a_{11}C'_{31} + a_{12}C'_{32} + a_{13}C'_{33}$$

où C'_{ij} sont les cofacteurs de la matrice A'. Mais

$$C'_{31} = C_{31}\,, \qquad C'_{32} = C_{32}\,, \quad \text{et} \quad C'_{33} = C_{33}\,.$$

On a donc

$$\det(A') = a_{11}C_{31} + a_{12}C_{32} + a_{13}C_{33} = 0.$$

□

En résumé, on a

$$a_{k1}C_{i1} + a_{k2}C_{i2} + \cdots + a_{kn}C_{in} = \begin{cases} \det(A) & \text{si} \quad i = k \\ 0 & \text{si} \quad i \neq k\,. \end{cases}$$

Considérons maintenant la matrice des cofacteurs

$$C = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{pmatrix}$$

et calculons

$$AC^T = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{pmatrix}$$

Dans le produit AC^T, l'élément de la i-ème ligne et de la j-ème colonne est

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}\,.$$

Par ce qui précéde, on a donc

$$AC^T = \begin{pmatrix} \det(A) & 0 & \dots & 0 \\ 0 & \det(A) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \det(A) \end{pmatrix} = \det(A) \cdot I.$$

On en déduit que si $\det(A) \neq 0$, c'est-à-dire si A est inversible, alors on a

$$A^{-1} = \frac{1}{\det(A)} C^T\,.$$

On a ainsi une formule explicite pour calculer A^{-1}. On appelle C^T la *matrice adjointe de* A. Elle est notée

$$\mathrm{adj}(A)\,.$$

Théorème 3.35. *Soit A une matrice carrée avec* $\det(A) \neq 0$. *On a alors*

$$\boxed{A^{-1} = \frac{1}{\det(A)}\,\mathrm{adj}(A).}$$

Exemple 3.36.

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

On a $\det(A) = 2\,.$

La matrice formée des M_{ij} est

$$M = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

La matrice des signes est

$$A = \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

La matrice des cofacteurs est

$$C = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

On a donc

$$\mathrm{adj}(A) = C^T = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Donc

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

3.3.3 Systèmes linéaires : règle de Cramer

Le théorème suivant, appelé règle de Cramer, donne une formule explicite pour la solution de certaines systèmes d'équations linéaires.

Soit

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \qquad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{cases}$$

un système d'équations linéaires à n équations et n inconnues. Ce système peut aussi s'écrire $AX = B$ où

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

La matrice A est appelée *la matrice des coefficients du système* et la matrice B est appelée *le second membre.*

Définissons A_j par

$$A_j = \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,j-1} & b_2 & a_{2,j+1} & \dots & a_{2n} \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & b_n & a_{n,j+1} & \dots & a_{nn} \end{pmatrix}$$

$$\uparrow$$

$j$ème colonne

Autrement dit, A_j est la matrice obtenue en remplaèccant la j-ème colonne de A par le second membre B. La règle de Cramer va nous permettre de calculer la solution du système dans le cas où $\det(A) \neq 0$ en fonction des déterminants des matrices A et A_j.

Théorème 3.37 (Règle de Cramer)**.** *Soit*

$$AX = B$$

un système de n équations à n inconnues. Supposons que $\det(A) \neq 0$. *Alors l'unique solution du système est donnée par*

$$x_1 = \frac{\det(A_1)}{\det(A)},\ x_2 = \frac{\det(A_2)}{\det(A)}, \dots,\ x_n = \frac{\det(A_n)}{\det(A)}.$$

Démonstration. Nous avons supposé que

$$\det(A) \neq 0\,.$$

Donc A est inversible. Alors

$$X = A^{-1}B$$

est l'unique solution du système. D'autre part, nous avons vu que

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)\,.$$

Donc

$$X = \frac{1}{\det(A)}\operatorname{adj}(A)B.$$

Autrement dit,

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & \dots & C_{n1} \\ \vdots & & \vdots \\ C_{1n} & \dots & C_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$
$$= \frac{1}{\det(A)} \begin{pmatrix} C_{11}b_1 + C_{21}b_2 + \cdots + C_{n1}b_n \\ \vdots \\ C_{1n}b_1 + C_{2n}b_2 + \cdots + C_{nn}b_n \end{pmatrix},$$

c'est- à -dire

$$\begin{aligned} x_1 &= \frac{C_{11}b_1 + \cdots + C_{n1}b_n}{\det(A)} \\ x_i &= \frac{C_{1i}b_1 + \cdots + C_{ni}b_n}{\det(A)} \\ &\cdots \\ x_n &= \frac{C_{1n}b_1 + \cdots + C_{nn}b_n}{\det(A)}. \end{aligned}$$

Mais

$$b_1C_{1i} + \cdots + b_nC_{ni}$$

est le développement en cofacteurs de $\det(A_i)$ par rapport à sa i-ème colonne. Donc

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

□

Exemple 3.38. Résolvons le système suivant :

$$\begin{cases} x_1 & + & & & 2x_3 & = & 6 \\ -3x_1 & + & 4x_2 & + & 6x_3 & = & 30 \\ -x_1 & - & 2x_2 & + & 3x_3 & = & 8. \end{cases}$$

On a

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{pmatrix} \qquad A_1 = \begin{pmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{pmatrix}$$
$$A_2 = \begin{pmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{pmatrix}.$$

et

$$\det(A) = 44 \qquad \det(A_1) = -40$$
$$\det(A_2) = 72 \qquad \det(A_3) = 152.$$

La solution est alors

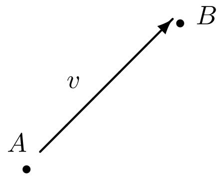
$$\begin{aligned} x_1 &= \tfrac{\det(A_1)}{\det(A)} &= -\tfrac{40}{44} &= -\tfrac{10}{11} \\ x_2 &= \tfrac{\det(A_2)}{\det(A)} &= \tfrac{72}{44} &= \tfrac{18}{11} \\ x_3 &= \tfrac{\det(A_3)}{\det(A)} &= \tfrac{152}{44} &= \tfrac{38}{11}. \end{aligned}$$

Chapitre 4

Calcul vectoriel dans le plan et dans l'espace

4.1 Définitions et règles de calcul

Un vecteur est un segment orienté dans le plan ou dans l'espace de dimension 3.



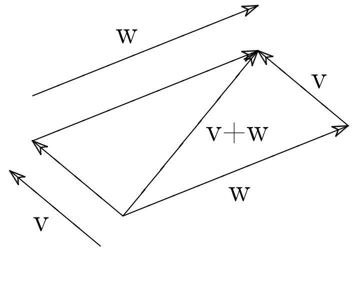
$$v = \overrightarrow{AB}$$

On appelle *module du vecteur* la longueur du segment AB. Le *support du vecteur* v est par définition la droite passant par A et B.

Deux vecteurs ont la *même direction* si leurs supports sont parallèles. Deux vecteurs ayant la même direction ont le *même sens* s'ils sont orientés de la même façon :

Deux vecteurs sont dits équivalents si l'on peut les superposer par une translation. Par la suite, deux vecteurs équivalents seront considérés comme égaux. Un vecteur est ainsi déterminé par son module, sa direction et son sens.

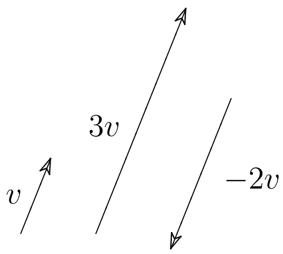
Dans le calcul vectoriel, on pourra donc faire des translations sans changer le vecteur. On définit *la somme de deux vecteurs* v et w par la règle du parallelogramme :



On place l'origine de w sur l'extrémité de v. Le vecteur $v+w$ est alors le segment orienté joignant l'origine de v à l'extrémité de w. Remarquons que $v + w = w + v$.

Le *produit d'un vecteur v par un scalaire k* est le vecteur kv défini par les propriétés suivantes :
- son module est égal à $|k|$ fois le module de v
- sa direction est celle de v
- son sens est celui de v si $k > 0$ et le sens opposé si $k < 0$.

Exemple 4.1.



L'opposé du vecteur v est le vecteur $-v$ et la *différence de deux vecteurs v et w* est définie par $v - w = v + (-w)$.

4.1.1 Systèmes de coordonnées

Si l'on choisit un système de coordonnées pour le plan (resp. pour l'espace), un vecteur peut alors s'écrire en composantes :

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \qquad \text{respectivement} \quad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

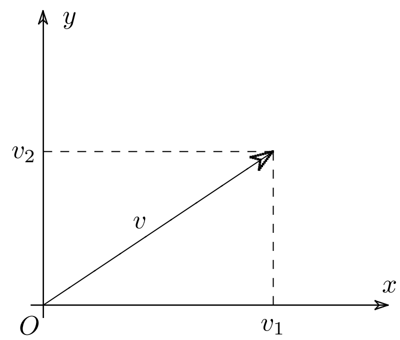


Figure 6.2

Dans cette représentation, l'origine du vecteur est toujours le point $O = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, intersection des axes de coordonnées x et y.

Dans le cas de l'espace à 3 dimensions, on choisit toujours un système d'axes orienté positivement comme le montre la figure ci-dessous :

La somme de deux vecteurs et le produit d'un vecteur par un scalaire se calculent comme suit (nous donnons les formules pour des vecteurs de l'espace, le cas du plan étant similaire) :

Si $v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ v_3 \end{smallmatrix}\right)$ et $w = \left(\begin{smallmatrix} w_1 \\ w_2 \\ w_3 \end{smallmatrix}\right)$ alors

$$v + w = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix},$$

$$kv = \begin{pmatrix} kv_1 \\ kv_2 \\ kv_3 \end{pmatrix}$$

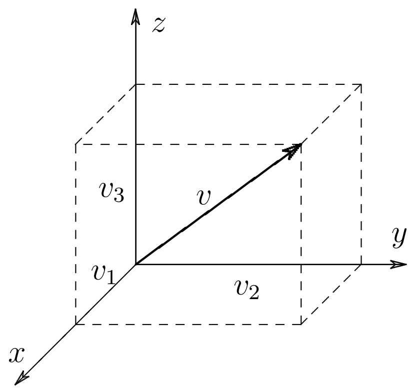


Figure 6.4 : $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$.

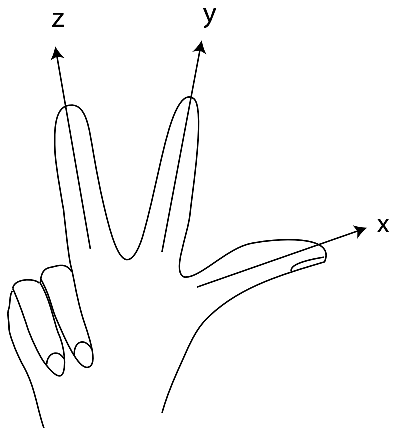


Figure 6.5 : $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ a l'orientation positive.

et
$$-v = \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}.$$

4.1.2 Propriétés du calcul vectoriel

Théorème 4.2. *(a)* $u + v = v + u$

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

(c) $u + 0 = 0 + u = u$

(d) $u + (-u) = 0$

(e) $k(\ell u) = (k\ell)u$

(f) $k(u + v) = ku + kv$

(g) $(k + \ell)u = ku + \ell u$

Soit v un vecteur. On note $\| v \|$ son module. C'est un nombre positif ou nul qui est aussi appelé la norme de v. Si $u = \left(\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right)$ alors
$$\| u \| = \sqrt{u_1^2 + u_2^2}$$
et, de même, si $u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right)$ alors
$$\| u \| = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$
Ceci découle du théorème de Pythagore.

Un vecteur de norme 1 est appelé *vecteur unité*.

4.2 Le produit scalaire

On va définir le produit scalaire de deux vecteurs u et v dans le plan ou dans l'espace de dimension 3. Le produit scalaire de deux vecteurs est un scalaire. Soient $u = \left(\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right)$ deux vecteurs dans le plan. On définit le produit scalaire par
$$u \bullet v = u_1 v_1 + u_2 v_2\,.$$
Soient $u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} v_1, v_2, v_3 \end{smallmatrix}\right)$ deux vecteurs dans l'espace de dimension 3. On définit le produit scalaire par
$$u \bullet v = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Le produit scalaire a les propriétés suivantes :

Théorème 4.3. *(1)* $u \bullet v = v \bullet u$

(2) $\lambda u \bullet v = \lambda(u \bullet v)$

(3) $(u + v) \bullet w = u \bullet w + v \bullet w$

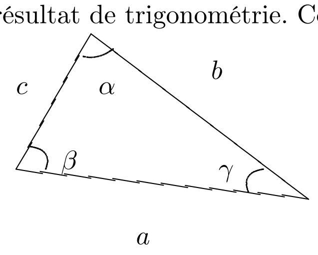
(4) $u \bullet u \geq 0$

(5) $u \bullet u = 0$ *si et seulement si* $u = 0$.

(6) $u \bullet u = \| u \|^2$

Démonstration. Ces propriétés sont des conséquences immédiates de la définition. Par exemple, si $u = \left(\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right)$ alors
$$u \bullet u = u_1^2 + u_2^2.$$
Mais
$$\| u \| = \sqrt{u_1^2 + u_2^2}$$
et donc
$$u \bullet u = \| u \|^2,$$
ce qui démontre (6). □

Nous rappelons maintenant un résultat de trigonométrie. Considérons le triangle suivant :

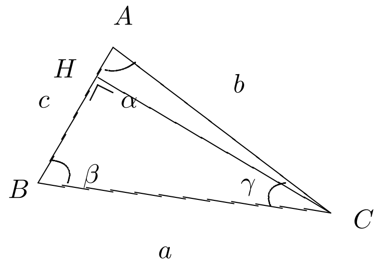


On a alors

Théorème 4.4 (Théorème du cosinus). *Soient a, b, c les côtés d'un triangle et α, β, γ ses angles comme dans la figure ci-dessus. Alors*

$$a^2 = b^2 + c^2 - 2\,b\,c\,\cos(\alpha).$$

Démonstration. Soit H le pied de la hauteur issue du sommet C.



Appliquons le théorème de Pythagore dans le triangle rectangle BCH :

$$a^2 = BH^2 + CH^2.$$

Or, $CH = b\,\sin(\alpha)$ et $BH = c - b\,\cos(\alpha)$. Ainsi

$$\begin{aligned} a^2 &= c^2 - 2\,b\,c\,\cos(\alpha) + b^2\,\cos^2(\alpha) + b^2\,\sin^2(\alpha) \\ &= b^2 + c^2 - 2\,b\,c\,\cos(\alpha) \end{aligned}$$

□

Théorème 4.5. *Soient u et v deux vecteurs non nuls, et soit θ l'angle qu'ils forment. Alors*

$$u \bullet v = \parallel u \parallel \cdot \parallel v \parallel \cdot \cos(\theta)$$

Démonstration. On a

$$\begin{aligned} \parallel v - u \parallel^2 &= (v-u) \bullet (v-u) \\ &= v \bullet v + u \bullet u - 2u \bullet v \\ &= \parallel v \parallel^2 + \parallel u \parallel^2 - 2u \bullet v\,. \end{aligned}$$

Par le théorème du cosinus, on a :

$$\parallel v - u \parallel^2 = \parallel v \parallel^2 + \parallel u \parallel^2 - 2 \parallel u \parallel \parallel v \parallel \ \cos(\theta)\,.$$

Donc

$$u \bullet v = \parallel u \parallel \parallel v \parallel \ \cos(\theta)\,.$$

Ceci démontre le théorème. □

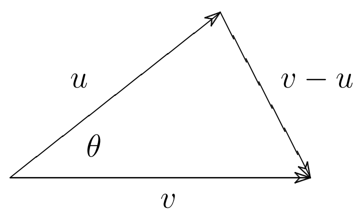


Figure 6.8

Théorème 4.6. *Soient u et v deux vecteurs non nuls. Alors u et v sont orthogonaux si et seulement si*

$$u \bullet v = 0\,.$$

Démonstration. Comme

$$u \bullet v = \| u \| \cdot \| v \| \cdot \cos(\theta)\,.$$

on a $u \bullet v = 0$ si et seulement si $\cos(\theta) = 0$ et donc si et seulement si $\theta = \frac{\pi}{2}$ et donc si et seulement si u et v sont orthogonaux. □

Remarque 4.7. On convient en général que le vecteur nul est orthogonal à tous les vecteurs. Ainsi, l'énoncé du théorème 4.6 reste vrai, même si l'un des deux vecteurs est nul, bien que l'angle θ entre u et v ne soit pas défini.

4.2.1 Projection orthogonale

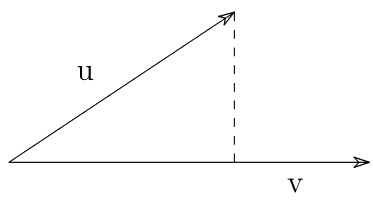


Figure 6.9

La projection (orthogonale) de u sur v est notée

$$\mathrm{proj}_v\, u.$$

Le théorème suivant nous donne le lien entre la projection et le produit scalaire :

Théorème 4.8.

$$proj_v\, u = \frac{u \bullet v}{\| v \|^2} v$$

Posons $w_1 = \mathrm{proj}_v\, u$. Comme w_1 est parallèle à v, on peut l'écrire

$$w_1 = \ell v$$

pour un certain scalaire ℓ. On a :

$$u = w_1 + w_2\,.$$

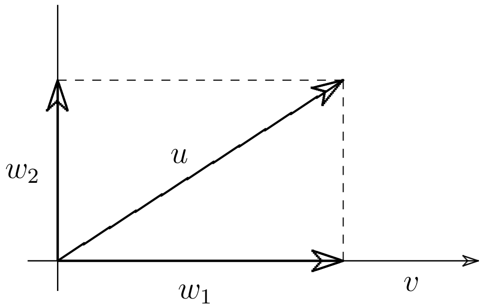


Figure 6.10
Démonstration.

Calculons le produit scalaire avec v. On a :

$$\begin{aligned} u \bullet v &= (w_1 + w_2) \bullet v = \\ &= w_1 \bullet v + w_2 \bullet v\,. \end{aligned}$$

Mais $w_2 \bullet v = 0$, car w_2 et v sont orthogonaux. Il reste alors

$$\begin{aligned} u \bullet v &= w_1 \bullet v = (\ell v) \bullet v = \\ &= \ell\ v \bullet v = \ell\ \parallel v \parallel^2 \end{aligned}$$

d'où

$$\ell = \frac{u \bullet v}{\parallel v \parallel^2}.$$

On a donc

$$\mathrm{proj}_v\, u = \frac{u \bullet v}{\parallel v \parallel^2}\, v.$$

□

Théorème 4.9. *Soit θ l'angle formé par les vecteurs u et v. Alors*

$$\parallel \mathit{proj}_v\, u \parallel = \parallel u \parallel \cdot |\cos(\theta)|\ .$$

Démonstration. Par les théorèmes 4.8 et 4.5, on a

$$\begin{aligned} \parallel \mathrm{proj}_v\, u \parallel &= \frac{|u \bullet v|}{\parallel v \parallel^2} \parallel v \parallel = \frac{|u \bullet v|}{\parallel v \parallel} \\ &= \frac{\parallel u \parallel \parallel v \parallel\ \cos(\theta)}{\parallel v \parallel} = \parallel u \parallel \cdot |\cos(\theta)|. \end{aligned}$$

□

4.3 Le produit vectoriel (cross product)

Le produit vectoriel associe à deux vecteurs de l'espace u et v un troisième vecteur, noté $u \times v$, et défini de la façon suivante :

Définition 4.10. Soient $u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ v_3 \end{smallmatrix}\right)$ deux vecteurs. Leur produit vectoriel est le vecteur $u \times v$ défini par

$$
\begin{aligned}
u \times v &= \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix} \\
&= \begin{pmatrix} \begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix} \\ -\begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix} \\ \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \end{pmatrix}.
\end{aligned}
$$

Posons

$$
i = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad j = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \text{et} \qquad k = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.
$$

Alors

$$
u \times v = \begin{vmatrix} i & u_1 & v_1 \\ j & u_2 & v_2 \\ k & u_3 & v_3 \end{vmatrix}.
$$

Le produit vectoriel satisfait les propriétés suivantes :

Théorème 4.11. *Si u, v et w sont des vecteurs dans l'espace de dimension 3, on a :*

(a) $u \bullet (u \times v) = 0$ — *$u \times v$ est orthogonal à u*

(b) $v \bullet (u \times v) = 0$ — *$u \times v$ est orthogonal v*

(c) $\| u \times v \|^2 = \| u \|^2 \cdot \| v \|^2 - (u \bullet v)^2$ — *identité de Lagrange.*

(d) $u \times (v \times w) = (u \bullet w)v - (u \bullet v)w$

(e) $(u \times v) \times w = (u \bullet w)v - (v \bullet w)u$

Démonstration. (a) Soient $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Alors

$$
\begin{aligned}
u \bullet (u \times v) &= \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \bullet \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix} = \\
&= u_1(u_2v_3 - u_3v_2) + u_2(u_3v_1 - u_1v_3) + u_3(u_1v_2 - u_2v_1) = \\
&= u_1u_2u_3 - u_1u_3v_2 + u_2u_3v_1 - u_1u_2v_3 + u_1u_3v_2 - u_2u_3v_1 = 0\,.
\end{aligned}
$$

(b) calcul similaire à (a)

(c) On a

$$
\| u \times v \|^2 = (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2
$$

et

$$
\| u \|^2 \; \| v \|^2 - (u \bullet v)^2 = (u_1^2 + u_2^2 + u_3^3)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2
$$

et un calcul direct montre que les deux termes de droites sont égaux.

(d) Les égalités (d) et (e) se montrent de manière similaire.

□

Le produit vectoriel est bilinéaire et anti-symétrique. En d'autres termes, on a

Théorème 4.12. *(a)* $u \times v = -(v \times u)$

(b) $u \times (v + w) = (u \times v) + (u \times w)$

(c) $(u + v) \times w = (u \times w) + (v \times w)$

(d) $k(u \times v) = (ku) \times v = u \times (kv)$

(e) $u \times 0 = 0 \times u = 0$

(f) $u \times u = 0\,.$

La notion de produit vectoriel est liée à celle de colinéarité par le théorème suivant :

Théorème 4.13. *Soient u et v deux vecteurs non nuls de l'espace de dimension 3. Les affirmations (1) et (2) sont équivalentes :*

(1) u et v sont colinéaires (c'est-à-dire $u = \ell v$)

(2) $u \times v = 0$.

Démonstration. (1) $\Longrightarrow$ (2) : Supposons que $u = \ell v$. Alors

$$u \times v = (\ell v) \times v = \ell(v \times v) = 0.$$

ce qui démontre (2).

Montrons maintenant l'implication inverse (2) $\Longrightarrow$ (1) : Soient $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ et $u \times v = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}$. Supposons que $u \times v = 0$. On a donc

$$\begin{aligned} u_2v_3 - u_3v_2 &= 0 \\ u_3v_1 - u_1v_3 &= 0 \\ u_1v_2 - u_2v_1 &= 0\,. \end{aligned}$$

1er cas : Si $u_1 \neq 0$ et $u_2 = u_3 = 0$, les équations deviennent

$$\begin{aligned} 0 &= 0 \\ -u_1v_3 &= 0 \\ u_1v_2 &= 0\,. \end{aligned}$$

Comme $u_1 \neq 0$, ceci entraîne $v_2 = v_3 = 0$ et donc

$$\begin{aligned} u &= \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix} \\ v &= \begin{pmatrix} v_1 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Comme v est non nul, on a $v_1 \neq 0$, d'où $u = \ell v$ avec $\ell = \frac{u_1}{v_1}$ ce qui démontre (1).

2ème cas : Supposons maintenant que $u_1 \neq 0$ et $u_2 \neq 0$. Si $v_2 = 0$ la 1-ère équation devient

$$u_2v_3 = u_3v_2 = 0.$$

Comme $u_2 \neq 0$, ceci entraîne $v_3 = 0$. Comme par hypothèse v est non nul, on doit avoir $v_1 \neq 0$. Alors par la 3ème équation, on a

$$u_1v_2 = u_2v_1 \neq 0$$

ce qui implique $v_2 \neq 0$ ce qui est absurde. Ainsi, on a donc $v_1 \neq 0$ et $v_2 \neq 0$.

Posons

$$\ell = \frac{u_1}{v_1}\,.$$

Alors la 3-ème équation donne

$$\frac{u_2}{v_2} = \ell$$

et par la 2ème équation on a

$$u_3v_1 = u_1v_3$$

ce qui implique

$$u_3 = \frac{u_1}{v_1}v_3 = \ell v_3.$$

En conclusion, on a

$$u = \ell \cdot v$$

ce qui démontre (2). □

4.3.1 Interprétation géométrique du produit vectoriel

Nous avons vu que

$$\| u \times v \|^2 = \| u \|^2 \| v \|^2 - (u \bullet v)^2$$

(identité de Lagrange). Soit θ l'angle formé par u et v. Alors on a :

$$u \bullet v = \| u \| \| v \| \cos(\theta)\,.$$

On a donc

$$\begin{aligned} \| u \times v \|^2 &= \| u \|^2 \| v \|^2 - \| u \|^2 \| v \|^2 \cos^2\theta \\ &= \| u \|^2 \| v \|^2 (1 - \cos^2\theta) \\ &= \| u \|^2 \| v \|^2 \sin^2\theta\,. \end{aligned}$$

On obtient donc

Théorème 4.14.

$$\| u \times v \| = \| u \| \| v \| \sin(\theta)$$

Démonstration. Comme

$$0 \leq \theta \leq \pi$$

on a

$$\sin(\theta) \geq 0\,.$$

et donc l'égalité cherchée. □

Considérons maintenant un parallélogramme dont les côtés sont les vecteurs u et v :

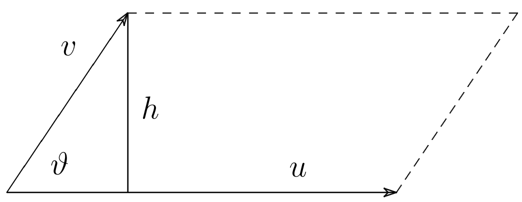


Figure 6.12 : $h = \| v \| \sin(\theta)$

L'aire A de ce parallélogramme se calcule de la façon suivante :

$$\begin{aligned} A &= (\text{base})\cdot(\text{hauteur}) \\ &= \| u \| \| v \| \sin(\theta) \\ &= \| u \times v \|\,. \end{aligned}$$

On obtient donc le théorème suivant qui donne une interprétation géométrique du produit vectoriel de deux vecteurs :

Théorème 4.15. *La norme $\| u \times v \|$ est égale à l'aire du parallèlogramme déterminé par u et v. En résumé, $u \times v$ est un vecteur perpendiculaire à u et v, et de longueur (norme) égale à l'aire du parallèlogramme déterminé par u et v. De plus, l'orientation du triplet $(u, v, u \times v)$ est positive*

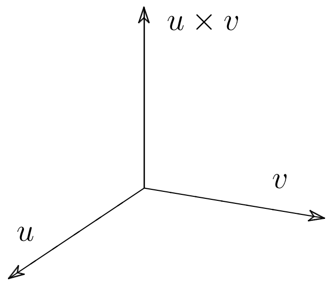


Figure 6.13

4.4 Le produit mixte (triple product)

Définition 4.16. Soient u, v et w des vecteurs de l'espace de dimension 3. On définit le produit mixte des vecteurs u, v et w par

$$[u,\, v,\, w] = u \bullet (v \times w)\,.$$

C'est un scalaire. Si $u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right)$, $v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ v_3 \end{smallmatrix}\right)$ et $w = \left(\begin{smallmatrix} w_1 \\ w_2 \\ w_3 \end{smallmatrix}\right)$ alors

$$\begin{aligned}
[u, v, w] &= u \bullet (v \times w) \\
&= u \bullet \left(\begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} i - \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} j + \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} k \right) \\
&= \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} u_1 - \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} u_2 + \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} u_3 \\
&= \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.
\end{aligned}$$

Théorème 4.17. *Soient u, v et w trois vecteurs de l'espace. Alors*

(1) $[u, v, w] = [w, u, v] = [v, w, u] = -[v, u, w] = -[u, w, v] = -[w, v, u]$

(2) $[\lambda u, v, w] = \lambda [u, v, w]$ *pour tout scalaire* λ.

(3) $[u, v, w] = 0$ *si et seulement s'il existe des scalaires* α, β, γ *non tous nuls tels que* $\alpha u + \beta v + \gamma w = 0$.

Démonstration. (1) et (2) découlent directement de la définition. Montrons alors la propriété (3). Supposons que

$$\alpha u + \beta v + \gamma w = 0$$

avec α, β, γ non tous nuls. Sans perte de généralité, on peut supposer $\alpha \neq 0$.

Alors

$$u = \lambda v + \mu w$$

avec

$$\lambda = -\frac{\beta}{\alpha} \qquad \text{et} \qquad \mu = -\frac{\gamma}{\alpha},$$

et

$$\begin{aligned}
[u, v, w] &= (\lambda v + \mu w) \bullet (v \times w) \\
&= \lambda \underbrace{v \bullet (v \times w)}_{0} + \mu \underbrace{w \bullet (v \times w)}_{0} \\
&= 0\,.
\end{aligned}$$

Réciproquement, supposons que $[u, v, w] = 0$. On a donc

$$(u \times v) \bullet w = 0$$

c'est-à-dire que $u \times v$ est orthogonal à w. Mais $u \times v$ est aussi orthogonal à u et à v. Ceci entraîne que u, v et w sont coplanaires et donc que l'on peut écrire

$$w = \alpha u + \beta v$$

ce qui termine la démonstration. □

Le produit mixte peut également être interprété géométriquement ce qui est l'objet du théorème suivant.

Théorème 4.18. *(1) Soit $u = \left(\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right)$ deux vecteurs de $\mathbb{R}^2$. Alors la valeur absolue du déterminant*

$$\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}$$

est égal à l'aire du parallèlogramme défini par les vecteurs u et v.

(2) Soient $u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ u_3 \end{smallmatrix}\right)$, $v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ v_3 \end{smallmatrix}\right)$ et $w = \left(\begin{smallmatrix} w_1 \\ w_2 \\ w_3 \end{smallmatrix}\right)$ trois vecteurs de $\mathbb{R}^3$. Alors la valeur absolue du déterminant

$$\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}$$

est égal au volume du parallélépipède déterminé par ces trois vecteurs

Démonstration. (1) On considère u et v comme vecteurs de l'espace de dimension 3 :

$$u = \left(\begin{smallmatrix} u_1 \\ u_2 \\ 0 \end{smallmatrix}\right) \qquad \text{et} \qquad v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ 0 \end{smallmatrix}\right).$$

Alors

$$u \times v = \begin{vmatrix} i & u_1 & v_1 \\ j & u_2 & v_2 \\ k & 0 & 0 \end{vmatrix} = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} k\,.$$

L'aire du parallélogramme déterminé par u et v est

$$\begin{aligned} \| u \times v \| &= \left| \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} k \right| \\ &= \left| \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \right| \cdot \| k \| = \left| \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \right|. \end{aligned}$$

(2) Prenons le parallélogramme déterminé par v et w comme base du parallélépipède déterminé par u, v et w. L'aire de la base est donc

$$\| v \times w \|$$

et la hauteur du parallélépipède est la projection orthogonale de u sur $v \times w$.
On a

$$h = \| \operatorname{proj}_{v \times w} u \| = \frac{|u \bullet (v \times w)|}{\| v \times w \|}$$

et le volume V du parallélépipède est alors

$$\begin{aligned} V &= (\text{aire de la base}) \cdot (\text{hauteur}) \\ &= \| v \times w \| \frac{|u \bullet (v \times w)|}{\| v \times w \|} \\ &= |u \bullet (v \times w)| \end{aligned}$$

qui est donc bien la valeur absolue du déterminant

$$\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

□

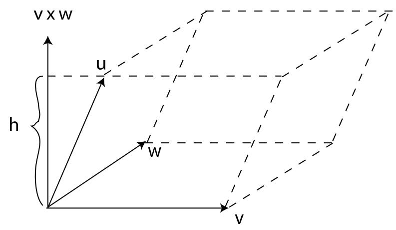


Figure 6.14 : h = hauteur

4.5 Droites et plans dans l'espace de dimension 3

4.5.1 Equation du plan passant par un point P_0 et ayant vecteur normal n

Soit $P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ un point de l'espace de dimension 3 et $n = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$ un vecteur.

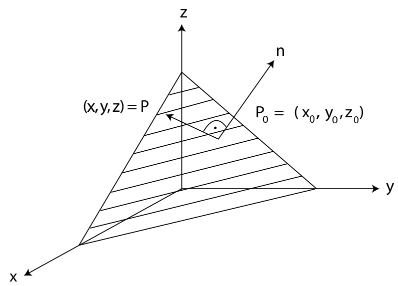


Figure 6.15

Le plan passant par P_0 et ayant n comme vecteur normal est formé des points P tels que le vecteur $\overrightarrow{P_0P}$ est orthogonal au vecteur n. On a donc

$$\overrightarrow{P_0P} \bullet n = 0\,.$$

Si $P = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ alors $\overrightarrow{P_0P} = \begin{pmatrix} X-x_0 \\ Y-y_0 \\ Z-z_0 \end{pmatrix}$ et la condition $\overrightarrow{P_0P} \bullet n = 0$ s'écrit

$$\begin{pmatrix} X-x_0 \\ Y-y_0 \\ Z-z_0 \end{pmatrix} \bullet \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = 0$$

ou encore

$$n_1(X - x_0) + n_2\,(Y - y_0) + n_3\,(Z - z_0) = 0\,.$$

Exemple 4.19. L'équation du plan passant par le point $P_0 = \begin{pmatrix} 2 \\ -5 \\ 6 \end{pmatrix}$ et perpendiculaire au vecteur $n = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ est

$$(X - 2) + 3(Y + 5) - 2(Z - 6) = 0$$

ou encore

$$X + 3Y - 2Z + 25 = 0.$$

Théorème 4.20. *Soient a, b, c, d des scalaires tels que $(a, b, c) \neq (0, 0, 0)$. Alors l'équation*

$$aX + bY + cZ + d = 0$$

est l'équation d'un plan ayant comme vecteur normal le vecteur

$$n = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Démonstration. Par hypothèse, les scalaires a, b, c sont non tous nuls. Nous pouvons supposer que $a \neq 0$. Alors l'équation

$$aX + bY + cZ + d = 0$$

peut être réécrite comme

$$a\left(X + \frac{d}{a}\right) + bY + cZ = 0\,.$$

Mais ceci est l'équation du plan passant par le point $\left(\begin{smallmatrix} -d/a \\ 0 \\ 0 \end{smallmatrix}\right)$ et ayant comme vecteur normal le vecteur $\left(\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right)$. □

4.5.2 Droites dans l'espace de dimension 3

Soit L la droite passant par le point $P_0 = (x_0, y_0, z_0)$ et parallèle au vecteur $v = (a, b, c)$. Alors L est constituée des points $P = (X, Y, Z)$ tels que le vecteur $\overrightarrow{P_0P}$ est parallèle au vecteur v. Autrement dit :

$$\overrightarrow{P_0P} = \lambda v$$

pour un scalaire λ. On a donc

$$\begin{pmatrix} X - x_0 \\ Y - y_0 \\ Z - z_0 \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \\ \lambda c \end{pmatrix}$$

ou, de manière équivalente,

$$\begin{cases} X - x_0 &= \lambda a \\ Y - y_0 &= \lambda b \\ Z - z_0 &= \lambda c. \end{cases}$$

Ce système est est appelé *système d'équations paramétriques de la droite L.*

Théorème 4.21. *La distance D entre le point $P_0 = \left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)$ et le plan d'équation*

$$aX + bY + cZ + d = 0$$

est donnée par

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

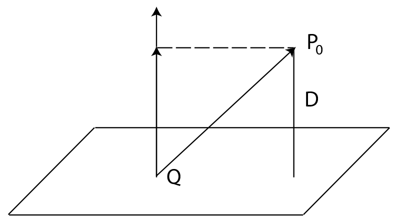


Figure 6.16

Démonstration.

Soit $Q = \left(\begin{smallmatrix} x_1 \\ y_1 \\ z_1 \end{smallmatrix}\right)$ un point du plan. On place le vecteur normal n au point Q. La distance de P_0 au plan est égale à la norme de la projection orthogonale du vecteur $\overrightarrow{QP_0}$ sur le vecteur n :

Donc

$$D = \parallel \mathrm{proj}_n \overrightarrow{QP_0} \parallel = \frac{|\overrightarrow{QP_0} \bullet n|}{\parallel n \parallel}.$$

On a $\overrightarrow{QP_0} = \left(\begin{smallmatrix} x_0 - x_1 \\ y_0 - y_1 \\ z_0 - z_1 \end{smallmatrix}\right)$ et ainsi

$$\overrightarrow{QP_0} \bullet n = \begin{pmatrix} x_0 - x_1 \\ y_0 - y_1 \\ z_0 - z_1 \end{pmatrix} \bullet \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1).$$

Comme

$$\parallel n \parallel = \sqrt{a^2 + b^2 + c^2},$$

on a

$$D = \frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)|}{\sqrt{a^2 + b^2 + c^2}}.$$

Or $Q = \left(\begin{smallmatrix} x_1 \\ y_1 \\ z_1 \end{smallmatrix}\right)$ est un point du plan, ce qui entraîne que $ax_1 + by_1 + cz_1 + d = 0$ et donc que $d = -ax_1 - by_1 - cz_1$. On obtient finalement

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

□

Chapitre 5

Espaces euclidiens et applications linéaires

5.1 Espaces de dimension n

5.1.1 Définitions et notations

L'ensemble des nombres réels est noté $\mathbb{R}$. L'ensemble des nombres réels $\mathbb{R}$ est souvent représenté par une droite. C'est un espace de dimension 1. Le plan est formé des couples $\left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right)$ de nombres réels. Il est noté $\mathbb{R}^2$. L'espace de dimension 3 est constitué des triplets de nombres réels $\left(\begin{smallmatrix} a_1 \\ a_2 \\ a_3 \end{smallmatrix}\right)$. Il est noté $\mathbb{R}^3$. Le symbole $\left(\begin{smallmatrix} a_1 \\ a_2 \\ a_3 \end{smallmatrix}\right)$ a deux interprétations géométriques :

1. Un point

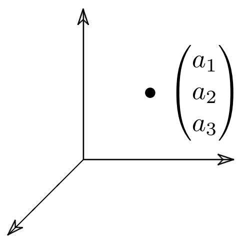


Figure 7.1

2. Un vecteur

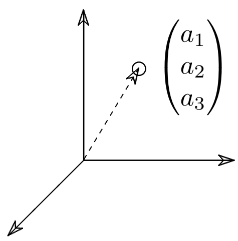


Figure 7.2

On généralise ces notions en considérant aussi des espaces de dimension n pour tout entier positif $n = 1, 2, 3, 4, \ldots$. Les éléments de l'espace de dimension n sont les n-uples $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ de nombres réels. L'espace de dimension n est noté $\mathbb{R}^n$. Comme en dimension 3, le n-uple $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ dénote aussi bien le point que le vecteur de l'espace de dimension n.

Définition 5.1 (Somme de deux vecteurs)**.** Soient $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ deux vecteurs de $\mathbb{R}^n$. Leur somme est par définition le vecteur

$$u + v = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}.$$

Définition 5.2 (Produit d'un vecteur par un scalaire)**.** Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ un vecteur et λ un scalaire. Alors

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

Le vecteur nul de $\mathbb{R}^n$ est le vecteur $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ un vecteur de $\mathbb{R}^n$. Alors son opposé est le vecteur $-u = \begin{pmatrix} -u_1 \\ \vdots \\ -u_n \end{pmatrix}$.

Théorème 5.3. *Soient* $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ *et* $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ *des vecteurs de* $\mathbb{R}^n$*. Alors :*

(a) $u + v = v + u$
(b) $u + (v + w) = (u + v) + w$
(c) $u + 0 = 0 + u = u$
(d) $u + (-u) = 0$
(e) $\lambda(\mu u) = (\lambda\mu)u$
(f) $\lambda(u + v) = \lambda u + \lambda v$
(g) $(\lambda + \mu)u = \lambda u + \mu u$
(h) $1 \cdot u = u$

5.1.2 Produit scalaire

Soient $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ deux vecteurs de $\mathbb{R}^n$. On définit leur *produit scalaire* par

$$u \bullet v = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

C'est un scalaire. Remarquons que cette définition généralise la notion de produit scalaire dans le plan $\mathbb{R}^2$ et dans l'espace $\mathbb{R}^3$.

Théorème 5.4. *Soient u, v et w des vecteurs de $\mathbb{R}^n$ et soit λ un scalaire. Alors on a :*

(a) $u \bullet v = v \bullet u$
(b) $(u + v) \bullet w = u \bullet w + v \bullet w$
(c) $(\lambda u) \bullet v = \lambda(u \bullet v)$
(d) $v \bullet v \geq 0.$
(e) $v \bullet v = 0$ *si et seulement si* $v = 0.$

5.1.3 Norme et distance dans $\mathbb{R}^n$

Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ un vecteur. On définit la *norme* de u (ou norme euclidienne de u) par

$$\| u \| = \sqrt{u \bullet u} = \sqrt{u_1^2 + \cdots + u_n^2}.$$

Soient $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ deux vecteurs. La *distance* (ou distance euclidienne) entre u et v est définie par

$$d(u, v) = \| u - v \| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

Exemple 5.5. Soient $u = \begin{pmatrix} 1 \\ 6 \\ -4 \\ 3 \end{pmatrix}$ et $v = \begin{pmatrix} 3 \\ 7 \\ 2 \\ -2 \end{pmatrix}$. Alors leur distance dans $\mathbb{R}^4$ est

$$\begin{aligned} d(u, v) &= \sqrt{(1-3)^2 + (6-7)^2 + (-4-2)^2 + (3+2)^2} \\ &= \sqrt{4 + 1 + 36 + 25} = \sqrt{66}. \end{aligned}$$

Théorème 5.6. *(Inégalité de Cauchy-Schwarz)*
Soient u et v des vecteurs de $\mathbb{R}^n$. Alors on a

$$|u \bullet v| \leq \| u \| \cdot \| v \| .$$

Théorème 5.7. *Soient u et v des vecteurs de $\mathbb{R}^n$, et soit λ un scalaire. Alors on a :*

(a) $\| u \| \geq 0$

(b) $\| u \| = 0$ *si et seulement si* $u = 0$

(c) $\| \lambda u \| = |\lambda| \cdot \| u \|$

(d) $\| u + v \| \leq \| u \| + \| v \|$ *(Inégalité du triangle)*

Démonstration. (a) et (b) découlent des points (d) et (e) du théorème 5.4.

(c) Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$. On a $\lambda u = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix}$ et donc

$$\begin{aligned} \| \lambda u \| &= \sqrt{(\lambda u_1)^2 + \cdots + (\lambda u_n)^2} \\ &= \sqrt{\lambda^2 u_1^2 + \cdots + \lambda^2 u_n^2} = \sqrt{\lambda^2} \sqrt{u_1^2 + \cdots + u_n^2} \\ &= |\lambda| \sqrt{u_1^2 + \cdots + u_n^2} = |\lambda| \ \| u \| . \end{aligned}$$

(d) Soient $u = \left({u_1 \atop dotsu_n} \right)$ et $v = \left({v_1 \atop v_n} \right)$. On a

$$\begin{aligned} \| u + v \|^2 &= (u + v) \bullet (u + v) \\ &= u \bullet u + v \bullet v + 2u \bullet v \\ &= \| u \|^2 + \| v \|^2 + 2u \bullet v. \end{aligned}$$

Remarquons que $u \bullet v \leq |u \bullet v|$ car, pour n'importe quel scalaire a, on a $a \leq |a|$. On obtient alors

$$\| u + v \|^2 \leq \| u \|^2 + \| v \|^2 + 2 \| u \| \ \| v \| = (\| u \| + \| v \|)^2$$

et donc

$$\| u + v \| \leq \| u \| + \| v \|$$

□

Corollaire 5.8. *Soient u, v et w des vecteurs dans $\mathbb{R}^n$, et soit λ un scalaire. Alors*

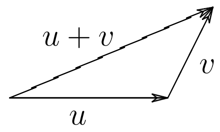


Figure 7.3

(a) $d(u, v) \geq 0$

(b) $d(u, v) = 0$ *si et seulement si* $u = v$.

(c) $d(u, v) = d(v, u)$

(d) $d(u, v) \leq d(u, w) + d(w, v)$ *(Inégalité du triangle)*

Théorème 5.9. *Soient u et v des vecteurs de $\mathbb{R}^n$. Alors*

$$u \bullet v = \frac{1}{4} \parallel u+v \parallel^2 - \frac{1}{4} \parallel u-v \parallel^2 .$$

Démonstration.

$$\begin{aligned} \parallel u+v \parallel^2 &= (u+v) \bullet (u+v) = \parallel u \parallel^2 + \parallel v \parallel^2 + 2u \bullet v \\ \parallel u-v \parallel^2 &= (u-v) \bullet (u-v) = \parallel u \parallel^2 + \parallel v \parallel^2 - 2u \bullet v \end{aligned}$$

Donc

$$\parallel u+v \parallel^2 - \parallel u-v \parallel^2 = 4u \bullet v$$

d'où

$$u \bullet v = \frac{1}{4} \parallel u+v \parallel^2 - \frac{1}{4} \parallel u-v \parallel^2 .$$

□

Définition 5.10. On dit que deux vecteurs u et v de $\mathbb{R}^n$ sont *orthogonaux* si

$$u \bullet v = 0.$$

Théorème 5.11 (Théorème de Pythagore dans $\mathbb{R}^n$)**.** *Soient u et v deux vecteurs de $\mathbb{R}^n$ orthogonaux. Alors*

$$\parallel u+v \parallel^2 = \parallel u \parallel^2 + \parallel v \parallel^2 .$$

Démonstration.

$$\begin{aligned} \parallel u+v \parallel^2 &= (u+v) \bullet (u+v) \\ &= \parallel u \parallel^2 + \parallel v \parallel^2 + 2u \bullet v \\ &= \parallel u \parallel^2 + \parallel v \parallel^2 \qquad \text{car } u \bullet v = 0. \end{aligned}$$

□

5.1.4 Représentation matricielle des vecteurs de $\mathbb{R}^n$

Soit $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ un vecteur de $\mathbb{R}^n$. On l'appelle «vecteur colonne», et on le considère naturellement u comme une matrice de taille $n \times 1$. Parfois, on rencontre aussi des «vecteurs ligne» : on peut voir u comme une matrice $1 \times n$, de la forme $u = (u_1, \ldots, u_n)$. En fait, le vecteur ligne correspondant à u est le transposé u^T du vecteur colonne u.

Les opérations de somme et de produit par un scalaire définies ci-dessus se transposent parfaitement à l'écriture matricielle et l'on retrouve ainsi les opérations définies sur les matrices introduites au chapitre 2 :

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

et

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

5.1.5 Formule matricielle du produit scalaire

Soient

$$u=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \qquad \text{et} \qquad v=\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

deux vecteurs. On a

$$v^T=\begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$$

et ainsi

$$v^Tu=(v_1\dots v_n)\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}=u_1v_1+\dots+u_nv_n=u\bullet v.$$

On a donc

$$u\bullet v=v^Tu\,.$$

Exemple 5.12.

$$u=\begin{pmatrix} 1 \\ -3 \\ 6 \\ 4 \end{pmatrix}, \quad v=\begin{pmatrix} 2 \\ -3 \\ 5 \\ 0 \end{pmatrix}$$

$$u\bullet v=v^Tu=(2\ -3\,5\,0)\begin{pmatrix} 1 \\ -3 \\ 6 \\ 4 \end{pmatrix}=(41)=41.$$

Soit A une matrice de taille $n\times n$ et u un vecteur colonne (i.e. une matrice $n\times 1$). Le produit matriciel Au est une matrice $n\times 1$, donc un vecteur colonne. On peut alors considérer le produit scalaire de Au avec un autre vecteur v,

$$(Au)\bullet v$$

On a ainsi

$$\begin{aligned}(Au)\bullet v&=v^T(Au)\\&=(v^TA)u=(A^Tv)^Tu=\\&=u\bullet(A^Tv).\end{aligned}$$

On obtient donc la formule

Théorème 5.13.

$$(Au)\bullet v=u\bullet(A^Tv).$$

On verra plus tard (matrices orthogonales) la signification de cette équation.

Exemple 5.14. Posons

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad u = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} \quad \text{et} \quad v = \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}.$$

Alors

$$Au = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 10 \\ 5 \end{pmatrix}$$

et

$$A^T v = \begin{pmatrix} 1 & 2 & -1 \\ -2 & 4 & 0 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} -7 \\ 4 \\ -1 \end{pmatrix}.$$

On obtient ainsi

$$(Au) \bullet v = (v^T)(Au) = \begin{pmatrix} -2 & 0 & 5 \end{pmatrix} \begin{pmatrix} 7 \\ 10 \\ 5 \end{pmatrix} = -14 + 0 + 25 = 11$$

et

$$u \bullet (A^T v) = (A^T v)^T u = \begin{pmatrix} -7 & 4 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} = 7 + 8 - 4 = 11.$$

5.1.6 Multiplication des matrices et produit scalaire

Soient $A = (a_{ij})$ une matrice de taille $m \times r$, et $B = (b_{ij})$ une matrice de taille $r \times n$. Nous avons vu que l'on peut former le produit matriciel AB. On obtient une matrice de taille $m \times n$. L'élément d'indice ij de la matrice AB est

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj}\,.$$

Remarquons que ceci est aussi le produit

$$\begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{ir} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{pmatrix}$$

Autrement dit, c'est le produit scalaire du i-ème vecteur ligne de A avec le j-ème vecteur colonne de B. Notons $l_1, \ldots, l_m$ les vecteurs ligne de A, et $c_1, \ldots, c_n$ les vecteurs colonne de B. On a alors

$$AB = \begin{pmatrix} \ell_1 \bullet c_1 & \ell_1 \bullet c_2 & \ldots & \ell_1 \bullet c_n \\ \ell_2 \bullet c_1 & \ell_2 \bullet c_2 & \ldots & \ell_2 \bullet c_n \\ \vdots & \vdots & & \vdots \\ \ell_m \bullet c_1 & \ell_m \bullet c_2 & \ldots & \ell_m \bullet c_n \end{pmatrix}$$

En particulier, un système d'équations linéaires peut être exprimé grâce au produit scalaire comme suit : soit $Ax = b$ un système linéaire où A est la matrice des coefficients du système de taille $m \times n$,

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

est le second membre et

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

le vecteur colonne des inconnues.

Soient $\ell_1, \ldots, \ell_m$ les lignes de la matrice A. Ce sont des vecteurs ligne (matrices de taille $1 \times n$). Alors le système peut être exprimé comme

$$\begin{pmatrix} \ell_1 \bullet x \\ \ell_2 \bullet x \\ \vdots \\ \ell_m \bullet x \end{pmatrix} = b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

5.2 Applications linéaires

5.2.1 Rappels sur les applications

Soient A et B deux ensembles. Une application $f : A \longrightarrow B$ associe à tout élément a de A un unique élément $f(a)$ de B. L'élément $f(a)$ est appelé l'image de a par f (ou la valeur de f en a). L'ensemble A est appelé le domaine de définition de f. Le sous-ensemble $f(A)$ de l'ensemble B formés des éléments $f(a)$, où a est un élément de A, s'appelle l'image de A par f. Lorsque $A = \mathbb{R}$, on dit que f est une application (ou fonction) d'une variable réelle. Lorsque $A = \mathbb{R}^n$, on dit que f est une application (ou fonction) de n variables réelles. Lorsque $B = \mathbb{R}$, on dit que f est à valeurs réelles.

Exemples

(1)

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$f(x) = x^2$$

est une fonction d'une variable réelle à valeurs réelles.

(2)

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$$

est une fonction à valeurs réelles, de n variables réelles.

(3)

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$f(x) = 3x$$

est une fonction d'une variable réelle à valeurs réelles.

Soient

$$f_1 : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\vdots$$

$$f_m : \mathbb{R}^n \longrightarrow \mathbb{R}$$

m fonctions de n variables réelles à valeurs réelles. On peut construire une application

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

définie par

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)).$$

5.2.2 Applications linéaires

Définition 5.15. Une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ définie par $f(x_1, \dots, x_n) = (w_1, \dots, w_m)$ est dite *linéaire* si

$$\begin{array}{ccccccccc} w_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n \\ w_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ w_m & = & a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n \,. \end{array}$$

En notation matricielle, on a

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

ou encore

$$f(x) = w = Ax \qquad \text{avec} \qquad A = (a_{ij}).$$

A est appelée *la matrice de l'application* f.

Exemple 5.16. La fonction $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ définie par

$$\begin{aligned} w_1 &= -2x_1 + 5x_2 + 2x_3 - 7x_4 \\ w_2 &= 4x_1 + 2x_2 - 3x_3 + 3x_4 \\ w_3 &= 7x_1 - 3x_2 + 9x_3 \end{aligned}$$

peut être exprimée sous forme matricielle comme suit :

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & 2 & -7 \\ 4 & 2 & -3 & 3 \\ 7 & -3 & 9 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Notation :

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ une application linéaire. On note $[f]$ la matrice de f (appelée aussi matrice standard de f). Soit A une matrice $m \times n$. On note $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ l'application linéaire

$$f_A(x) = Ax \qquad \text{pour tout } x \in \mathbb{R}^n.$$

L'application linéaire induite par la matrice identité

$$f_I(x) = Ix = x$$

est appelée *application identité*, et notée I ou Id. L'application linéaire induite par la matrice nulle 0 de taille $m \times n$

$$f_0(x) = 0x = 0$$

est appelée *application nulle*, et notée 0.

Remarque 5.17. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ une application linéaire. Alors $f(0) = 0$. En effet, si A est la matrice standard de f, on a $f(0) = A0 = 0$.

5.2.3 Quelques exemples d'applications linéaires

Réflexion par rapport à l'axe Oy

La fonction

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) &\mapsto \left(\begin{smallmatrix} -x \\ y \end{smallmatrix}\right) \end{aligned}$$

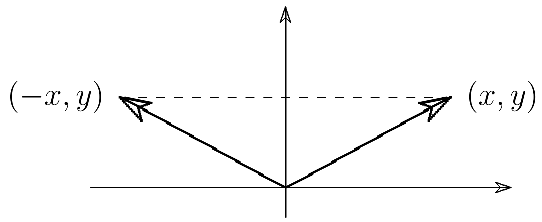


Figure 7.6

est la réflexion par rapport à l'axe Oy et sa matrice est

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

car

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

Réflexion par rapport à l'axe Ox

La réflexion par rapport à l'axe des x est donnée par la matrice

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

Réflexion par rapport à la droite $y = x$

La réflexion par rapport à la droite $y = x$ est donnée par

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \left(\begin{smallmatrix} x \\ y \end{smallmatrix} \right) &\mapsto \left(\begin{smallmatrix} y \\ x \end{smallmatrix} \right) \end{aligned}$$

et sa matrice est

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

Réflexions dans l'espace

L'application

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix} \right) &\mapsto \left(\begin{smallmatrix} x \\ y \\ -z \end{smallmatrix} \right) \end{aligned}$$

est la réflexion par rapport au plan Oxy. C'est une application linéaire et sa matrice est

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

De même, la réflection par rapport au plan Oxz est donnée par la matrice

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

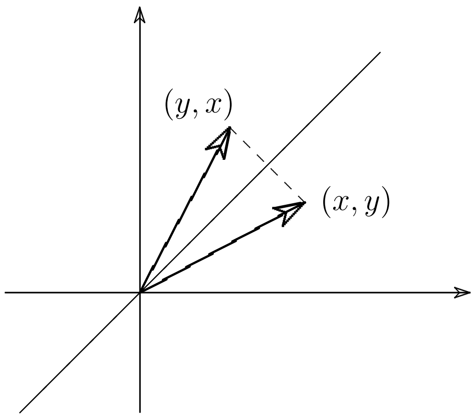


Figure 7.7

et la réflection par rapport au plan Oyz par la matrice

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

Projections orthogonales

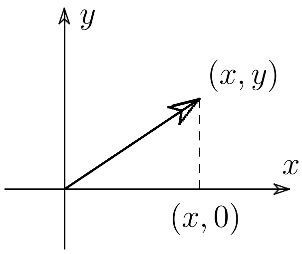


Figure 7.8

L'application

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \left(\begin{smallmatrix} x \\ y \end{smallmatrix} \right) &\mapsto \left(\begin{smallmatrix} x \\ 0 \end{smallmatrix} \right) \end{aligned}$$

est la projection orthogonale sur l'axe Ox. C'est une application linéaire donnée par la matrice

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

L'application linéaire

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix} \right) &\mapsto \left(\begin{smallmatrix} x \\ y \\ 0 \end{smallmatrix} \right) \end{aligned}$$

est la projection orthogonale sur le plan Oxy et sa matrice est

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

De même, la projection orthogonale sur le plan Oxz est donnée par la matrice standard

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et la projection orthogonale $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ sur le plan Oyz par la matrice standard

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

5.2.4 Rotations

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la rotation d'angle θ. On a

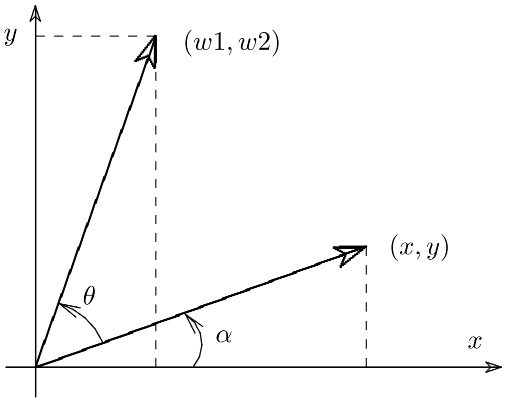


Figure 7.9 : Rotation d'angle θ

$$\begin{aligned} x &= r\cos\alpha \\ y &= r\sin\alpha \end{aligned}$$

Après rotation d'angle θ, on obtient :

$$\begin{aligned} w_1 &= r\cos(\alpha+\theta), \\ w_2 &= r\sin(\alpha+\theta)\,. \end{aligned}$$

En appliquant des formules trigonométriques, on obtient

$$\begin{aligned} w_1 &= r\cos\alpha\cos\theta \,-\, r\sin\alpha\sin\theta \\ w_2 &= r\cos\alpha\sin\theta \,+\, r\sin\alpha\cos\theta \end{aligned}$$

c'est-à-dire

$$\begin{aligned} w_1 &= x\ \cos\theta - y\sin\,\theta \\ w_2 &= x\ \sin\,\theta + y\cos\,\theta\,. \end{aligned}$$

Autrement dit, la rotation d'angle θ est donnée par la matrice standard

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

5.2.5 Composition d'applications linéaires

Soient

$$f_A : \mathbb{R}^m \longrightarrow \mathbb{R}^r$$

et

$$f_B : \mathbb{R}^r \longrightarrow \mathbb{R}^n$$

deux applications linéaires induites respectivement par les matrices A et B. Considérons leur composition

$$f_B \circ f_A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

définie par $(f_B \circ f_A)(x) = f_B(f_A(x))$.

C'est une application linéaire et on a

$$\begin{aligned}(f_B \circ f_A)(x) &= f_B(f_A(x)) = \\ &= B(Ax) = (BA)x\end{aligned}$$

ce qui implique que

$$f_B \circ f_A = f_{BA}.$$

Autrement dit, la matrice associée à la composition de deux applications linéaires est égale au produit de leurs matrices standard.

Exemple 5.18. Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

la réflection par rapport à la droite $y = x$, et soit

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

la projection orthogonale sur l'axe des y.

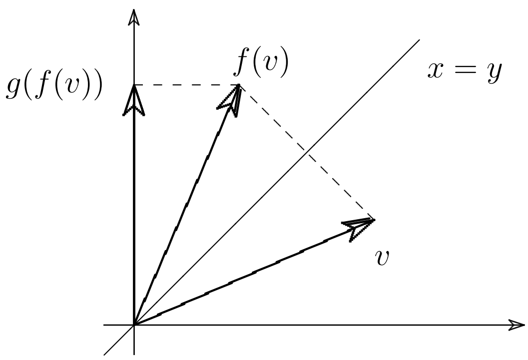


Figure 7.10 : Composition de f puis de g

La matrice standard de f est

$$[f] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

et la matrice standard de g est

$$[g] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

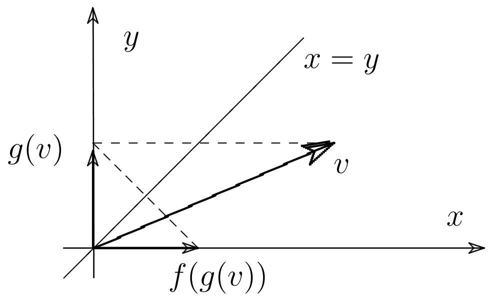


Figure 7.11 : Composition de g puis de f

On a alors

$$[f \circ g] = [f][g] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

et

$$[g \circ f] = [g][f] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Remarquons que

$$[g \circ f] = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = [f \circ g]$$

ce qui montre que la composition d'applications linéaires n'est pas commutative en général.

5.3 Propriétés des applications linéaires

Définition 5.19. Soient E_1 et E_2 deux ensembles et soit

$$f : E_1 \longrightarrow E_2$$

une application. On dit que f est *injective* si $f(x) = f(y)$ implique $x = y$. Autrement dit, deux éléments distincts ont des images distinctes.

On dit que f est *surjective* si pour tout $z \in E_2$, il existe $x \in E_1$ avec $f(x) = z$. Autrement dit, $f(E_1) = E_2$.

On dit que f est *bijective* si elle est injective et surjective. Autrement dit, f est bijective si, pour tout $z \in E_2$, il existe un unique $x \in E_1$ avec $f(x) = z$.

Soit A une matrice $n \times n$. Nous avons vu (théorème 2.32) que les propriétés suivantes sont équivalentes :

(1) A est inversible

(2) Pour toute matrice colonne w, de taille $n \times 1$, le système

$$Ax = w$$

a une solution.

D'autre part, si pour toute matrice colonne w de taille $n \times 1$ le système

$$Ax = w$$

a une solution, alors A est inversible. En effet, vu que pour tout w le système a une solution, il existe $x_1, x_2, \ldots, x_n$ des matrices colonne telles que

$$Ax_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad Ax_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad Ax_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Ainsi,

$$A \underbrace{(x_1\, x_2 \ldots x_n)}_{=C} = I_n$$

et $\det(AC) = \det(A)\det(C) = 1$. Le déterminant de A est donc non nul, ce qui est équivalent à dire que A est inversible. D'où le théorème suivant :

Théorème 5.20. *Soit A une matrice $n \times n$. Les propriétés suivantes sont équivalentes :*

(1) A est inversible

(2) Pour toute matrice colonne w de taille $n \times 1$, le système

$$Ax = w$$

a une solution

(3) Pour toute matrice colonne w de taille $n \times 1$, le système

$$Ax = w$$

a exactement une solution.

Soit $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'application linéaire de matrice standard A. Remarquons que la propriété (2) ci-dessus est équivalente à

(2') f_A est surjective

et que (3) équivalent à

(3') f_A est bijective.

On a donc le théorème suivant :

Théorème 5.21. *Soit A une matrice de taille $n \times n$, et soit $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'application linéaire associée. Les propriétés suivantes sont équivalentes :*

(1) f_A est injective

(2) f_A est surjective

(3) f_A est bijective

(4) A est inversible.

Démonstration. L'équivalence des points (2), (3) et (4) découle du théorème précédent.
Il suffit donc de montrer que (1) est équivalent à (4) par exemple.
Supposons f_A injective. Soit $x \in \mathbb{R}^n$ tel que $f_A(x) = 0$. Alors $Ax = 0$. Or, on a aussi $A0 = 0$. Comme f_A est injective, on a nécessairement $x = 0$. Autrement dit, le système $Ax = 0$ a une unique solution. Le théorème 2.32 permet de conclure que la matrice A est inversible.
Réciproquement, supposons A inversible. Soient $x_1, x_2 \in \mathbb{R}^n$ tels que $f_A(x_1) = f_A(x_2)$, c'est-à-dire tels que $Ax_1 = Ax_2$. On a alors $A(x_1 - x_2) = 0$. A nouveau par le théorème 2.32, on en déduit que $x_1 = x_2$. L'application f_A est donc injective. □

Exemple 5.22. Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

la projection sur l'axe des x. Alors f n'est pas injective. En effet,

$$f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right) \qquad \text{pour tout } y \in \mathbb{R}^2$$

La matrice standard de f est

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

Elle n'est pas inversible, car $\det([f]) = 0$. L'application f n'est donc pas surjective : ceci se vérifie aisément car aucun point en-dehors de l'axe des x n'est dans l'image de f.

Soit
$$f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
une application linéaire associée à une matrice A inversible. Alors
$$f_{A^{-1}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
est aussi une application linéaire et on a
$$\begin{aligned} f_A(f_{A^{-1}}(x)) &= AA^{-1}x = Ix = x \\ f_{A^{-1}}(f_A(x)) &= A^{-1}Ax = Ix = x \end{aligned}$$
pour tout x dans $\mathbb{R}^n$. L'application $f_{A^{-1}}$ est appelée l'application inverse de f_A.

On dit qu'une application linéaire $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est inversible si $f = f_A$, où A est une matrice inversible. Si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est inversible, on note
$$f^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
son inverse. Par le théorème 5.21, une application linéaire est inversible si et seulement si elle est bijective.

Exemple 5.23. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^2$ la rotation d'angle θ. Alors $f^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ est la rotation d'angle $-\theta$.

On a
$$\begin{aligned} [f] &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ \left[f^{-1}\right] &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix}. \end{aligned}$$

Théorème 5.24. *Une application*
$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$
est linéaire si et seulement si pour tous les vecteurs u, v de $\mathbb{R}^n$ et pour tout scalaire λ, on a

(1) $f(u+v) = f(u) + f(v)$

(2) $f(\lambda u) = \lambda f(u)$.

Démonstration. Supposons $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ linéaire, et soit A sa matrice standard. On a
$$f(u+v) = A(u+v) = Au + Av = f(u) + f(v)$$
et
$$f(\lambda u) = A(\lambda u) = \lambda Au = \lambda f(u).$$
Réciproquement, supposons (1) et (2). Notons d'abord que (1) implique que
$$f(v_1 + v_2 + \cdots + v_r) = f(v_1) + f(v_2) + \cdots + f(v_r).$$

Posons
$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$
Soit A la matrice $n \times n$ dont les colonnes sont
$$f(e_1), f(e_2), \ldots, f(e_n)$$

et

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

Alors

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

et donc

$$\begin{aligned} Ax &= A(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) \\ &= x_1 A e_1 + x_2 A e_2 + \cdots + x_n A e_n \\ &= x_1 f(e_1) + x_2 f(e_2) + \cdots + x_n f(e_n) \\ &= f(x_1 e_1) + f(x_2 e_2) + \cdots + f(x_n e_n) \\ &= f(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) = f(x). \end{aligned}$$

On a alors que $f = f_A$. C'est une application linéaire. □

Définition 5.25. Les vecteurs

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

sont appelés les *vecteurs de base standard de* $\mathbb{R}^n$.

La démonstration précédente a ainsi montré :

Théorème 5.26. *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ une application linéaire, et soient $e_1, \ldots, e_n$ les vecteurs de base standard de $\mathbb{R}^n$. Alors la matrice standard de f est donnée par*

$$[f] = \begin{pmatrix} f(e_1) & f(e_2) & \cdots & f(e_n) \end{pmatrix}.$$

Chapitre 6

Espaces vectoriels

6.1 Définition et premières propriétés

Rappels et notations

1. Soit E un ensemble. La notation $x \in E$ signifie «x est un élément de E» ou «x appartient à E». Par exemple, $\lambda \in \mathbb{R}$ signifie «λ est un nombre réel».
2. Soient E et E' deux ensembles. La notation $E \subset E'$ signifie «E est contenu dans E'» ou «E est un sous-ensemble de E'».
3. Soient E, E' deux ensembles. Alors l'ensemble des couples

$$E \times E' = \{(e, e') \,|\, e \in E,\, e' \in E'\}$$

s'appelle le produit cartésien de E et E'.

Définition 6.1 (Espace vectoriel)**.** Soit V un ensemble muni de deux opérations : une addition

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (u, v) & \longmapsto & u + v \end{array}$$

et une multiplication par les scalaires

$$\begin{array}{ccc} \mathbb{R} \times V & \longrightarrow & V \\ (\lambda, u) & \longmapsto & \lambda u. \end{array}$$

On dit que V, muni de ces deux opérations, est un *espace vectoriel* si pour tout $u, v, w \in V$ et tout $\lambda, \mu \in \mathbb{R}$, on a les propriétés suivantes :

(1) $u + v = v + u$
(2) $u + (v + w) = (u + v) + w$
(3) Il existe $0 \in V$ tel que $0 + u = u + 0 = u$ pour tout $u \in V$.
(4) Pour tout $u \in V$, il existe $-u \in V$ tel que $u + (-u) = 0$
(5) $\lambda(u + v) = \lambda u + \lambda v$
(6) $(\lambda + \mu)u = \lambda u + \mu u$
(7) $\lambda(\mu u) = (\lambda\mu)u$
(8) $1 \cdot u = u$

Exemple 6.2. L'ensemble $V = \mathbb{R}^n$ muni des opérations habituelles d'addition de vecteurs et de produit d'un vecteur par un scalaire est un espace vectoriel.

Exemple 6.3. Soit V l'ensemble des matrices de taille $n \times m$, muni de l'addition des matrices et de la multiplication d'une matrice par un scalaire. Alors V est un espace vectoriel.

Exemple 6.4. Soit V l'ensemble des fonctions

$$f : \mathbb{R} \longrightarrow \mathbb{R}\,.$$

On définit sur V une opération d'addition :

$$(f+g)(x) = f(x) + g(x)$$

et une opération de multiplication par les scalaires

$$(\lambda f)(x) = \lambda f(x) \qquad \text{pour tout } \lambda \in \mathbb{R}.$$

Ceci en fait un espace vectoriel. Les propriétés sont faciles à vérifier. Par exemple, on définit la fonction nulle

$$\underline{0} : \mathbb{R} \longrightarrow \mathbb{R}$$

par

$$\underline{0}(x) = 0 \qquad \text{pour tout } x \in \mathbb{R}.$$

Alors $f + \underline{0} = f$ pour toute fonction f, et on est donc légitimé à écrire «0» pour la fonction $\underline{0}$. Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction, on définit

$$-f : \mathbb{R} \longrightarrow \mathbb{R}$$

par

$$(-f)(x) = -f(x)$$

Les autres propriétés se vérifient facilement.

Exemple 6.5. Tout plan passant par l'origine est un espace vectoriel (par rapport aux opérations habituelles sur les vecteurs). Le plan est donné par une équation de la forme

$$ax + by + cz = 0$$

où a, b et c sont des scalaires non tous nuls.

Notons V ce plan. Soient $\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)$ et $\left(\begin{smallmatrix} x_1 \\ y_1 \\ z_1 \end{smallmatrix}\right)$ deux éléments de V. Autrement dit,

$$\begin{array}{rcl} ax_0 + by_0 + cz_0 & = & 0 \\ ax_1 + by_1 + cz_1 & = & 0. \end{array}$$

Alors $\left(\begin{smallmatrix} x_0+x_1 \\ y_0+y_1 \\ z_0+z_1 \end{smallmatrix}\right)$ est aussi dans V car on a

$$a(x_0 + x_1) + b(y_0 + y_1) + c(z_0 + z_1) = 0.$$

Les autres propriétés sont aussi faciles à vérifier.

Remarque 6.6. Attention ! Un plan ne contenant pas l'origine n'est pas un espace vectoriel.

Théorème 6.7. *Soit V un espace vectoriel, $v \in V$ et $\lambda \in \mathbb{R}$. Alors on a :*

(a) $0 \cdot v = 0$.

(b) $\lambda \cdot 0 = 0$

(c) $(-1) \cdot v = -v$

(d) Si $\lambda \cdot v = 0$, alors $\lambda = 0$ ou $v = 0$.

6.2 Sous-espaces vectoriels

Définition 6.8. Un sous-ensemble W d'un espace vectoriel V est un sous-espace vectoriel de V si W est un espace vectoriel par rapport aux opérations de V.

Théorème 6.9. *Soit V un espace vectoriel, et soit W un sous-ensemble de V. Alors W est un sous-espace vectoriel de V si et seulement si les conditions suivantes sont vérifiées :*

(a) $0 \in W$;

(b) *si* $u, v \in W$, *alors* $u + v \in W$;

(c) *si* $u \in W$ *et* $\lambda \in \mathbb{R}$, *alors* $\lambda u \in W$.

Démonstration. Soit W un sous-espace vectoriel de V. Alors W satisfait aux axiomes qui définissent un espace vectoriel, en particulier on a (a) (b) et (c). Réciproquement, si W est un sous-ensemble de V vérifiant (a), (b) et (c), alors il est facile de vérifier que W est un espace vectoriel. □

Terminologie : Soit W un sous-ensemble d'un espace vectoriel V. Si la propriété (b) ci-dessus est satisfaite, on dit que W est fermé par rapport à l'addition. Si (c) est satisfait, on dit que W est fermé par rapport à la multiplication par les scalaires.

Remarque 6.10. Soit V un espace vectoriel et W un sous-ensemble de V. Pour montrer que W est un sous-espace vectoriel de V, il suffit de vérifier les propriétés (b) et (c) du théorème ci-dessous. En effet, en prenant $\lambda = 0$ dans (c), on a directement que $0 \in W$.
Il est par contre souvent utile de commencer par vérifier (a) ; ainsi, si $0 \notin W$, on en déduit immédiatemment que W n'est pas un sous-espace vectoriel.

Exemple 6.11. Soit M_n l'espace vectoriel de toutes les matrices de taille $n \times n$ Soit S_n le sous-ensemble des matrices symétriques. Alors S_n est un sous-espace vectoriel de M_n. Il suffit en effet de vérifier que la matrice nulle est symétrique, que la somme de deux matrices symétriques est encore symétrique et finalement que le produit d'une matrice symétrique par un scalaire est une matrice symétrique.

Exemple 6.12. Soit V l'espace vectoriel des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ et soit $\mathcal{P}_n$ le sous-ensemble des fonctions polynomiales de degré au plus n, autrement dit des fonctions de la forme

$$p : \mathbb{R} \longrightarrow \mathbb{R}$$

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

avec $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Alors $\mathcal{P}_n$ est un sous-espace vectoriel de V. En effet :

(a) la fonction nulle est une fonction polynomiale ;

(b) soient $p, q \in \mathcal{P}_n$

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_n x^n \end{aligned}$$

Alors

$$(p + q)(x) = p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n.$$

Donc $p + q \in \mathcal{P}_n$;

(2) soit $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathcal{P}_n$ et $\lambda \in \mathbb{R}$. Alors

$$(\lambda p)(x) = \lambda p(x) = \lambda a_0 + \lambda a_1\ x + \cdots + \lambda a_n x^n.$$

Donc $\lambda p \in \mathcal{P}_n$.

Exemple 6.13. Notons $C(\mathbb{R})$ l'ensemble des fonctions continues $f : \mathbb{R} \longrightarrow \mathbb{R}$. Alors $C(\mathbb{R})$ est un espace vectoriel. C'est un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$. Il suffit de vérifier que la somme de deux fonctions continues est continue et que si $f \in C(\mathbb{R})$ alors λf est continue pour tout $\lambda \in \mathbb{R}$.

Notons $C^1(\mathbb{R})$ l'ensemble des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ qui sont dérivables et de première dérivée continue. Alors $C^1(\mathbb{R})$ est un sous-espace vectoriel de $C(\mathbb{R})$. Notons $C^i(\mathbb{R})$ l'ensemble des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ qui sont dérivables i fois et dont la i-ème dérivée est continue. On a les inclusions

$$C^i(\mathbb{R}) \subset C^{i-1}(\mathbb{R}) \subset \cdots \subset C(\mathbb{R}).$$

$C^i(\mathbb{R})$ est un sous-espace vectoriel de $C(\mathbb{R})$ (et aussi de $C^{i-1}(\mathbb{R})$).

Notons $C^\infty(\mathbb{R})$ l'ensemble des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ qui sont indéfiniment dérivables et dont toutes les dérivées sont continues. Alors $C^\infty(\mathbb{R})$ est un sous-espace vectoriel de $C(\mathbb{R})$. On a les inclusions d'espaces vectoriels

$$\mathcal{P}_n \subset C^\infty(\mathbb{R}) \subset \cdots \subset C(\mathbb{R}).$$

Remarque 6.14. Soit $\{e_1, e_2\}$ la base standard de $\mathbb{R}^2$. Soit d_1 (respectivement d_2) la droite passant par l'origine et de vecteur directeur e_1 (respectivement e_2). La réunion de d_1 et de d_2 n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. En effet, le vecteur

$$e_1 + e_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

n'appartient ni à d_1, ni à d_2.

6.2.1 Espace des solutions d'un système d'équations linéaires homogènes

Un autre exemple d'espace vectoriel est donné par l'ensemble des solutions d'un système linéaire homogène. Soit $Ax = 0$ un système de m équations à n inconnues :

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

On a alors

Théorème 6.15. *Soit $Ax = 0$ un système d'équations linéaires homogènes à n variables. Alors l'ensemble des vecteurs solution est un sous-espace vectoriel de $\mathbb{R}^n$.*

Démonstration. Soit W l'ensemble des vecteurs solution. Le vecteur 0 est un élément de W. Vérifions que W est fermé par rapport à l'addition et par rapport à la multiplication par un scalaire.

Si x et x' sont des vecteurs-solution, alors $Ax = 0$ et $Ax' = 0$ et donc $A(x + x') = Ax + Ax' = 0$ ce qui montre que W est fermé par rapport à l'addition. On a aussi $A(\lambda x) = \lambda Ax = \lambda 0 = 0$ ce qui prouve que W est aussi fermé par rapport à la multiplication par un scalaire. □

Exemple 6.16. Considérons le système

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Les solutions de ce système sont

$$\begin{aligned} x &= 2s - 3t \\ y &= s \\ z &= t \end{aligned}$$

ce qui donne

$$x = 2y - 3z$$

d'où

$$x - 2y + 3z = 0.$$

L'ensemble des solutions est donc un plan passant par l'origine. Nous avons vu que ceci est un espace vectoriel.

6.3 Combinaison linéaire

Définition 6.17. Un vecteur w est appelé combinaison linéaire des vecteurs $v_1, v_2, \ldots, v_r$ si

$$w = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r$$

avec $\lambda_1 \in \mathbb{R}, \ldots, \lambda_r \in \mathbb{R}$.

Exemple 6.18. Soient

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

les vecteurs de la base standard de $\mathbb{R}^3$ et

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

un vecteur quelconque de $\mathbb{R}^3$. On a

$$x = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ce qui montre que x est une combinaison linéaire de e_1, e_2 et e_3. Ainsi, tout vecteur de $\mathbb{R}^3$ est combinaison linéaire de e_1, e_2 et e_3.

Exemple 6.19. Soient $u = \left(\begin{smallmatrix} 1 \\ 2 \\ -1 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} 6 \\ 4 \\ 2 \end{smallmatrix}\right)$ deux vecteurs de $\mathbb{R}^3$. Montrons que $w = \left(\begin{smallmatrix} 9 \\ 2 \\ 7 \end{smallmatrix}\right)$ est combinaison linéaire de u et v. On cherche donc λ et μ tels que

$$\begin{aligned} \left(\begin{smallmatrix} 9 \\ 2 \\ 7 \end{smallmatrix}\right) &= \lambda \left(\begin{smallmatrix} 1 \\ 2 \\ -1 \end{smallmatrix}\right) + \mu \left(\begin{smallmatrix} 6 \\ 4 \\ 2 \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} \lambda \\ 2\lambda \\ -\lambda \end{smallmatrix}\right) + \left(\begin{smallmatrix} 6\mu \\ 4\mu \\ 2\mu \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} \lambda+6\mu \\ 2\lambda+4\mu \\ -\lambda+2\mu \end{smallmatrix}\right). \end{aligned}$$

On a donc

$$\begin{cases} 9 & = \lambda + 6\mu \\ 2 & = 2\lambda + 4\mu \\ 7 & = -\lambda + 2\mu. \end{cases}$$

Une solution de ce système est par exemple

$$\lambda = -3, \quad \mu = 2,$$

ce qui implique que w est combinaison linéaire de u et v. On vérifie que l'on a

$$\begin{pmatrix} 9 \\ 2 \\ 7 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}.$$

Exemple 6.20. Soient $u = \left(\begin{smallmatrix} 1 \\ 2 \\ -1 \end{smallmatrix}\right)$ et $v = \left(\begin{smallmatrix} 6 \\ 4 \\ 2 \end{smallmatrix}\right)$. Montrons que $w = \left(\begin{smallmatrix} 4 \\ -1 \\ 8 \end{smallmatrix}\right)$ n'est pas une combinaison linéaire de u et v. L'égalité

$$\begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}$$

donne le système

$$\begin{cases} 4 & = \lambda + 6\mu \\ -1 & = 2\lambda + 4\mu \\ 8 & = -\lambda + 2\mu. \end{cases}$$

Or ce système n'a aucune solution.

Théorème 6.21. *Soient $v_1, \dots, v_r$ des vecteurs d'un espace vectoriel V. Alors l'ensemble W des combinaisons linéaires de $v_1, \dots, v_r$ est un sous-espace vectoriel de V.*

Démonstration. Le vecteur 0 est dans W car $0 = 0 \cdot v_1 + \cdots + 0 \cdot v_r$. Vérifions que W est fermé pour l'addition et pour la multiplication par un scalaire. Soient $u, v \in W$. Alors, par définition de W,

$$\begin{aligned} u &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r \\ v &= \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_r v_r \,, \end{aligned}$$

où $\lambda_1, \dots, \lambda_r$ et $\mu_1, \dots, \mu_r$ sont des scalaires. Donc

$$u + v = (\lambda_1 + \mu_1)v_1 + (\lambda_2 + \mu_2)v_2 + \cdots + (\lambda_r + \mu_r)v_r \,.$$

Donc $u + v$ est aussi une combinaison linéaire de $v_1, \dots, v_r$. Ceci implique que $u + v \in W$. Pour tout $\alpha \in \mathbb{R}$, on a

$$\alpha u = (\alpha\lambda_1)v_1 + \cdots + (\alpha\lambda_r)v_r \,.$$

Donc αu est combinaison linéaire de $v_1, \dots, v_r$. Ceci implique que $\alpha u \in W$. Donc W est bien un sous-espace vectoriel de V. □

Définition 6.22. Le sous-espace W du théorème précédent est appelé le sous-espace vectoriel de v engendré par $v_1, \dots, v_r$. Si l'on pose $S = \{v_1, \dots, v_n\}$, alors W est noté

$$W = \mathcal{L}(S) = \mathcal{L}(v_1, v_2, \dots, v_r).$$

Théorème 6.23. *Soit V un espace vectoriel, et soient $v_1, \dots, v_r \in V$. Soit W le sous-espace vectoriel de V engendré par $v_1, \dots, v_r$. Alors W est le plus petit sous-espace vectoriel de V contenant $v_1, \dots, v_r$.*

Démonstration. On a clairement que $v_1 \in W, v_2 \in W, \dots, v_r \in W$. Soit W' un autre sous-espace vectoriel de V qui contient $v_1, \dots, v_r$. Comme W' est un sous-espace vectoriel de V, il est fermé par addition et par multiplication par un scalaire. Donc W' contient toutes les combinaisons linéaires de $v_1, \dots, v_r$ ce qui implique que $W \subset W'$. □

Exemple 6.24. Soient v_1 et v_2 deux vecteurs non colinéaires de $\mathbb{R}^3$. Alors l'espace vectoriel W engendré par v_1 et v_2 est un plan.

Exemple 6.25. Soient v_1 et v_2 deux vecteurs colinéaires de $\mathbb{R}^3$. Alors $v_2 = \lambda v_1$. L'espace vectoriel engendré par v_1 et v_2 est une droite.

Exemple 6.26. Soit $\mathcal{P}_n$ l'espace vectoriel des fonctions polynomiales de degré $\leq n$. Alors les polynômes $1, x, \dots, x^n$ engendrent $\mathcal{P}_n$.

Théorème 6.27. *Soit V un espace vectoriel, et soient $S = \{v_1, v_2, \dots, v_r\}$ et $S' = \{v'_1, v'_2, \dots, v'_s\}$ deux ensembles de vecteurs de V. On a*

$$\mathcal{L}(v_1, v_2, \dots, v_r) = \mathcal{L}(v'_1, v'_2, \dots, v'_s)$$

si et seulement si tout vecteur de S est une combinaison linéaire de vecteurs de S' et tout vecteur de S' est une combinaison linéaire de vecteurs de S.

Démonstration. C'est une conséquence immédiate de la définition de $\mathcal{L}(S)$ et de $\mathcal{L}(S')$. □

6.4 Indépendance linéaire

Définition 6.28. Soit $S = \{v_1, \ldots, v_r\}$ un ensemble non vide de vecteurs d'un espace vectoriel. L'équation

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$$

a au moins une solution, notamment la solution triviale $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. Si cette solution est unique, on dira que S est un *ensemble linéairement indépendant*, ou que les vecteurs $v_1, \ldots v_r$ sont *linéairement indépendants*.

Sinon, on dira que S est un ensemble *linéairement dépendant*, ou que les vecteurs $v_1, \ldots, v_r$ sont *linéairement dépendants*.

Remarque 6.29. Tout vecteur non nul est linéairement indépendant. En effet, $\lambda v = 0$ implique $\lambda = 0$ ou $v = 0$.
Le vecteur nul est linéairement dépendant car $\lambda \cdot 0 = 0$ pour tout $\lambda \in \mathbb{R}$. Plus généralement, tout ensemble contenant le vecteur nul est linéairement dépendant. En effet, soient $v_1, \ldots, v_{r-1}$ des vecteurs quelconques et $v_r = 0$. Alors

$$0 \cdot v_1 + \cdots + 0 \cdot v_{r-1} + \lambda \cdot 0 = 0$$

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

Exemple 6.30. Soient $v_1 = \left(\begin{smallmatrix} 2 \\ -1 \\ 0 \\ 3 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} 1 \\ 2 \\ 5 \\ -1 \end{smallmatrix}\right)$ et $v_3 = \left(\begin{smallmatrix} 7 \\ -1 \\ 5 \\ 8 \end{smallmatrix}\right)$. Alors l'ensemble $S = \{v_1, v_2, v_3\}$ est linéairement dépendant, car

$$3v_1 + v_2 - v_3 = 0.$$

Exemple 6.31. Les polynômes $p_1(x) = 1 - x$, $p_2(x) = 5 + 3x - 2x^2$ et $p_3(x) = 1 + 3x - x^2$ forment un ensemble linéairement dépendant, car

$$3p_1 - p_2 + 2p_3 = 0.$$

Exemple 6.32. Soient

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

les vecteurs de la base standard de $\mathbb{R}^3$. Alors l'ensemble $\{e_1, e_2, e_3\}$ est linéairement indépendant. En effet, supposons que $\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0$. Alors

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ce qui donne

$$\begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

et donc

$$\lambda_1 = 0\,, \quad \lambda_2 = 0\,, \quad \lambda_3 = 0.$$

Exemple 6.33. Les polynômes $1, x, \ldots, x^n$ forment un ensemble linéairement indépendant de $\mathcal{P}_n$.

En effet, supposons qu'il existe $a_0, \ldots, a_n \in \mathbb{R}$ tels que

$$a_0 + a_1 x + \cdots + a_n x^n$$

soit la fonction nulle 0. Ceci implique que pour tout $x_0 \in \mathbb{R}$, on a

$$a_0 + a_1 x_0 + \cdots + a_n x_0^n = 0\,.$$

Mais un polynôme de degré n a au plus n racines. On doit donc avoir

$$a_0 = a_1 = \cdots = a_n = 0\,.$$

Théorème 6.34. *Un ensemble S est linéairement dépendant si et seulement si au moins l'un des vecteurs de S est une combinaison linéaire des autres vecteurs de S.*

Démonstration. Soit $S = \{v_1, v_2, \dots, v_r\}$ un ensemble linéairement dépendant. Alors il existe $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ non tous nuls tels que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = 0.$$

Comme les $\lambda_1, \dots, \lambda_r$ ne sont pas tous nuls, il en existe au moins un qui est non nul. Supposons $\lambda_i \neq 0$, avec $1 \leq i \leq r$. On a

$$\lambda_i v_i = -\lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_r v_r.$$

Comme $\lambda_i \neq 0$, ceci implique

$$v_i = \left(-\frac{\lambda_1}{\lambda_i}\right) v_1 + \cdots + \left(-\frac{\lambda_{i-1}}{\lambda_i}\right) v_{i-1} + \left(-\frac{\lambda_{i+1}}{\lambda_i}\right) v_{i+1} + \cdots + \left(-\frac{\lambda_r}{\lambda_i}\right) v_r$$

et alors v_i est une combinaison linéaire des autres vecteurs.

Réciproquement, supposons que l'un des vecteurs — disons v_i — soit combinaison linéaire des autres vecteurs. Alors on a

$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_r v_r.$$

Ceci implique

$$\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} - v_i + \lambda_{i+1} v_{i+1} + \cdots + \lambda_r v_r = 0.$$

Donc $v_1, \dots, v_r$ sont linéairement dépendants. □

6.4.1 Interprétation géométrique de la dépendance linéaire

- Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, deux vecteurs sont linéairement dépendants si et seulement s'ils sont colinéaires.
- Dans $\mathbb{R}^3$, trois vecteurs sont linéairement dépendants si et seulement s'ils sont sur le même plan (coplanaires).

Théorème 6.35. *Soit $S = \{v_1, v_2, \dots, v_r\}$ un sous-ensemble de $\mathbb{R}^n$. Si S contient plus de n éléments, alors S est linéairement dépendant.*

Démonstration. Supposons que

$$v_1 = \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{pmatrix}, \quad v_2 = \begin{pmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{pmatrix}, \quad \dots, \quad v_r = \begin{pmatrix} v_{r1} \\ v_{r2} \\ \vdots \\ v_{rn} \end{pmatrix}.$$

L'équation

$$x_1 v_1 + x_2 v_2 + \cdots + x_r v_r = 0$$

donne alors le système suivant

$$\begin{array}{rcl} v_{11}x_1 + v_{21}x_2 + \cdots + v_{r1}x_r & = & 0 \\ v_{12}x_1 + v_{22}x_2 + \cdots + v_{r2}x_r & = & 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & & \\ v_{1n}x_1 + v_{2n}x_2 + \cdots + v_{rn}x_r & = & 0 \end{array}$$

C'est un système homogène de n équations à r inconnues. Lorsque $r > n$, ce système a des solutions non triviales ce qui montre que la famille S est linéairement dépendante. □

6.5 Bases et dimension

Définition 6.36 (Base d'un espace vectoriel)**.** Soit V un espace vectoriel, et $S = \{v_1, v_2, \ldots, v_n\}$ un sous-ensemble de V. Alors S est une *base de* V si et seulement si les deux conditions suivantes sont vérifiées :

(1) l'ensemble S est linéairement indépendant ;

(2) l'ensemble S engendre V, c'est-à-dire $V = \mathcal{L}(S)$.

Théorème 6.37. *Si $S = \{v_1, v_2, \ldots, v_n\}$ est une base de l'espace vectoriel V, alors tout vecteur $v \in V$ s'exprime de façon unique comme combinaison linéaire d'éléments de S :*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

avec $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ déterminés de façon unique.

Démonstration. Par définition, S engendre V, donc pour tout $v \in V$ il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Il reste à montrer l'unicité des $\lambda_1, \lambda_2, \ldots, \lambda_n$. Soient $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{R}$ tels que

$$v = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n.$$

Alors on a

$$(\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \cdots + (\lambda_n - \mu_n)v_n = 0.$$

Comme $S = \{v_1, \ldots, v_n\}$ est linéairement indépendant, ceci implique

$$\lambda_1 - \mu_1 = 0, \quad \lambda_2 - \mu_2 = 0, \quad \ldots, \quad \lambda_n - \mu_n = 0$$

et donc

$$\lambda_1 = \mu_1, \quad \lambda_2 = \mu_2, \quad \ldots, \lambda_n = \mu_n.$$

□

Définition 6.38 (Coordonnées par rapport à une base)**.** Les $\lambda_1, \ldots, \lambda_n$ du théorème précédent sont appelés *les coordonnées de v dans la base S.*

Exemple 6.39. Les vecteurs v_1 et v_2 de la figure ci-dessous forment une base du plan car on a $v = av_1 + bv_2$ pour tout vecteur v du plan où a et b sont des nombres réels uniquement déterminés.

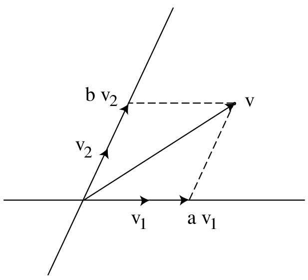


Figure 8.4

Exemple 6.40. Soient $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ les vecteurs de la base standard de $\mathbb{R}^3$. Alors $S = \{e_1, e_2, e_3\}$ est une base de $\mathbb{R}^3$ car S est linéairement indépendant et engendre $\mathbb{R}^3$.

Exemple 6.41. Base standard de $\mathbb{R}^n$: les vecteurs

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

forment une base de $\mathbb{R}^n$, appelée la base standard de $\mathbb{R}^n$.

Exemple 6.42. Soient $v_1 = \left(\begin{smallmatrix} 1 \\ 2 \\ 1 \end{smallmatrix}\right)$, $v_2 = \left(\begin{smallmatrix} 2 \\ 9 \\ 0 \end{smallmatrix}\right)$ et $v_3 = \left(\begin{smallmatrix} 3 \\ 3 \\ 4 \end{smallmatrix}\right)$. Montrons que l'ensemble $S = \{v_1, v_2, v_3\}$ est une base de $\mathbb{R}^3$. Montrons d'abord que S engendre $\mathbb{R}^3$. Soit $a = \left(\begin{smallmatrix} a_1 \\ a_2 \\ a_3 \end{smallmatrix}\right)$ un vecteur quelconque de $\mathbb{R}^3$. On cherche $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que

$$a = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3.$$

Ceci se reformule comme suit :

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} &= \lambda_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 9 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 + 2\lambda_2 + 3\lambda_3 \\ 2\lambda_1 + 9\lambda_2 + 3\lambda_3 \\ \lambda_1 + 4\lambda_3 \end{pmatrix}. \end{aligned}$$

Ceci conduit au système suivant :

$$\left\{ \begin{array}{ccccccc} \lambda_1 & + & 2\lambda_2 & + & 3\lambda_3 & = & a_1 \\ 2\lambda_1 & + & 9\lambda_2 & + & 3\lambda_3 & = & a_2 \\ \lambda_1 & & & + & 4\lambda_3 & = & a_3. \end{array} \right.$$

Il reste à montrer que ce système a une solution $\lambda_1, \lambda_2, \lambda_3$. Pour montrer que S est linéairement indépendant, il faut montrer que l'unique solution de

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

est

$$\lambda_1 = \lambda_2 = \lambda_3 = 0\,.$$

Ceci conduit à montrer que le système

$$\left\{ \begin{array}{ccccccc} \lambda_1 & + & 2\lambda_2 & + & 3\lambda_3 & = & 0 \\ 2\lambda_1 & + & 9\lambda_2 & + & 3\lambda_3 & = & 0 \\ \lambda_1 & & & + & 4\lambda_3 & = & 0 \end{array} \right.$$

a une unique solution

$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Remarquons que les deux systèmes ont la même matrice de coefficients. On peut donc montrer simultanément que S engendre $\mathbb{R}^3$ et que S est linéairement indépendant en montrant que la matrice des coefficients est inversible. Cette matrice est

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}$$

et son déterminant vaut $\det(A) = -1$. Donc S est une base de $\mathbb{R}^3$.

Remarque 6.43. L'exemple précédent se généralise de la façon suivante : pour montrer que n vecteurs de $\mathbb{R}^n$ forment une base de $\mathbb{R}^n$, il suffit de montrer la chose suivante : la matrice A constituée des composantes de ces vecteurs (chaque vecteur formant une colonne de A) est de déterminant non nul.

Exemple 6.44. Base standard de $\mathcal{P}_n$: la famille $S = \{1, x, \ldots, x^n\}$ est une base de $\mathcal{P}_n$, appelée base standard de $\mathcal{P}_n$. En effet, nous avons déjà vu que $\mathcal{L}(S) = \mathcal{P}_n$ et que S est linéairement indépendant.

Exemple 6.45. Base standard de M_{22}. On note M_{22} l'ensemble des matrices 2×2. Nous avons vu que M_{22} est un espace vectoriel. Soient

$$\begin{array}{rclcrcl} M_1 & = & \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \quad & M_2 & = & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ \\ M_3 & = & \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} & \quad & M_4 & = & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{array}$$

L'ensemble $S = \{M_1, M_2, M_3, M_4\}$ est une base de l'espace vectoriel M_{22}, appelée base standard de M_{22}.

• Montrons d'abord que $\mathcal{L}(S) = M_{22}$. Soit $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ un élément quelconque de M_{22}. On a

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} \\ &= a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= aM_1 + bM_2 + cM_3 + dM_4 \end{aligned}$$

ce qui montre que $\mathcal{L}(S) = M_{22}$.

• Montrons maintenant que S est linéairement indépendant. Pour cela, supposons que $\lambda_1 M_1 + \lambda_2 M_2 + \lambda_3 M_3 + \lambda_4 M_4 = 0$. Ceci implique

$$\begin{aligned} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} &= \lambda_1\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \lambda_3\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \lambda_4\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \lambda_2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \lambda_3 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \lambda_4 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} \end{aligned}$$

ce qui implique $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Donc S est linéairement indépendant.

Définition 6.46. Soit V un espace vectoriel non nul. On dit que V est de *dimension finie* s'il existe un ensemble fini S de V qui est une base de V. S'il n'existe aucun tel ensemble, alors on dit que V est de *dimension infinie.*

Exemple 6.47. Les espaces vectoriels $\mathbb{R}^n$, M_{22}, $\mathcal{P}_n$ sont de dimension finie alors que $C^\infty(\mathbb{R})$, $C^1(\mathbb{R})$, $C(\mathbb{R})$ sont de dimension infinie.

Théorème 6.48. *Soit V un espace vectoriel de dimension finie, et soit $S = \{v_1, \ldots, v_n\}$ une base de V. Alors*

(1) Tout sous-ensemble de V ayant plus de n éléments est linéairement dépendant.

(2) Aucun sous-ensemble de V ayant moins de n éléments n'engendre V.

Démonstration. (1) Soit $S' = \{w_1, \ldots, w_m\}$ avec $w_1, \ldots, w_m \in V$ et $m > n$. Montrons que S' est linéairement dépendant. Comme $S = \{v_1, v_2, \ldots, v_n\}$ est une base de V, on a

$$\begin{array}{rcl} w_1 & = & a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ w_2 & = & a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ & & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_m & = & a_{1m}v_1 + a_{2m}v_2 + \cdots + a_{nm}v_n \end{array}$$

Pour montrer que S' est linéairement dépendant, on doit trouver des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_m$ non tous nuls tels que $\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_m w_m = 0$. On peut reécrire cette équation comme suit

$$\begin{array}{rl} & (\lambda_1 a_{11} + \lambda_2 a_{12} + \cdots + \lambda_m a_{1m}) v_1 + \\ + & (\lambda_1 a_{21} + \lambda_2 a_{22} + \cdots + \lambda_m a_{2m}) v_2 + \\ + & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + & (\lambda_1 a_{n1} + \lambda_2 a_{n2} + \cdots + \lambda_m a_{nm}) v_n = 0. \end{array}$$

Par l'indépendance linéaire de S, on obtient

$$\begin{array}{c} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1m}\lambda_m = 0 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2m}\lambda_m = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\lambda_1 + a_{n2}\lambda_2 + \cdots + a_{nm}\lambda_m = 0. \end{array}$$

C' est un système homogène qui a plus d'inconnues que d'équations. Il a donc des solutions non triviales $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ telles que $\lambda_1 w_1 + \cdots + \lambda_m w_m = 0$ ce qui montre que $S' = \{w_1, \ldots, w_m\}$ est linéairement dépendant.

La démonstration de (2) est similaire. □

Corollaire 6.49. *Toutes les bases d'un espace vectoriel de dimension finie ont le même nombre d'éléments.*

Définition 6.50 (Dimension d'un espace vectoriel)**.** La dimension d'un espace vectoriel de dimension finie V, notée $\dim(V)$, est par définition le nombre d'éléments d'une base de V. L'espace vectoriel $V = \{0\}$ est de dimension 0.

Exemple 6.51. Par les exemples qui précédent, on a

1. $\dim(\mathbb{R}^n) = n$ (Exemple 6.41) ;
2. $\dim(\mathcal{P}_n) = n + 1$ (Exemple 6.44) ;
3. $\dim(M_{22}) = 4$ (Exemple 6.45) ;
4. Plus généralement, si M_{mn} est l'espace vectoriel des matrices $m \times n$, on a $\dim(M_{mn}) = mn$. En effet, l'ensemble

$$M_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \vdots \\ \vdots & & & \vdots \\ 0 & \ldots & \ldots & 0 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 \end{pmatrix}$$

$$M_n = \begin{pmatrix} 0 & \ldots & 0 & 1 \\ \vdots & & & \vdots \\ 0 & \ldots & \ldots & 0 \end{pmatrix} \qquad M_{n+1} = \begin{pmatrix} 0 & \ldots & 0 \\ 1 & & \vdots \\ 0 & & \vdots \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

$$M_{mn} = \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ \vdots & & 0 \\ 0 & \ldots & 1 \end{pmatrix}$$

est une base de M_{mn} et posséde mn éléments.

Exemple 6.52. On a vu la base standard $\{1, x, \ldots, x^n\}$ des polynômes de degré au plus n. Voici une autre base, souvent commode pour les calculs : on fixe des nombres $\lambda_0, \lambda_1, \ldots, \lambda_n$ distincts, et on considère la base

$$\begin{aligned}\{&(x-\lambda_0)(x-\lambda_1)\ldots(x-\lambda_{n-1}),\\ &\qquad(x-\lambda_0)(x-\lambda_1)\ldots(x-\lambda_{n-2})(x-\lambda_n),\ldots,\\ &\qquad(x-\lambda_0)\ldots(x-\lambda_i)\ldots(x-\lambda_{i+2})\ldots(x-\lambda_n),\ldots,\\ &\qquad\qquad(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_n)\}\end{aligned}$$

où à chaque fois on prend tous les facteurs $(x - \lambda_i)$ sauf un.

Cette base est trés utile parce que le i-iéme vecteur s'annulle en tous les λ_j pour $j \neq i$.

Exemple 6.53. Nous avons vu que l'ensemble des solutions d'un système d'équations linéaires homogène est un espace vectoriel.

On considère le système

$$\left\{\begin{array}{rcrcrcrcrcl} 2x_1 & + & 2x_2 & - & x_3 & & & + & x_5 & = & 0\\ -x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & + & x_5 & = & 0\\ x_1 & + & x_2 & - & 2x_3 & & & - & x_5 & = & 0\\ & & & & x_3 & + & x_4 & + & x_5 & = & 0.\end{array}\right.$$

On vérifie que la solution générale de ce système est

$$x_1 = -s - t \quad x_2 = s \quad x_3 = -t \quad x_4 = 0 \quad x_5 = t\,.$$

Donc les vecteurs solution s'écrivent sous la forme

$$\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{pmatrix} = \begin{pmatrix} -s-t\\ s\\ -t\\ 0\\ t\end{pmatrix} = \begin{pmatrix} -s\\ s\\ 0\\ 0\\ 0\end{pmatrix} + \begin{pmatrix} -t\\ 0\\ -t\\ 0\\ t\end{pmatrix} = s\begin{pmatrix} -1\\ 1\\ 0\\ 0\\ 0\end{pmatrix} + t\begin{pmatrix} -1\\ 0\\ -1\\ 0\\ 1\end{pmatrix}.$$

Ceci montre que les vecteurs

$$v_1 = \begin{pmatrix} -1\\ 1\\ 0\\ 0\\ 0\end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} -1\\ 0\\ -1\\ 0\\ 1\end{pmatrix}$$

engendrent l'espace des solutions du système. D'autre part, on vérifie que v_1 et v_2 sont linéairement indépendants. Donc $\{v_1, v_2\}$ est une base de l'espace des solutions du système. Ceci montre que cet espace vectoriel est de dimension 2.

Théorème 6.54. *Soit V un espace vectoriel et S un sous-ensemble fini et non vide de V. Alors*

(a) si S est linéairement indépendant et si $v \notin \mathcal{L}(S)$, alors $S \cup \{v\}$ est encore linéairement indépendant ;

(b) si $v \in S$ est une combinaison linéaire d'autres vecteurs de S, alors

$$\mathcal{L}(S \backslash \{v\}) = \mathcal{L}(S)$$

où $S \backslash \{v\}$ désigne l'ensemble obtenu en retirant v de S.

Théorème 6.55. *Soit V un espace vectoriel de dimension n. Soit S un sous-ensemble de V contenant exactement n vecteurs. Alors les affirmations suivantes sont équivalentes :*

(1) S est une base de V ;

(2) S engendre V ;

(3) S est linéairement indépendant.

Démonstration. Il est clair que (1) implique (2) et (3). Montrons que (2) implique (1). Pour cela, il suffit de montrer que (2) implique (3). Autrement dit, il suffit de montrer que que si $V = \mathcal{L}(S)$ alors S est linéairement indépendant. On le montre par l'absurde. Supposons S linéairement dépendant. Alors il existe $v \in S$ qui est une combinaison linéaire des autres vecteurs de S. Par le théorème précédent, ceci implique que

$$\mathcal{L}(S\backslash\{v\}) = \mathcal{L}(S) = V.$$

Mais $S\backslash\{v\}$ contient seulement $n-1$ vecteurs, ce qui contredit l'hypothèse $\dim(V) = n$.

Montrons que (3) implique (1). Il suffit de montrer que (3) implique (2). Autrement dit, il suffit de montrer que si S est linéairement indépendant, alors S engendre V. Supposons que ce ne soit pas le cas. Alors il existe $v \in V$ tel que $v \notin \mathcal{L}(S)$. Par le théorème précédent, ceci implique que $S \cup \{v\}$ est linéairement indépendant. Mais $S \cup \{v\}$ contient $n+1$ vecteurs ce qui contredit à nouveau l'hypothèse sur la dimension de V. □

Théorème 6.56. *Soit V un espace vectoriel de dimension finie et soit S un sous-ensemble fini de V.*

(a) Si S engendre V, alors on peut rétrécir S en une base de V : il existe $v_1, \dots, v_r \in S$ tels que $S\backslash\{v_1, \dots, v_r\}$ soit une base de V.

(b) Si S est linéairement indépendant, alors on peut compléter S en une base de V : il existe $w_1, \dots, w_k \in V$ tels que $S \cup \{w_1, \dots, w_k\}$ soit une base de V.

Démonstration. Pour (a), on procède par récurrence sur $\dim V - \#S$. Supposons que S ne soit pas linéairement indépendant (sinon on a déjà gagné). On a alors une combinaison linéaire $c_1u_1 + \cdots + c_nu_n = 0$ avec des $u_i \in S$, et au moins un des c_i, disons c_1, non nul. On pose alors $v_1 = u_1$ et $S' = S \setminus \{v_1\}$. Par le théorème 6.54(a), S' engendre aussi V et est plus petit ; on peut donc, par récurrence, enlever $v_2, \dots, v_r$ à S' pour obtenir une base de V ; et donc enlever $v_1, \dots, v_r$ à S pour obtenir une base de V.

Pour (b), on procède par récurrence sur $\#S - \dim V$. Supposons que S n'engendre pas V. Soit $w_1 \in V \setminus \mathcal{L}(S)$, et $S' = S \cup \{w_1\}$. Alors, par le théorème 6.54(b), S' est toujours linéairement indépendant, mais plus grand, donc par récurrence peut se compléter en une base $S' \cup \{w_2, \dots, w_k\}$ de V. Ainsi $S \cup \{w_1, \dots, w_k\}$ est une base de V. □

Théorème 6.57. *Soit V un espace vectoriel de dimension finie, et soit W un sous-espace vectoriel de V. Alors*

$$\dim(W) \leq \dim(V).$$

De plus, si $\dim(W) = \dim(V)$*, alors $W = V$.*

6.6 Espace des lignes et colonnes d'une matrice

Définition 6.58. Soit A une matrice de taille $m \times n$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Les vecteurs de $\mathbb{R}^n$

$$\begin{aligned} \ell_1 &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix}, \\ \ell_2 &= \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ \ell_m &= \begin{pmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \end{aligned}$$

sont appelés vecteurs ligne de A, et les vecteurs de $\mathbb{R}^m$

$$c_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad c_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \dots, \quad c_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

sont appelés vecteurs colonne de A.

Définition 6.59. Soit A une matrice de taille $m \times n$. Le sous-espace de $\mathbb{R}^n$ engendré par les vecteurs ligne de A est appelé *espace des lignes de A*, et le sous-espace de $\mathbb{R}^m$ engendré par les vecteurs colonne de A est appelé *espace des colonnes de A*.

Ainsi la dimension de l'espace des lignes de A est le nombre de lignes linéairement indépendantes de A ; la dimension de l'espace des colonnes est le nombre de colonnes linéairement indépendantes de A. On verra bientôt que ces dimensions sont égales.

Définition 6.60 (Noyau (ou Nilespace) d'une matrice)**.** L'espace des solutions du système d'équations homogène

$$Ax = 0$$

est appelé *noyau de A*, ou nilespace de A. Le noyau (nilespace) de A est un sous-espace vectoriel de $\mathbb{R}^n$. Il est noté $\mathrm{Ker}(A)$.

Théorème 6.61. *Soit A une matrice $m \times n$. Alors*

1. *L'espace des colonnes de A est $\mathbb{R}^m$ si et seulement si f_A est surjective ;*
2. *Le noyau* $\ker(A) = 0$ *si et seulement si f_A est injective ;*
3. *$Ax = b$ admet une solution si et seulement si b appartient á l'espace des colonnes de A ;*
4. *L'espace des lignes de A est l'espace des colonnes de A^T ;*
5. *Soit x_0 un vecteur solution du système $Ax = b$, et soit $\{v_1, v_2, \dots, v_k\}$ une base du noyau de A. Alors l'ensemble des solutions du système $Ax = b$ est*

$$\{x_0 + c_1 v_1 + \cdots + c_k v_k \,:\, c_i \in \mathbb{R}\}.$$

Démonstration. 1. L'espace des colonnes est l'image de l'application f_A. Ainsi l'espace des colonnes est $\mathbb{R}^m$ si et seulement si l'image de f_A est $\mathbb{R}^m$, c'est-à-dire si et seulement si f_A est surjective.

2. Le noyau de A est l'ensemble des $x \in \mathbb{R}^n$ tels que $Ax = 0$. Ainsi le noyau de A est 0 si et seulement si $Ax = 0$ n'admet que $x = 0$ comme solution, ce qui veut dire que A est injective.
3. $Ax = b$ admet une solution si et seulement si b est dans l'image de f_A, qui est l'espace des colonnes de A.
4. C'est évident.
5. On calcule d'abord

$$A(x_0 + c_1 v_1 + \cdots + c_k v_k) = Ax_0 + c_1 A v_1 + \cdots + c_k A v_k = b + 0 + \cdots + 0 = b,$$

ce qui montre que $x_0 + c_1 v_1 + \cdots + c_k v_k$ est une solution quels que soient $c_1, \dots, c_k$.
Ensuite, soit x_1 une autre solution à $Ax = b$. On a alors $A(x_1 - x_0) = b - b = 0$, donc $x_1 - x_0 \in \ker(A)$. On peut donc écrire $x_1 - x_0 = c_1 v_1 + \cdots + c_k v_k$ dans la base de $\ker(A)$. On a donc $x_1 = x_0 + c_1 v_1 + \cdots + c_k v_k$ pour certains $c_1, \dots, c_k \in \mathbb{R}$.

□

Théorème 6.62. *Les opérations élémentaires sur les lignes ne changent pas le noyau de la matrice.*

Démonstration. Soit A une matrice. Le noyau de A est par définition l'espace vectoriel des solutions du système $Ax = 0$. Nous avons vu que les opérations élémentaires sur A ne changent pas les solutions du système $Ax = 0$. Ceci implique que les opérations élémentaires ne changent pas le noyau de A. □

Théorème 6.63. *Les opérations élémentaires sur les lignes ne changent pas l'espace des lignes de la matrice.*

Démonstration. Soit A une matrice, et soient $\ell_1, \ldots, \ell_m$ ses vecteurs ligne. Soit B une matrice obtenue à partir de A par une opération élémentaire sur les lignes. Si l'opération consiste à permuter deux lignes de A, alors B a les mêmes lignes que A, donc l'espace des lignes est le même. Si l'opération consiste à multiplier une ligne de A par un scalaire non nul, ou à ajouter un multiple d'une ligne à une autre, alors les lignes de B sont des combinaisons linéaires de $\ell_1, \ldots, \ell_m$. Donc l'espace des lignes de B est contenu dans l'espace des lignes de A. Comme les opérations élémentaires sont réversibles, on voit que l'espace des lignes de A est contenu dans l'espace des lignes de B. Donc ces deux espaces vectoriels sont égaux. $\square$

Théorème 6.64. *Soit A une matrice. Alors les lignes non nulles de la matrice échelonnée réduite forment une base de l'espace des lignes de A.*

Démonstration. La matrice échelonnée réduite a autant de 1 directeurs que de lignes non nulles. Les 1 directeurs sont dans des colonnes différentes. Donc les vecteurs ligne non nuls de la matrice échelonnée réduite sont linéairement indépendants. Comme ils engendrent également l'espace des lignes, ils forment une base de cet espace. $\square$

Exemple 6.65.

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

est une matrice échelonnée réduite. Les lignes non nulles de A sont

$$\ell_1 = (1\,1\,0\,0\,1\,0), \quad \ell_2 = (0\,0\,1\,0\,1\,0) \quad \text{et} \quad \ell_3 = (0\,0\,0\,1\,0\,0).$$

Elles sont linéairement indépendantes. En effet, soient $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que

$$\lambda_1 \ell_1 + \lambda_2 \ell_2 + \lambda_3 \ell_3 = 0\,.$$

Alors on a

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

Donc $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Application. Nous donnons la méthode pour extraire une base d'une famille de vecteurs.
Soient

$$x = \begin{pmatrix} 2 \\ 3 \\ 8 \\ 19 \end{pmatrix}, \quad y = \begin{pmatrix} -1 \\ -10 \\ -21 \\ -69 \end{pmatrix}, \quad z = \begin{pmatrix} 1 \\ 11 \\ 23 \\ 76 \end{pmatrix}.$$

Nous voulons extraire une base de cette famille de vecteurs. Nous écrivons ces trois vecteurs en ligne dans une matrice 3×4. On obtient

$$A = \begin{pmatrix} 2 & 3 & 8 & 19 \\ -1 & -10 & -21 & -69 \\ 1 & 11 & 23 & 76 \end{pmatrix}.$$

Par le théorème 6.64, les lignes non nulles de la matrice échelonnée réduite forment une base de l'espace des lignes de la matrice A. Il suffit donc d'echelonner A et de la réduire pour trouver une base.

$$\begin{pmatrix} 2 & 3 & 8 & 19 \\ -1 & -10 & -21 & -69 \\ 1 & 11 & 23 & 76 \end{pmatrix}$$

$\ell_2 \longrightarrow 2\ell_2 + \ell_1$

$$\begin{pmatrix} 2 & 3 & 8 & 19 \\ 0 & -17 & -34 & -119 \\ 1 & 11 & 23 & 76 \end{pmatrix}$$

$\ell_3 \longrightarrow 2\ell_3 - \ell_1$

$$\begin{pmatrix} 2 & 3 & 8 & 19 \\ 0 & -17 & -34 & -119 \\ 0 & 19 & 38 & 133 \end{pmatrix}$$

$\ell_3 \longrightarrow 17\ell_3 + 19\ell_2$

$$\begin{pmatrix} 2 & 3 & 8 & 19 \\ 0 & -17 & -34 & -119 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\ell_2 \longrightarrow -\frac{1}{17}\ell_2$

$$\begin{pmatrix} 2 & 3 & 8 & 19 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\ell_1 \longrightarrow \ell_1 - 3\ell_2$

$$\begin{pmatrix} 2 & 0 & 2 & -2 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\ell_1 \longrightarrow \frac{1}{2}\ell_1$

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Une base de l'espace engendré par les vecteurs x, y et z est

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 7 \end{pmatrix} \right\}.$$

En fait, on peut également prendre comme vecteurs de base les lignes non nulles de la matrice échelonnée (non nécessairement réduite).

Théorème 6.66. *Soit A une matrice de taille $m \times n$. Alors la dimension de l'espace des lignes de A est égale à la dimension de l'espace des colonnes de A. Cette dimension est appelée le rang de la matrice A et est notée* $\mathrm{rg}(A)$.

Démonstration. Soit

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Soient $\ell_1, \dots, \ell_m$ les lignes de A et $c_1, \dots, c_n$ les colonnes de A. Supposons que la dimension de l'espace des lignes de A soit k, et soit $\{b_1, \dots, b_k\}$ une base de cet espace. On peut donc écrire chaque ligne comme combinaison linéaire des vecteurs de cette base :

$$\begin{aligned} \ell_1 &= \lambda_{11} b_1 + \cdots + \lambda_{1k} b_k \\ & \dots\dots\dots\dots\dots\dots \\ \ell_m &= \lambda_{m1} b_1 + \cdots + \lambda_{mk} b_k \,. \end{aligned}$$

Ceci donne le système

$$\begin{array}{rcl} a_{1j} & = & \lambda_{11}b_{1j} + \cdots + \lambda_{1k}b_{kj} \\ a_{2j} & = & \lambda_{21}b_{2j} + \cdots + \lambda_{2k}b_{kj} \\ & & \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{mj} & = & \lambda_{m1}b_{1j} + \cdots + \lambda_{mk}b_{kj} \end{array}$$

avec

$$b_j = (b_{j1} \ldots b_{jn}) \,.$$

On a donc

$$c_j = b_{1j} \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \vdots \\ \lambda_{m1} \end{pmatrix} + \cdots + b_{kj} \begin{pmatrix} \lambda_{1k} \\ \lambda_{2l} \\ \vdots \\ \lambda_{mk} \end{pmatrix} .$$

Ainsi l'espace des colonnes de A est engendré par k vecteurs. La dimension de cet espace est donc au plus égale à k. Ce raisonnement est valable pour n'importe quelle matrice et en particulier pour la transposée de A. Or, les colonnes de A^T sont les lignes de A, et les lignes de A^T sont les colonnes de A. Gréce au raisonnement précédent appliquée à A^T, on voit que la dimension de l'espace des lignes de A est au plus égale à la dimension de l'espace des colonnes de A. Ces deux dimensions sont donc égales. □

Définition 6.67. Soit A une matrice. La *nullité de A* est par définition la dimension du noyau de A, et est noté $\mathrm{null}(A)$.

Théorème 6.68 (Théorème du rang)**.** *Soit A une matrice à n colonnes. Alors*

$$\mathrm{rg}(A) + \mathrm{null}(A) = n.$$

Démonstration. Le système linéaire homogène

$$Ax = 0$$

a n variables. Donc la somme du nombre des variables directrices et du nombre des variables libres est n. Or, le nombre de variables directrices est égal au nombre de 1 directeurs dans la forme échelonnée réduite de A. Ce nombre est égal au rang de (A). D'autre part, le nombre de variables libres est égal au nombre de paramétres de la solution générale du système, ce qui est égal à la dimension de l'espace des solutions du système, qui est par définition égal à la nullité de A. On a donc bien

$$\mathrm{rg}(A) + \mathrm{null}(A) = n.$$

□

Théorème 6.69. *Soit A une matrice de taille $n \times n$. Alors les affirmations suivantes sont équivalentes :*

(1) *A est inversible*

(2) $\mathrm{rg}(A) = n$

(3) $\mathrm{null}(A) = 0$.

Démonstration. Nous avons vu que A est inversible si et seulement si sa forme échelonnée réduite est la matrice identité I_n. Ceci se passe si et seulement si toutes les variables sont directrices. Mais le nombre de variables directrices est égal au rang de A. On a donc (1) $\Leftrightarrow$ (2). L'équivalence (2) $\Leftrightarrow$ (3) résulte du théorème précédent. □

6.7 Changements de bases

Soit V un espace vectoriel, et soit

$$B = \{v_1, v_2, \ldots, v_n\}$$

une base de V. Nous avons vu que tout $v \in V$ s'écrit de façon unique

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

avec $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. Les $\lambda_1, \ldots, \lambda_n$ sont appelés les coordonnées de v dans la base B, et on note

$$(v)_B = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

6.7.1 Changement de bases en 2 dimensions

Soient

$$B = \{u_1, u_2\} \qquad \text{et} \qquad B' = \{u'_1,\ u'_2\}$$

deux bases du même espace vectoriel. Supposons que

$$(u'_1)_B = \begin{pmatrix} a \\ b \end{pmatrix}$$

et

$$(u'_2)_B = \begin{pmatrix} c \\ d \end{pmatrix}$$

autrement dit que

$$\begin{aligned} u'_1 &= au_1 + bu_2 \\ u'_2 &= cu_1 + du_2. \end{aligned}$$

Soit $v \in V$ tel que

$$(v)_{B'} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix},$$

c'est-à-dire que

$$v = \lambda_1 u'_1 + \lambda_2 u'_2 .$$

Nous aimerions trouver les coordonnées de v par rapport à la base B. On a

$$\begin{aligned} v &= \lambda_1(au_1 + bu_2) + \lambda_2(cu_1 + du_2) \\ &= (\lambda_1 a + \lambda_2 c)u_1 + (\lambda_1 b + \lambda_2 d)u_2 \end{aligned}$$

On a donc

$$(v)_B = \begin{pmatrix} \lambda_1 a + \lambda_2 c \\ \lambda_1 b + \lambda_2 d \end{pmatrix}.$$

Ceci peut également s'écrire

$$\begin{aligned} (v)_B &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \\ &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} (v)_{B'} . \end{aligned}$$

La matrice

$$P_{B,B'} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

est la matrice de changement de bases (ou de passage de bases) de B' à B.

6.7.2 Dimension quelconque

Soit V un espace vectoriel de dimension n, et soient

$$B = \{u_1, \ldots, u_n\} \qquad \text{et} \qquad B' = \{u'_1, \ldots, u'_n\}$$

deux bases de V. Alors

$$(v)_B = P_{B,B'} \cdot (v)_{B'}$$

où $P_{B,B'}$ est la matrice dont les colonnes sont les coordonnées de $u'_1, \ldots, u'_n$ par rapport à la base B ; autrement dit,

$$P_{B,B'} = \begin{pmatrix} (u'_1)_B & (u'_2)_B & \ldots & (u'_n)_B \end{pmatrix}.$$

On appelle $P_{B,B'}$ la matrice de changement de base (ou de passage de base) de B' à B.

Exemple 6.70. Soient $B = \{u_1, u_2\}$ et $B' = \{u'_1, u'_2\}$ deux bases de $\mathbb{R}^2$, avec

$$\begin{array}{rclcrcl} u_1 & = & \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), & \qquad & u_2 & = & \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right); \\ u'_1 & = & \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right), & & u'_2 & = & \left(\begin{smallmatrix}2\\1\end{smallmatrix}\right). \end{array}$$

On a

$$\begin{array}{rcl} u'_1 & = & u_1 + u_2 \\ u'_2 & = & 2u_1 + u_2 . \end{array}$$

Donc

$$(u'_1)_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ (u'_2)_B = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

La matrice de changement de base de B' à B est ainsi

$$P_{B,B'} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Théorème 6.71. *Soit V un espace vectoriel de dimension finie, et soient B et B' deux bases de V. Soit $P_{B,B'}$ la matrice de changement de base de B' à B. Alors $P_{B,B'}$ est inversible et la matrice de changement de base de B à B' est $P^{-1}_{B,B'}$.*

Démonstration. Soit $P_{B',B}$ la matrice de changement de base de B à B'. Montrons que

$$P_{B,B'} P_{B',B} = I .$$

Soit $B = \{u_1, u_2, \ldots, u_n\}$, et supposons que

$$P_{B,B'} P_{B',B} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & & \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix}.$$

Nous avons

$$(x)_B = P_{B,B'} (x)_{B'}$$

et

$$(x)_{B'} = P_{B',B} (x)_B$$

pour tout $x \in V$. On a donc

$$(x)_B = P_{B,B'} P_{B',B} (x)_B.$$

Pour $x = u_1$, on obtient

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

d'où

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}.$$

Le même raisonnement avec $x = u_2, \dots, u_n$ nous donne

$$\begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Ceci montre que

$$P_{B,B'} P_{B',B} = I.$$

□

Théorème 6.72. *Soit V un espace vectoriel de dimension n et soient B, B', B'' trois bases de V. Alors*

$$P_{B,B'}\ P_{B',B''} = P_{B,B''}.$$

Démonstration. Soit $x \in V$. On a

$$(x)_B = P_{B,B'}(x)_{B'}$$

et

$$(x)_{B'} = P_{B',B''}(x)_{B''}$$

d'où

$$(x)_B = P_{B,B'}\ P_{B',B''}(x)_{B''}.$$

Ainsi, la matrice de changement de base de B'' à B est $P_{B,B'}\ P_{B',B''}$. Autrement dit,

$$P_{B,B''} = P_{B,B'}\ P_{B',B''}.$$

□

Application. Soit V un espace vectoriel de dimension n. Soient B' et B'' deux bases de V et C une autre base de V. Si on connaét les coordonnées des vecteurs de B et de B' dans C, on peut calculer la matrice $P_{B,B'}$ de la maniére suivante : Ecrivons $\left(P_{C,B} \ : \ P_{C,B'} \right)$. Comme $P_{C,B}$ est inversible (son inverse est $P_{B,C}$), elle est équivalente par lignes à I_n. On obtient donc, par la méthode de Gauss, que

$$\left(I_n \ : \ (P_{C,B})^{-1}\ P_{C,B'} \right)$$

ou, autrement dit

$$\left(I_n \ : \ \underbrace{(P_{B,C})\ P_{C,B'}}_{P_{B,B'}} \right).$$

La matrice C désigne souvent la matrice canonique. Dans ce cas, si on connaét les coordonnées de B et B' dans la base canonique, on calcule directement la matrice de passage de B' à B.

Exemple 6.73. Soit $V = \mathbb{R}^3$ et C sa base canonique. Définissons

$$B = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \right\}$$

et

$$B' = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}.$$

On a alors

$$P_{C,B} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{et} \quad P_{C,B'} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

On applique maintenant la méthode de Gauss pour calculer $P_{C,B}^{-1}$ et donc $P_{B,B'} = P_{C,B}^{-1} P_{C,B'}$:

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

$\ell_2 \longrightarrow \ell_2 - \ell_1$

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

$\ell_1 \longrightarrow \ell_1 + 3\ell_3$
$\ell_2 \longrightarrow -\ell_2 + \ell_3$
$\ell_3 \longrightarrow -\ell_3$

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

D'où :

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

Chapitre 7

Produits scalaires généralisés

7.1 Définition et premières propriétés

Définition 7.1. Soit V un espace vectoriel. Alors un produit scalaire généralisé sur V est par définition une application

$$\begin{aligned} <,>: V \times V &\longrightarrow \mathbb{R} \\ (u,v) &\mapsto < u,v > \end{aligned}$$

ayant les propriétés

(1)	$< u,v >=< v,u >$	symétrie
(2)	$< u+v,w >=< u,w > + < v,w >$	additivité
(3)	$< \lambda u,v >= \lambda < u,v >$	homogénéïté
(4)	$< v,v > \geq 0$	positivité
(5)	$< v,v >= 0 \iff v = 0.$	

pour tout $u, v, w \in V$ et $\lambda \in \mathbb{R}$.

Exemple 7.2. Produit scalaire euclidien sur $\mathbb{R}^n$. Dans cet exemple, on a

$$V = \mathbb{R}^n$$

et

$$< u,v >= u \bullet v$$

le produit scalaire euclidien (produit scalaire habituel) de $\mathbb{R}^n$. Nous avons déjà vu que les propriétés (1) - (5) sont vérifiées dans cet exemple.

Exemple 7.3. Soient $u = \binom{u_1}{u_2}$ et $v = \binom{v_1}{v_2}$ deux vecteurs de $\mathbb{R}^2$. Posons

$$< u,v >= 3u_1v_1 + 2u_2v_2.$$

Alors $<,>$ est un produit scalaire généralisé sur $\mathbb{R}^2$. Vérifions les propriétés (1) - (5) :

(1) On a

$$\begin{aligned} < u,v > &= 3u_1v_1 + 2u_2v_2 \\ &= 3v_1u_1 + 2v_2u_2 \\ &= < v,u > . \end{aligned}$$

(2) Soit $w = \binom{w_1}{w_2}$. Alors

$$\begin{aligned} < u+v,w > &= 3(u_1+v_1)w_1 + 2(u_2+v_2)w_2 \\ &= (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) \\ &= < u,w > + < v,w > . \end{aligned}$$

(3)

$$\begin{aligned} <\lambda u, v> &= 3(\lambda u_1)v_1 + 2(\lambda u_2)v_2 \\ &= \lambda(3u_1v_1 + 2u_2v_2) \\ &= \lambda < u, v > . \end{aligned}$$

(4)

$$\begin{aligned} < v, v > &= 3v_1v_1 + 2v_2v_2 \\ &= 3v_1^2 + 2v_2^2 \geq 0. \end{aligned}$$

(5)

$$\begin{aligned} < v, v > &= 3v_1^2 + 2v_2^2 = 0 \\ &\Longleftrightarrow v_1 = 0 \quad v_2 = 0 \\ &\Longleftrightarrow v = 0\,. \end{aligned}$$

Définition 7.4. Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Alors on définit la norme (ou longueur) d'un vecteur $v \in V$ par la formule

$$\| v \| = \sqrt{< v, v >}$$

La distance entre $u, v \in V$ est même par définition

$$d(u, v) = \| u - v \| \ .$$

Exemple 7.5. Norme et distance dans $\mathbb{R}^n$. Lorsque $V = \mathbb{R}^n$ et $<,>$ est le produit scalaire habituel $< u, v >= u \bullet v$, on retrouve les notions habituelles de norme et de distance dans $\mathbb{R}^n$.

Exemple 7.6. Soit $V = \mathbb{R}^2$. On obtient des notions de norme et de distance différentes selon le produit scalaire généralisé choisi. Soient $u = \binom{1}{0}$ et $v = \binom{0}{1}$. Si l'on choisit le produit scalaire habituel sur $\mathbb{R}^2$, on a

$$\| u \| = \sqrt{u_1^2 + u_2^2} = \sqrt{1+0} = 1$$

et

$$\begin{aligned} d(u, v) &= \| u - v \| = \| \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) \| \\ &= \sqrt{1^2 + (-1)^2} = \sqrt{2}. \end{aligned}$$

Si l'on choisit le produit scalaire généralisé $< u, v >= 3u_1v_1 + 2u_2v_2$ on obtient

$$\begin{aligned} \| u \| &= \sqrt{< u, u >} \\ &= \sqrt{3u_1^2 + 2u_2^2} \\ &= \sqrt{3+0} = \sqrt{3}. \end{aligned}$$

et

$$d(u, v) = \| u - v \| = \sqrt{< \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) >} = \sqrt{3+2} = \sqrt{5}.$$

Définition 7.7 (Cercle unité (ou sphère unité))**.** Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Alors le *cercle unité*, ou *sphère unité* de V est par définition l'ensemble des $u \in V$ tels que

$$\| u \| = 1\,.$$

Ce sont les éléments de V dont la distance à l'origine est égale à 1.

Exemple 7.8. Soit $V = \mathbb{R}^2$ et $<,>$ le produit scalaire habituel. Soit $u = \binom{x}{y}$. Alors

$$\| u \| = \sqrt{x^2 + y^2}$$

et l'équation du cercle unité est alors $\sqrt{x^2 + y^2} = 1$ ou

$$x^2 + y^2 = 1.$$

Exemple 7.9. Soit $V = \mathbb{R}^2$ muni du produit scalaire généralisé

$$< u, v >= 3u_1v_1 + 2u_2v_2.$$

Soit $w = (x, y)$. Alors

$$\| w \| = \sqrt{3x^2 + 2y^2}$$

ce qui donne $\sqrt{3x^2 + 2y^2} = 1$ ou

$$3x^2 + 2y^2 = 1$$

comme équation pour le cercle unité.
Le graphe de cette équation est une ellipse.

Définition 7.10. Soit A une matrice de taille $n \times n$ inversible. Le produit scalaire généralisé associé à A est, par définition, le produit scalaire sur $\mathbb{R}^n$ qui à $u, v \in \mathbb{R}^n$ associe

$$Au \bullet Av$$

(le produit scalaire habituel de Au et de Av).

Soient $u, v \in \mathbb{R}^n$ représentés comme vecteurs colonne. Nous avons vu que

$$Au \bullet Av = (Av)^T Au$$

et ceci conduit à une nouvelle formulation de ce produit scalaire généralisé :

$$< u, v >= v^T A^T Au.$$

Notons que lorsque $A = I$, on obtient le produit scalaire habituel $< u, v >= u \bullet v$.

Exemple 7.11. Soit

$$A = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{pmatrix}.$$

Alors le produit scalaire généralisé associé à A est

$$\begin{aligned} < u, v > &= (v_1v_2) \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ &= (v_1v_2) \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ &= 3u_1v_1 + 2u_2v_2. \end{aligned}$$

On retrouve ainsi le produit scalaire de l'exemple 7.3.

Théorème 7.12. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Alors, pour tout $u, v, w \in V$ et tout $\lambda \in \mathbb{R}$, on a*
(a) $< 0, v >=< v, 0 >= 0$
(b) $< u, \lambda v >= \lambda < u, v >$
(c) $< u, v + w >=< u, v > + < u, w >$.

Démonstration. Ceci découle immédiatement des propriétés du produit scalaire :
(a) $< 0, v >= 0 < 0, v >= 0$ et $< v, 0 >=< 0, v >= 0$.
(b) $< u, \lambda v >=< \lambda v, u >= \lambda < v, u >= \lambda < u, v >$.
(c) $< u, v + w >=< v + w, u >=< v, u > + < w, u >=< u, v > + < u, w >$.

□

7.2 Angles et orthogonalité

Théorème 7.13 (Inégalité de Cauchy-Schwarz)**.** *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soient $u, v \in V$. Alors on a*

$$|<u,v>| \leq \| u \| \cdot \| v \| \ .$$

Démonstration. Si $u = 0$, alors on a

$$<u,v> = \| u \| = 0 \,,$$

donc l'affirmation est vraie. Supposons que $u \neq 0$, et posons

$$\begin{aligned} a &= <u,u> \\ b &= 2<u,v> \\ c &= <v,v> \ . \end{aligned}$$

On a, pour tout $t \in \mathbb{R}$,

$$0 \leq <tu+v, tu+v> = <u,u> t^2 + 2<u,v> t + <v,v> = at^2 + bt + c.$$

Donc

$$at^2 + bt + c \geq 0$$

pour tout $t \in \mathbb{R}$. Ceci implique que le polynôme quadratique

$$aX^2 + bX + c$$

a soit une racine double, soit aucune racine réelle.

Son discriminant est donc négatif ou nul : $b^2 - 4ac \leq 0$. Remplaçant a par $<u,u>$, b par $2<u,v>$ et c par $<v,v>$, on obtient

$$4<u,v>^2 - 4<u,u><v,v> \leq 0$$

ce qui implique

$$<u,v>^2 \leq <u,u> \cdot <v,v>$$

donc

$$|<u,v>| \ \leq \ \sqrt{<u,u><v,v>} = \| u \| \cdot \| v \| \ .$$

□

Théorème 7.14. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Alors on a, pour tout $u, v \in V$ et pour tout $\lambda \in \mathbb{R}$*

(a) $\| u \| \geq 0$

(b) $\| u \| = 0 \iff u = 0$

(c) $\| \lambda u \| = |\lambda| \ \| u \|$

(d) $\| u + v \| \leq \| u \| + \| v \|$ *inégalité du triangle*

Démonstration. La démonstration est la même que dans le cas du produit scalaire habituel. □

Corollaire 7.15. *Soit V un espace vectoriel muni d'un produit scalaire généralisé. Alors, pour tout $u, v, w \in V$ et tout $\lambda \in \mathbb{R}$, on a*

(a) $d(u,v) \geq 0$

(b) $d(u,v) = 0 \iff u = v$

(c) $d(u,v) = d(v,u)$

(d) $d(u,v) \leq d(u,w) + d(w,v)$ *inégalité du triangle*

Démonstration. Ceci est une conséquence immédiate des définitions et du théorème précédent. □

7.2.1 Angle formé par deux vecteurs

Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soient $u, v \in V$. Alors par l'inégalité de Cauchy-Schwarz, on a si $u \neq 0$ et $v \neq 0$

$$\left(\frac{< u, v >}{\| u \| \cdot \| v \|}\right)^2 \leq 1$$

donc

$$-1 \leq \frac{< u, v >}{\| u \| \cdot \| v \|} \leq 1 \,.$$

Il existe donc un unique angle θ, avec $0 \leq \theta \leq \pi$, tel que

$$\cos\theta = \frac{< u, v >}{\| u \| \, \| v \|} \,.$$

On appelle θ l'angle formé par les vecteurs u et v. Remarquons que dans le cas du produit scalaire habituel de $\mathbb{R}^n$, on retrouve la notion habituelle d'angle entre deux vecteurs.

Définition 7.16. Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soient $u, v \in V$. On dit que u et v sont orthogonaux si et seulement si

$$\textit{polynme} < u, v >= 0 \,.$$

Exemple 7.17. Soit $\mathcal{P}_2$ l'espace vectoriel des fonctions polynomiales de degré au plus 2. On définit un produit scalaire généralisé sur $\mathcal{P}_2$ en posant

$$< p, q >= \int_{-1}^{1} p(x)q(x)dx$$

pour tout $p, q \in \mathcal{P}_2$. Les axiomes de produit scalaire généralisé sont faciles à vérifier. On a par exemple la positivité : pour tout $p \in \mathcal{P}_2$, on a

$$< p, p >= \int_{-1}^{1} p^2(x)dx \geq 0 \,.$$

Soient

$$p(x) = x$$

et

$$q(x) = x^2$$

alors p et q sont orthogonaux. En effet

$$\begin{aligned} < p, q > &= \int_{-1}^{1} x \cdot x^2 \, dx = \int_{-1}^{1} x^3 dx \\ &= 0 \,. \end{aligned}$$

Théorème 7.18 (Théorème de Pythagore généralisé)**.** *Soit V un espace vectoriel muni d'un produit scalaire généralisé. Soient $u, v \in V$, et supposons que u et v soient orthogonaux. Alors*

$$\| u + v \|^2 = \| u \|^2 + \| v \|^2 \,.$$

Démonstration. Comme u et v sont orthogonaux, on a

$$< u, v >= 0 \,.$$

Donc

$$\begin{aligned} \| u + v \|^2 &= < u + v, u + v > \\ &= \| u \|^2 + \| v \|^2 + 2 < u, v > \\ &= \| u \|^2 + \| v \|^2 \,. \end{aligned}$$

□

Exemple 7.19. Soit $\mathcal{P}_2$ l'espace vectoriel des fonctions polynomiales de degré au plus 2, muni du produit scalaire généralisé défini dans l'exemple 7.17.

Soient $p(x) = x$ et $q(x) = x^2$. Nous avons vu que p et q sont orthogonaux. Calculons $\| p \|^2$, $\| q \|^2$, et $\| p+q \|^2$:

$$\| p \|^2 =< p, p >= \int_{-1}^{1} x^2 dx = \frac{2}{3} \qquad \| q \|^2 =< q, q >= \int_{-1}^{1} x^4 dx = \frac{2}{5}$$

$$\| p+q \|^2 =< p+q, p+q >= \int_{-1}^{1} (x+x^2)(x+x^2) dx = \int_{-1}^{1} (x^4 + 2x^3 + x^2) dx = \frac{16}{15}$$

On a bien

$$\| p \|^2 + \| q \|^2 = \| p+q \|^2$$

Définition 7.20. Soit V un espace vectoriel muni d'un produit scalaire généralisé, et soit W un sous-espace vectoriel de V. On dit que $u \in V$ est orthogonal à W si u est orthogonal à tout élément de W. L'ensemble des $u \in V$ qui sont orthogonaux à W est appelé le complément orthogonal de W. Il est noté $W^\perp$.

Théorème 7.21. *Soit V un espace vectoriel muni d'un produit scalaire généralisé, et soit W un sous-espace vectoriel de V. Alors on a*

(a) $W^\perp$ est un sous-espace vectoriel de V.

(b) $W \cap W^\perp = \{0\}$.

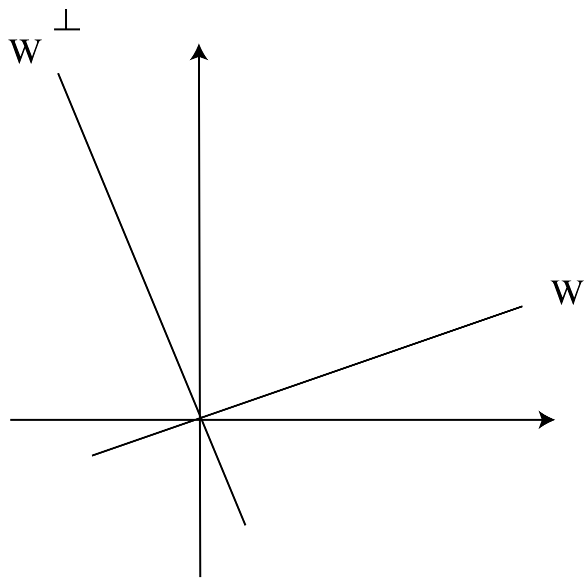


Figure 9.1 : $V = \mathbb{R}^2 \qquad < u, v >= u \bullet v$

Démonstration. (a) On a $< 0, w >= 0$ pour tout $w \in W$, donc $0 \in W^\perp$. Montrons que $W^\perp$ est fermé par addition et par multiplication par les scalaires. Soient

$$u, v \in W^\perp .$$

Alors

$$< u, w >=< v, w >= 0$$

pour tout $w \in W$. On a donc

$$< u+v, w >=< u, w > + < v, w >= 0 + 0 = 0$$

pour tout $w \in W$ ce qui montre que

$$u + v \in W^{\perp}.$$

De même, si $\lambda \in \mathbb{R}$, alors

$$< \lambda u, w >= \lambda < u, w >= 0$$

pour tout $w \in W$ et donc $\lambda u \in W^{\perp}$.

(b) Soit $v \in W \cap W^{\perp}$. Alors

$$< v, v >= 0$$

et donc

$$v = 0.$$

□

Théorème 7.22. *Soit A une matrice de taille $m \times n$. Alors*

(a) Par rapport au produit scalaire habituel de $\mathbb{R}^n$, le complément orthogonal de l'espace des lignes de A est le nilespace (noyau) de A.

(b) Par rapport au produit scalaire habituel de $\mathbb{R}^m$, le complément orthogonal de l'espace des colonnes de A est le nilespace de A^T.

Démonstration. (a) Soit W l'espace des lignes de A. Soit $v \in W^{\perp}$. Soient $\ell_1, \dots, \ell_m$ les vecteurs ligne de A. Alors

$$\ell_1 \bullet v = \ell_2 \bullet v = \cdots = \ell_m \bullet v = 0\,.$$

Or, nous avons vu que le système linéaire

$$Ax = 0$$

peut être exprimé par

$$\begin{pmatrix} \ell_1 & \bullet & x \\ \ell_2 & \bullet & x \\ & \vdots & \\ \ell_m & \bullet & x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Comme $\ell_1 \bullet v = \ell_2 \bullet v = \cdots = \ell_m \bullet v = 0$, le vecteur v est une solution de ce système. Par définition, ceci implique que v est dans le nilespace de A.
Réciproquement, supposons que v est dans le nilespace de A, autrement dit

$$Av = 0\,.$$

Alors on a

$$\ell_1 \bullet v = \ell_2 \bullet v = \cdots = \ell_m \bullet v = 0\,.$$

Soit $w \in W$. Alors w est combinaison linéaire de $\ell_1, \ell_2, \dots, \ell_m$., c'est-à-dire que l'on a

$$w = \lambda_1 \ell_1 + \lambda_2 \ell_2 + \cdots + \lambda_m \ell_m$$

avec $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$. On a donc

$$\begin{aligned} w \bullet v &= (\lambda_1 \ell_1 + \lambda_2 \ell_2 + \cdots + \lambda_m \ell_m) \bullet v \\ &= \lambda_1(\ell_1 \bullet v) + \lambda_2(\ell_2 \bullet v) + \cdots + \lambda_m(\ell_m \bullet v) \\ &= 0 \end{aligned}$$

Donc $v \in W^{\perp}$. Ceci démontre la partie (a) du théorème.

(b) On applique (a) à la matrice A^T : en effet, l'espace des colonnes de A est égal à l'espace des lignes de A^T.

□

7.3 Bases orthogonales et méthode de Gram-Schmidt

Définition 7.23. Soit V un espace vectoriel muni d'un produit scalaire généralisé . Soit

$$S = \{v_1, \ldots, v_n\}$$

un sous-ensemble de V. On dit que S est un ensemble orthogonal si v_i est orthogonal à v_j pour tout $i \neq j$. On dit que S est un ensemble orthonormé si S est orthogonal et si

$$\| v_i \| = 1$$

pour tout $i = 1, \ldots, n$. Un ensemble orthogonal qui est une base est appelé une base orthogonale ; un ensemble orthonormé qui est une base est appelé une base orthonormée.

Théorème 7.24. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$, et soit $S = \{v_1, \ldots, v_n\}$ une base orthonormée de V. Soit $u \in V$. Alors*

$$u = < u, v_1 > v_1 + \cdots + < u, v_n > v_n.$$

Démonstration. Comme S est une base, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que

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

On a

$$< u, v_i > = < \lambda_1 v_1 + \cdots + \lambda_n v_n, v_i > = \lambda_1 < v_1, v_i > + \cdots + \lambda_n < v_n, v_i > .$$

Or, comme $< v_j, v_i > = 0$ si $i \neq j$ et $< v_i, v_i > = \| v_i \|^2 = 1$, on a

$$< u, v_i > = \lambda_i$$

et par conséquent

$$u = < u, v_1 > v_1 + < u, v_2 > v_2 + \cdots + < u, v_n > v_n\,.$$

□

On appelle les scalaires

$$< u, v_1 >, < u, v_2 >, \ldots, < u, v_n >$$

les coordonnées de u par rapport à la base orthonormée $\{v_1, \ldots, v_n\}$.

Théorème 7.25. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soit $S = \{v_1, v_2, \ldots, v_n\}$ une base orthogonale de V et u un vecteur de V. Alors*

$$u = \frac{< u, v_1 >}{\| v_1 \|^2} v_1 + \frac{< u, v_2 >}{\| v_2 \|^2} v_2 + \cdots + \frac{< u, v_n >}{\| v_n \|^2} v_n.$$

Démonstration. La base

$$S' = \left\{ \frac{v_1}{\| v_1 \|}, \quad \frac{v_2}{\| v_2 \|}, \ldots, \frac{v_n}{\| v_n \|} \right\}$$

est orthonormée. Par le théorème précédent, on a donc

$$u = \left\langle u, \frac{v_1}{\| v_1 \|} \right\rangle \frac{v_1}{\| v_1 \|} + \cdots + \left\langle u, \frac{v_n}{\| v_n \|} \right\rangle \frac{v_n}{\| v_n \|}$$

et alors

$$u = \frac{< u, v_1 >}{\| v_1 \|^2} v_1 + \cdots + \frac{< u, v_n >}{\| v_n \|^2} v_n.$$

□

Théorème 7.26. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soit $S = \{v_1, \ldots, v_n\}$ un sous-ensemble orthogonal de V, avec $v_i \neq 0$ pour tout $i = 1, \ldots, n$. Alors S est une famille linéairement indépendante.*

Démonstration. Supposons que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

avec $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. On a

$$< \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, v_i >=< 0, v_i >= 0$$

pour tout $i = 1, \ldots, n$. Par linéarité du produit scalaire, on obtient

$$\lambda_1 < v_1, v_i > + \lambda_2 < v_2, v_i > + \cdots + \lambda_n < v_n, v_i >= 0.$$

Mais comme $< v_j, v_i >= 0$ si $i \neq j$, on obtient

$$\lambda_i < v_i, v_i >= 0$$

ce qui implique, comme $v_i \neq 0$, que $\lambda_i = 0$ pour tout $i = 1, \ldots, n$. Donc S est linéairement indépendant. □

Théorème 7.27. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soit $S = \{v_1, v_2, \ldots, v_n\}$ un sous-ensemble orthonormé (qui n'est pas forcément une base ; il peut être plus petit) de V et soit $u \in V$. Posons*

$$w_1 =< u, v_1 > v_1 + < u, v_2 > v_2 + \cdots + < u, v_n > v_n$$

et $w_2 = u - w_1$. Notons $W = \mathcal{L}(S)$, le sous-espace vectoriel de V engendré par S. Alors $w_1 \in W$ et $w_2 \in W^\perp$.

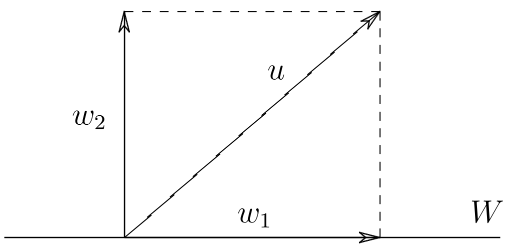


Figure 9.2 : $u = w_1 + w_2$

Notation : Le vecteur $w_1 \in W$ est appelé *la projection orthogonale de u sur W*, et est noté

$$proj_W(u).$$

Démonstration. Il est clair que $w_1 \in W$, puisque w_1 est une combinaison linéaire de $v_1, \ldots, v_n$. Montrons que $w_2 \in W^\perp$. On a

$$< w_2, v_i >=< u - w_1, v_i >=< u, v_i > - < w_1, v_i >$$

pour tout $i = 1, \ldots, n$. On a

$$\begin{aligned} < w_1, v_i > &= \langle < u, v_1 > v_1 + \cdots + < u, v_n > v_n, v_i \rangle \\ &= < u, v_1 >< v_1, v_i > + \cdots + < u, v_n >< v_n, v_i > \\ &= 0 + \cdots + 0 + < u, v_i >< v_i, v_i > + 0 + \cdots + 0 \\ &= < u, v_i > \cdot 1 =< u, v_i > \end{aligned}$$

pour tout $i = 1, \ldots, n$. On a donc

$$< w_2, v_i >=< u, v_i > - < w_1, v_i >= 0$$

pour tout $i = 1, \ldots, n$ ce qui montre que w_2 est orthogonal à $S = \{v_1, v_2, \ldots, v_n\}$. Comme S engendre W, on a que w_2 est orthogonal à W. Donc

$$w_2 \in W^{\perp}.$$

□

Corollaire 7.28. *Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soit W un sous-espace vectoriel de dimension finie de V. Alors tout $u \in V$ s'écrit de manière unique comme*

$$u = w_1 + w_2$$

avec $w_1 \in W$ et $w_2 \in W^{\perp}$. De plus, $w_1 = \text{proj}_W(u)$.

Exemple 7.29. Soit $V = \mathbb{R}^3$, $<,>$ le produit scalaire habituel, et soit W le sous-espace engendré par

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \text{et} \qquad v_2 = \begin{pmatrix} -4/5 \\ 0 \\ 3/5 \end{pmatrix}.$$

On vérifie que $\| v_1 \| = \| v_2 \| = 1$, et que $< v_1, v_2 >= v_1 \bullet v_2 = 0$. Donc $S = \{v_1, v_2\}$ est un ensemble orthonormé. Soit $u = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$. Alors

$$w_1 = \text{proj}_W(u) =< u, v_1 > v_1 + < u, v_2 > v_2.$$

Comme $< u, v_1 >= 1$ et $< u, v_2 >= -\frac{1}{5}$, ceci donne

$$\text{proj}_W(u) = v_1 - \frac{1}{5} v_2 = \begin{pmatrix} 4/25 \\ 1 \\ -3/25 \end{pmatrix}.$$

La décomposition orthogonale de u par rapport à W est donc

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4/25 \\ 1 \\ -3/25 \end{pmatrix} + \begin{pmatrix} 21/25 \\ 0 \\ 28/25 \end{pmatrix}.$$

On vérifie que $\left(\begin{smallmatrix} 21/25 \\ 0 \\ 28/25 \end{smallmatrix}\right)$ est orthogonal à v_1 et à v_2 et donc qu'il est bien dans $W^{\perp}$.

Théorème 7.30. *Soit V un espace vectoriel de dimension finie muni d'un produit scalaire généralisé $<,>$. Alors V a une base orthonormée.*

Démonstration. On va démontrer le théorème en construisant une base orthonormée. La méthode de construction utilisée ici s'appelle **méthode de Gram-Schmidt** ou **procédé de Gram-Schmidt**. Soit $S = \{u_1, \ldots, u_n\}$ une base quelconque de V. On construit une base orthonormée en n étapes.

1ère étape : On pose

$$v_1 = \frac{u_1}{\| u_1 \|}$$

ce qui nous donne $\| v_1 \| = 1$.

2ère étape : Soit W_1 le sous-espace de V engendré par v_1. On pose

$$w_2 = u_2 - \text{proj}_{W_1}(u_2).$$

Alors

$$< w_2, v_1 >= 0$$

car w_2 est orthogonal à W_1, donc à v_1. Vérifions que $w_2 \neq 0$. En effet, si $w_2 = 0$, alors

$$u_2 = \text{proj}_{W_1}(u_2)\,.$$

Ceci implique que u_2 et v_1 sont linéairement dépendants et donc que u_2 et u_1 le sont également. Ceci n'est pas possible car u_1 et u_2 font partie d'une base de V. La norme de w_2 est donc non nulle. On pose alors

$$v_2 = \frac{w_2}{\| w_2 \|}\,.$$

On a $\| v_2 \| = 1$, $\| v_1 \| = 1$ et $< v_1, v_2 > = 0$.

k-ème étape Supposons construits les vecteurs

$$v_1, v_2, \dots, v_{k-1}$$

tels que

$$\| v_1 \| = \| v_2 \| = \cdots = \| v_{k-1} \| = 1,$$

que $< v_i, v_j > = 0$ si $i \neq j$ et que le sous-espace W_{k-1} engendré par $v_1, v_2, \dots, v_{k-1}$ soit égal au sous-espace engendré par $u_1, \dots, u_{k-1}$. On pose

$$w_k = u_k - \text{proj}_{W_{k-1}}(u_k).$$

Montrons que $w_k \neq 0$. Si $w_k = 0$, alors $u_k = \text{proj}_{W_{k-1}}(u_k)$, donc $u_k \in W_{k-1} = \mathcal{L}(u_1, \dots, u_{k-1})$. Ceci contredit l'hpothèse que $u_1, \dots, u_k$ font partie d'une base de V. Donc $w_k \neq 0$. Comme $w_k = u_k - \text{proj}_{W_{k-1}}(u_k)$, le vecteur w_k est orthogonal à W_{k-1} et en particulier à $v_1, \dots, v_{k-1}$. On pose finalement

$$v_k = \frac{w_k}{\| w_k \|}\,.$$

On a ainsi construit $v_1, \dots, v_k$ tels que

$$\| v_1 \| = \cdots = \| v_k \| = 1,$$

et que $< v_i, v_j > = 0$ si $i \neq j$ et $\mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(u_1, \dots, u_k)$. On termine ce procédé lorsque l'on a obtenu v_n. □

Exemple 7.31. On considère $\mathbb{R}^3$ muni du produit scalaire habituel. Soit $S = \{u_1, u_2, u_3\}$, avec $u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ une base de $\mathbb{R}^3$. On applique la méthode de Gram-Schmidt pour obtenir une base orthonormée.

1ère étape : $w_1 = u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

2ème étape : soit $W_1 = \mathcal{L}(w_1) = \mathcal{L}(u_1)$. On pose

$$\begin{aligned} w_2 &= u_2 - \text{proj}_{W_1}(u_2) = \\ &= u_2 - \frac{< u_2, w_1 >}{\| w_1 \|^2} w_1 \end{aligned}$$

Comme $< u_2, w_1 > = < u_2, u_1 > = 2$ et $\| w_1 \|^2 = \| u_1 \|^2 = 3$, on obtient

$$w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}.$$

3ème étape : soit

$$W_2 = \mathcal{L}(w_1, w_2) = \mathcal{L}(u_1, u_2).$$

Posons

$$w_3 = u_3 - \text{proj}_{W_2}(u_3) = u_3 - \frac{< u_3, w_1 >}{\| w_1 \|^2} w_1 - \frac{< u_3, w_2 >}{\| w_2 \|^2} w_2 = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix}.$$

On a donc

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix}.$$

Alors $\{w_1, w_2, w_3\}$ est une base orthogonale de $\mathbb{R}^3$. On obtient une base orthonormée en posant

$$v_1 = \frac{w_1}{\| w_1 \|}, \quad v_2 = \frac{w_2}{\| w_2 \|} \qquad \text{et} \qquad v_3 = \frac{w_3}{\| w_3 \|}.$$

Ceci nous donne

$$v_1 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \quad v_2 = \begin{pmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

7.4 Matrices orthogonales

7.4.1 Définition et Propriétés

Définition 7.32. Soit A une matrice carrée. On dit que A est une matrice orthogonale si

$$A^{-1} = A^T$$

Remarque : A est une matrice orthogonale si et seulement si

$$A^T A = A A^T = I .$$

Exemple 7.33. Les matrices de rotation sont orthogonales. Soit θ un angle. Alors la matrice de la rotation d'angle θ de $\mathbb{R}^2$ est

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Cette matrice est orthogonale. En effet, l'identité

$$\sin^2\theta + \cos^2\theta = 1$$

montre que l'on a

$$A^T A = I.$$

Théorème 7.34. *Soit A une matrice $n \times n$. Les trois propriétés suivantes sont équivalentes :*

(a) A est orthogonale.

(b) Pour tout $x \in \mathbb{R}^n$, on a $\| Ax \| = \| x \|$.

(c) Pour tout $x, y \in \mathbb{R}^n$, on a $Ax \bullet Ay = x \bullet y$.

En d'autres termes, une matrice orthogonale conserve les normes et les angles.

Démonstration. (a) $\Longrightarrow$ (b) Supposons A orthogonale. On a

$$\| Ax \|^2 = Ax \bullet Ax = x \bullet (A^T A)x = x \bullet x = \| x \|^2 .$$

(b) $\Longrightarrow$ (c) Supposons que

$$\| Ax \| = \| x \|$$

pout tout $x \in \mathbb{R}^n$. On a

$$\begin{aligned} Ax \bullet Ay &= \frac{1}{4} \| Ax + Ay \|^2 - \frac{1}{4} \| Ax - Ay \|^2 = \\ &= \frac{1}{4} \| A(x+y) \|^2 - \frac{1}{4} \| A(x-y) \|^2 = \\ &= \frac{1}{4} \| x+y \|^2 - \frac{1}{4} \| x-y \|^2 \\ &= x \bullet y . \end{aligned}$$

(c) $\Longrightarrow$ (a) On considère (c) avec $x = e_i$ et $y = e_j$ des vecteurs de la base standard. Alors $x \bullet y$ est soit 0 (si $i \neq j$) soit 1 (si $i = j$), et est le coefficient I_{ij} de la matrice identité. D'autre part,

$$Ax \bullet Ay = e_j^T A^T A e_i$$

est le coefficient (i, j) de la matrice $A^T A$. Il suit que $A^T A = I$, et donc que A est orthogonale. $\square$

Remarque 7.35. L'inverse d'une matrice orthogonale est orthogonale. En effet on a

$$\begin{aligned} A \text{ orthogonale } &\iff A^{-1} = A^T \\ &\iff (A^{-1})^{-1} = (A^T)^{-1} \\ &\iff (A^{-1})^{-1} = (A^{-1})^T \\ &\iff A^{-1} \text{est orthogonale.} \end{aligned}$$

7.4.2 Changement de bases orthonormées

Théorème 7.36. *Soit V un espace vectoriel de dimension finie muni d'un produit scalaire généralisé. Soit P la matrice de changement de base d'une base orthonormée vers une autre base orthonormée. Alors P est une matrice orthogonale.*

Démonstration. Soient $B = \{e_1, \ldots, e_n\}$ et $B' = \{e'_1, \ldots, e'_n\}$ deux bases orthonormées de V et soit $P_{B,B'}$ la matrice de changement de base de B' à B. On a

$$u = \sum_{i=1}^{n} u_i e_i = \sum_{i=1}^{n} u'_i e'_i.$$

Comme B et B' sont des bases orthonormées, on a

$$\| u \| = \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} u_i'^2.$$

De plus, on a

$$(u)_B = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = P_{B,B'} \cdot \begin{pmatrix} u'_1 \\ \vdots \\ u'_n \end{pmatrix} = P_{B,B'} \cdot (u)_{B'},$$

$$\sum_{i=1}^{n} u_i^2 = (u)_B^T \cdot (u)_B \qquad \text{et} \qquad \sum_{i=1}^{n} u_i'^2 = (u)_{B'}^T \cdot (u)_{B'}$$

ce qui implique que

$$(u)_{B'}^T \cdot (u)_{B'} = \| u \| = (u)_B^T \cdot (u)_B = (P_{B,B'} \cdot (u)_{B'})^T \cdot P_{B,B'} \cdot (u)_{B'}^T = (u)_{B'}^T \cdot P_{B,B'}^T P_{B,B'} \cdot (u)_{B'}.$$

Comme ceci est vrai pour tout vecteur u, on en déduit que $P_{B,B'}^T P_{B,B'} = I$ ce qui montre que $P_{B,B'}$ est orthogonale. $\square$

7.4.3 Décomposition Q-R : application du théorème 7.30

Soit A une matrice $n \times n$ ayant ses colonnes linéairement indépendantes. Alors il existe une matrice orthogonale Q (par rapport au produit scalaire standard sur $\mathbb{R}^n$) et une matrice triangulaire supérieure R, telles que :

$$A = QR.$$

Plus précisément, Q est la matrice dont les colonnes sont les vecteurs obtenus en appliquant le procédé de Gram-Schmidt aux vecteurs colonne de A.
On procède comme suit :
Soit

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

On appelle c_i la i-ème colonne de A. Soit $\{c'_1, \ldots, c'_n\}$ la base obtenue en appliquant Gram-Schmidt à $\{c_1, \ldots, c_n\}$ et Q la matrice dont les colonnes sont $c'_1, \ldots, c'_n$.

$$Q = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix}$$

Comme $\{c'_1, \ldots, c'_n\}$ est une base orthonormée, on a

$$\begin{aligned} c_1 &= < c_1, c'_1 > c'_1 + \cdots + < c_1, c'_n > c'_n \\ &\vdots \\ c_n &= < c_n, c'_1 > c'_1 + \cdots + < c_n, c'_n > c'_n \end{aligned}$$

c'est-à-dire

$$\begin{aligned} c_1 &= < c_1, c'_1 > \begin{pmatrix} q_{11} \\ \vdots \\ q_{n1} \end{pmatrix} + \cdots + < c_1, c'_n > \begin{pmatrix} q_{1n} \\ \vdots \\ q_{nn} \end{pmatrix} \\ &\vdots \\ c_n &= < c_n, c'_1 > \begin{pmatrix} q_{11} \\ \vdots \\ q_{n1} \end{pmatrix} + \cdots + < c_n, c'_n > \begin{pmatrix} q_{1n} \\ \vdots \\ q_{nn} \end{pmatrix}. \end{aligned}$$

Ainsi, on voit que

$$A = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix} \underbrace{\begin{pmatrix} < c_1, c'_1 > & < c_2, c'_1 > & \ldots & < c_n, c'_1 > \\ \vdots & & & \vdots \\ < c_1, c'_n > & < c_2, c'_n > & \ldots & < c_n, c'_n > \end{pmatrix}}_{=R}$$

La matrice R est triangulaire. En effet, par Gram-Schmidt, pour tout $k > 1$, c'_k est orthogonal à l'espace engendré par $c_1, \ldots, c_{k-1}$. Ainsi, $R_{ij} =< c_j, c'_i >= 0$ pour tout $i > j$, et R est bien triangulaire supérieure.

La décomposition d'une matrice en un produit d'une matrice orthogonale et d'une matrice triangulaire s'appelle la *décomposition QR*.

Exemple 7.37. Calculons la décomposition QR de la matrice

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

En appliquant le procédé de Gram-Schmidt aux vecteurs colonne de A, on obtient

$$c'_1 = \begin{pmatrix} \frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{pmatrix} \quad \text{et} \quad c'_2 = \begin{pmatrix} \frac{\sqrt{5}}{5} \\ \frac{-2\sqrt{5}}{5} \end{pmatrix}$$

De plus,

$$\begin{aligned} < c_1, c'_1 > &= \sqrt{5} \\ < c_2, c'_1 > &= \frac{6\sqrt{5}}{5} \\ < c_1, c'_2 > &= 0 \\ < c_2, c'_2 > &= \frac{3\sqrt{5}}{5} \end{aligned}$$

Ainsi,

$$Q = \begin{pmatrix} \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & \frac{-2\sqrt{5}}{5} \end{pmatrix} \quad \text{et} \quad R = \begin{pmatrix} \sqrt{5} & \frac{6\sqrt{5}}{5} \\ 0 & \frac{3\sqrt{5}}{5} \end{pmatrix}.$$

7.5 La méthode des moindres carrés

Théorème 7.38. *Soit V un espace vectoriel muni d'un produit scalaire généralisé, et soit W un sous-espace de V. Soit $u \in V$. Alors*

$$\| u - \mathrm{proj}_W(u) \| < \| u - w \|$$

pour tout $w \in W$, $w \neq \mathrm{proj}_W(u)$. On dit que $\mathrm{proj}_W(u)$ est la meilleure approximation de u par des vecteurs de W.

Démonstration. Pour tout $w \in W$, on a

$$u - w = (u - \mathrm{proj}_W(u)) + (\mathrm{proj}_W(u) - w)\ .$$

On a

$$\mathrm{proj}_W(u) - w \in W$$

et

$$u - \mathrm{proj}_W(u) \in W^\perp$$

donc ces deux termes sont orthogonaux l'un à l'autre. Par le théorème de Pythagore généralisé, on a

$$\| u - w \|^2 = \| u - \mathrm{proj}_W(u) \|^2 + \| \mathrm{proj}_W(u) - w \|^2\ .$$

Si $w \neq \mathrm{proj}_W(u)$, alors

$$\| \mathrm{proj}_W(u) - w \|^2 > 0$$

, d'où

$$\| u - w \|^2 > \| u - \mathrm{proj}_w(u) \|^2$$

et donc

$$\| u - w \| > \| u - \mathrm{proj}_w(u) \|\ .$$

□

7.5.1 Solution approximative d'un système d'équations linéaires

Soit $Ax = b$ un système d'équations linéaires qui est incompatible, c'est-à-dire qui n'a pas de solution. Nous allons voir une méthode pour donner une solution approximative de ce système. Cette méthode s'appelle la méthode des moindres carrés (least squares). Le principe de la méthode est le suivant : étant donné un système $Ax = b$ de m équations et n inconnues, on se propose de trouver un vecteur x tel que la norme

$$\| Ax - b \|$$

soit minimale. Un tel vecteur x est appelé une solution au sens des moindres carrés. Cette terminologie vient du fait que l'on minimise la norme euclidienne de l'erreur commise. En effet, posons $e = Ax - b$. Si le système admettait une solution, on aurait $e = 0$. Ce terme mesure donc l'erreur commise par l'approximation faite. Si $e = (e_1, \dots, e_m)$, alors

$$\| e \|^2 = e_1^2 + e_2^2 + \cdots + e_m^2\,.$$

On désire que ces carrés soient aussi petits que possible.

Soit $Ax = b$ un système de m équations linéaires à n inconnues. Alors A est de taille $m \times n$. Soit W l'espace des colonnes de A. C'est un sous-espace vectoriel de $\mathbb{R}^m$. Nous avons vu que le système a une solution si et seulement si $b \in W$. Lorsque $b \notin W$, le mieux que l'on puisse faire est de trouver une approximation de b par un vecteur de W. Or, nous avons vu que la meilleure approximation de b par un vecteur de W est

$$\mathrm{proj}_W(b)\,.$$

On doit donc résoudre le système

$$Ax = \mathrm{proj}_W(b).$$

On pourrait donc calculer $\mathrm{proj}_W(b)$, et résoudre le système $Ax = \mathrm{proj}_W(b)$, mais il y a une meilleure méthode. En effet, si x est tel que $Ax = \mathrm{proj}_W(b)$, alors le vecteur

$$b - Ax = b - \mathrm{proj}_W(b)$$

est orthogonal à W. Mais W est l'espace des colonnes de A et $W^\perp$ est donc le nilespace de A^T. On doit donc avoir

$$A^T(b - Ax) = 0$$

autrement dit

$$A^T Ax = A^T b.$$

Définition 7.39. Le système

$$A^T Ax = A^T b$$

est appelé le *système normal* associé au système $Ax = b$. Les solutions du système normal sont les solutions au sens des moindres carrés (solutions approximatives) du système de départ $Ax = b$.

Théorème 7.40. *Pour tout système d'équations linéaires*

$$Ax = b$$

le système normal associé

$$A^T Ax = A^T b$$

a au moins une solution. Toutes les solutions du système normal sont des solutions au sens des moindres carrés du système $Ax = b$. De plus, si W est l'espace des colonnes de A et si x est une solution de moindres carrés de $Ax = b$, alors

$$\mathrm{proj}_W b = Ax\,.$$

Théorème 7.41. *Soit A une matrice de taille $m \times n$. Supposons que les vecteurs colonne de A soient linéairement indépendantes. Alors la matrice $A^T A$ est inversible.*

Démonstration. Pour montrer que $A^T A$ est inversible, il suffit de montrer que le système

$$A^T Ax = 0$$

a une unique solution (la solution triviale). Soit x une solution de ce système. Alors Ax est dans le nilespace de A^T. Mais Ax est aussi dans l'espace des colonnes de A. Or nous avons vu que ces deux espaces sont orthogonaux l'un à l'autre. Leur intersection est donc nulle. Ceci implique que

$$Ax = 0.$$

Comme nous avons supposé que les vecteurs colonne de A sont linéairement indépendants, ceci implique que $x = 0$. □

Théorème 7.42. *Soit A une matrice dont les vecteurs colonne sont linéairement indépendants. Alors tout système linéaire*

$$Ax = b$$

a une unique solution au sens des moindres carrés qui est donnée par

$$x = (A^T A)^{-1} A^T b\,.$$

Démonstration. Ceci découle directements des deux théorèmes précédents. □

Exemple 7.43. Trouver les solutions de moindres carrés du système $Ax = b$ suivant :

$$\begin{aligned} x_1 + x_2 &= 6 \\ 2x_1 - x_2 &= 1 \\ x_1 - 3x_2 &= -1 \end{aligned}$$

On a

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 1 & -3 \end{pmatrix} \quad b = \begin{pmatrix} 6 \\ 1 \\ -1 \end{pmatrix}$$

Les vecteurs colonne de A étant linéairement indépendants, le système a une unique solution au sens des moindres carrés. On a

$$A^T A = \begin{pmatrix} 6 & -5 \\ -5 & 11 \end{pmatrix}$$

et

$$A^T b = \begin{pmatrix} 7 \\ 8 \end{pmatrix}.$$

Le système normal $A^T Ax = A^T b$ est donc

$$\begin{aligned} 6x_1 - 5x_2 &= 7 \\ -5x_1 + 11x_2 &= 8. \end{aligned}$$

dont l'unique solution est

$$x_1 = \frac{117}{39} \qquad x_2 = \frac{83}{39}.$$

Chapitre 8

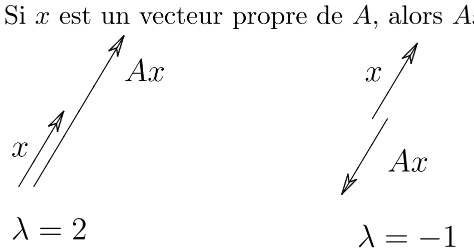
Diagonalisation des matrices

8.1 Définitions et premières propriétés

Définition 8.1. Soit A une matrice $n \times n$. On dit que $x \in \mathbb{R}^n$, $x \neq 0$, est un vecteur propre de A s'il existe $\lambda \in \mathbb{R}$ tel que

$$Ax = \lambda x\,.$$

Le scalaire λ est appelé valeur propre de A. On dit que x est un vecteur propre associé (ou correspondant) à la valeur propre λ. Si x est un vecteur propre de A, alors Ax est colinéaire à x :



Exemple 8.2. Soit

$$A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}.$$

Alors $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ est un vecteur propre de A correspondant à la valeur propre $\lambda = 3$, car

$$Ax = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3x.$$

Soit A une matrice $n \times n$. On peut réécrire l'égalité

$$Ax = \lambda x$$

comme

$$Ax = \lambda I x,$$

ce qui est équivalent à

$$(\lambda I - A)x = 0\,.$$

Si λ est une valeur propre de A, alors cette équation a une solution $x \neq 0$. Ceci implique que

$$\det(\lambda I - A) = 0.$$

Cette dernière équation est appelée *l'équation caractéristique de* A et $p(\lambda) = \det(\lambda I - A)$ est appelé *le polynôme caractéristique de* A. Les racines λ de l'équation caractéristique sont les valeurs propres de A. Si A est une matrice $n \times n$, alors l'équation caractéristique de A est de degré n et le coefficient de λ^n est 1. Autrement dit, le polynôme caractéristique d'une matrice $n \times n$ est de la forme

$$p(\lambda) = \det(\lambda I - A) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n.$$

On sait qu'un polynôme de degré n a au plus n racines distinctes ce qui montre qu'une matrice $n \times n$ a au plus n valeurs propres distinctes.

Exemple 8.3. Cherchons les valeurs propres et les vecteurs propres de la matrice

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

Son équation caractéristique est

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{pmatrix} = \lambda^2 - 3\lambda - 10 = 0.$$

Les valeurs propres de A sont donc les racines de

$$\lambda^2 - 3\lambda - 10 = 0.$$

qui sont polynôme

$$\lambda = -2 \qquad \text{et} \qquad \lambda = 5.$$

Cherchons maintenant les vecteurs propres de A. Un vecteur

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

est vecteur propre de A si et seulement si

$$Ax = \lambda x$$

autrement dit si

$$(\lambda I - A)x = 0.$$

Ceci nous donne

$$\begin{pmatrix} \lambda - 1 & -3 \\ -4 & \lambda - 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Pour $\lambda = -2$, on obtient

$$\begin{pmatrix} -3 & -3 \\ -4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Les solutions de ce système sont

$$x_1 = -t, \qquad x_2 = t.$$

L'ensemble des vecteurs propres de A correspondant à la valeur propre $\lambda = -2$ est donc

$$\left\{ \begin{pmatrix} -t \\ t \end{pmatrix} \mid t \in \mathbb{R}, t \neq 0 \right\}$$

Pour $\lambda = 5$, on obtient le système

$$\begin{pmatrix} 4 & -3 \\ -4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

dont les solutions sont

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{3}{4}t \\ t \end{pmatrix}.$$

L'ensemble des vecteurs propres de A correspondant à la valeur propre $\lambda = 5$ est donc

$$\left\{ \begin{pmatrix} \frac{3}{4}t \\ t \end{pmatrix} \mid t \in \mathbb{R}, t \neq 0 \right\}$$

Théorème 8.4. *Soit A une matrice triangulaire. Alors les valeurs propres de A sont les éléments de la diagonale de A.*

Démonstration. Soit A une matrice triangulaire supérieure

$$A = \begin{pmatrix} a_{11} & * & \dots & \dots & * \\ 0 & a_{22} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & * \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

Le polynôme caractéristique de A est

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - a_{11} & * & \dots & \dots & * \\ 0 & \lambda - a_{22} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & * \\ 0 & 0 & \dots & 0 & \lambda - a_{nn} \end{pmatrix} = (\lambda - a_{11})(\lambda - a_{22}) \dots (\lambda - a_{nn})$$

ce qui montre que les valeurs propres de A sont

$$\lambda = a_{11}, \quad \lambda = a_{22}, \quad \dots \quad , \lambda = a_{nn}\,.$$

La démonstration est analogue lorsque A est triangulaire inférieure. □

8.1.1 Calcul des vecteurs propres

Soit A une matrice carrée et λ une valeur propre de A. Les vecteurs propres de A correspondant à la valeur propre λ sont les vecteurs $x \neq 0$ qui satisfont l'équation $Ax = \lambda x$. De façon équivalente, ces vecteurs sont les vecteurs non nuls de l'espace des solutions de

$$(\lambda I - A)x = 0\,.$$

L'ensemble de ces vecteurs propres est appelé *l'espace propre de A correspondant à la valeur propre* λ et est noté E_λ.

Nous venons donc de voir que l'espace propre E_λ est le nilespace de la matrice $\lambda I - A$.

Exemple 8.5. Trouver des bases pour les espaces propres de la matrice

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

L'équation caractéristique de A est

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 = 0\,.$$

Les valeurs propres de A sont donc

$$\lambda = 1 \qquad \text{et} \qquad \lambda = 2\,.$$

Il y a donc deux espaces propres.

Pour calculer E_2 (l'espace propre associé à la valeur propre 2), il faut trouver le noyau de $2I - A$, c'est-à-dire le noyau de

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

On obtient

$$x_1 = -s, \quad x_2 = t, \quad x_3 = s.$$

Les vecteurs propres de A correspondant à la valeur propre $\lambda = 2$ sont donc les vecteurs non nuls de la forme

$$x = \begin{pmatrix} -s \\ t \\ s \end{pmatrix} = \begin{pmatrix} -s \\ 0 \\ s \end{pmatrix} + \begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix} = s\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Comme

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

sont linéairement indépendants, ces vecteurs forment une base de l'espace propre E_2.

Pour $\lambda = 1$, on calcule le noyau de $I - A$:

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

Les solutions sont

$$x_1 = -2s, \quad x_2 = s, \quad x_3 = s$$

ce qui nous permet de conclure que les vecteurs propres de A correspondant à la valeur propre $\lambda = 1$ sont les vecteurs non nuls de la forme

$$s\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

L'ensemble

$$\left\{\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\right\}$$

constitue ainsi une base possible de E_1

Théorème 8.6. *Soit A une matrice carrée, et soit k un entier. Si λ est une valeur propre de A, et x un vecteur propre correspondant, alors λ^k est une valeur propre de A^k, et x est un vecteur propre correspondant à λ^k.*

polynôme

Démonstration. Supposons d'abord que k est positif. On a

$$A^k x = A^{k-1}(Ax) = A^{k-1}(\lambda x) = \lambda A^{k-1}x = \lambda A^{k-2}(Ax) = \lambda^2 A^{k-2}x = \cdots = \lambda^k x.$$

Si k est négatif, et si A est inversible, on a alors $A^{-1}x = \lambda^{-1}x$ en divisant l'équation $Ax = \lambda x$ par $A^{-1}\lambda^{-1}$; le même calcul donne $A^k x = \lambda^k x$. □

En particulier, les valeurs propres de A^k sont précisément les puissances k-ièmes des valeurs propres de A. Remarquons toutefois que A^k peut avoir plus de vecteurs propres que A. Un exemple simple est la matrice $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$; l'espace propre E_0 de A est de dimension 1, mais comme $A^2 = 0$ l'espace propre E_0 de A^2 est de dimension 2.

Lemme 8.7. *Soit $f = a_n X^n + \cdots + a_1 X + a_0$ un polynôme à coefficients réels. Alors $X = 0$ est une racine de $f = 0$ si et seulement si $a_0 = 0$.*

Démonstration. Si 0 est une racine de $f = 0$, alors en remplaçant X par 0 on trouve $a_0 = 0$. Réciproquement, si $a_0 = 0$, alors on vérifie que 0 est bien une solution de l'équation $f = 0$. □

Théorème 8.8. *Soit $A = (a_{ij})$ une matrice $n \times n$ et $\det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ son polynôme caractéristique. Alors*

$$c_0 = (-1)^n \det(A) \qquad \textit{et} \qquad c_{n-1} = -tr(A).$$

Démonstration. Montrons la premiére égalité. Remplaçons λ par 0 dans le polynôme caractéristique de A. On obtient ainsi

$$\det(-A) = c_0.$$

Comme $\det(-A) = (-1)^n \det(A)$, on obtient l'égalité voulue.
Montrons la deuxième égalité.

$$\det(\lambda I - n) = \det \begin{pmatrix} \lambda - a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \lambda - a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & a_{n(n-1)} & \lambda - a_{nn} \end{pmatrix}.$$

Si on développe ce déterminant par rapport à la premiére ligne par exemple, on remarque que le seul terme qui contient des puissances de λ supérieures à $n-2$ est

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).$$

On a donc

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) = \lambda^n + (-a_{11} - a_{22} - \cdots - a_{nn})\lambda^{n-1} + \dots$$

Le coefficient de λ^{n-1} dans le polynôme caractéristique est donc bien $-\text{trace}(A)$. □

Corollaire 8.9. *Une matrice carrée A est inversible si et seulement si $\lambda = 0$ n'est pas valeur propre de A.*

Démonstration. la matrice A est inversible si et seulement si $\det(A) \neq 0$. Or, par le théorème précédent, ceci est équivalent à dire que $\lambda = 0$ n'est pas un zéro du polynôme caractéristique, c'est-à-dire pas une valeur propre de A. □

Notons que si $\lambda = 0$ est une valeur propre de A, alors l'espace propre associé E_0 n'est rien d'autre que le nilespace de A. En résumé,

$$E_0 = \text{Ker} A$$

si 0 est valeur propre.

8.2 Diagonalisation

Définition 8.10. Soit A une matrice carrée. On dit que A est diagonalisable s'il existe une matrice inversible P telle que $P^{-1}AP$ soit une matrice diagonale.

Théorème 8.11. *Soit A une matrice $n \times n$. Alors les propriétés suivantes sont équivalentes :*

(a) A est diagonalisable ;

(b) A a n vecteurs propres linéairement indépendants.

Démonstration. $(a) \Longrightarrow (b)$ Comme A est diagonalisable, il existe une matrice inversible

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix},$$

telle que $P^{-1}AP$ soit diagonale. On a donc $P^{-1}AP = D$, avec

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Comme $P^{-1}AP = D$, on a $AP = PD$. Donc

$$AP = PD = \begin{pmatrix} p_{11} & \dots & p_{1n} \\ \vdots & & \vdots \\ p_{n_1} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \dots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \dots & \lambda_n p_{2n} \\ \vdots & \vdots & & \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \dots & \lambda_n p_{nn} \end{pmatrix}.$$

Soient $p_1, p_2, \dots, p_n$ les vecteurs colonne de P. Alors les colonnes de AP sont

$$\lambda_1 p_1, \quad \lambda_2 p_2, \quad \dots, \quad \lambda_n p_n.$$

Mais on sait que les colonnes de AP sont

$$Ap_1, \quad Ap_2, \quad \dots, \quad Ap_n.$$

Donc

$$Ap_1 = \lambda_1 p_1,\ Ap_2 = \lambda_2 p_2,\ \dots,\ Ap_n = \lambda_n p_n\,.$$

Comme P est inversible, les vecteurs colonne $p_1, p_2, \dots, p_n$ sont linéairement indépendants. Par les formules ci-dessus, $p_1, p_2, \dots, p_n$ sont des vecteurs propres de A correspondant aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$. Donc A a n vecteurs propres linéairement indépendants.

$(b) \implies (a)$ Supposons que A ait n vecteurs propres linéairement indépendants $p_1, p_2, \dots, p_n$, avec valeurs propres correspondantes $\lambda_1, \lambda_2, \dots, \lambda_n$. Posons

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}.$$

Les colonnes de la matrice P sont les vecteurs $p_1, p_2, \dots, p_n$. Les colonnes du produit AP sont les vecteurs $Ap_1, Ap_2, \dots, Ap_n$. Mais comme

$$Ap_1 = \lambda_1 p_1, Ap_2 = \lambda_2 p_2,\ \dots,\ Ap_n = \lambda_n p_n,$$

on a

$$\begin{aligned} AP &= \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \dots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \dots & \lambda_n p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \dots & \lambda_n p_{nn} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \\ &= PD, \end{aligned}$$

avec

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

On a donc :

$$AP = PD\,.$$

Comme les vecteurs colonne de P sont linéairement indépendants, la matrice P est inversible et l'on obtient finalement

$$P^{-1}AP = D\,,$$

ce qui montre que A est diagonalisable. □

8.2.1 Méthode pour diagonaliser une matrice

1. Trouver n vecteurs propres linéairement indépendants, $p_1, p_2, \ldots, p_n$.
2. Construire la matrice P ayant $p_1, p_2, \ldots, p_n$ comme vecteurs colonne.
3. La matrice $P^{-1}AP$ est diagonale. Les éléments diagonaux sont les valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ correspondant respectivement aux vecteurs propres $p_1, p_2, \ldots, p_n$.

Exemple 8.12. Diagonalisons la matrice

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

L'équation caractéristique de A est

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

et les valeurs propres sont alors $\lambda = 1$ et $\lambda = 2$.

Une base de E_2 est donnée par

$$p_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

alors qu'une base de E_1 est donnée par le vecteur

$$p_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

Ces 3 vecteurs propres étant linéairement indépendants, la matrice A est diagonalisable. On pose alors

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

et on a

$$P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Théorème 8.13. *Si $v_1, v_2, \ldots, v_n$ sont des vecteurs propres d'une matrice A correspondant à des valeurs propres distinctes $\lambda_1, \lambda_2, \ldots, \lambda_n$, alors $\{v_1, v_2, \ldots, v_n\}$ est un ensemble linéairement indépendant.*

Démonstration. Nous allons démontrer ce théorème par l'absurde.

Supposons donc que les vecteurs propres $v_1, v_2, \ldots, v_n$ sont linéairement dépendants. Soit r le plus grand entier tel que $\{v_1, v_2, \ldots, v_r\}$ soit linéairement indépendant. On a $1 \leq r < n$. Il existe donc des scalaires $c_1, c_2, \ldots, c_{r+1}$ non tous nuls tels que

$$c_1 v_1 + c_2 v_2 + \cdots + c_{r+1} v_{r+1} = 0 . \qquad (*)$$

Multiplions cette équation par A. On obtient

$$c_1 A v_1 + c_2 A v_2 + \cdots + c_{r+1} A v_{r+1} = 0 .$$

Comme

$$Av_1 = \lambda_1 v_1,\ Av_2 = \lambda_2 v_2, \ldots, Av_{r+1} = \lambda_{r+1} v_{r+1} ,$$

on obtient

$$c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_{r+1}\ \lambda_{r+1}\ v_{r+1} = 0 .$$

En multipliant $(*)$ par λ_{r+1}, on a

$$c_1 \lambda_{r+1} v_1 + c_2 \lambda_{r+1} v_2 + \cdots + c_{r+1} \lambda_{r+1} v_{r+1} = 0 .$$

En soustrayant cette équation de la dernière, on obtient

$$c_1(\lambda_1 - \lambda_{r+1})v_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})v_r = 0\,.$$

Comme les $\lambda_1, \ldots, \lambda_n$ sont distincts, on a

$$c_1 = c_2 = \cdots = c_r = 0\,.$$

On a donc $c_{r+1}v_{r+1} = 0$. Comme $v_{r+1} \neq 0$, on a $c_{r+1} = 0$ ce qui conduit à

$$c_1 = c_2 = \cdots = c_{r+1} = 0.$$

Ceci contredit l'hypothèse que $c_1, c_2, \ldots, c_{r+1}$ sont non tous nuls. $\square$

Corollaire 8.14. *Soit A une matrice $n \times n$. Si A a n valeurs propres distinctes, alors A est diagonalisable.*

Démonstration. C'est une conséquence immédiate des théorèmes 8.13 et 8.11. $\square$

Exemple 8.15. Soit $A = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$. Par le théorème 8.8, le polynôme carcatéristique de A est

$$\lambda^2 - \mathrm{trace}(A)\lambda + \det(A) = \lambda^2 + 4.$$

A n'est pas diagonalisable car son polynôme n'a pas de racines dans $\mathbb{R}$.

Exemple 8.16. Soit $A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$. Comme A est triangulaire, ses valeurs propres apparaissent sur la diagonale. La matrice a donc une unique valeur propre qui est égale à 2. On sait, par le théorème 8.11, que A est diagonalisable si et seulement si A a deux vecteurs propres linéairement indépendants. Calculons l'espace E_2 : on doit résoudre le système

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix}, \quad x, y \in \mathbb{R}.$$

Il est équivalent à $y = 0$. Ainsi, $E_2 = \left\{ \begin{pmatrix} s \\ 0 \end{pmatrix} \mid s \in \mathbb{R} \right\}$ qui est de dimension 1. A n'a donc pas 2 vecteurs linéairement indépendants , elle n'est donc pas diagonalisable.

8.3 Matrices symétriques et diagonalisation

Comme nous allons le voir, les matrices symétriques sont des matrices ayant la propriété d'être toujours diagonalisables. De plus, il existe une base de vecteurs propres orthonormée.

Théorème 8.17. *Soit A une matrice $n \times n$, symétrique. Alors A possède n vecteurs propres linéairement indépendants. De plus, on peut choisir ces vecteurs propres de façon qu'ils forment une base orthonormale. Une matrice carrée symétrique est donc diagonalisable.*

On va faire appel au lemme suivant, dont la démonstration nécessite l'usage de nombres complexes.

Lemme 8.18. *Soit A une matrice $n \times n$, symétrique. Alors A possède une valeur propre réelle.*

Preuve du théorème 8.17. Il s'agit de montrer : pour une matrice A symétrique ($A^T = A$), il existe une matrice P orthogonale ($P^TP = I$) telle que $P^{-1}AP$ est diagonale. En effet, la matrice P en question a pour colonnes les vecteurs propres de A, et ces derniers forment une base orthonormale si et seulement si P est orthogonale.

On procède par récurrence, le cas $n = 1$ étant trivial. Par le lemme, soit λ_1 une valeur propre de A, et soit x un vecteur propre associé. On peut supposer x de norme 1, soit $x^Tx = 1$.

Par le procédé de Gram-Schmidt, on peut compléter x en une base orthonormale $(x, y_2, \ldots, y_n)$ de $\mathbb{R}^n$. On note Y la matrice $n \times (n-1)$ dont les colonnes sont $y_1, \ldots, y_n$. On a donc $Y^TY = 1$ et $x^TY = 0$ et $Y^Tx = 0$, par orthogonalité de la base.

La matrice $B = Y^T AY$ est une matrice symétrique, car $B^T = Y^T A^T (Y^T)^T = B$. De plus, elle est de taille $(n-1) \times (n-1)$. Ainsi, par récurrence, il existe une matrice orthogonale Q telle que $Q^{-1}BQ$ soit diagonale, disons de coéfficients diagonaux $\lambda_2, \dots, \lambda_n$.

On pose alors $P = (x|YQ)$; c'est la matrice carrée dont la première colonne est x et dont les $n-1$ colonnes suivantes sont la matrice YQ. On calcule :

$$P^T P = \begin{pmatrix} x^T \\ \hline (YQ)^T \end{pmatrix} \left(\begin{array}{c|c} x & YQ \end{array} \right) = \left(\begin{array}{c|c} x^T x & x^T YQ \\ \hline Q^T Y^T x & Q^T Y^T YQ \end{array} \right) = \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \end{pmatrix} = I_n,$$

donc P est une matrice orthogonale. De plus,

$$\begin{aligned} P^{-1}AP = P^T AP &= \left(\begin{array}{c|c} x^T Ax & x^T AYQ \\ \hline Q^T Y^T Ax & Q^T Y^T AYQ \end{array} \right) = \left(\begin{array}{c|c} x^T \lambda_1 x & (Ax)^T YQ \\ \hline Q^T Y^T \lambda_1 x & Q^T BQ \end{array} \right) \\ &= \begin{pmatrix} \lambda_1 & \lambda_1 x^T YQ \\ \lambda_1 Q^T Y^T x & \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \end{aligned}$$

est diagonale. □

Théorème 8.19. *Deux vecteurs propres associés à deux valeurs propres distinctes sont orthogonaux (pour le produit scalaire usuel).*

Démonstration. Soient A une matrice carrée et λ, μ deux valeurs propres de A distinctes. Soient encore v un vecteur propre associé à la valeur propre λ et w un vecteur propre associé à la valeur propre μ. On a

$$(Av) \bullet w = (\lambda v) \bullet w = \lambda(v \bullet w) = \lambda\, w^T v.$$

D'autre part,

$$v \bullet (Aw) = (Aw)^T v = w^T A^T v = w^T Av = w^T \lambda v = \lambda w^T v.$$

La troisième égalité découle du fait que A est symétrique. On a donc $(Av) \bullet w = v \bullet (Aw)$, c'est-à-dire $\lambda\, v \bullet w = \mu\, v \bullet w$. Ainsi

$$(\lambda - \mu)\, v \bullet w = 0.$$

Comme $\lambda \neq \mu$, on a nécessairement $v \bullet w = 0$. Les vecteurs v et w sont donc bien orthogonaux. □

Définition 8.20. Soit A une matrice carrée. On dit que A est *orthogonalement diagonalisable* si elle est diagonalisable et s'il existe une base de vecteurs propres qui est orthonormée.

Par le théorème précédent, toute matrice symétrique est orthogonalement diagonalisable.
Nous donnons maintenant un exemple d'une diagonalisation orthogonale.

Exemple 8.21. Soit $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Le polynôme caractéristique est

$$\det(\lambda I - A)$$

c'est-à-dire

$$\det \begin{pmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda - 1 \end{pmatrix}.$$

On trouve ainsi

$$\lambda^3 - 3\lambda^2.$$

Les valeurs propres de A sont donc 0 et 3. Calculons les espaces propres. Pour E_0, on doit résoudre le système

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

qui est équivalent à $x+y+z=0$. Ainsi,

$$E_0 = \left\{ s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \middle| s,t \in \mathbb{R} \right\}$$

Comme A est diagonalisable et que E_0 est de dimension 2, alors E_3 est nécessairement de dimension 1. On a clairement $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in E_3$. Ainsi

$$E_3 = \left\{ s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \middle| s \in \mathbb{R} \right\}$$

On remarque qu'on a bien $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \perp \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \perp \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. Appliquons le procédé de Gram-Schmidt pour extraire une base orthogonale de E_0. Posons $f_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ et $f_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

$$\begin{aligned} f_2' &= f_2 - \frac{f_1 \bullet f_2}{f_1 \bullet f_1} f_1 \\ &= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \end{aligned}$$

Les vecteurs, $g_1 = \frac{f_1}{||f_1||} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ et $g_2 = \frac{f_2'}{||f_2'||} = \begin{pmatrix} -\frac{\sqrt{2}}{2\sqrt{3}} \\ -\frac{\sqrt{2}}{2\sqrt{3}} \\ \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$ forment une base orthonormée de E_0. Pour E_3, il suffit de prendre $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ comme vecteur de base de norme 1. On vérifie que

$$\underbrace{\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{\sqrt{2}}{2\sqrt{3}} & -\frac{\sqrt{2}}{2\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}}_{P^T} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}}_{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Chapitre 9

Applications linéaires

9.1 Définitions et exemples

Nous avons étudié la notion d'application linéaire

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m .$$

Cette notion se généralise à des espaces vectoriels quelconques.

Définition 9.1. Soient V et W des espaces vectoriels. On dit qu'une application

$$f : V \longrightarrow W$$

est une application linéaire si et seulement si les deux propriétés suivantes sont vérifiées :

(a) $f(u+v) = f(u) + f(v)$ pout tout $u, v \in V$;

(b) $f(\lambda u) = \lambda f(u)$ pour tout $u \in V$, $\lambda \in \mathbb{R}$.

Exemple 9.2. Voici quelques exemples d'applications linéaires :

(1) **Applications linéaires**

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$
$$f(x) = Ax$$

pour une certaine matrice A.

(2) L'**application nulle**

$$f : V \longrightarrow W$$
$$f(u) = 0$$

pour tout $u \in V$.

(3) L'**application identité**

$$f : V \longrightarrow V$$
$$f(u) = u$$

pour tout $u \in V$.

(4) La **symétrie centrale**

$$S : V \longrightarrow V$$
$$S(v) = -v$$

pour tout $v \in V$.

(5) L'**homothétie** de rapport k

$$H : V \longrightarrow V$$
$$H(v) = k\,v$$

pour tout $v \in V$.

(6) La **projection orthogonale**. Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$, et soit W un sous-espace vectoriel de V. Nous avons défini la projection orthogonale

$$f : V \longrightarrow W$$

$$f(v) = \mathrm{proj}_W(v)\,.$$

Si $S = \{w_1, w_2, \ldots, w_r\}$ est une base orthonormée de W, alors

$$f(v) = \mathrm{proj}_W(v) = < v, w_1 > w_1 + \cdots + < v, w_r > w_r\,.$$

C'est une application linéaire. En effet, vérifions les deux propriétés :

$$\begin{aligned} f(u+v) &= < u+v, w_1 > w_1 + \cdots + < u+v, w_r > w_r \\ &= < u, w_1 > w_1 + < v, w_1 > w_1 + \cdots + < u, w_r > w_r + < v, w_r > w_r \\ &= < u, w_1 > w_1 + \cdots + < u, w_r > w_r + < u, w_1 > w_1 + \cdots + < v, w_r > w_r \\ &= f(u) + f(v). \end{aligned}$$

$$\begin{aligned} f(\lambda u) &= < \lambda u, w_1 > w_1 + \cdots + < \lambda u, w_r > w_r \\ &= \lambda(< u, w_1 > w_1 + \cdots + < u, w_r > w_r) \\ &= \lambda f(u). \end{aligned}$$

Donc f est bien une application linéaire.

(7) Rappelons que l'on note $\mathcal{P}_n$ l'ensemble des fonctions polynomiales de degré $\leq n$: $p(x) = a_n x^n + \cdots + a_1 x + a_0$. On définit une application $f : \mathcal{P}_n \longrightarrow \mathcal{P}_{n+1}$ par

$$f(p) = xp(x) = a_n x^{n+1} + \cdots + a_1 x^2 + a_0 x.$$

C'est une application linéaire car on a :

$$\begin{aligned} f(p_1 + p_2) &= x(p_1(x) + p_2(x)) \\ &= xp_1(x) + xp_2(x) \\ &= f(p_1) + f(p_2) \\ &\text{et} \\ f(\lambda p) &= x(\lambda p(x)) \\ &= \lambda(xp(x)) \\ &= \lambda f(p)\,. \end{aligned}$$

(8) Soient $a, b \in \mathbb{R}$. Définissons l'application $T : \mathcal{P}_n \longrightarrow \mathcal{P}_n$ par

$$T(f) = f(ax+b).$$

C'est bien une application linéaire car :

$$\begin{aligned} T(f_1 + f_2) &= (f_1 + f_2)(ax+b) \\ &= f_1(ax+b) + f_2(ax+b) \\ &= T(f_1) + T(f_2) \end{aligned}$$

et

$$\begin{aligned} T(\lambda f) &= (\lambda f)(ax+b) \\ &= \lambda(f(ax+b)) \\ &= \lambda T(f). \end{aligned}$$

(9) L'**application linéaire définie par un produit scalaire généralisé**. Soit V un espace vectoriel muni d'un produit scalaire généralisé $<,>$. Soit $v_0 \in V$. On définit une application

$$f : V \longrightarrow \mathbb{R}$$

en posant

$$f(v) = < v, v_0 > \ .$$

Alors f est une application linéaire. En effet, on a :

$$f(u+v) = < u+v,\, v_0 > = < u, v_0 > + < v,\, v_0 > = f(u) + f(v)\,,$$

et

$$f(\lambda v) = < \lambda v,\, v_0 > = \lambda < v,\, v_0 > = \lambda f(v)\,.$$

(10) La **dérivation**. Soit $V = \mathcal{C}^1(\mathbb{R})$ l'espace vectoriel des fonctions $f : \mathbb{R} \longrightarrow \mathbb{R}$ dérivables avec première dérivée continue et $W = \mathcal{C}^0(\mathbb{R})$ l'espace vectoriel des fonctions continues réelles à variable réelle. Soit

$$D : \mathcal{C}^1(\mathbb{R}) \longrightarrow \mathcal{C}^0(\mathbb{R})$$

l'application définie par

$$D(f) = f'$$

où f' est la première dérivée de f. Alors D est une application linéaire car la dérivation est une opération linéraire.

(11) L'**intégration**. Soit $V = \mathcal{C}^0(\mathbb{R})$ et $W = \mathcal{C}^1(\mathbb{R})$. Soit

$$\begin{array}{rcl} J : \mathcal{C}^0(\mathbb{R}) & \longrightarrow & \mathcal{C}^1(\mathbb{R}) \\ f(t) & \mapsto & \displaystyle\int_0^x f(t)dt \end{array}$$

L'application J est linéaire car $\int_0^x (f(t) + g(t))dt = \int_0^x f(t)dt + \int_0^x g(t)dt$ et $\int_0^x (\lambda\, f(t))dt = \lambda \int_0^x f(t)dt$ pour toutes fonctions f et g et pour tout $\lambda \in \mathbb{R}$.

(12) Soit $M_n(\mathbb{R})$ l'ensemble des matrices $n \times n$. Alors

$$\begin{array}{rcl} tr : M_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ A & \mapsto & \mathrm{trace}(A) \end{array}$$

est une application linéaire.

Exemple 9.3. Donnons un exemple d'une application qui n'est pas linéaire. Soit

$$f : M_{22} \longrightarrow \mathbb{R}$$

la fonction définie par

$$f(A) = \det(A)\,.$$

On a

$$\det(\lambda A) = \lambda^2 \det(A)$$

et $\lambda^2 \det(A)$ est différent de $\lambda \det(A)$ dès que $\lambda \neq 1$ et $\det(A) \neq 0$. De plus, en général

$$\det(A+B) \neq \det(A) + \det(B)$$

ce qui montre que cette application n'est pas linéaire.

9.1.1 Propriétés des applications linéaires

Soit $f : V \longrightarrow W$ une application linéaire. Pour tout $v_1, v_2 \in V$ et pour tout $\lambda_1, \lambda_2 \in \mathbb{R}$, on a

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2)\,.$$

Plus généralement, pour tout $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ et tout $v_1, \dots, v_n \in V$, on a

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

On dit que f préserve les combinaisons linéaires.

Théorème 9.4. *Soit $f : V \longrightarrow W$ une application linéaire. Alors on a les propriétés suivantes :*

(a) $f(0) = 0$

(b) $f(-v) = -f(v)$ *pour tout* $v \in V$

(c) $f(v - w) = f(v) - f(w)$ *pour tout* $v, w \in V$.

Démonstration. (a) Soit $v \in V$. On a $0 \cdot v = 0$. Comme f est linéaire, on a

$$f(0) = f(0 \cdot v) = 0 f(v) = 0\,.$$

(b) Pour tout $v \in V$, on a

$$f(-v) = f((-1)v) = (-1)f(v) = -f(v)\,.$$

Pour tout $v, w \in V$, on a

$$v - w = v + (-1)w,$$

donc

$$\begin{aligned} f(v-w) &= f(v + (-1)w) \\ &= f(v) + f((-1)w) \\ &= f(v) + (-1)f(w) \\ &= f(v) - f(w)\,. \end{aligned}$$

□

9.1.2 Expression d'une application linéaire dans une base

Soit $f : V \longrightarrow W$ une application linéaire, et soit $\{v_1, \dots, v_n\}$ une base de l'espace vectoriel V. Alors f est déterminée par les images des éléments de la base, c'est-à-dire par l'ensemble

$$\{f(v_1), \dots, f(v_n)\}.$$

En effet, si $v \in V$, alors $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. On a donc

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_n v_n) \\ &= \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)\,. \end{aligned}$$

Exemple 9.5. Soit $\{v_1, v_2, v_3\}$ une base de $\mathbb{R}^3$, avec $v_1 = (1,1,1), v_2 = (1,1,0)$ et $v_3 = (1,0,0)$. Soit

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

l'application linéaire telle que

$$f(v_1) = (1,0), f(v_2) = (2,-1), f(v_3) = (4,3)\,.$$

Soit

$$v = (2,-3,5)\,.$$

Alors
$$v = 5v_1 - 8v_2 + 5v_3\,,$$
donc
$$\begin{array}{rcl} f(v) &=& f(5v_1 - 8v_2 + 5v_3) = \\ &=& 5f(v_1) - 8f(v_2) + 5f(v_3) \\ &=& 5(1,0) - 8(2,-1) + 5(4,3) \\ &=& (5 - 16 + 20, 8 + 15) \\ &=& (9, 23)\,. \end{array}$$

Définition 9.6. Soient $f_1 : V_1 \longrightarrow V_2$ et $f_2 : V_2 \longrightarrow V_3$ deux applications linéaires. Alors on définit la composition de f_2 avec f_1, notée $f_2 \circ f_1$ par
$$\begin{array}{rl} (f_2 \circ f_1) : V_1 & \longrightarrow V_3 \\ v & \mapsto f_2(f_1(v)) \end{array}$$
pour tout $v \in V_1$.

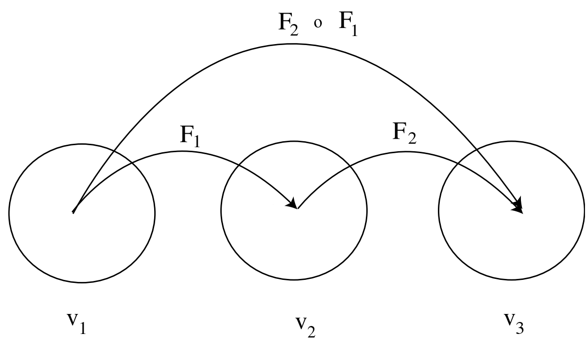


Figure 11.1

Théorème 9.7. *Si $f_1 : V_1 \longrightarrow V_2$ et $f_2 : V_2 \longrightarrow V_3$ sont des applications linéaires, alors $(f_2 \circ f_1) : V_1 \longrightarrow V_3$ est aussi une application linéaire.*

Démonstration. Soient $u, v \in V_1$. Alors on a
$$\begin{array}{rcl} (f_2 \circ f_1)(u+v) &=& f_2(f_1(u+v)) \\ &=& f_2(f_1(u) + f_1(v)) \\ &=& f_2(f_1(u)) + f_2(f_1(v)) \\ &=& (f_2 \circ f_1)(u) + (f_2 \circ f_1)(v)\,. \end{array}$$
Soient $v \in V_1$ et $\lambda \in \mathbb{R}$. Alors
$$\begin{array}{rcl} (f_2 \circ f_1)(\lambda v) &=& f_2(f_1(\lambda v)) = \\ &=& f_2(\lambda f_1(v)) = \\ &=& \lambda f_2(f_1(v)) = \\ &=& \lambda (f_2 \circ f_1)(v)\,. \end{array}$$
Donc $f_2 \circ f_1$ est bien une application linéaire. □

9.2 Noyau et image d'une application linéaire

Définition 9.8. Soit $f : V \longrightarrow W$ une application linéaire. Le **noyau de f**, noté $\mathrm{Ker}(f)$, est par défintion l'ensemble des $v \in V$ tels que

$$f(v) = 0\,.$$

L'image de f , notée $\mathrm{Im}(f)$, est par définition l'ensemble des $w \in W$ tel qu'il existe $v \in V$ avec

$$f(v) = w\,.$$

Exemple 9.9. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ définie par $f(x) = Ax$, où A est une matrice $m \times n$. Alors $\mathrm{Ker}(f)$ est égal au nilespace de A, et $\mathrm{Im}\ (f)$ est égal à l'espace des colonnes de A.

Exemple 9.10. Soit $0 : V \longrightarrow W$ l'application linéaire nulle. Alors $\mathrm{Ker}(0) = V$ et $\mathrm{Im}(0) = \{0\}$.

Exemple 9.11. Soit $f : V \longrightarrow V$ l'application identité. Alors $\mathrm{Ker}(f) = \{0\}$ et $\mathrm{Im}\ (f) = V$.

Exemple 9.12. Soit $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la projection orthogonale sur le plan xy. Alors $\mathrm{Ker}(f)$ est égal à l'axe des z et l'image de f est égale au plan xy.

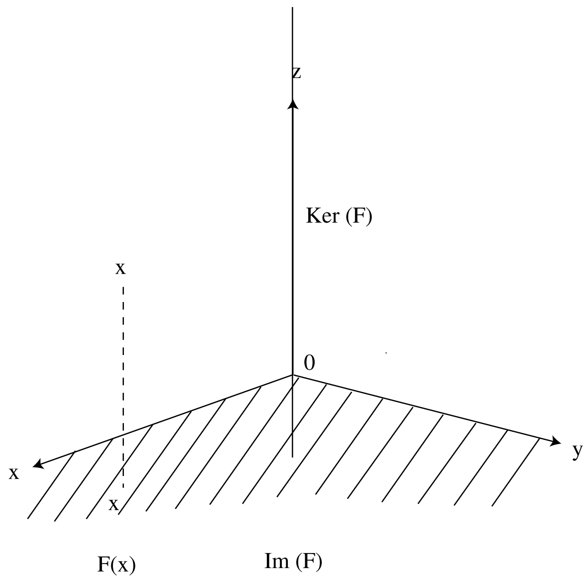


Figure 11.2

Théorème 9.13. *Soit $f : V \longrightarrow W$ une application linéaire. Alors*

(a) $\mathrm{Ker}(f)$ *est un sous-espace vectoriel de V.*

(b) $\mathrm{Im}(f)$ *est un sous-espace vectoriel de W.*

Démonstration. (a) On a $0 \in \mathrm{Ker}(f)$. Supposons que $u, v \in \mathrm{Ker}(f)$. Alors $f(u) = f(v) = 0$. On a $f(u+v) = f(u) + f(v) = 0$ ce qui montre que $u + v \in \mathrm{Ker}(f)$. Soit $\lambda \in \mathbb{R}$. Alors $f(\lambda v) = \lambda f(v) = 0$, car $f(v) = 0$. Ceci implique que $\lambda v \in \mathrm{Ker}(f)$. Ceci montre que $\mathrm{Ker}(f)$ est un sous-espace vectoriel de V.

(b) On a $f(0) = 0$, donc $0 \in \mathrm{Im}(f)$. Soient $w_1, w_2 \in \mathrm{Im}(f)$. Alors il existe $v_1, v_2 \in V$ tels que

$$\begin{aligned} f(v_1) &= w_1, \\ f(v_2) &= w_2\,. \end{aligned}$$

Mais alors

$$\begin{aligned} f(v_1+v_2) &= f(v_1)+f(v_2) \\ &= w_1+w_2\,, \end{aligned}$$

donc $w_1 + w_2 \in \mathrm{Im}(f)$. Soient $w \in \mathrm{Im}(f)$ et $\lambda \in \mathbb{R}$. Soit $v \in V$ tel que $f(v) = w$. Alors $f(\lambda v) = \lambda f(v) = \lambda w$. Donc $\lambda w \in \mathrm{Im}(f)$. Ceci implique que $\mathrm{Im}(f)$ est un sous-espace vectoriel de W. □

Exemple 9.14. Soit n un entier. Considérons l'application linéaire

$$\begin{aligned} \varphi : \mathcal{P}_n &\longrightarrow \mathcal{P}_{n+1} \\ f &\mapsto xf. \end{aligned}$$

Etudions d'abord le noyau de ϕ : soit $f = a_n x^n + \cdots + a_0 \in \mathcal{P}_n$ tel que $xf = 0$. Alors

$$a_n x^{n+1} = \cdots + a_1 x^2 + a_0 x = 0.$$

Ainsi, $a_i = 0$ pour tout $i \in \{0, \ldots, n\}$ et $f = 0$. Le noyau de ϕ est donc nul.
L'espace $\mathrm{Im}(\phi)$ est l'ensemble des polynômes de $\mathcal{P}_{n+1}$ sans terme constant. Une base de cet espace est $\{x, x^2, \ldots, x^{n+1}\}$.

Lemme 9.15. *Soit $f : V \longrightarrow W$ une application linéaire. Alors f est injective si et seulement si* $\mathrm{Ker}(f) = \{0\}$.

Démonstration. Supposons f injective. On a, par le théorème 9.4, que $f(0) = 0$. Mais, l'injectivité implique que pour tout $v \in V$, $v \neq 0$, on a $f(v) \neq 0$. Ainsi, le noyau de f est nul.
Réciproquement, supposons que $\mathrm{Ker}(f) = 0$, et soient $v_1, v_2 \in V$ avec $f(v_1) = f(v_2)$. Alors

$$0 = f(v_1) - f(v_2) = f(v_1) + f(-v_2) = f(v_1 - v_2).$$

Ainsi, $v_1 - v_2 \in \mathrm{Ker}(f)$ et $v_1 = v_2$. L'application f est donc bien injective. □

Définition 9.16. Soit $f : V \longrightarrow W$ une application linéaire. La dimension de $\mathrm{Ker}(f)$ est appelée *la nullité de* f, et est notée $\mathrm{null}(f)$. La dimension de $\mathrm{Im}(f)$ est appelée le rang de f, et est notée $\mathrm{rg}(f)$.

Ces définitions sont compatibles avec celles données pour les matrices comme le montre le théorème suivant :

Théorème 9.17. *Soit A une matrice $m \times n$, et soit*

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

définie par

$$f(x) = Ax.$$

Alors

(a) $\mathrm{null}(f) = \mathrm{null}(A)$

(b) $\mathrm{rg}(f) = \mathrm{rg}(A)$.

Théorème 9.18 (Théorème du rang)**.** *Soient V un espace vectoriel de dimension finie et $f : V \longrightarrow W$ une application linéaire. Alors*

$$\mathrm{rg}(f) + \mathrm{null}(f) = \dim(V)\,.$$

Démonstration. (i) Si $\mathrm{null}(f) = 0$, alors $\mathrm{Ker}(f) = \{0\}$ et, par le lemme 9.15, f est injective. Soit $\{e_1, \ldots, e_n\}$ une base de l'espace V. La famille $\{f(e_1), \ldots, f(e_n)\}$ est une base de l'image de f. En effet, c'est une famille linéairement indépendante car si $\lambda_1 f(e_1) + \cdots + \lambda_n f(e_n) = 0$, la linéarité permet de dire que $f(\lambda_1 e_1 + \cdots + \lambda_n e_n) = 0$. Mais l'injectivité implique que $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$ et comme $\{e_1, \ldots, e_n\}$ est une base, on a finalement $\lambda_1 = \cdots = \lambda_n = 0$. Il reste à montrer que c'est une famille génératrice. Soit $w \in \mathrm{Im}(f)$. Il existe $v \in V$ tel que $f(v) = w$. Mais v s'écrit $\mu_1 e_1 + \cdots + \mu_n e_n$, pour certains $\mu_1, \ldots, \mu_n \in \mathbb{R}$. Ainsi, $w = f(\mu_1 e_1 + \cdots + \mu_n e_n) = \mu_1 f(e_1) + \cdots + \mu_n f(e_n)$.
On a ainsi $\mathrm{rg}(f) = \dim(V)$ et donc bien $\mathrm{rg}(f) + (f) = \dim(V)$.

(ii) Si $\mathrm{rg}(f) = 0$, cela signifie que $\mathrm{Im}(f) = \{0\}$. Autrement dit, tout vecteur de V est envoyé sur 0 par l'application f. Ainsi, $\mathrm{Ker}(f) = V$ et $\mathrm{null}(f) = \dim(V)$.

(iii) Supposons maintenant que $\mathrm{null}(f), \mathrm{rg}(f) \neq 0$. Posons $r = (f)$ et soit $\{e_1, \ldots, e_r\}$ une base de $\mathrm{Ker}(f)$. Pour montrer le théorème, il suffit de montrer que l'image de f est de dimension $n-r$. Comme $\{e_1, \ldots, e_n\}$ est une base de
$Ker(f)$, c'est en particulier une famille de vecteurs linéairement indépendants. Il existe donc $e_{r+1}, \ldots, e_n \in V$ tels que $\{e_1, \ldots, e_n\}$ soit une base de V. La famille $B = \{f(e_{r+1}), \ldots, f(e_n)\}$ est une base de $\mathrm{Im}(f)$. En effet,

i) si $\lambda_{r+1} f(e_{r+1}) + \cdots + \lambda_n f(e_n) = 0$, alors, par linéarité, on a que $\lambda_{r+1} e_{r+1} + \cdots + \lambda_n e_n \in Ker(f)$. Ce vecteur s'exprime alors dans la base du noyau :

$$\lambda_{r+1} e_{r+1} + \cdots + \lambda_n e_n = \mu_1 e_1 + \cdots + \mu_r e_r \quad \text{pour certains } \mu_1, \ldots, \mu_r \in \mathbb{R}$$

ou, autrement dit,

$$-\mu_1 e_1 + \cdots - \mu_r e_r + \lambda_{r+1} e_{r+1} + \cdots + \lambda_n e_n = 0.$$

Ceci implique en particulier que $\lambda_{r+1} = \cdots = \lambda_n = 0$. B est donc linéairement indépendante.

ii) soit $w \in \mathrm{Im}(f)$. Il existe $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tels que

$$w = f(\alpha_1 e_1 + \cdots + \alpha_n e_n).$$

Mais cette égalité implique que

$$w = \alpha_1 \underbrace{f(e_1)}_{=0} + \cdots + \alpha_r \underbrace{f(e_r)}_{=0} + \alpha_{r+1} f(e_{r+1}) + \cdots + \alpha_n f(e_n).$$

Ainsi, tout élément de l'image de f est une combinaison linéaire des vecteurs de B.

□

Théorème 9.19. *Soit V un espace vectoriel de dimension finie, et soit*

$$f : V \longrightarrow V$$

une application linéaire. Alors les affirmations suivantes sont équivalentes :

(a) f *est injective*

(b) $\mathrm{Ker}(f) = \{0\}$

(c) $\mathrm{null}(f) = 0$

(d) $\mathrm{rg}(f) = n = \dim(V)$.

9.3 Applications linéaires inversibles

Nous avons déjà défini l'inverse d'une application linéaire f_A donnée par une matrice inversible A. Nous allons maintenant généraliser cette notion pour toutes les applications linéaires.

Définition 9.20. Soit $f : V \longrightarrow W$ une application linéaire injective. On définit

$$f^{-1} : \mathrm{Im}(f) \longrightarrow V$$

de la manière suivante :
Soit $w \in \mathrm{Im}(f)$. Il existe $v \in V$ tel que $w = f(v)$. De plus, comme f est injective, le vecteur v est unique. On définit alors

$$f^{-1}(w) = v.$$

L'application f^{-1} est appelée l'inverse de f.

Remarque 9.21. Dans le définition ci-dessus, on définit f^{-1} uniquement sur l'espace $\mathrm{Im}(f)$. Si $\mathrm{Im}(f) = W$, c'est-à-dire si f est aussi surjective, alors on peut définir $f^{-1} : W \longrightarrow V$. On dit alors que f est inversible.

Remarque 9.22. Une application linéaire $f : V \longrightarrow W$ est donc dite inversible si et seulement si elle est bijective.

Théorème 9.23. *Soit $f : V \longrightarrow W$ une application linéaire injective. Alors*
(1) $(f^{-1} \circ f)(v) = v$ pour tout $v \in V$.
(2) $(f \circ f^{-1})(w) = w$ pour tout $w \in \mathrm{Im}(f)$.
(3) f^{-1} est une application linéaire.

Démonstration. (1) Ce point découle directement de la définition de l'inverse.
(2) Soit $w \in \mathrm{Im}(f)$. Il existe un unique $v \in V$ tel que $f(v) = w$. On a donc

$$(f \circ f^{-1})(w) = f(\underbrace{f^{-1}(w)}_{=v}) = f(v) = w.$$

(3) Soient $w_1, w_2 \in \mathrm{Im}(f)$ et $\lambda \in \mathbb{R}$. Alors

$$\begin{aligned} f^{-1}(w_1 + w_2) &= f^{-1}(f(v_1) + f(v_2)) \quad \text{pour certains } v_1, v_2 \in V \\ &= f^{-1}(f(v_1 + v_2)) \\ &= v_1 + v_2. \end{aligned}$$

Or, $f^{-1}(w_i) = v_i$ pour $i = 1, 2$. Ainsi

$$f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2).$$

La propriété $f^{-1}(\lambda w_1) = \lambda f^{-1}(w_1)$ se montre de manière analogue.

□

Théorème 9.24. *Soit $f : V \longrightarrow W$ une application linéaire injective. Si $g : \mathrm{Im}(f) \longrightarrow V$ est une application telle que $(g \circ f)(v) = v$ pour tout $v \in V$, alors $g = f^{-1}$. De plus, on a également $(f \circ g)(w) = w$ pour tout $w \in \mathrm{Im}(f)$.*

Démonstration. Soit $w \in \mathrm{Im}(f)$. Par hypothèse, on a

$$(g \circ f)(f^{-1}(w)) = f^{-1}(w) \qquad (\star)$$

Mais, par le théorème précédent, cette expression est égale à $g(w)$. Ainsi, $g(w) = f^{-1}(w)$ pour tout $w \in \mathrm{Im}(f)$, c'est-à-dire que $g = f^{-1}$.
En appliquant f des deux côtés de l'égalité $(\star)$, on obtient $(f \circ g)(w) = w$ pour tout $w \in \mathrm{Im}(f)$. □

Cas particulier

Soit V un espace vectoriel de dimension finie et $f : V \longrightarrow V$ une application injective. Alors, par le théorème du rang, $\mathrm{rg}(f) = \dim(V)$ et donc, $\mathrm{Im}(f) = V$. Dans ce cas, l'inverse de f est défini sur tout V.

Théorème 9.25. *Soient $f : V_1 \longrightarrow V_2$ et $g : V_2 \longrightarrow V_3$ deux applications linéaires injectives. Alors $g \circ f$ est aussi injective et*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Démonstration. Soit $v_1 \in V_1$ avec $g(f(v_1)) = 0$. On utilise succesivement l'injectivité de g et de f pour conclure que $v_1 = 0$. La composition $g \circ f$ est donc bien injective.
Soit $v_3 \in \mathrm{Im}(g \circ f)$. Posons $u = (g \circ f)^{-1}(v_3)$. Montrons que $u = (f^{-1} \circ g^{-1})(v_3)$. On a

$$(g \circ f)(u) = (g \circ f)((g \circ f)^{-1}(v_3)) = v_3.$$

Ainsi,

$$g^{-1}((g \circ f)(u)) = f(u) = g^{-1}(v_3)$$

et

$$f^{-1}(f(u)) = u = f^{-1}(g^{-1}(v_3)) = (f^{-1} \circ g^{-1})(v_3).$$

□

9.4 Matrice d'une application linéaire

Soient V et W des espaces vectoriels avec $n = \dim(V)$ et $m = \dim(W)$. Soit $B = \{u_1, \dots u_n\}$ une base de V et $B' = \{v_1, \dots, v_m\}$ une base de W. On sait que les n vecteurs $f(u_1), \dots, f(u_n)$ déterminent l'application linéaire f. Notons $[f(u_i)]_{B'}$, les coordonnées de $f(u_i)$ dans la base B'. C'est un vecteur colonne de taille m. Considérons alors la matrice $m \times n$ suivante :

$$A = \begin{pmatrix} [f(u_1)]_{B'} & \dots & [f(u_n)]_{B'} \end{pmatrix}$$

Cette matrice est notée

$$[f]_{B',B}$$

et est appelée la matrice de f par rapport aux bases B et B'. Si $V = W$ et $B = B'$, on note $[f]_B$ au lieu de $[f]_{B,B'}$.

Théorème 9.26. *Soient V_1, V_2 et V_3 des espaces vectoriels de dimensions finies munis des bases B_1, B_2 et B_3 respectivement. Soient $f_1 : V_1 \longrightarrow V_2$ et $f_2 : V_2 \longrightarrow V_3$ des applications linéaires. Alors*

$$[f_2 \circ f_1]_{B_3,B_1} = [f_2]_{B_3,B_2}[f_1]_{B_2,B_1}.$$

Démonstration. Posons $n_1 = \dim(V_1)$, $n_2 = \dim(V_2)$, $n_3 = \dim(V_3)$ et

$$\begin{aligned} B_1 &= \{e_1, \dots, e_{n_1}\} \\ B_2 &= \{b_1, \dots, b_{n_2}\} \\ B_3 &= \{\beta_1, \dots, \beta_{n_3}\}. \end{aligned}$$

Ecrivons

$$[f_2]_{B_3,B_2} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n_2} \\ \vdots & & \vdots \\ \lambda_{n_31} & \cdots & \lambda_{n_3n_2} \end{pmatrix} \qquad \text{et} \qquad [f_1]_{B_2,B_1} = \begin{pmatrix} \mu_{11} & \cdots & \mu_{1n_1} \\ \vdots & & \vdots \\ \mu_{n_21} & \cdots & \mu_{n_2n_1} \end{pmatrix}.$$

On a $(f_2 \circ f_1)(e_1) = f_2(f_1(e_1)) = f_2(\mu_{11}b_1 + \cdots + \mu_{n_21}b_{n_2}) = \mu_{11}f_2(b_1) + \cdots + \mu_{n_21}f_2(b_{n_2})$ $= \mu_{11}[\lambda_{11}\beta_1 + \cdots + \lambda_{n_31}\beta_{n_3}] + \cdots + \mu_{n_21}[\lambda_{1n_2}\beta_1 + \cdots + \lambda_{n_3n_2}\beta_{n_3}]$.
Ainsi, la première colonne de $[f_2 \circ f_1]_{B_3,B_1}$ est

$$\begin{pmatrix} \mu_{11}\lambda_{11} + \cdots + \mu_{n_21}\lambda_{1n_2} \\ \mu_{11}\lambda_{21} + \cdots + \mu_{n_21}\lambda_{2n_2} \\ \vdots \\ \mu_{11}\lambda_{n_31} + \cdots + \mu_{n_21}\lambda_{n_3n_2} \end{pmatrix}.$$

Mais ceci est aussi la première colonne de la matrice $[f_2]_{B_3,B_2}[f_1]_{B_2,B_1}$. En faisant la même chose avec les autres colonnes, on remarque que $[f_2]_{B_3,B_2}[f_1]_{B_2,B_1}$ et $[f_2 \circ f_1]_{B_3,B_1}$ sont deux matrices ayant leurs colonnes égales. On a donc bien l'égalité cherchée. □

Théorème 9.27. *Soit $f : V \longrightarrow V$ une application linéaire, et soit B une base de V. Alors les conditions suivantes sont équivalentes :*

(a) l'application f est inversible ;

(b) La matrice $[f]_B$ est inversible.

Démonstration. Si f est inversible, alors, par le théorème précédent, $[f]_B[f^{-1}]_B = [f \circ f^{-1}]_B = [I]_B$. Ainsi, $[f]_B$ est inversible et son inverse est $[f^{-1}]_B$.
Réciproquement, si $[f]_B$ est inversible, il existe une matrice $[g]_B$ telle que $[f]_B[g]_B = I = [g]_B[f]_B$. Par le théorème ci-dessus, $[f \circ g]_B = I = [g \circ f]_B$, et les applications $f \circ g$ et $g \circ f$ sont toutes deux égales à l'identité. Donc f est bien inversible. □

Théorème 9.28. *Soient V un espace vectoriel de dimension finie, B et B' des bases de V et $P_{B,B'}$ la matrice de changement de base de B' à B. Soit $f : V \longrightarrow V$ une application linéaire. Alors*

$$[f]_{B'} = P_{B,B'}^{-1}[f]_B P_{B,B'}.$$

Démonstration. Soit $x \in V$. Il faut montrer que $P^{-1}_{B,B'}[f]_B P_{B,B'}\,(v)_{B'} = (f(v))_{B'}$. Rappelons que $P^{-1}_{B,B'} = P_{B',B}$ et que $P_{B,B'}\,(v)_{B'} = (v)_B$. On a alors

$$\begin{aligned}
P^{-1}_{B,B'}[f]_B P_{B,B'}\,(v)_{B'} &= P_{B',B}[f]_B\,(v)_B \\
&= P_{B',B}\,([f]_B(v)_B) \\
&= P_{B',B}\,(f(v))_B \\
&= (f(v))_{B'}.
\end{aligned}$$

□

Chapitre 10

Applications multilinéaires et tenseurs

La théorie des tenseurs offre un langage mathématique simple et efficace pour décrire des phénomènes naturels, comme la trajectoire d'un satellite, la circulation de la chaleur dans un corps ou encore la focre électromagnétique d'un électron. L'objectif du présent chapitre est d'introduire la notion de tenseur ainsi que certaines propriétés associées. Celles-ci permettent, par exemple, d'expliquer les relations entre des quantités physiques et de prédire leur évolution future.

Dans toute la suite, on considère un espace vectoriel V.

Nous commençons ce chapitre en introduisant, dans le premier paragraphe, les tenseurs d'ordre $(0,1)$ et ceux d'ordre $(1,0)$: ce sont des formes linéaires. Dans les paragraphes suivants, nous introduirons les tenseurs d'ordre $(0,m)$ (dits aussi *m fois covariants*) puis les tenseurs d'ordre $(m,0)$ (dits *m fois contravariants*) avant de définir les tenseurs mixtes (p,q) : ce sont des formes multilinéaires. L'étude du comportement des tenseurs lors de changements de bases permettra d'expliquer les terminologies *covariant* et *contravariant* (paragraphe 10.6). Ce chapitre clôturera avec un paragraphe sur les champs tensoriels.

10.1 Formes linéaires

10.1.1 Formes linéaires sur V : tenseurs d'ordre (0,1)

Définition 10.1. On appelle *forme linéaire* ou *tenseur d'ordre* $(0,1)$ ou encore *tenseur covariant d'ordre* 1 sur V toute application linéaire :

$$f : V \longrightarrow \mathbb{R}.$$

Exemple 10.2.

1. Si $V = \mathbb{R}^n$, l'application de i-ième projection $\pi_i : (x_1, ..., x_n) \mapsto x_i$ est une forme linéaire sur V.
2. Si V est l'espace vectoriel des matrices carrées de taille n, alors l'application trace $M \mapsto \mathrm{Tr}(M)$ définit une application linéaire sur V.
3. Si V est un espace vectoriel muni d'un produit scalaire généralisé $<,>$ et si v_0 un élément de V, alors l'application $f : V \to \mathbb{R}$ définie par $f(v) =< v, v_0 >$ est une forme linéaire sur V.

10.1.2 Espace dual, bases duales

Définition 10.3. On appelle *espace dual* de V l'ensemble, noté V^*, des formes linéaires sur V.

On a donc :
$$V^* = \{f : V \longrightarrow \mathbb{R} \parallel f \ \ \textit{est linéaire}\ \}.$$

Théorème 10.4. *L'espace dual V^* est muni d'une structure d'espace vectoriel.*

Démonstration. On vérifie facilement que l'addition des formes linéaires et leur multiplication par un scalaire satisfont aux propriétés des espaces vectoriels. □

Si V est un espace vectoriel de dimension finie (voir Chapitre 6 pour la notion de dimension), alors V^* est aussi de dimension finie, et leurs dimensions sont égales. Plus précisément, chaque base de V détermine une base de V^* de la façon suivante :

Théorème 10.5. *On suppose que l'espace V est de dimension finie, avec $dim(V) = n$.*

Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de V. Soit $\mathcal{B}^*$ la famille $(\alpha_1, ..., \alpha_n)$ de V^* où les $\alpha_i : V \to \mathbb{R}$ sont des applications linéaires sur V données par :

$$\alpha_i(e_j) = \delta_{ij}$$

et étendues linéairement à V.

Alors, la famille $\mathcal{B}^*$ forme une base de l'espace dual V^*, appelée base duale de la base $\mathcal{B}$.

En particulier, on a bien :
$$dim(V) = dim(V^*).$$

Rappelons que le symbole δ_{ij} désigne le symbole de Kronecker, c'est-à-dire la fonction qui vaut 1 si $i = j$ et 0 sinon.

Démonstration. Soit f une forme linéaire sur V. Puisque toute forme linéaire sur V est déterminée par ses valeurs prises sur les e_i, on a :
$$f = \sum_{i=1}^{n} f_i \alpha_i$$
avec $f_i = f(e_i)$. Ceci montre que les α_i engendrent l'espace dual V^*.

D'autre part, on vérifie facilement que les α_i sont linéairement indépendantes. L'ensemble $\mathcal{B}^*$ forme donc une base de l'espace V^*.

□

Remarque 10.6. Si f est une application linéaire sur V, alors, avec les notations précédentes, on a $f = \sum_i f_i \alpha_i$ et les coordonnées f_i sont appelées les *coordonnées covariantes* de f.

Exemple 10.7. Considérons la base suivante de $V = \mathbb{R}^2$:

$$\mathcal{B} = \{e_1 = (2, 1)\ ,\ e_2 = (3, 1)\}.$$

Nous allons déterminer la base $\mathcal{B}^*$ de V^* qui est duale de $\mathcal{B}$, c'est-à-dire identifier les formes linéaires sur $\mathbb{R}^2$, notées α_1 et α_2, telles que :

$$\alpha_1(e_1) = 1,\ \alpha_1(e_2) = 0,\ \text{ et }\ \alpha_2(e_1) = 0,\ \alpha_2(e_2) = 0.$$

On écrit ces applications sous la forme :

$$\alpha_1(x, y) = ax + by \ \text{ et }\ \alpha_2(x, y) = cx + dy.$$

Il s'agit donc de déterminer les réels a, b, c, d.

Les relations précédentes impliquent :

$$\begin{cases} \alpha_1(e_1) = \alpha_1(2,1) = 2a + b = 1 \\ \alpha_1(e_2) = \alpha_1(3,1) = 3a + b = 0 \end{cases}$$

d'où $a = -1$ et $b = 3$. Puis :

$$\begin{cases} \alpha_2(e_1) = \alpha_2(2,1) = 2c + d = 0 \\ \alpha_2(e_2) = \alpha_2(3,1) = 3c + d = 1 \end{cases}$$

d'où $c = 1$ et $d = -2$.

La base duale est donc :

$$\mathcal{B}^* = \{\alpha_1 : (x,y) \mapsto -x + 3y \,;\, \alpha_2 : (x,y) \mapsto x - 2y\}.$$

Le théorème 10.5 implique un résultat plus fort :

Théorème 10.8. *Si V est un espace de dimension finie, alors les espaces V et V^* sont isomorphes.*

Démonstration. Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de V et soit $\mathcal{B}^* = (\alpha_1, ..., \alpha_n)$ la base duale de V^* correspondante, donnée par le théorème 10.5. D'après ce théorème, il est évident que l'application :

$$\begin{array}{cccc} \Theta_{\mathcal{B}} : & V & \longrightarrow & V^* \\ & e_i & \mapsto & \alpha_i \end{array}$$

est un isomorphisme entre V et V^*.
Notons que cet isomorphisme dépend de la base $\mathcal{B}$ choisie : en ce sens, il n'est pas canonique. □

Il est intéressant de noter les relations suivantes entre une base $\mathcal{B} = (e_1, ..., e_n)$ de V et sa base duale $\mathcal{B}^* = (\alpha_1, ..., \alpha_n)$ de V^*. Ces relations expriment les coordonnées de tout vecteur v (resp. forme linéaire f) dans la base $\mathcal{B}$ (resp. $\mathcal{B}^*$) :

$$\begin{array}{lcl} \forall v \in V, \quad u & = & \alpha_1(v)e_1 + \alpha_2(v)e_2 + ... + \alpha_n(v)e_n \,; \\ \forall f \in V^*, \quad f & = & f(e_1)\alpha_1 + f(e_2)\alpha_2 + ... + f(e_n)\alpha_n. \end{array}$$

10.1.3 Formes linéaires sur V^* : tenseurs d'ordre $(1,0)$

L'espace dual V^* de V étant encore un espace vectoriel, on peut définir son espace dual. Celui-ci est noté V^{**} : c'est l'espace vectoriel formé de toutes les formes linéaires sur V^*. Si l'on applique deux fois le théorème 10.8, on voit clairement que les espaces V et V^{**} sont isomorphes, lorsque V est de dimension finie. En fait, on a mieux : l'isomorphisme est cette fois canonique, dans le sens qu'il ne dépend pas de la base choisie sur V. C'est ce que nous montrons dans le théorème qui suit.

Avant cela, il est utile de faire la remarque suivante : si v est un vecteur de V, l'application :

$$\widehat{v} := f \mapsto f(v)$$

est une forme linéaire sur V^*, c'est-à-dire un élément de V^{**}. On définit ainsi une application de l'espace V dans son bidual V^{**}, donnée par :

$$\begin{array}{cccc} \iota : & V & \longrightarrow & V^{**} \\ & v & \mapsto & \widehat{v}. \end{array}$$

Le théorème est le suivant :

Théorème 10.9. *L'application ι est linéaire et injective. De plus, lorsque l'espace V est de dimension finie, ι est un isomorphisme.*

Démonstration. Pour montrer la linéarité, il faut montrer que $\iota(k \cdot v + w) = k \cdot \iota(v) + \iota(w)$ pour tout $k \in \mathbb{R}$ et pour tout $v, w \in V$. Il faut donc montrer que $\widehat{k \cdot v + w} = k \cdot \widehat{v} + \widehat{w}$. Ce qui est vrai car

$$(\widehat{k \cdot v + w})(f) = f(k \cdot v + w) = k \cdot f(v) + f(w) = k \cdot \widehat{v}(f) + \widehat{w}(f)$$

puisque f est linéaire.

Supposons maintenant que $\iota(v) = 0$, c'est-à-dire que $\widehat{v}(f) = 0$ pour tout $f \in V^*$. Ceci implique que $f(v) = 0$ pour tout $f \in V^*$ et donc que $v = 0$. Ceci démontre bien l'injectivité de ι.

Enfin, si l'on calcule la base duale de $\mathcal{B}^*$, on obtient une base de V^{**}, notée :

$$\mathcal{B}^{**} = \{\widehat{e_1}, \widehat{e_2}, \ldots, \widehat{e_n}\}$$

et l'isomorphisme (cf. notations de la preuve du théorème 10.8)

$$V \xrightarrow{\theta_{\mathcal{B}}} V^* \xrightarrow{\theta_{\mathcal{B}^*}} V^{**}$$

envoie e_i sur $\widehat{e_i}$. Cet isomorphisme n'est donc rien d'autre que ι (en particulier, ι ne dépend plus de $\mathcal{B}$). Ceci prouve la dernière assertion du théorème. □

Dans la suite, nous supposerons l'espace V de dimension finie et nous identifierons souvent les espaces V et V^{**}. Les vecteurs de V sont appelés tenseurs d'ordre $(1, 0)$, ou encore tenseurs *contravariants* d'ordre 1.

Avec cette identification, remarquons que si $(\alpha_1, ..., \alpha_n)$ est une base de V^*, duale d'une base $(e_1, ..., e_n)$ de V, alors $(e_1, ..., e_n)$ est aussi une base de V^{**}, duale de la base $(\alpha_1, ..., \alpha_n)$.

10.2 Formes multilinéaires sur V : tenseurs d'ordre $(0, m)$

10.2.1 Formes bilinéaires sur V : tenseurs d'ordre $(0, 2)$

Définition 10.10 (Forme bilinéaire)**.** Une application

$$f : V \times V \longrightarrow \mathbb{R}$$

est dite *bilinéaire* si elle est linéaire en chacun des facteurs, autrement dit, si :

(i) $f(ku + v, w) = kf(u, w) + f(v, w)$ et si

(ii) $f(u, kv + w) = kf(u, v) + f(u, w)$.

L'ensemble des formes bilinéaires sur V est un espace vectoriel où l'addition et la multiplication par un scalaire sont données par

$$(f + g)(u, v) = f(u, v) + g(u, v)$$

et

$$(\lambda f)(u, v) = \lambda f(u, v).$$

Cet espace vectoriel est noté $\mathrm{Bil}(V, \mathbb{R})$.

Si $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ est une base de V, on peut considérer les formes bilinéaires particulières

$$\alpha_i \otimes \alpha_j \qquad\qquad 1 \leq i \leq n,\ 1 \leq j \leq n$$

définies par

$$(\alpha_i \otimes \alpha_j)(e_k, e_l) = \delta_{ik} \cdot \delta_{jl}.$$

L'ensemble des $\alpha_i \otimes \alpha_j$ forme alors une base de $\mathrm{Bil}(V, \mathbb{R})$. Les coordonnées f_{ij} d'une forme bilinéaire f quelconque dans cette base sont simplement données par l'évaluation de f au couple (e_i, e_j). Explicitement, on a

$$f_{ij} = f(e_i, e_j)$$

et

$$f = \sum_{i,j} f_{ij} \cdot (\alpha_i \otimes \alpha_j).$$

Ainsi, si une base $\mathcal{B}$ de V est choisie, une forme bilinéaire sur V n'est rien d'autre que la donnée d'une matrice $F = (f_{ij})$ de dimension $n \times n$ où $n = \dim V$. Si $[u]_{\mathcal{B}}$ (resp. $[v]_{\mathcal{B}}$) désigne le vecteur des coordonnées de u (resp. v) dans la base $\mathcal{B}$, alors $f(u, v)$ se calcule par le produit matriciel suivant :

$$f(u, v) = [u]_{\mathcal{B}}^T \cdot F \cdot [v]_{\mathcal{B}}.$$

Une forme bilinéaire sur V est aussi appelée un *tenseur d'ordre* $(0, 2)$ ou encore un *tenseur 2 fois covariant*. Cette terminologie provient du comportement de la forme lors d'un changement de base que nous verrons dans la section 10.6.

Remarque 10.11. La notation $\otimes$ ci-dessus peut être prise ici pour une simple notation.

Exemple 10.12. Soit $V = \mathbb{R}^n$. Alors le produit scalaire usuel est une forme bilinéaire (symétrique). Sa matrice dans la base canonique est I_n.

Un produit scalaire généralisé est une forme bilinéaire (symétrique). Il est représenté par une matrice A symétrique et définie positive.

10.2.2 Tenseurs d'ordre $(0, m)$

Définition 10.13 (Forme multilinéaire)**.** Soit V un espace vectoriel, et soit

$$f : \underbrace{V \times \cdots \times V}_{m} \longrightarrow \mathbb{R}$$

une application. On dit que f est une forme multilinéaire ou tenseur d'ordre $(0, m)$ si pour tout $1 \leq i \leq m$, on a :

$$f(v_1, \ldots, \alpha v_i + \beta w_i, \ldots, v_m) = \alpha f(v_1, \ldots, v_i, \ldots, v_m) + \beta f(v_1, \ldots, w_i, \ldots, v_m)$$

pour tout $v_1, \ldots, v_m, w_1, \ldots, w_m \in V$ et tout $\alpha, \beta \in \mathbb{R}$. Une telle application est appelée tenseur d'ordre $(0, m)$, ou tenseur *covariant* d'ordre m. Pour $m = 1$, on retrouve les formes linéaires sur V, et pour $m = 2$ les formes bilinéaires sur V.

Exemple 10.14 (Produit mixte)**.** Le produit mixte

$$\begin{gathered} \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R} \\ f(x, y, z) = [x, y, z] = x \cdot (y \times z) \end{gathered}$$

est un tenseur d'ordre $(0, 3)$.

10.2.3 Quelques interprétations physiques

Avant de poursuivre, voici quelques interpétations physiques des tenseurs d'ordre $(0, m)$ que nous venons de rencontrer.

1. Les tenseurs d'ordre $(0, 0)$ sont les applications linéaires $\mathbb{R} \to \mathbb{R}$. Ces dernières étant précisément de la forme $x \mapsto ax$ pour un scalaire $a \in \mathbb{R}$, les tenseurs d'ordre $(0, 0)$ peuvent donc être naturellement identifiés aux scalaires de $\mathbb{R}$. En physique, ils représentent des quantités chiffrées telles la masse d'un satellite, la température d'un corps ou la charge d'un électron.
2. Les tenseurs d'ordre $(0, 1)$ sont les formes linéaires $V \to \mathbb{R}$. Un exemple de tel tenseur est donné par l'étude de la trajectoire d'un bateau à voile, qui se dirige selon un vecteur unitaire e. Considérons la force du vent exercée sur ce bateau : elle est représentée par un vecteur f perpendiculaire à la voile du bateau. Seule la composante $e\dot{f}$ (produit scalaire) propulse le bateau à l'avant. L'application $f \mapsto e\dot{f}$ est un tenseur d'ordre $(0, 1)$. Le bateau avancera d'autant plus vite que la quantité $e\dot{f}$ est grande.
3. Les tenseurs d'ordre $(0, 2)$ sont les formes bilinéaires $V \times V \to \mathbb{R}$. Le calcul du moment d'une force appliquée sur une clef plate pour visser un boulon fournit un exemple.

10.3 Formes multilinéaires sur V^* : tenseurs d'ordre $(m, 0)$

10.3.1 Une remarque sur les tenseurs d'ordre $(1, 0)$

Nous avons vu à la section 10.1.3 que V et V^{**} sont isomorphes si V est de dimension finie (ce que nous supposons toujours ici). Ainsi un vecteur $v \in V$ peut être vu comme une forme linéaire $\widehat{v}$ sur V^* :

$$\begin{aligned} \widehat{v} : V^* &\longrightarrow \mathbb{R} \\ f &\mapsto f(v). \end{aligned}$$

C'est pourquoi, on identifie les tenseurs d'ordre $(1, 0)$ avec les vecteurs. En physique, la direction que suit la trajectoire d'un électron à un moment donné est représentée par un vecteur et fournit donc un exemple de tenseur 1 fois contravariant.

10.3.2 Formes bilinéaires sur V^* : tenseurs d'ordre $(2, 0)$

Considérons un espace vectoriel V muni d'une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ et son dual V^* muni de la base duale $\mathcal{B}^* = \{\alpha_1, \ldots \alpha_n\}$.

On se donne une forme bilinéaire

$$f : V^* \times V^* \longrightarrow \mathbb{R}.$$

Cette forme est entièrement déterminée par ses valeurs sur les couples (α_i, α_j).

Rappelons que l'espace vectoriel V^{**} des formes linéaires sur V^* est isomorphe canoniquement à V et admet pour base l'ensemble :

$$\mathcal{B}^{**} = \{\widehat{e_1}, \ldots, \widehat{e_n}\}$$

où les $\widehat{e_i}$ sont donnés par :

$$\widehat{e_i}(\alpha_j) = \alpha_j(e_i) = \delta_{ij}.$$

Notons $\widehat{e_i} \otimes \widehat{e_j}$ la forme bilinéaire sur $V^* \times V^*$ définie par

$$\widehat{e_i} \otimes \widehat{e_j}(\alpha_k, \alpha_l) = \delta_{ik} \cdot \delta_{jl}.$$

On montre de manière similaire à ce qui a été fait précédemment que l'ensemble

$$\mathcal{C} = \{\widehat{e_i} \otimes \widehat{e_j}\}_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

est une base de l'espace vectoriel des formes bilinéaires sur $V^* \times V^*$.

Dans cette base, la forme bilinéaire f s'écrit

$$f = \sum_i \sum_j f_{ij} \cdot (\widehat{e_i} \otimes \widehat{e_j})$$

où $f_{ij} = f(\alpha_i, \alpha_j)$. Le choix de $\mathcal{B}$ étant fait (et donc aussi celui de $\mathcal{B}^*$ et de $\mathcal{C}$), la forme f est représentée par une matrice $F = (f_{ij})$.

Les formes bilinéaires sur $V^* \times V^*$ sont appelées tenseurs d'ordre $(2, 0)$, ou encore tenseurs *contravariants* d'ordre 2.

10.3.3 Tenseurs d'ordre $(m, 0)$

En généralisant ceci, on peut poser la définition suivante :

Définition 10.15 (Tenseur d'ordre $(m, 0)$)**.** Une application multilinéaire

$$\underbrace{V^* \times V^* \times \cdots \times V^*}_{m} \longrightarrow \mathbb{R}$$

est appelée un *tenseur d'ordre* $(m, 0)$ ou *tenseur m fois contravariant* ou encore tenseur contravariant d'ordre m.

À nouveau, la terminologie de tenseur contravariant vient du comportement d'une telle application lors d'un changement de bases (voir paragraphe 10.6).

10.4 Tenseurs mixtes d'ordre (p, q)

10.4.1 Tenseurs d'ordre (p, q)

Il existe des tenseurs mixtes, à savoir des tenseurs d'ordre (p, q) qui sont donc (en copiant les terminologies précédentes) p fois contravariants et q fois covariants.

Définition 10.16 (Tenseur d'ordre (p, q))**.** Soient p et q des entiers ≥ 0. Un tenseur d'ordre (p, q) est une application multilinéaire :

$$f : \underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q} \longrightarrow \mathbb{R}.$$

10.4.2 Exemple des tenseurs d'ordre $(1, 1)$

Pour illustrer la notion de tenseur d'ordre $(1, 1)$, considérons une application bilinéaire :

$$f : V^* \times V \longrightarrow \mathbb{R}.$$

Nous allons montrer qu'une telle application n'est rien d'autre qu'une application linéaire :

$$\phi : V \longrightarrow V.$$

On commence par la construction inverse. Etant donnée une application linéaire :

$$\phi : V \longrightarrow V,$$

on peut définir une application bilinéaire :

$$f : V^* \times V \longrightarrow \mathbb{R}$$

en posant :

$$f(\alpha, u) = \widehat{\phi(u)}(\alpha) = \alpha(\phi(u)).$$

On vérifie facilement que f est bilinéaire et que l'association $\phi \mapsto f$ est injective. Vérifions ici l'injectivité. Supposons que f soit l'application triviale $f(\alpha, u) = 0$ pour tout $u \in V$ et pour tout $\alpha \in V^*$. Par définition de f, ceci implique que $\alpha(\phi(u)) = 0$ pour tout u et tout α. Mais ceci entraéne que $\phi(u)$ doit être nul pour tout $u \in V$ et donc que ϕ est l'application nulle.

Pour des raisons de dimensions, l'injectivité implique la surjectivité et on a donc montré que donner une application bilinéaire :

$$f : V^* \times V \longrightarrow \mathbb{R}$$

est équivalent à donner une application linéaire :

$$\phi : V \longrightarrow V.$$

Ce qui est remarquable, c'est que les matrices de f et de ϕ sont les mêmes, une base $\mathcal{B}$ de V étant choisie. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de V et $\mathcal{B}^* = \{\alpha_1, \ldots, \alpha_n\}$ sa base duale. La matrice de f dans les bases $\mathcal{B}^*$ et $\mathcal{B}$ est définie par $F_{\mathcal{B}^*,\mathcal{B}} = F = (f_{ij})$ avec

$$f_{ij} = f(e_i, \alpha_j).$$

Mais, d'autre part, la j-ème colonne de la matrice $[\phi]_{\mathcal{B}}$ est donnée par l'image de e_j exprimée dans la base $\mathcal{B}$. Ainsi, on a

$$[\phi]_{ij} = \alpha_i(\phi(e_j)) = f(\alpha_i, e_j) = f_{ij}$$

ce qui prouve que

$$[\phi]_{\mathcal{B}} = F_{\mathcal{B}^*,\mathcal{B}}.$$

10.5 Opérations sur les tenseurs

Sur les tenseurs de même ordre, on a une opération d'addition et de multiplication par les scalaires qui en font un espace vectoriel.

Nous avons déjà vu cette structure d'espace vectoriel pour les tenseurs d'ordre $(0,1)$ et d'ordre $(0,2)$, c'est-à-dire sur V^* et sur l'espace des formes bilinéaires. La généralisation à tout tenseur se fait de manière évidente.

On a aussi une opération de multiplication entre tenseurs. Nous commençons par l'exemple du produit de deux formes linéaires :

Exemple 10.17. Soient $f, g \in V^*$ deux tenseurs d'ordre $(0,1)$,

$$\begin{array}{rcl} f & : & V \longrightarrow \mathbb{R} \\ g & : & V \longrightarrow \mathbb{R}\,. \end{array}$$

on définit leur produit par

$$b : V \times V \longrightarrow \mathbb{R}$$

$$b(u,v) = f(u)g(v).$$

C'est une forme bilinéaire, c'est-à-dire un tenseur d'ordre $(0,2)$.

Plus généralement, on peut multiplier un tenseur d'ordre (p,q) avec un tenseur d'ordre (r,s) et le résultat est un tenseur d'ordre $(p+r, q+s)$. Voici comment le produit est défini. Soit :

$$f : \underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q} \longrightarrow \mathbb{R}$$

un tenseur d'ordre (p,q) et soit :

$$g : \underbrace{V^* \times \cdots \times V^*}_{r} \times \underbrace{V \times \cdots \times V}_{s} \longrightarrow \mathbb{R}$$

un tenseur d'ordre (r,s). On définit leur produit :

$$f \cdot g : \underbrace{V^* \times \cdots \times V^*}_{p+r} \times \underbrace{V \times \cdots \times V}_{q+s} \longrightarrow \mathbb{R}$$

par :

$$(f \cdot g)(a_1, \ldots, a_{p+r}, x_1, \ldots x_{q+s}) = f(a_1, \ldots, a_p, x_1, \ldots x_q) g(a_{p+1}, \ldots a_{p+r}, x_{q+1}, \ldots x_{q+ss}).$$

Remarque 10.18. L'application identité $\mathrm{Id} : \mathbb{R} \to \mathbb{R}$ peut être considérée comme un tenseur d'ordre $(0,0)$. Il joue alors le rôle d'élément neutre pour le produit défini ci-dessus.

10.6 Changement de bases

Dans ce paragraphe, nous allons étudier le comportement d'un tenseur (ou plutôt de ses coordonnées) lors d'un changement de bases. Il s'agit ici d'introduire des propriétés importantes des tenseurs liées aux changements de coordonnées, très utiles par exemple en physique lorsque l'on change de système référent. D'autre part, les relations obtenues nous permettront d'éclairer les notions de tenseurs "covariants" et "contravariants".

Soient donc deux bases de V :

$$\mathcal{B} = \{e_1, \ldots, e_n\} \qquad \text{et} \qquad \mathcal{B}' = \{e'_1, \ldots, e'_n\}.$$

D'après le paragraphe 6.7, la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$ est la matrice :

$$P_{\mathcal{B}\mathcal{B}'} = (\lambda_{ij})$$

dont la j-ème colonne repésente les coordonnées de e'_j dans la base $\mathcal{B}$. En d'autres termes

$$\boxed{e'_j = \sum_{i=1}^{n} \lambda_{ij} e_i} \tag{10.1}$$

ou encore, matriciellement (avec un petit abus de notation)

$$\begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{pmatrix} = P^T_{\mathcal{B}\mathcal{B}'} \cdot \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}. \tag{10.2}$$

D'après le théorème 6.71, la matrice $P_{\mathcal{B}\mathcal{B}'}$ est inversible, et l'on a :

$$P^{-1}_{\mathcal{B}\mathcal{B}'} = P_{\mathcal{B}'\mathcal{B}}.$$

Dans la suite, nous noterons γ_{ij} les coefficients de la matrice $P_{\mathcal{B}'\mathcal{B}}$, de sorte que l'on a :

$$e_i = \sum_{k=1}^{n} e'_k.$$

Le but de cette section est de d'étudier le changement de coordonnées des tenseurs.

10.6.1 Cas des tenseurs d'ordre $(1, 0)$ (vecteurs de V)

La modification des coordonnées d'un vecteur par un changement de bases a déjà été étudiée au chapitre 6, paragraphe 6.7). Rappelons que si v est un vecteur de V et si $[v]_{\mathcal{B}}$ (resp. $[v]_{\mathcal{B}'}$) désigne le vecteur colonne dont les composantes sont les coordonnées de v dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$), on a :

$$\boxed{[v]_{\mathcal{B}} = P_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}'}}, \tag{10.3}$$

ou encore :

$$\boxed{[v]_{\mathcal{B}'} = P^{-1}_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}}}, \tag{10.4}$$

chaque v_i (resp. v'_i) désignant la i-ème composante de v dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$).

10.6.2 Cas des tenseurs d'ordre $(0, 1)$ (formes linéaires sur V)

Considérons maintenant une forme linéaire $f \in V^*$. Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de V et soit $\mathcal{B}^* = (\alpha_1, ..., \alpha_n)$ la base de V^*, duale de V.

Matriciellement, en écrivant les vecteurs de V dans la base $\mathcal{B}$, l'application f est de la forme :

$$f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \begin{pmatrix} f_1 & \dots & f_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Le vecteur ligne $[f]_{\mathcal{B}^*} = \begin{pmatrix} f_1 & \dots & f_n \end{pmatrix}$ est la matrice de f dans la base $\mathcal{B}^*$, donnée par $f_i = f(e_i)$. En d'autres termes, on a :

$$f = \sum_{i=1}^{n} f_i \alpha_i.$$

De même, si $\mathcal{B}' = (e'1, \dots, e'_n)$ est une autre base de V, de base duale $\mathcal{B}'^* = (\alpha'_1, \dots, \alpha' n)$, et si on note $[f]_{\mathcal{B}'^*}$ la matrice de f dans cette base, ses coefficients sont donnés par :

$$f'_i = f(e'_i),$$

de sorte que :

$$f = \sum_{i=1}^{n} f'_i \alpha'_i.$$

Il s'agit d'écrire la matrice de passage de la base $\mathcal{B}'^*$ á la base $\mathcal{B}^*$ et de voir les modifications sur les coefficients de la matrice de f.

Pour cela, déterminons d'abord les coefficients, notés μ_{ij} de la matrice de passage $P_{\mathcal{B}^*\mathcal{B}'^*}$, dont la j-ème colonne est donnée par les coefficients de l'application α'_j dans la base $\mathcal{B}^*$.

Des relations $\alpha'_j = \sum_i \mu_{ij} \alpha_i$ et $\alpha_i(e_j) = \delta_{ij}$, on déduit :

$$\alpha'_j(e_i) = \mu_{ij},$$

c'est-à-dire :

$$\begin{array}{rcl} \mu_{ij} & = & \alpha'_j(\sum_k \gamma_{ki} e'_k) \\ & = & \gamma_{ji}, \end{array}$$

où les γ_{ij} sont les coefficients de la matrice $P_{\mathcal{B}'\mathcal{B}} = P^{-1}_{\mathcal{B}\mathcal{B}'}$.

On a donc prouvé :

$$\boxed{P_{\mathcal{B}^*\mathcal{B}'^*} = (P^{-1}_{\mathcal{B}\mathcal{B}'})^T.} \tag{10.5}$$

On peut maintenant calculer les coordonnées de f dans la base $\mathcal{B}'^*$, i.e. trouver les f'_j tels que

$$f = \sum_j f'_j \alpha'_j.$$

Cela s'écrit facilement, en utilisant la relation $\alpha'_i(e'_j) = \delta_{ij}$:

$$\begin{array}{rcl} f'_j & = & f(e'_j) \\ & = & \sum_i \lambda_{ij} f(e_i) \\ & = & \sum_i \lambda_{ij} f(e_i) \\ & = & \sum_i \lambda_{ij} f_i. \end{array}$$

On a donc :

$$\boxed{[f]_{\mathcal{B}'^*} = P^T_{\mathcal{B}\mathcal{B}'} \cdot [f]_{\mathcal{B}^*}}. \tag{10.6}$$

On constate que les formes linéaires sur V se transforment « **comme** » les vecteurs de base tandis que les vecteurs se transforment selon la règle **inverse** (comparer (10.6) avec (10.2) et (10.4)).

C'est pour cette raison que les vecteurs (ou les formes linéaires sur V^*) sont appelés des tenseurs *contravariants* d'ordre 1 et les formes linéaires sur V sont appelées des tenseurs *covariants* d'ordre 1.

Le qualificatif "**contr**avariant" concernerait donc les tenseurs dont les composantes se transforment **contr**airement à celles des vecteurs de base. Poursuivons notre investigation...

10.6.3 Cas des tenseurs d'ordre $(0,2)$ (formes bilinéaires sur V)

On considère une forme bilinéaire f sur V et sa matrice F (resp. F') dans la base $\mathcal{B}$ (resp, $\mathcal{B}'$). On peut montrer que l'on a la relation

$$\boxed{F' = P_{\mathcal{B}\mathcal{B}'}^T \cdot F \cdot P_{\mathcal{B}\mathcal{B}'}.} \tag{10.7}$$

Cette équation est du même type que la relation (10.6). Les formes bilinéaires suivent les mêmes règles que les formes linéaires lors d'un changement de bases. C'est pourquoi, on dit aussi que les formes bilinéaires sur V sont des tenseurs *covariants* d'ordre 2.

10.6.4 Cas des tenseurs $(2,0)$ (formes bilinéaires sur V^*)

On considère un tenseur d'ordre $(2,0)$, c'est-à-dire une forme bilinéaire sur V^*. Notons F (resp. F') sa matrice dans la base $\mathcal{B}^*$ (resp, $\mathcal{B}'^*$). On a la relation :

$$\boxed{F' = P_{\mathcal{B}\mathcal{B}'}^{-1} \cdot F \cdot (P_{\mathcal{B}\mathcal{B}'}^{-1})^T.} \tag{10.8}$$

On constate que c'est la règle inverse que pour un tenseur d'ordre $(0,2)$ (comparer avec 10.7). En revanche, c'est une loi similaire à celle qui régit le changement de coordonnées d'un vecteur. Un tenseur d'ordre $(2,0)$ est donc dit 2 fois *contravariant.*

10.6.5 Cas des tenseurs d'ordre $(1,1)$

Lors d'un changement de bases, on sait comment la matrice d'une application $\phi : V \longrightarrow V$ change. Si Φ (resp. Φ') est la matrice de ϕ dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$), alors on a

$$\Phi' = P_{\mathcal{B}\mathcal{B}'}^{-1} \cdot \Phi \cdot P_{\mathcal{B}\mathcal{B}'}.$$

Via l'équivalence vue au paragraphe 10.4.2, et tenant compte du fait que $\Phi = F$, on obtient la manière dont les coefficients d'un tenseur d'ordre $(1,1)$ varient lors d'un changement de bases. C'est la même règle que pour une application linéaire. Plus précisément, si

$$f : V^* \times V \longrightarrow \mathbb{R}$$

est une application bilinéaire de matrice F (resp. F') dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$), alors on a

$$\boxed{F' = P_{\mathcal{B}\mathcal{B}'}^{-1} \cdot F \cdot P_{\mathcal{B}\mathcal{B}'}}. \tag{10.9}$$

On constate que la matrice de changement de base apparaît à droite de F mais que c'est son inverse qui apparait à gauche. C'est pourquoi, un tenseur d'ordre $(1,1)$ est dit 1 fois covariant et 1 fois contravariant (comparer la relation (10.9) avec l'équation (10.8) qui donne la règle pour un tenseur 2 fois covariant.)

10.6.6 Cas des tenseurs d'ordre (p, q)

Les formules précédentes se généralisent encore... et expliquent pourquoi un tenseur d'ordre (p, q) est dit *p fois contravariant et q fois covariant.*

10.7 Champs tensoriels

10.7.1 Définitions

Jusqu'ici, nous n'avons considéré que des tenseurs isolés. En physique, il est plus fréquent de rencontrer un champ tensoriel, c'est-à-dire, la donnée d'un tenseur en tout point de l'espace (ou d'une partie de celui-ci) ou en plusieurs instants.

On considère donc l'espace affine $E = \mathbb{R}^n$ et en tout point y de l'espace on considère l'espace vectoriel $V = \mathbb{R}^n$.

Définition 10.19 (Champ tensoriel)**.** Un *champ tensoriel* est la donnée en tout point y de l'espace d'un tenseur d'ordre (p, q). Alors que le tenseur varie en fonction du point $y \in E$, l'ordre (p, q), quant à lui, est indépendant de y.

Exemple 10.20. Considérons un fluide localisé dans une partie de l'espace E. La vitesse du fluide en chaque point de l'espace est un tenseur d'ordre $(0, 1)$.

En chaque point $y \in \mathbb{R}^3$, on se donne une base :

$$\mathcal{B}(y) = \{e_1(y), \dots e_n(y)\}$$

dans laquelle on exprimera le champ tensoriel.

Par exemple, on pourrait considérer la base canonique en chaque point y : dans ce cas, la base ne dépend pas de la position y.

10.7.2 Changements de coordonnées

Supposons que l'on se donne un changement de coordonnées, i.e. en termes mathématiques, une fonction :

$$\begin{array}{cccc} \Phi : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & (x_1, \dots, x_n) & \mapsto & \left\{ \begin{array}{c} u_1 = u_1(x_1, \dots, x_n) \\ \vdots \\ u_n = u_n(x_1, \dots, x_n) \end{array} \right. \end{array}$$

qui soit de classe C^1, bijective et dont l'inverse est aussi de classe C^1.

L'application Φ transforme, en chaque point y, la base $\mathcal{B}(y)$ en une nouvelle base

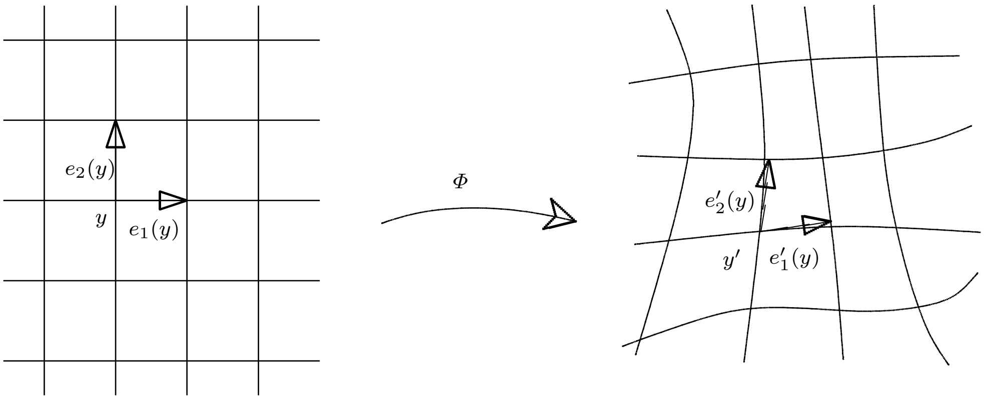
$$\mathcal{B}'(y) = \{e'_1(y), \dots, e'_n(y)\}.$$

Le dessin ci-dessous illustre la situation dans le cas $E = \mathbb{R}^2$.

Explicitement, en chaque point y de l'espace $\mathbb{R}^n$, on a un changement de bases qui est donné par

$$e'_i(y) = \sum_j \frac{\partial u_j}{\partial x_i}(y) e_j(y).$$

Notons la dépendance en y dans la formule ci-dessus.



Définissons la matrice :

$$\Lambda = (\lambda_{ij}) := \left(\frac{\partial u_i}{\partial x_j}(y)\right).$$

C'est la matrice de changement de base au point y. Son inverse est donné par

$$\Lambda^{-1} = \Gamma = (\gamma_{ij}) = \left(\frac{\partial x_i}{\partial u_j}(y)\right).$$

Ces deux matrices dépendent du point y.

10.7.3 Cas d'un champ tensoriel d'ordre $(1, 0)$ (champ vectoriel)

Admettons que l'on ait maintenant un champ vectoriel, i.e. la donnée d'un vecteur v de $\mathbb{R}^n$ en tout point y de l'espace $\mathbb{R}^n$. Un champ vectoriel (ou champ tensoriel d'ordre $(1, 0)$) est donc la donnée d'une application :

$$\begin{aligned} E = \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ y &\mapsto v(y) \end{aligned}$$

qui à tout point y associe un vecteur $v(y)$.

Notons $v_i(y)$ (resp. $v'_i(y)$) les coordonnées de $v(y)$ dans la base $\mathcal{B}(y)$ (resp. $\mathcal{B}'(y)$). On montre alors qu'on a la relation suivante :

$$v'_i(y) = \sum_j \frac{\partial u_i}{\partial x_j}(y) v_j(y)$$

ou matriciellement

$$[v'] = \Gamma \cdot [v].$$

Cette règle est similaire à celle obtenue en (10.3), avec comme seule différence que la matrice de changement de bases dépend du point y et est donnée par la matrice jacobienne $\Gamma = \Lambda^{-1}$.

On parle dans ce cas de champ *contravariant*.

10.7.4 Cas d'un champ tensoriel d'ordre $(0, 1)$

Soit w un champ de formes linéaires, i.e. la donnée en tout point y de l'espace d'une forme linéaire

$$w(y) : \mathbb{R}^n \longrightarrow \mathbb{R}.$$

On peut, en tout point y, considérer la base duale de $\mathcal{B}(y)$ que l'on note $\mathcal{B}^*(y)$. Comme précédemment, si

$$\mathcal{B}^*(y) = \{\alpha_i(y)\}$$

alors chaque forme linéaire $w(y)$ s'écrit

$$w(y) = \sum_i w_i(y) \cdot \alpha_i(y).$$

Le changement de coordonnées donné par Φ induit un changements des coordonnées $w_i(y)$ qui suit la règle suivante :

$$w'_i(y) = \sum_j \frac{\partial x_j}{\partial u_i}(y) w_j(y)$$

qui s'écrit matricellement

$$[w'] = \Lambda^T \cdot [w]$$

ou encore

$$\boxed{[w] = (\Lambda^{-1})^T \cdot [w']} \tag{10.10}$$

Ce résultat n'est pas surprenant. L'équation (10.10) est semblable à la relation (10.6), la matrice $P_{\mathcal{B}\mathcal{B}'}$ ayant été remplaçée par la matrice Λ.

On parle dans ce cas d'un champ *covariant*.

10.7.5 Cas d'un champ quelconque

On peut généraliser ce qui précède au cas d'un champ tensoriel d'ordre (p, q) quelconque. Les changements de coordonnées se font de la même manière que pour un tenseur d'ordre (p, q) à la différence près, à nouveau, que la matrice de changement de bases dépend du point y (et de Φ). On remplace ainsi la matrice

$$P_{\mathcal{B}\mathcal{B}'}$$

par la matrice

$$\Lambda = \left(\frac{\partial u_i}{\partial x_j} \right).$$

Bibliographie

[1] H. Anton and C. Rorres. *Elementary linear Algebra.* Applications Version. John Wiley & Sons, Inc. New York, 2000.

[2] R. Dalang and A. Caabouni. *Algèbre linéaire, Aide-mémoire, exercices et applications.*

[3] D. Danielson. *Vectors and Tensors in Enginee and Physics.* Addison-Wesley Publishing Company, 1992.

[4] T. Liebling. *Algèbre linéaire : une introduction pour ingénieurs.*

Index

Index des notations