LMAT 1131

ALGEBRE LINEAIRE

E.M. Vitale

École de Mathématique - Faculté des Sciences

Institut de Recherche en Mathématique et Physique

Université catholique de Louvain

“Mais à quoi sert de faire des maths si on ne peut pas compter les uns sur les autres.”
Youssoupha, *L'amour*, Bomaye Musik, 2012.

“Il n'y a jamais mauvais élève, seulement mauvais enseignant.”
Jackie Chan, *The Karate Kid*, Columbia Pictures, 2010.

“Va savoir pourquoi une descente vue d'en bas ressemble tellement à une montée.”
Goofy, *Le super-héros est fatigué*, Disney Studio (foreign market stories), 1969.

Table des matières

Chapitre 1

Introduction

Le premier objectif du cours est d'étudier les *systèmes d'équations linéaires.* Par exemple

$$\mathcal{S}\colon \left\{ \begin{array}{rcl} 3x - 2y + \frac{1}{4}z & = & 5 \\ -\frac{1}{2}x + y + \frac{1}{\sqrt{2}}z & = & 0 \end{array} \right. \qquad \underline{1.1}$$

est un système de deux équations linéaires en trois variables x, y et z.

1.1 Remarque. Le mot *linéaire* signifie que dans chaque équation la partie qui contient les variables est un polynôme de premier degré. Dans le système 1.1 c'est bien le cas : $3x-2y+\frac{1}{4}z$ et $-\frac{1}{2}x+y+\frac{1}{\sqrt{2}}z$ sont deux polynômes de premier degré par rapport aux variables x, y et z. Par contre, l'équation $3e^x + 2y = 5$ n'est pas linéaire à cause du terme exponentielle e^x, et l'équation $3xy + 2y = 5$ n'est pas linéaire à cause du terme xy, qui est du second degré.

En face d'un système $\mathcal{S}$ comme le 1.1 on peut se poser les questions suivantes :

1. Le système $\mathcal{S}$ admet-il des solutions ?
2. Si oui, combien de solutions admet-il ?
3. Si non, peut-on trouver des solutions approchées ?

Pour poser ces questions correctement et y répondre en toute généralité, nous introduirons des *structures algébriques* :

1. les sous-espaces vectoriels et les espaces vectoriels,
2. les applications linéaires,
3. les corps commutatifs,
4. les espaces euclidiens,
5. les formes quadratiques.

L'objectif principal du cours sera donc d'étudier ces structures algébriques qui (avec d'autres, comme les groupes, les anneaux, les catégories, . . .) reviennent dans plusieurs cours de mathématique et de physique.

Chapitre 2

Combinaisons linéaires

Le chapitre 2 section par section

Dans ce chapitre nous allons introduire les opérations sur les vecteurs de $\mathbb{R}^n$ dont on a besoin pour étudier l'ensemble des solutions d'un système d'équations linéaires.

1. Un système d'équations linéaires à coefficients réels peut être impossible (pas de solutions), déterminé (une solution unique) ou indéterminé (une infinité de solutions).
2. Une solution d'un système à n variables est un vecteur de $\mathbb{R}^n$, c'est-à-dire une n-uple $\vec{x} = (x_1, x_2, \ldots, x_n)$ de nombres réels.
3. On peut effectuer la somme $\vec{x} + \vec{y}$ de vecteurs et le produit $\alpha \cdot \vec{x}$ d'un scalaire par un vecteur. Ces deux opérations permettent de construire les combinaisons linéaires $\alpha \cdot \vec{x} + \beta \cdot \vec{y}$.

2.1 Systèmes d'équations linéaires

Un système d'équations linéaires est donné par un nombre fini d'équations linéaires, chaque équation linéaire contient un nombre fini de variables. Par exemple

$$\mathcal{S}\colon \left\{ \begin{array}{rcl} 3x - 2y + \frac{1}{4}z & = & 5 \\ -\frac{1}{2}x + y + \frac{1}{\sqrt{2}}z & = & 0 \end{array} \right. \qquad \underline{2.1}$$

est un système de deux équations linéaires en trois variables x, y et z. Pour pouvoir travailler avec plus de variables, il convient de remplacer

$$x\ y\ z \quad \text{par} \quad x_1\ x_2\ x_3\,.$$

De cette façon le système $\underline{2.1}$ s'écrit comme

$$\mathcal{S}\colon \left\{ \begin{array}{rcl} 3x_1 - 2x_2 + \frac{1}{4}x_3 & = & 5 \\ -\frac{1}{2}x_1 + x_2 + \frac{1}{\sqrt{2}}x_3 & = & 0 \end{array} \right. \qquad \underline{2.2}$$

Voici d'autres exemples :

$$\mathcal{S}\colon \begin{cases} 3x_1 + x_2 - x_3 &= 0 \\ 4x_2 + 7x_3 &= 1 \\ -x_1 - x_2 + x_3 &= \frac{1}{2} \end{cases} \quad \text{trois équations en trois variables}$$

$$\mathcal{S}\colon \left\{ -x_1 + \tfrac{2}{3}x_2 = \tfrac{3}{5} \right. \quad \text{une équation en deux variables}$$

2.1 Notation. La forme générale d'un système linéaire est :

$$\mathcal{S}\colon \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots\ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{cases} \qquad \underline{2.3}$$

qui est un système avec m équations linéaires en n variables.

Dans le système 2.3 :

- les $x_1, x_2, \ldots, x_n$ sont les *variables*,
- les $a_{11}, a_{12}, \ldots, a_{mn}$ sont les *coefficients* réels,
- les $b_1, b_2, \ldots, b_m$ sont les *termes indépendants*.

Si nous notons par $\mathbb{R}$ l'ensemble des nombres réels, nous avons :

- $a_{ij} \in \mathbb{R} \ \ (i = 1, \ldots, m, \ \ j = 1, \ldots, n)$
- $b_i \in \mathbb{R} \ \ (i = 1, \ldots, m)$

2.2 Définition. Une *solution* d'un système $\mathcal{S}$ à n variables est un n-*uple* $(x_1, x_2, \ldots, x_n)$ de nombres réels qui est solution de chaque équation contenue dans le système $\mathcal{S}$.

2.3 Exemple. Le triplet $(0, 0, 0)$ n'est pas solution du système 2.2 car il est solution de la deuxième équation (si on remplace x_1 par 0, x_2 par 0 et x_3 par 0 dans $-\frac{1}{2}x_1 + x_2 + \frac{1}{\sqrt{2}}x_3 = 0$, on obtient $0 = 0$) mais pas de la première (si on remplace x_1 par 0, x_2 par 0 et x_3 par 0 dans $3x_1 - 2x_2 + \frac{1}{4}x_3 = 5$, on obtient $0 = 5$).
Par contre le triplet $(\frac{5}{2}, \frac{5}{4}, 0)$ est solution du système 2.2 car il est solution tant de la première que de la deuxième équation.

2.4 Remarque. Trois cas peuvent se présenter pour un système $\mathcal{S}$ d'équations linéaires :

1. Le système $\mathcal{S}$ n'a pas de solutions. On dit que $\mathcal{S}$ est un système *impossible* (ou sur-déterminé).
2. Le système $\mathcal{S}$ admet une solution unique. On dit que $\mathcal{S}$ est un système *déterminé*.
3. Le système $\mathcal{S}$ admet une infinité de solutions. On dit que $\mathcal{S}$ est un système *indéterminé* (ou sous-déterminé).

Pour trouver facilement un exemple de chaque type, on peut utiliser un peu de géométrie : l'ensemble des points du plan dont les coordonnées (x_1, x_2) sont solution d'une équation linéaire en deux variables

$$a_{11}x_1 + a_{12}x_2 = b_1$$

est une droite. Dès lors, résoudre un système d'équations linéaires en deux variables revient à chercher les points d'intersection entre les droites qui correspondent aux différentes équations.

2.5 Exemple.

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

   Les deux droites sont parallèles et distinctes, elles n'ont pas de points d'intersection. Le système est donc impossible.

2. $$\left\{\begin{array}{rcl} x_1 - x_2 & = & 0 \\ x_1 + 2x_2 & = & 6 \end{array}\right.$$

   Les deux droites sont sécantes, elles ont un seul point d'intersection qui est le point de coordonnées $(2, 2)$. Le système est déterminé, l'unique solution est le couple $(x_1 = 2, x_2 = 2)$.

3. $$\left\{\begin{array}{rcl} x_1 - x_2 & = & 3 \\ -5x_1 + 5x_2 & = & -15 \end{array}\right.$$

   On a deux fois la même droite et donc tout point de la droite est point d'intersection. Le système est indéterminé.

2.6 Remarque. Attention ! Il ne suffit pas de comparer le nombre d'équations avec le nombre de variables pour savoir si le système est impossible, déterminé ou indéterminé. Un système avec plus d'équations que de variables peut être indéterminé :

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

et un système avec plus de variables que d'équations peut être impossible :

$$\left\{\begin{array}{rcl} x_1 + x_2 + x_3 & = & 0 \\ x_1 + x_2 + x_3 & = & 1 \end{array}\right.$$

Ce qu'il faut faire c'est comparer le nombre de variables avec le nombre d'équations *linéairement indépendantes.* Ceci sera expliqué dans les prochains chapitres.

2.7 Remarque. La Remarque 2.4 est valable si on cherche des solutions réelles (ou rationnelles, ou complexes, ...) mais elle n'est pas valable si on utilise d'autres ensembles numériques. On reviendra sur ce point dans le Chapitre 12.

2.2 Les vecteurs de $\mathbb{R}^n$

Si $\mathcal{S}$ est un système de m équations linéaires en n variables, on a dit qu'une solution de $\mathcal{S}$ est un n-uple de nombres réels. Dans la suite, plutôt que n-uple de nombres réels nous dirons *vecteur de* $\mathbb{R}^n$ et nous utiliserons les notations suivantes :

$$\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \ , \ \ x_i \in \mathbb{R} \ \ (i = 1, \ldots, n) \, .$$

Autrement dit :

$$\mathbb{R}^n = \{\vec{x} = (x_1, x_2, \ldots, x_n) \mid \ x_i \in \mathbb{R} \ \ (i = 1, \ldots, n)\} \, .$$

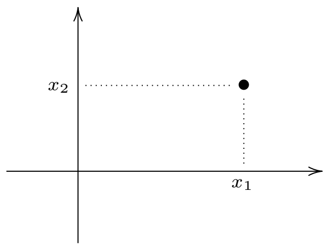
Si $\vec{x} = (x_1, x_2, \ldots, x_n)$, on appelle x_i la i-ème *composante* du vecteur $\vec{x}$.

Par exemple :

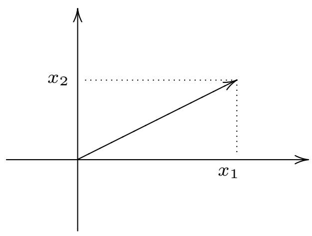
- Si $n = 1$ on a simplement $\mathbb{R}^1 = \mathbb{R}$.
- Si $n = 2$ on a $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$.

2.8 Remarque. On peut représenter un vecteur $\vec{x} = (x_1, x_2)$ de $\mathbb{R}^2$ (ou $\vec{x} = (x_1, x_2, x_3)$ de $\mathbb{R}^3$) de deux façons différentes :

1. $\vec{x}$ est un point de coordonnées (x_1, x_2):



2. $\vec{x}$ est une flèche de composantes (x_1, x_2):



2.9 Notation. Si $\mathcal{S}$ est un système de m équations linéaires en n variables, l'ensemble de ses solutions est un sous-ensemble de $\mathbb{R}^n$. On va le noter $\mathrm{Sol}(\mathcal{S})$:

$$\mathrm{Sol}(\mathcal{S}) \subseteq \mathbb{R}^n \, .$$

2.3 Opérations sur les vecteurs de $\mathbb{R}^n$

Pour pouvoir étudier Sol($\mathcal{S}$) il faut d'abord apprendre à effectuer des calculs avec les vecteurs de $\mathbb{R}^n$.

2.10 Définition. Soit $\vec{x} = (x_1, x_2, \ldots, x_n)$ et $\vec{y} = (y_1, y_2, \ldots, y_n)$ deux vecteurs de $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$.

1. Somme de deux vecteurs : on pose
$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \qquad \vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$
2. Produit d'un scalaire par un vecteur : on pose
$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n \qquad \lambda \cdot \vec{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

Voici les propriétés qu'on peut utiliser en effectuant des calculs avec les vecteurs :

2.11 Propriété. *Pour tout $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ et pour tout $\alpha, \beta \in \mathbb{R}$, on a :*

1. $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$
2. $\vec{x} + \vec{0} = \vec{x} = \vec{0} + \vec{x}$, *où* $\vec{0} = (0, 0, \ldots, 0) \in \mathbb{R}^n$
3. $\vec{x} + \vec{y} = \vec{y} + \vec{x}$
4. $\vec{x} + (-\vec{x}) = \vec{0}$, *où* $-\vec{x} = -1 \cdot \vec{x}$
5. $(\alpha + \beta) \cdot \vec{x} = \alpha \cdot \vec{x} + \beta \cdot \vec{x}$
6. $\alpha \cdot (\vec{x} + \vec{y}) = \alpha \cdot \vec{x} + \alpha \cdot \vec{y}$
7. $\alpha \cdot (\beta \cdot \vec{x}) = (\alpha\beta) \cdot \vec{x}$
8. $1 \cdot \vec{x} = \vec{x}$.

Preuve. La preuve de cette propriété est un exercice facile. □

2.12 Remarque.

1. On peut résumer la Propriété 2.11 en disant que $\mathbb{R}^n$ est un *espace vectoriel réel* par rapport aux opérations définies dans la Définition 2.10. Dans le Chapitre 6 on reviendra sur la définition générale d'espace vectoriel.
2. Les propriétés associatives 1. et 7. nous permettent de simplifier l'usage des parenthèses : on écrira simplement $\vec{x} + \vec{y} + \vec{z}$ et $\alpha\beta\vec{x}$.
3. De même, on écrira $\vec{x} - \vec{y}$ en lieu de $\vec{x} + (-\vec{y})$.

En utilisant les deux opérations $\vec{x} + \vec{y}$ et $\alpha \cdot \vec{x}$ nous pouvons créer des combinaisons linéaires :

2.13 Définition. Une *combinaison linéaire* de $\vec{x}, \vec{y} \in \mathbb{R}^n$ est un vecteur de type
$$\alpha \cdot \vec{x} + \beta \cdot \vec{y} \quad \text{avec} \quad \alpha, \beta \in \mathbb{R}\,.$$
Plus généralement, une combinaison linéaire de $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in \mathbb{R}^n$ est un vecteur de type
$$\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2 + \ldots \alpha_k \cdot \vec{x}_k \quad \text{avec} \quad \alpha_i \in \mathbb{R} \;\; (i = 1, \ldots, k)$$
Les scalaires α_i sont appelés les coefficients de la combinaison linéaire.

2.14 Remarque. Dans la définition de combinaison linéaire (2.13) il faut disposer de k vecteurs $\vec{x}_1, \ldots, \vec{x}_k$, avec $k \geq 1$. Pour compléter cette définition, nous pouvons considérer que la combinaison linéaire vide, c'est-à-dire la combinaison linéaire de zéro vecteurs, est le vecteur nul $\vec{0}$.

Chapitre 3

Sous-espaces vectoriels

Le chapitre 3 section par section

Dans ce chapitre nous introduisons les sous-espaces vectoriels de $\mathbb{R}^n$, ainsi que la notion de base et de dimension d'un sous-espace vectoriel.

1. L'exemple principal de sous-espace vectoriel est donné par l'ensemble des solutions d'un système homogène.
2. Un sous-espace vectoriel de $\mathbb{R}^n$ est un sous-ensemble de $\mathbb{R}^n$ qui est fermé par combinaisons linéaires.
3. Une base d'un sous-espace vectoriel V est une famille génératrice minimale de V, ou, de façon équivalente, une famille génératrice de V formée par des vecteurs linéairement indépendants.
4. Deux bases d'un même sous-espace vectoriel contiennent le même nombre de vecteurs. Ce nombre est la dimension du sous-espace vectoriel.
5. Si on connait la dimension d'un sous-espace vectoriel, la recherche d'une base devient plus facile car il suffit de vérifier que la famille est libre ou qu'elle est génératrice.

3.1 Systèmes homogènes

Un *système homogène* est un système où tous les termes indépendants sont nuls :

$$\mathcal{S}_0 : \left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0 \\ \ldots\ldots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0 \end{array} \right. \qquad \underline{3.1}$$

Dans le Chapitre 2 on a introduit les opérations $\vec{x} + \vec{y}$ et $\alpha \cdot \vec{x}$ afin de pouvoir énoncer la propriété suivante :

3.1 Propriété. *Soit $\mathcal{S}_0$ un système homogène en n variables et soit*

$$\mathrm{Sol}(\mathcal{S}_0) \subseteq \mathbb{R}^n$$

l'ensemble de ses solutions. On a :

1. $\vec{0} \in \mathrm{Sol}(\mathcal{S}_0)$
2. *Si* $\vec{x}, \vec{y} \in \mathrm{Sol}(\mathcal{S}_0)$ *alors* $\vec{x} + \vec{y} \in \mathrm{Sol}(\mathcal{S}_0)$
3. *Si* $\vec{x} \in \mathrm{Sol}(\mathcal{S}_0)$ *et* $\alpha \in \mathbb{R}$ *alors* $\alpha \cdot \vec{x} \in \mathrm{Sol}(\mathcal{S}_0)$.

Preuve. Voir plus loin (Corollaire 7.14). (Cette preuve, bien que facile, est fastidieuse à faire à l'état actuel de nos connaissances. Comme beaucoup d'autres preuves, elle sera vue plus loin, quand on disposera d'outils efficaces comme le calcul matriciel et les applications linéaires.) □

3.2 Remarque.

1. le point 1. de la Propriété 3.1 dit qu'un système homogène admet toujours au moins la solution triviale $\vec{0} = (0, 0, \ldots, 0)$.
2. Les points 2. et 3. permettent, une fois obtenues des solutions $\vec{x}, \vec{y}$ de $\mathcal{S}_0$, d'en construire d'autres par combinaison linéaire.
3. La Propriété 3.1 n'est pas valable si le système n'est pas homogène.

3.2 Sous-espaces vectoriels

On peut extraire de la Propriété 3.1 la définition suivante :

3.3 Définition. Soit V un sous-ensemble de $\mathbb{R}^n$. On dit que V est un *sous-espace vectoriel de* $\mathbb{R}^n$ s'il remplit les conditions suivantes :

1. $\vec{0} \in V$
2. Si $\vec{x}, \vec{y} \in V$ alors $\vec{x} + \vec{y} \in V$
3. Si $\vec{x} \in V$ et $\alpha \in \mathbb{R}$ alors $\alpha \cdot \vec{x} \in V$.

Un s.e.v. de $\mathbb{R}^n$ est donc un sous-ensemble de $\mathbb{R}^n$ qui est fermé dans $\mathbb{R}^n$ par combinaisons linéaires (en particulier, un tel sous-ensemble de $\mathbb{R}^n$ contient nécessairement le vecteur nul $\vec{0}$). En utilisant la Définition 3.3 on peut exprimer la Propriété 3.1 en disant que pour tout système homogène $\mathcal{S}_0$ en n variables, l'ensemble des solutions $\mathrm{Sol}(\mathcal{S}_0)$ est un s.e.v. de $\mathbb{R}^n$. Cette propriété admet une réciproque, qui est notre premier *théorème de représentation* :

3.4 Propriété. *Pour tout s.e.v.* V *de* $\mathbb{R}^n$ *on peut trouver un système homogène* $\mathcal{S}_0$ *en* n *variables tel que* $V = \mathrm{Sol}(\mathcal{S}_0)$.

Preuve. Voir plus loin (Exercice 4.19). □

Les propriétés 3.1 et 3.4 montrent le lien entre la notion de s.e.v. et la résolution d'un système homogène. On va donc approfondir l'étude des s.e.v. par des exercices.

3.5 Exercice.

1. Soit V, W deux s.e.v. de $\mathbb{R}^n$. Montrer que $V \cap W$ est le plus grand s.e.v. de $\mathbb{R}^n$ contenu dans V et dans W.
2. Soit V, W deux s.e.v. de $\mathbb{R}^n$. On pose :

$$V + W = \{\vec{v} + \vec{w} \mid \vec{v} \in V\ ,\ \vec{w} \in W\}$$

   Montrer que $V + W$ est le plus petit s.e.v. de $\mathbb{R}^n$ qui contient V et W.
3. Soit $\{\vec{x}_i\}_{i \in I}$ une famille de vecteurs de $\mathbb{R}^n$. On dénote par

$$\langle \vec{x}_i \mid i \in I \rangle$$

   l'ensemble de toutes les combinaisons linéaires des vecteurs $\vec{x}_i$. Montrer que $\langle \vec{x}_i \mid i \in I \rangle$ est le plus petit s.e.v. de $\mathbb{R}^n$ qui contient les vecteurs $\vec{x}_i$. Le s.e.v. $\langle \vec{x}_i \mid i \in I \rangle$ est appelé le *s.e.v. engendré* par les vecteurs $\vec{x}_i$.
4. Soit V, W deux s.e.v. de $\mathbb{R}^n$. Montrer que $V \cup W$ est un s.e.v. de $\mathbb{R}^n$ si et seulement si l'un des deux s.e.v. est contenu dans l'autre.

D'après la Propriété 3.1 un système homogène $\mathcal{S}_0$ admet toujours au moins une solution, la solution triviale $\vec{x} = \vec{0}$. Le cas plus intéressant est si $\mathcal{S}_0$ admet une infinité de solutions. Le problème qu'on se pose est alors :

> Peut-on extraire de $\mathrm{Sol}(\mathcal{S}_0)$ un nombre fini de solutions de sorte que toute autre solution de $\mathcal{S}_0$ soit combinaison linéaire des solutions séléctionnées ?

Voyons cela d'abord sur un exemple facile.

3.6 Exemple. Considérons le système homogène

$$\mathcal{S}_0 \colon 2x_1 - x_2 = 0 \quad \text{une équation en deux variables}$$

$\mathcal{S}_0$ admet une infinité de solutions qu'on peut écrire sous la forme

$$(x_1, 2x_1) \quad \text{avec } x_1 \in \mathbb{R} \text{ arbitraire}$$

Mais $(x_1, 2x_1) = x_1 \cdot (1, 2)$. Si on pose $\vec{v} = (1, 2)$ on a alors

$$\mathrm{Sol}(\mathcal{S}_0) = \{(x_1, 2x_1) \mid x_1 \in \mathbb{R}\} = \{x_1 \cdot (1, 2) \mid x_1 \in \mathbb{R}\} = \langle \vec{v}\, \rangle$$

Autrement dit : les solutions de $\mathcal{S}_0$ sont exactement les multiples du vecteur $\vec{v}$.

3.3 Familles libres, familles génératrices, bases

Introduisons ici la notion de base d'un sous-espace vectoriel, d'abord comme famille génératrice minimale et ensuite comme famille libre et génératrice.

3.7 Définition. Soit V un s.e.v. de $\mathbb{R}^n$ et soit $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$. Les vecteurs $\vec{v}_1, \ldots, \vec{v}_k$ forment une *famille génératrice* de V si :

1. $\vec{v}_1, \ldots, \vec{v}_k \in V$,

2. on peut écrire tout vecteur de V comme combinaison linéaire de $\vec{v}_1, \dots, \vec{v}_k$:

$$\forall\, \vec{v} \in V\ \exists\, \alpha_1, \dots, \alpha_k \in \mathbb{R} \quad \text{tels que} \quad \vec{v} = \alpha_1 \cdot \vec{v}_1 + \dots + \alpha_k \cdot \vec{v}_k\,.$$

3.8 Définition. Soit V un s.e.v. de $\mathbb{R}^n$ et soit $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$. Les vecteurs $\vec{v}_1, \dots, \vec{v}_k$ forment une *base* de V s'ils forment une famille génératrice *minimale.* Ici minimale signifie que si on retire même un seul des vecteurs $\vec{v}_1, \dots, \vec{v}_k$, les vecteurs qui restent ne sont plus une famille génératrice de V. Autrement dit :

1. $\vec{v}_1, \dots, \vec{v}_k \in V$,
2. $\forall\, \vec{v} \in V\ \exists\, \alpha_1, \dots, \alpha_k \in \mathbb{R}$ tels que $\vec{v} = \alpha_1 \cdot \vec{v}_1 + \dots + \alpha_k \cdot \vec{v}_k$,
3. $\forall\ i = 1, \dots, k$ la famille $\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k$ n'est pas une famille génératrice de V.

3.9 Exemple. Les n vecteurs de $\mathbb{R}^n$

$$\vec{e}_1 = (1, 0, \dots, 0)\,,\ \vec{e}_2 = (0, 1, \dots, 0)\,, \dots, \vec{e}_n = (0, \dots, 1)$$

forment une base de $\mathbb{R}^n$ qu'on appelle *base canonique.*

Preuve. 1. famille génératrice :

$$\vec{x} = (x_1, x_2, \dots, x_n) = x_1 \cdot \vec{e}_1 + \dots + x_n \cdot \vec{e}_n$$

2. famille génératrice minimale : retirons un des $\vec{e}_i$, pour fixer les idées retirons $\vec{e}_1$. Les vecteurs $\vec{e}_2, \dots, \vec{e}_n$ ne forment pas une famille génératrice de $\mathbb{R}^n$ car toute combinaison linéaire de $\vec{e}_2, \dots, \vec{e}_n$ aura 0 à la première composante (et donc par exemple $\vec{e}_1$ lui-mème ne peut pas s'exprimer comme combinaison linéaire de $\vec{e}_2, \dots, \vec{e}_n$. □

La base canonique n'est pas l'unique base de $\mathbb{R}^n$. Par exemple, il est facile de vérifier que les vecteurs $\vec{v}_1 = (1, 1)$ et $\vec{v}_2 = (1, -1)$ forment une base de $\mathbb{R}^2$. On peut donc avoir plusieurs bases pour un même s.e.v. V de $\mathbb{R}^n$, mais elles seront toutes formées par un même nombre de vecteurs (on ne peut pas avoir une base de V formée par 4 vecteurs et une deuxième base de V formée par 5 vecteurs). Avant de montrer cela il convient de retravailler la définition de base.

3.10 Définition. Soit $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$.

1. $\vec{v}_1, \dots, \vec{v}_k$ sont *linéairement dépendants* ou *liés* si au moins l'un d'entre eux est combinaison linéaire des autres.
2. Dans le cas contraire $\vec{v}_1, \dots, \vec{v}_k$ sont *linéairement indépendants* ou *libres.*

3.11 Exercice.

1. Soit $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$. Les condition suivantes sont équivalentes :
   (a) $\vec{v}_1, \dots, \vec{v}_k$ sont linéairement indépendants,
   (b) si $\lambda_1 \cdot \vec{v}_1 + \dots + \lambda_k \cdot \vec{v}_k = \vec{0}$ alors $\lambda_i = 0$ pour tout $i = 1, \dots, k$,

   (c) si $\alpha_1 \cdot \vec{v}_1 + \ldots + \alpha_k \cdot \vec{v}_k = \beta_1 \cdot \vec{v}_1 + \ldots + \beta_k \cdot \vec{v}_k$ alors $\alpha_i = \beta_i$ pour tout $i = 1, \ldots, k$.

2. Si $\vec{v}_1, \ldots, \vec{v}_k$ est une famille libre, alors chaque $\vec{v}_i$ est différent de $\vec{0}$.
   Une famille $\{\vec{v}\}$ qui contient un seul vecteur est libre ssi $\vec{v} \neq \vec{0}$.
   La famille vide (celle qui ne contient aucun vecteur) est libre.

En utilisant la notion de famille libre, nous pouvons reformuler la définition de base d'un sous-espace vectoriel.

3.12 Propriété. *Soit V un s.e.v. de $\mathbb{R}^n$ et soit $\vec{v}_1, \ldots, \vec{v}_k \in V$. Les conditions suivantes sont équivalentes :*

1. *$\vec{v}_1, \ldots, \vec{v}_k$ forment une base de V, c'est-à-dire, une famille génératrice minimale de V,*
2. *$\vec{v}_1, \ldots, \vec{v}_k$ forment une famille libre et génératrice de V,*
3. *$\vec{v}_1, \ldots, \vec{v}_k$ forment une famille libre* maximale *de V, c'est-à-dire, la famille $\vec{v}_1, \ldots, \vec{v}_k$ est libre et pour tout vecteur $\vec{v} \in V$, la famille $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ n'est pas libre,*
4. *tout vecteur $\vec{v} \in V$ s'écrit de façon unique comme combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_k$.*

Preuve. A titre d'exercice, montrons l'équivalence entre la condition 2 et la condition 3.
$2 \Rightarrow 3$: soit $\vec{v} \in V$, nous devons montrer que la famille $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ n'est pas libre. Pour cela, il suffit de remarquer que $\vec{v}$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_k$ car $\vec{v}_1, \ldots, \vec{v}_k$ est une famille génératrice de V.
$3 \Rightarrow 2$: soit $\vec{v} \in V$, nous devons démontrer que $\vec{v}$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_k$. Puisque $\vec{v}_1, \ldots, \vec{v}_k$ est une famille libre maximale, la famille $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ n'est pas libre, c'est-à-dire que l'un de ces vecteurs est combinaison linéaire des autres. Deux cas peuvent se présenter :
- Soit $\vec{v}$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_k$, et dans ce cas la preuve est terminée.
- Soit un des $\vec{v}_i$, disons $\vec{v}_1$ pour fixer les idées, est combinaison linéaire des autres

$$\vec{v}_1 = \alpha_2 \cdot \vec{v}_2 + \ldots \alpha_k \cdot \vec{v}_k + \alpha \cdot \vec{v}$$

Puisque $\alpha \neq 0$ (car $\vec{v}_1, \ldots, \vec{v}_k$ est une famille libre), nous pouvons écrire

$$\vec{v} = \frac{1}{\alpha} \cdot \vec{v}_1 - \frac{\alpha_2}{\alpha} \cdot \vec{v}_2 - \ldots - \frac{\alpha_k}{\alpha} \cdot \vec{v}_k$$

et la preuve est terminée. □

3.4 Dimension

Nous allons démontrer par étapes que deux bases d'un même sous-espace vectoriel contiennent le même nombre de vecteurs. Ce nombre est la dimension du sous-espace vectoriel.

3.13 Lemme. *Soient m et n deux nombres entiers. Si $m > n$ alors m vecteurs de $\mathbb{R}^n$ sont linéairement dépendants.*

Preuve. Soit $\vec{v}_1, \ldots, \vec{v}_n$ une famille génératrice de $\mathbb{R}^n$ (par exemple la base canonique) et soit $\vec{x}_1, \ldots, \vec{x}_{n+1} \in \mathbb{R}^n$.
On peut écrire $\vec{x}_1$ comme combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_n$:

$$\vec{x}_1 = a_{11} \cdot \vec{v}_1 + \ldots + a_{1n} \cdot \vec{v}_n \qquad [3.2]$$

Si tous les coefficients sont nuls, alors $\vec{x}_1 = \vec{0}$ et donc les vecteurs $\vec{x}_1, \ldots, \vec{x}_{n+1}$ sont linéairement dépendants.
Si au moins un des coefficients est non nul, on peut supposer (quitte à changer l'ordre) que $a_{11} \neq 0$. Montrons que $\vec{x}_1, \vec{v}_2, \ldots, \vec{v}_n$ est une famille génératrice de $\mathbb{R}^n$: tout vecteur de $\mathbb{R}^n$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_n$, mais de [3.2] on tire que $\vec{v}_1$ est combinaison linéaire de $\vec{x}_1, \vec{v}_2, \ldots, \vec{v}_n$ et donc tout vecteur de $\mathbb{R}^n$ est combinaison linéaire de $\vec{x}_1, \vec{v}_2, \ldots, \vec{v}_n$.
On peut écrire $\vec{x}_2$ comme combinaison linéaire de $\vec{x}_1, \vec{v}_2, \ldots, \vec{v}_n$:

$$\vec{x}_2 = a_{21} \cdot \vec{x}_1 + a_{22}\vec{v}_2 \ldots + a_{2n} \cdot \vec{v}_n \qquad [3.3]$$

Si les coefficients a_{2i} sont nuls pour $i \geq 2$, alors $\vec{x}_2 = a_{21} \cdot \vec{x}_1$ et donc les vecteurs $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{n+1}$ sont linéairement dépendants.
Si au moins un des coeffients a_{2i} $(i \geq 2)$ est non nul, on peut supposer (quitte à changer l'ordre) que $a_{22} \neq 0$. Montrons que $\vec{x}_1, \vec{x}_2, \vec{v}_3, \ldots, \vec{v}_n$ est une famille génératrice de $\mathbb{R}^n$: tout vecteur de $\mathbb{R}^n$ est combinaison linéaire de $\vec{x}_1, \vec{v}_2, \ldots, \vec{v}_n$, mais de [3.3] on tire que $\vec{v}_2$ est combinaison linéaire de $\vec{x}_1, \vec{x}_2, \vec{v}_3, \ldots, \vec{v}_n$ et donc tout vecteur de $\mathbb{R}^n$ est combinaison linéaire de $\vec{x}_1, \vec{x}_2, \vec{v}_3, \ldots, \vec{v}_n$.
On peut répéter cet argument au plus n fois et on arrive à montrer que $\vec{x}_1, \ldots, \vec{x}_n$ est une famille génératrice de $\mathbb{R}^n$. Par conséquent, $\vec{x}_{n+1}$ est combinaison linéaire de $\vec{x}_1, \ldots, \vec{x}_n$ et donc $\vec{x}_1, \ldots, \vec{x}_n, \vec{x}_{n+1}$ sont linéairement dépendants. □

3.14 Lemme. *Soit V un s.e.v. de $\mathbb{R}^n$. Si $\vec{x}_1, \ldots, \vec{x}_k$ est une famille génératrice de V et $\vec{y}_1, \ldots, \vec{y}_l$ est une famille libre de vecteurs de V, alors $k \geq l$.*
(Autrement dit : si $\vec{x}_1, \ldots, \vec{x}_k$ est une famille génératrice de V, si $\vec{y}_1, \ldots, \vec{y}_l$ sont des vecteurs de V et si $l > k$, alors $\vec{y}_1, \ldots, \vec{y}_l$ sont linéairement dépendants.)

Preuve. On peut écrire chaque $\vec{y}_i$ comme combinaison linéaire des $\vec{x}_i$:

$$\vec{y}_1 = a_{11} \cdot \vec{x}_1 + \ldots + a_{1k} \cdot \vec{x}_k \, , \; \ldots , \; \vec{y}_l = a_{l1} \cdot \vec{x}_1 + \ldots + a_{lk} \cdot \vec{x}_k$$

En utilisant les coefficients des combinaisons linéaires on obtient l vecteurs de $\mathbb{R}^k$:

$$\vec{a}_1 = (a_{11}, \ldots, a_{1k}) \, , \; \ldots , \; \vec{a}_l = (a_{l1}, \ldots, a_{lk})$$

Supposons par l'absurde que $l > k$. Par le Lemme 3.13 les vecteurs $\vec{a}_1, \ldots, \vec{a}_l$ sont linéairement dépendants. On peut alors trouver des coefficients $\alpha_1, \ldots, \alpha_l$ non tous nuls et tels que

$$\vec{0} = \alpha_1 \cdot \vec{a}_1 + \ldots + \alpha_l \cdot \vec{a}_l = (\alpha_1 a_{11} + \ldots + \alpha_l a_{l1}, \ldots, \alpha_1 a_{1k} + \ldots + \alpha_l a_{lk})$$

Par conséquent,

$$\alpha_1 \cdot \vec{y}_1 + \ldots + \alpha_l \cdot \vec{y}_l = \alpha_1 a_{11}\vec{x}_1 + \ldots + \alpha_1 a_{1k}\vec{x}_k + \ldots + \alpha_l a_{l1}\vec{x}_1 + \ldots + \alpha_l a_{lk}\vec{x}_k =$$

$$= (\alpha_1 a_{11} + \ldots + \alpha_l a_{l1})\vec{x}_1 + \ldots + (\alpha_1 a_{1k} + \ldots + \alpha_l a_{lk})\vec{x}_k = 0 \cdot \vec{x}_1 + \ldots + 0 \cdot \vec{x}_k = \vec{0}$$

et donc les vecteurs $\vec{y}_1, \ldots, \vec{y}_l$ sont linéairement dépendants, ce qui est en contradiction avec l'hypothèse. □

3.15 Propriété. *Soit V un s.e.v. de $\mathbb{R}^n$ et soient $\vec{x}_1, \ldots, \vec{x}_k$ et $\vec{y}_1, \ldots, \vec{y}_l$ deux bases de V. Alors $k = l$.*

Preuve. Puisque $\vec{x}_1, \ldots, \vec{x}_k$ est une famille génératrice de V et $\vec{y}_1, \ldots, \vec{y}_l$ est une famille libre de vecteurs de V, par le Lemme 3.14 on a $k \geq l$.
Puisque $\vec{y}_1, \ldots, \vec{y}_l$ est une famille génératrice de V et $\vec{x}_1, \ldots, \vec{x}_k$ est une famille libre de vecteurs de V, par le Lemme 3.14 on a $l \geq k$. □

Grâce à la propriété précédente nous pouvons définir la *dimension* d'un s.e.v. :

3.16 Définition. Soit V un s.e.v. de $\mathbb{R}^n$ et soit $\vec{x}_1, \ldots, \vec{x}_k$ une base de V. On pose $\dim V = k$.

3.17 Remarque.

1. L'Exemple 3.9 montre que $\dim \mathbb{R}^n = n$.
2. Le s.e.v. $\{\vec{0}\,\}$ admet la famille vide comme base. On a donc $\dim\{\vec{0}\,\} = 0$.

3.5 Recherche d'une base

Voyons comment la connaissance de la dimension d'un sous-espace vectoriel peut aider pour trouver une base du sous-espace.

Le prochain exercice est une relecture du Lemme 3.14. En particulier, le premier point montre que la dimension d'un s.e.v. V est le plus petit nombre de vecteurs de V nécessaires pour reconstruire tout vecteur de V par combinaisons linéaires.

3.18 Exercice. Soit V un s.e.v. de $\mathbb{R}^n$ avec $\dim V = k$.

1. Si $\vec{x}_1, \ldots, \vec{x}_l$ est une famille génératrice de V, alors $l \geq k$. De plus, on peut extraire de $\vec{x}_1, \ldots, \vec{x}_l$ une base de V.
2. Si $\vec{x}_1, \ldots, \vec{x}_l$ est une famille libre de vecteurs de V, alors $l \leq k$. De plus, on peut compléter $\vec{x}_1, \ldots, \vec{x}_l$ en une base de V.

Preuve. Le Lemme 3.14 donne $l \geq k$ dans le premier cas et $l \leq k$ dans le deuxième cas.
1. Pour extraire une base de V de $\vec{x}_1, \ldots, \vec{x}_l$ il suffit d'examiner un par un les l vecteurs. On commence par $\vec{x}_l$: si $\vec{x}_l$ est combinaison linéaire des autres on peut le retirer (et $\vec{x}_1, \ldots, \vec{x}_{l-1}$ reste une famille génératrice de V), sinon on le

garde. Et ainsi de suite jusqu'à $\vec{x}_1$.
2. Soit $\vec{y}_1, \dots, \vec{y}_k$ une base de V. Comme dans la preuve du Lemme 3.13 on peut remplacer l des vecteurs de cette base par les vecteurs $\vec{x}_1, \dots, \vec{x}_l$ et (quitte à changer l'ordre) on obtient une nouvelle base $\vec{x}_1, \dots, \vec{x}_l, \vec{y}_{l+1}, \dots, \vec{y}_k$ de V. Attention : dans la preuve du Lemme 3.13 on montre à chaque étape que la nouvelle famille est une famille génératrice. Ici il faut aussi montrer que elle est libre. Voyons par exemple la première étape :

$$\vec{x}_1 = a_{11} \cdot \vec{y}_1 + a_{12} \cdot \vec{y}_2 + \dots + a_{1k} \cdot \vec{y}_k$$

avec $a_{11} \neq 0$. Alors la famille $\vec{x}_1, \vec{y}_2, \dots, \vec{y}_k$ est une famille génératrice de V (comme dans la preuve du Lemme 3.13) mais aussi libre : supposons d'avoir

$$\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{y}_2 + \dots + \alpha_k \cdot \vec{y}_k = \vec{0}$$

Si $\alpha_1 = 0$ il reste $\alpha_2 \cdot \vec{y}_2 + \dots + \alpha_k \cdot \vec{y}_k = \vec{0}$ et donc $\alpha_2 = \dots = \alpha_k = 0$ car $\vec{y}_2, \dots, \vec{y}_k$ sont linéairement indépendants.
Si $\alpha_1 \neq 0$ on a

$$\vec{x}_1 = -\frac{\alpha_2}{\alpha_1} \cdot \vec{y}_2 - \dots - \frac{\alpha_k}{\alpha_1} \cdot \vec{y}_k$$

d'où $a_{11} = 0$ en contradiction avec l'hypothèse. □

A priori pour déterminer si une famille de vecteurs $\vec{x}_1, \dots, \vec{x}_l$ est une base d'un s.e.v. V il faudrait montrer qu'elle est une famille libre et génératrice. Une conséquence intéressante de l'Exercice 3.18 est que si on connait déjà la dimension de V il suffit vérifier seulement une des deux conditions :

3.19 Exercice. Soit V un s.e.v. de $\mathbb{R}^n$ avec $\dim V = k$.

1. Si $\vec{x}_1, \dots, \vec{x}_k$ est une famille génératrice de V, alors elle est une base de V.
2. Si $\vec{x}_1, \dots, \vec{x}_k$ est une famille libre de vecteurs de V, alors elle est une base de V.

3.20 Exercice. Soient V et W deux s.e.v. de $\mathbb{R}^n$.

1. Si $V \subseteq W$, alors $\dim V \leq \dim W$.
2. Si $V \subseteq W$ et $\dim V = \dim W$, alors $V = W$.

On peut resumer ce chapitre en disant que pour connaître un s.e.v. de $\mathbb{R}^n$ il faut en déterminer une base et la dimension. Cela pose un problème : peut-on affirmer que tout s.e.v. de $\mathbb{R}^n$ admet une base ?

3.21 Propriété. *Tout s.e.v. V de $\mathbb{R}^n$ admet une base.*

Preuve. Si $V = \{\vec{0}\}$. la famille vide est une base de V.
Supposons maintenant $V \neq \{\vec{0}\}$. Soit $\vec{x}_1 \in V, \vec{x}_1 \neq \vec{0}$. Si tout vecteur de V est un multiple de $\vec{x}_1$, alors $\vec{x}_1$ est une base de V et la preuve est terminée.
S'il existe un vecteur $\vec{x}_2$ de V qui n'est pas multiple de $\vec{x}_1$, alors $\vec{x}_1, \vec{x}_2$ est une famille libre de vecteurs de V. Si tout vecteur de V est combinaison linéaire de

$\vec{x}_1$ et $\vec{x}_2$, alors $\vec{x}_1, \vec{x}_2$ forment une base de V et la preuve est terminée.
S'il existe un vecteur $\vec{x}_3$ de V qui n'est pas combinaison linéaire de $\vec{x}_1$ et $\vec{x}_2$, alors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ est une famille libre de vecteurs de V. Si tout vecteur de V est combinaison linéaire de $\vec{x}_1, \vec{x}_2$ et $\vec{x}_3$, alors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ forment une base de V et la preuve est terminée.
Et ainsi de suite ...
Ce processus s'arrête après au plus n étapes (car dans $\mathbb{R}^n$ on ne peut pas avoir plus de n vecteurs linéairement indépendants). A la dernière étape on aura donc k vecteurs

$$\vec{x}_1, \dots, \vec{x}_k \in V \quad (k \leq n)$$

linéairement indépendants et tels que tout autre vecteur de V est combinaison linéaire de $\vec{x}_1, \dots, \vec{x}_k$. Cela signifie que $\vec{x}_1, \dots, \vec{x}_k$ forment une base de V. $\square$

Chapitre 4

Rang et méthode de Gauss

Le chapitre 4 section par section

D'après la Propriété 3.21, il est en principe possible de déterminer une base (et donc la dimension) de tout s.e.v. de $\mathbb{R}^n$. La méthode expliquée dans la preuve de la Propriété 3.21 n'est pas efficace en pratique. Ce chapitre est consacré à une autre méthode (dite *méthode de Gauss*) pour déterminer la dimension et une base de l'espace des solutions d'un système homogène.

1. On peut extraire d'un système la matrice des coefficients et la matrice des coefficients complétée par la colonne des termes indépendants.
2. Par des opérations élémentaires, on peut transformer une matrice en une matrice échelonnée. Le système correspondant est équivalent au système de départ, mais plus simple à étudier.
3. L'espace engendré par les lignes et l'espace engendré par les colonnes d'une matrice sont invariants par opérations élémentaires et ils sont de la même dimension.
4. Une fois que la matrice des coefficients est échelonnée, on peut facilement déterminer la dimension et une base de $\mathrm{Sol}(\mathcal{S}_0)$.
5. Si le système $\mathcal{S}$ admet une solution $\vec{x}_0$, alors $\mathrm{Sol}(\mathcal{S}) = \vec{x}_0 + \mathrm{Sol}(\mathcal{S}_0)$.

4.1 Les matrices associées à un système

4.1 Définition. Deux systèmes $\mathcal{S}$ et $\mathcal{S}'$ sont *équivalents* s'ils admettent exactement les mêmes solutions : $\mathrm{Sol}(\mathcal{S}) = \mathrm{Sol}(\mathcal{S}')$.

4.2 Remarque. Une *opération élémentaire* sur un système $\mathcal{S}$ est une opération qui transforme $\mathcal{S}$ en un système $\mathcal{S}'$ équivalent à $\mathcal{S}$. Il y a trois types d'opérations élémentaires :

E_1 : Permuter deux équations.

E_2 : Multiplier une équation par un scalaire λ non nul (λ est appelé le facteur de l'opération élémentaire).

E_3 : Ajouter à une équation un multiple d'une autre équation.

Le méthode de Gauss consiste à transformer par une suite d'opérations élémentaires un système $\mathcal{S}$ en un système $\mathcal{S}'$ *échelonné*. Pour expliquer ce que système échelonné signifie il convient de considérer la *matrice associée* au système :

4.3 Définition. Une *matrice* à m lignes et n colonnes est un tableau de nombres réels qui contient m lignes, chaque ligne étant un vecteur à n composantes, et n colonnes, chaque colonne étant un vecteur à m composantes :

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

On écrira aussi

$$A = (a_{ij})_{i=1,\dots,m}^{j=1,\dots,n}$$

où a_{ij} est l'élément (ou entrée) de la matrice A qui se trouve au croisement entre la i-ème ligne et la j-ème colonne.
On note par

$$\mathbb{R}^{m\times n}$$

l'ensemble des matrices à m lignes et n colonnes à coefficients réels.

4.4 Remarque. Soit

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

un système à m équations et n variables. On peut extraire de $\mathcal{S}$ deux matrices :

1. La matrice des coefficients :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

2. La matrice complète :

$$(A \mid \vec{b}\,) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \in \mathbb{R}^{m\times(n+1)}$$

4.2 Matrices échelonnées

4.5 Définition.

1. Une matrice est *échelonnée* si le nombre d'entrées nulles au début de chaque ligne augmente strictement quand on passe d'une ligne à la ligne suivante.
2. Un système est *échelonné* si sa matrice complète est échelonnée.

4.6 Remarque. On peut effectuer les opérations élémentaires (voir 4.2) sur les lignes d'une matrice plutôt que sur les équations d'un système :

E_1 : Permuter deux lignes.

E_2 : Multiplier une ligne par un scalaire λ non nul.

E_3 : Ajouter à une ligne un multiple d'une autre ligne.

4.7 Propriété.

1. *On peut toujours transformer une matrice $A \in \mathbb{R}^{m \times n}$ en une matrice échelonnée $A' \in \mathbb{R}^{m \times n}$ par une suite finie d'opérations élémentaires sur les lignes.*
2. *On peut toujours transformer un système $\mathcal{S}$ en un système $\mathcal{S}'$ équivalent à $\mathcal{S}$ et échelonné par une suite finie d'opérations élémentaires.*

Preuve. Remarque préalable : si la i-ème colonne d'une matrice A est nulle et que la matrice A' est obtenue à partir de la matrice A par une opération élémentaire sur les lignes, la i-ème colonne de A' est nulle.
1. Par induction sur m :
Si $m = 1$, la matrice A est réduite à une seule colonne. Elle est donc échelonnée. Supposons que la propriété soit vraie pour toute matrice à $m - 1$ lignes et démontrons-la pour une matrice à m lignes.
Soit $A \in \mathbb{R}^{m \times n}$ et supposons que la première colonne soit non nulle (si la première colonne est nulle, on peut commencer par la deuxième colonne, la première restera nulle en appliquant toute opération élémentaire). Soit $a_{i1} \neq 0$. Quitte à appliquer une opération de type E_1 on peut supposer que $a_{11} \neq 0$ et, quitte à appliquer une opération de type E_2, on peut supposer $a_{11} = 1$. On peut maintenant tuer tous les éléments de type a_{i1} $(i = 2, \ldots, m)$. Pour cela il suffit de retirer à la i-ème ligne a_{i1}-fois la première ligne (ce qui est une opération de type E_3). Ainsi :

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \mapsto A' = \begin{pmatrix} 1 & a_{12} & \ldots & a_{1n} \\ 0 & b_{22} & \ldots & b_{2n} \\ \vdots & & & \\ 0 & b_{m2} & \ldots & b_{mn} \end{pmatrix}$$

On peut maintenant considérer la matrice

$$B = \begin{pmatrix} 0 & b_{22} & \ldots & b_{2n} \\ \vdots & & & \\ 0 & b_{m2} & \ldots & b_{mn} \end{pmatrix}$$

Puisque B contient $m-1$ lignes, par hypothèse inductive on peut la transformer en une matrice échelonnée B' par une suite d'opérations élémentaires (par la remarque préalable, la première colonne de B' sera nulle). On a finalement :

$$A \quad \mapsto A'' = \begin{pmatrix} 1 & a_{12} & \ldots & a_{1n} \\ & & B' & \end{pmatrix}$$

2.Puisque une opération élémentaire transforme un système en un système équivalent, le point 2. est une conséquence du point 1. □

4.8 Remarque. La matrice échelonnée A'' de la Propriété 4.7 n'est pas déterminée de façon unique. De même pour le système $\mathcal{S}'$.

4.3 Espace lignes et espace colonnes

4.9 Définition. Soit $A \in \mathbb{R}^{m\times n}$ (chaque ligne de A est donc un vecteur de $\mathbb{R}^n$ et chaque colonne est un vecteur de $\mathbb{R}^m$). On pose

$$\mathrm{Lign}(A) = \text{ s.e.v. de } \mathbb{R}^n \text{ engendré par les lignes de } A$$

$$\mathrm{Col}(A) = \text{ s.e.v. de } \mathbb{R}^m \text{ engendré par les colonnes de } A$$

4.10 Remarque. On sait qu'on peut appliquer les opérations élémentaires aux équations d'un système (voir 4.2) ou aux lignes d'une matrice (voir 4.6). Clairement, on peut les appliquer aussi aux colonnes d'une matrice. Plus en générale, on peut appliquer les trois opérations élémentaires à toute famille finie

$$\vec{v}_1, \ldots, \vec{v}_k$$

de vecteurs de $\mathbb{R}^n$.

4.11 Propriété. *Soit $\vec{v}_1, \ldots, \vec{v}_k$ une famille de vecteurs de $\mathbb{R}^n$ et soit $\vec{w}_1, \ldots, \vec{w}_k$ une nouvelle famille obtenue à partir de la première famille par une suite finie d'opérations élémentaires. Alors les deux familles engendrent le même s.e.v. :*

$$\langle \vec{v}_1, \ldots, \vec{v}_k \rangle = \langle \vec{w}_1, \ldots, \vec{w}_k \rangle \,.$$

Preuve. Par exemple, si un vecteur $\vec{x}$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_k$

$$\vec{x} = \alpha_1 \cdot \vec{v}_1 + \ldots + \alpha_k \cdot \vec{v}_k$$

et si $\vec{w}_1$ est obtenu par une opération de type E_3

$$\vec{w}_1 = \vec{v}_1 + \lambda \cdot \vec{v}_2$$

alors $\vec{x}$ est aussi combinaison linéaire de $\vec{w}_1, \vec{v}_2 \ldots, \vec{v}_k$:

$$\vec{x} = \alpha_1 \cdot (\vec{w}_1 - \lambda \cdot \vec{v}_2) + \alpha_2 \cdot \vec{v}_2 + \ldots + \alpha_k \cdot \vec{v}_k \,.$$

□

4.12 Corollaire. *Soit* $A \in \mathbb{R}^{m\times n}$.

1. *Soit* A' *la matrice obtenue à partir de* A *par une suite finie d'opérations élémentaires sur les lignes de* A. *Alors* $\mathrm{Lign}(A) = \mathrm{Lign}(A')$.
2. *Soit* A' *la matrice obtenue à partir de* A *par une suite finie d'opérations élémentaires sur les colonnes de* A. *Alors* $\mathrm{Col}(A) = \mathrm{Col}(A')$.

Le prochain corollaire nous donne une méthode facile pour déterminer la dimension de l'espace engendré par les lignes d'une matrice.

4.13 Corollaire. *Soit* $A \in \mathbb{R}^{m\times n}$ *et* A' *une matrice échelonnée obtenue à partir de* A *par une suite finie d'opérations élémentaires sur les lignes de* A. *On a :*

$$\dim \mathrm{Lign}(A) = \textit{ nombre de lignes non nulles de } A' .$$

Preuve. Si A' est à lignes échelonnées, ses lignes non nulles sont des vecteurs de $\mathbb{R}^n$ linéairement indépendants. Elles forment donc une base de $\mathrm{Lign}(A)$. □

4.14 Propriété. *Soit* $A \in \mathbb{R}^{m\times n}$ *:*

$$\dim \mathrm{Lign}(A) = \dim \mathrm{Col}(A) .$$

On pose $\dim \mathrm{Col}(A) = \mathrm{rang}\, A$.

Preuve. Voir plus loin (Propriété 9.29). □

4.4 Dimension et base de $\mathrm{Sol}(\mathcal{S}_0)$

On peut revenir au problème de départ de ce chapitre : étudier l'espace des solutions d'un système homogène.

4.15 Propriété. *Soit* $\mathcal{S}_0$ *un système homogène à* m *équations et* n *variables et soit* $A \in \mathbb{R}^{m\times n}$ *sa matrice des coefficents. On a :*

$$\dim \mathrm{Sol}(\mathcal{S}_0) = n - \mathrm{rang}\, A .$$

Preuve. Voir plus loin (Corollaire 9.19). □

4.16 Exemple. Considérons le système homogène à 3 équations et 5 variables

$$\mathcal{S}_0 : \left\{ \begin{array}{rcl} x_1 + x_2 + x_3 + x_4 + x_5 & = & 0 \\ x_1 + 3x_2 + 3x_3 - x_4 + x_5 & = & 0 \\ x_2 + x_3 - x_4 & = & 0 \end{array} \right.$$

La matrice des coefficients est

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 3 & -1 & 1 \\ 0 & 1 & 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 5}$$

On peut échelonner A par exemple par les opérations élémentaires

ligne 2 - ligne 1 , ligne 2 - 2· ligne 3 , ligne 2 échangée avec ligne 3

et on obtient la matrice

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

La matrice A' contient 2 lignes non nulles et donc $\operatorname{rang} A = 2$ (voir 4.13). Si on applique maintenant la Propriété 4.15 on a

$$\dim \operatorname{Sol}(\mathcal{S}_0) = n - \operatorname{rang} A = 5 - 2 = 3\,.$$

La matrice A' nous donne aussi un nouveau système homogène équivalent au système de départ :

$$\mathcal{S}_0' : \left\{ \begin{array}{rcl} x_1 + x_2 + x_3 + x_4 + x_5 & = & 0 \\ x_2 + x_3 - x_4 & = & 0 \\ 0 & = & 0 \end{array} \right.$$

Pour déterminer une base de $\operatorname{Sol}(\mathcal{S}_0)$ il suffit de trouver trois solutions linéairement indépendantes de $\mathcal{S}_0'$. On remarque que les variables x_3, x_4, x_5 n'apparaissent pas en tête des équations de $\mathcal{S}_0'$. On peut effectuer trois choix linéairement indépendants de ces variables et utiliser $\mathcal{S}_0'$ pour déterminer ensuite x_1 et x_2. Le plus simple est :

- On fixe $(x_3, x_4, x_5) = (1, 0, 0)$, on remplace dans $\mathcal{S}_0'$ et on trouve $(x_1, x_2) = (0, -1)$. Donc $\vec{v}_1 = (0, -1, 1, 0, 0) \in \operatorname{Sol}(\mathcal{S}_0)$.
- On fixe $(x_3, x_4, x_5) = (0, 1, 0)$, on remplace dans $\mathcal{S}_0'$ et on trouve $(x_1, x_2) = (-2, 1)$. Donc $\vec{v}_2 = (-2, 1, 0, 1, 0) \in \operatorname{Sol}(\mathcal{S}_0)$.
- On fixe $(x_3, x_4, x_5) = (0, 0, 1)$, on remplace dans $\mathcal{S}_0'$ et on trouve $(x_1, x_2) = (-1, 0)$. Donc $\vec{v}_3 = (-1, 0, 0, 0, 1) \in \operatorname{Sol}(\mathcal{S}_0)$.

Puisque $\vec{v}_1, \vec{v}_2, \vec{v}_3$ sont linéairmement indépendants, nous avons :

$$\dim \operatorname{Sol}(\mathcal{S}_0) = 3\ , \quad \operatorname{Sol}(\mathcal{S}_0) = \langle \vec{v}_1, \vec{v}_2, \vec{v}_3 \rangle$$

On peut aussi dire que la solution générale de $\mathcal{S}_0$ est

$$\alpha \cdot \vec{v}_1 + \beta \cdot \vec{v}_2 + \gamma \cdot \vec{v}_3 \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R} \text{ arbitraires}$$

ou bien que l'ensemble des solutions de $\mathcal{S}_0$ est donné par les vecteurs du type

$$\alpha \cdot \vec{v}_1 + \beta \cdot \vec{v}_2 + \gamma \cdot \vec{v}_3 \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R} \text{ arbitraires}$$

La dimension de l'espace $\operatorname{Sol}(\mathcal{S}_0)$ correspond donc au nombre de variables indéterminées dans la solution générale.

4.17 Remarque. Voyons à présent comment on peut justifier le fait que dans l'Exemple 4.16 nous avons pu choisir de façon arbitraire les variables x_3, x_4, x_5 pour trouver des solutions du système $\mathcal{S}_0$. Nous avons remarqué que l'espace $\mathrm{Col}(A)$ a dimension 2 (car $\mathrm{rang}\, A = 2$). Or, en regardant la matrice échelonnée A' on se rend compte que la première et la deuxième colonne sont linéairement indépendantes et forment donc une base de $\mathrm{Col}(A)$. Par conséquent, toute combinaison linéaire des trois dernières colonnes peut s'exprimer comme combinaison linéaire de la première et de la deuxième colonne : pour tout $x_3, x_4, x_5 \in \mathbb{R}$ on peut trouver $x_1, x_2 \in \mathbb{R}$ tels que

$$x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = x_3 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_4 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_5 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Ou encore : pour tout $x_3, x_4, x_5 \in \mathbb{R}$ on peut trouver $x_1, x_2 \in \mathbb{R}$ tels que

$$x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_4 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_5 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

C'est-à-dire, en effectuant les opérations composante par composante,

$$\begin{pmatrix} x_1 + x_2 + x_3 + x_4 + x_5 \\ x_2 + x_3 - x_4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Finalement : pour tout $x_3, x_4, x_5 \in \mathbb{R}$ on peut trouver $x_1, x_2 \in \mathbb{R}$ tels que $(x_1, x_2, x_3, x_4, x_5) \in \mathrm{Sol}(\mathcal{S}_0)$.

Pour terminer cette remarque, montrons sur un exemple qu'en général on ne peut pas choisir de façon arbitraire les variables qui apparaissent en tête des équations dans le système échelonné. Voici l'exemple :

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_2 + x_3 + x_4 &= 0 \end{cases}$$

Les variables x_1, x_2 apparaissent en tête, et si on fixe $x_1 = 1$ et $x_2 = 0$ le système n'a pas de solutions.

Avant de passer aux systèmes non homogènes, voyons la preuve de la Propriété 3.4.

4.18 Notation. Soit $f = \{\vec{f_1}, \ldots, \vec{f_n}\}$ une base de $\mathbb{R}^n$ et soit $\vec{x} \in \mathbb{R}^n$. On peut écrire $\vec{x}$ comme combinaison linéaire de $\vec{f_1}, \ldots, \vec{f_n}$, et cela d'une seule façon :

$$\vec{x} = \alpha_1 \cdot \vec{f_1} + \ldots + \alpha_n \cdot \vec{f_n}$$

Le vecteur

$${}_f\vec{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$$

est le *vecteur des coordonnées de $\vec{x}$ dans la base f.*

4.19 Exercice. Pour tout s.e.v. V de $\mathbb{R}^n$ on peut trouver un système homogène $\mathcal{S}_0$ en n variables tel que $V = \mathrm{Sol}(\mathcal{S}_0)$.

Preuve. Soit $\vec{f_1}, \ldots, \vec{f_k}$ une base de V. On peut la compléter en une base f de $\mathbb{R}^n$ (voir 3.18) :

$$f = \{\vec{f_1}, \ldots, \vec{f_k}, \vec{f_{k+1}}, \ldots, \vec{f_n}\}$$

Considérons la décomposition de $\vec{x} \in \mathbb{R}^n$ par rapport à la base f :

$$\vec{x} = x_1 \cdot \vec{f_1} + \ldots + x_k \cdot \vec{f_k} + x_{k+1} \cdot \vec{f_{k+1}} + \ldots + x_n \cdot \vec{x}_n$$

On a alors que $\vec{x} \in V$ ssi $x_{k+1} = \ldots = x_n = 0$ ssi ses coordonnées dans la base f

$$_f\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix}$$

sont solution du système homogène

$$\mathcal{S}_0 \colon \left\{ \begin{array}{rcl} x_{k+1} & = & 0 \\ \vdots & & \\ x_n & = & 0 \end{array} \right.$$

4.5 Structure de $\mathrm{Sol}(\mathcal{S})$

Dans ce chapitre on s'est pour l'instant occupé seulement des systèmes homogènes. Voyons à présent comment on peut utiliser ce que nous avons appris sur les systèmes homogènes pour traiter le cas des systèmes non homogènes.

4.20 Notation. Soit $\mathcal{S}$ un système à m équations et n variables, soit $A \in \mathbb{R}^{m \times n}$ sa matrice des coefficients et $\vec{b} \in \mathbb{R}^m$ le vecteur des termes indépendants. Le système homogène $\mathcal{S}_0$ associé à $\mathcal{S}$ est le système obtenu de $\mathcal{S}$ en remplacant chaque terme indépendant par 0. La matrice des coefficients de $\mathcal{S}_0$ est donc la matrice A elle-même.

4.21 Remarque.

1. Notons par $C_1, \ldots, C_n \in \mathbb{R}^m$ les n colonnes de la matrice A. Trouver une solution de $\mathcal{S}$ revient exactement à exprimer $\vec{b}$ comme combinaison linéaire de $C_1, \ldots, C_n$. En effet, il suffit de remarquer que le système $\mathcal{S}$ peut s'écrire comme

$$x_1 \cdot C_1 + \ldots + x_n \cdot C_n = \vec{b}$$

Pour voir cela, il faut effectuer les opérations composante par composante sur les vecteurs de $\mathbb{R}^m$ (on peut voir aussi la Remarque 5.11). Voici un exemple simple : considérons le système

$$\mathcal{S}\colon \left\{ \begin{array}{rcl} x_1 + 2x_2 + 3x_3 & = & 2 \\ 4x_1 + 5x_2 + 6x_3 & = & 8 \end{array} \right.$$

Si C_1, C_2, C_3 sont les trois colonnes de la matrice des coefficients et $\vec{b}$ est la colonne des termes indépendantes, nous avons

$$C_1 + 2C_2 - C_3 = \vec{b}$$

et donc $(1, 2, -1)$ est une solution de $\mathcal{S}$.

2. Il est clair aussi que
$$\operatorname{rang} A \leq \operatorname{rang}(A \mid \vec{b}\,)$$
car $\operatorname{Col}(A) \subseteq \operatorname{Col}(A \mid \vec{b}\,)$. Dans la prochaine propriété on utilisera le rang pour établir si un système admet au moins une solution.

4.22 Propriété. *Soit $\mathcal{S}$ un système avec matrice des coefficients $A \in \mathbb{R}^{m\times n}$ et vecteur des termes indépendants $\vec{b} \in \mathbb{R}^m$. Les conditions suivantes sont équivalents :*

1. S admet au moins une solution ;

2. $\operatorname{Col}(A) = \operatorname{Col}(A \mid \vec{b}\,)$ *(c'est-à-dire : le vecteur $\vec{b}$ est combinaison linéaire des colonnes de A) ;*

3. $\operatorname{rang} A = \operatorname{rang}(A \mid \vec{b}\,)$ *(c'est-à-dire :* $\dim \operatorname{Col}(A) = \dim \operatorname{Col}(A \mid \vec{b}\,)$*).*

Preuve. 1. $\Leftrightarrow$ 2. : Voir la première partie de la Remarque 4.21. (Pour une autre rédaction de cette preuve, voir le Corollaire 7.18.)
2. $\Leftrightarrow$ 3. : Puisque on a toujours $\operatorname{Col}(A) \subseteq \operatorname{Col}(A \mid \vec{b}\,)$, l'équivalence entre les conditions 2. et 3. est un cas particulier de l'Exercice 3.20. □

Si un système $\mathcal{S}$ admet au moins une solution, on peut déterminer l'ensemble $\operatorname{Sol}(\mathcal{S})$ en passant par le système homogène $\mathcal{S}_0$ associé à $\mathcal{S}$, de façon à pouvoir utiliser la méthode de Gauss.

4.23 Théorème. *Soit $\mathcal{S}$ un système et $\mathcal{S}_0$ le système homogène associé. Soit $\vec{x}_0$ une solution de $\mathcal{S}$ et soit $\vec{x}_1, \ldots, \vec{x}_k$ une base de l'espace* $\operatorname{Sol}(\mathcal{S}_0)$*. Les solutions de $\mathcal{S}$ sont exactement les vecteurs de la forme*

$$\vec{x}_0 + \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_k \cdot \vec{x}_k \quad \textit{avec } \alpha_1, \ldots, \alpha_k \in \mathbb{R} \textit{ arbitraires.}$$

Preuve. Voir plus loin (Exercice 5.12 ou Corollaire 7.16). □

4.24 Remarque. D'après la Propriété 4.14, pour tout matrice A nous avons

$$\dim \operatorname{Lign}(A) = \dim \operatorname{Col}(A)$$

Cela implique en particulier que pour déterminer le rang d'une matrice nous pouvons échelonner la matrice par des opérations élémentaires sur les lignes aussi bien que sur les colonnes. Néanmoins, il faut être attentifs au fait que, si la matrice A est la matrice des coefficients d'un système $\mathcal{S}$, échelonner A par des opérations élémentaires sur les colonnes peut faire passer du système $\mathcal{S}$ à un systme $\mathcal{S}'$ qui n'est pas équivalent à $\mathcal{S}$. Voici un exemple : considérons le système :

$$\mathcal{S}\colon \left\{ \begin{array}{rcl} x_1 + x_2 & = & 2 \\ x_1 - x_2 & = & 2 \end{array} \right.$$

qui admet $(x_1 = 2, x_2 = 0)$ comme unique solution. Si nous multiplions la première colonne par 2, nous obtenons le système

$$\mathcal{S}'\colon \left\{ \begin{array}{rcl} 2x_1 + x_2 & = & 2 \\ 2x_1 - x_2 & = & 2 \end{array} \right.$$

qui admet $(x_1 = 1, x_2 = 0)$ comme unique solution.

Chapitre 5

Calcul matriciel

Le chapitre 5 section par section

Dans le chapitre précédent nous avons vu que pour étudier un système d'équations linéaires on fait intervenir les matrices. Il faut donc apprendre à effectuer des opérations algébriques sur les matrices.

1. La somme de deux matrices et le produit d'un scalaire par une matrice se font comme pour les vecteurs de $\mathbb{R}^n$.
2. Si le nombre de colonnes d'une matrice A est égal au nombre de lignes d'une matrice B, on peut effectuer le produit $A \cdot B$. Ce produit est défini en utilisant le produit scalaire canonique de $\mathbb{R}^n$.
3. En utilisant le produit matriciel, on peut exprimer en forme compacte un système d'équations linéaires et, en particulier, voir le lien ente les solutions du système et les combinaisons linéaires des colonnes de la matrice des coefficients.
4. Le produit matriciel permet d'exprimer les opérations élémentaires sur une matrice, et aussi d'exprimer le lien entre les composantes d'un vecteur dans une base et les composantes du même vecteur dans une autre base.

5.1 Combinaisons linéaires de matrices

5.1 Définition. Soit $\lambda \in \mathbb{R}$ et

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad , \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

deux matrices à m lignes et n colonnes.

1. Somme de deux matrices : $+ : \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$

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

2. Produit d'un scalaire par une matrice : $\cdot : \mathbb{R} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$

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

5.2 Propriété. *L'ensemble $\mathbb{R}^{m\times n}$ des matrices à m lignes et n colonnes est un espace vectoriel réel (voir Chapitre 6) par rapport aux opérations définies en 5.1. Pour tout $A, B, C \in \mathbb{R}^{m\times n}$ et pour tout $\alpha, \beta \in \mathbb{R}$, on a :*

1. $(A + B) + C = A + (B + C)$
2. $A + 0 = A = 0 + A$, *où* 0 *est la matrice dont tous les éléments sont nuls*
3. $A + B = B + A$
4. $A + (-A) = 0$, *où* $-A = -1 \cdot A$
5. $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$
6. $\alpha \cdot (A + B) = \alpha A + \alpha \cdot B$
7. $\alpha \cdot (\beta \cdot A) = (\alpha\beta) \cdot A$
8. $1 \cdot A = A$.

Preuve. La preuve de cette propriété est un exercice facile. □

5.2 Produit matriciel

5.3 Définition. Produit scalaire canonique :

$$\mathbb{R}^n \times \mathbb{R}^n \xrightarrow{\cdot} \mathbb{R}$$

Soit $\vec{x} = (x_1, x_2, \dots, x_n), \vec{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$. On pose :

$$\vec{x} \cdot \vec{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1,\dots,n} x_iy_i$$

5.4 Propriété. *Pour tout $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ et pour tout $\alpha, \beta \in \mathbb{R}$, on a :*

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$
2. $(\alpha \cdot \vec{x} + \beta \cdot \vec{y}) \cdot \vec{z} = \alpha \cdot (\vec{x} \cdot \vec{z}) + \beta \cdot (\vec{y} \cdot \vec{z})$
3. *si* $\vec{x} \neq \vec{0}$, *alors* $\vec{x} \cdot \vec{x} > 0$

Preuve. La preuve de cette propriété est un exercice facile. □

5.5 Définition. Produit matriciel :

$$\mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p} \xrightarrow{\cdot} \mathbb{R}^{m\times p}$$

Soit $A = (a_{i,j})_{i=1,\ldots,m}^{j=1,\ldots,n} \in \mathbb{R}^{m\times n}$, $B = (b_{jk})_{j=1,\ldots,n}^{k=1,\ldots,p} \in \mathbb{R}^{n\times p}$. On pose :

$$A \cdot B = (c_{ik})_{i=1,\ldots,m}^{k=1,\ldots,p} \in \mathbb{R}^{m\times p}$$

c_{ik} est le produit scalaire entre la i-ème ligne de A et la k-ème colonne de B:

$$c_{ik} = (a_{i1}, a_{i2}, \ldots, a_{in}) \cdot (b_{1k}, b_{2k}, \ldots, b_{nk}) = \sum_{j=1,\ldots,n} a_{ij} b_{jk}$$

5.6 Remarque.

1. Pour effectuer le produit matriciel $A \cdot B$ il faut donc que le nombre de colonnes de la matrice A soit égal au nombre de lignes de la matrice B. Si cette condition est vérifiée, on dit que A est B sont *compatibles* ou *composables*. La condition de compatibilité entre deux matrices dépend de l'ordre. Par exemple si $A \in \mathbb{R}^{3\times 4}$ et $B \in \mathbb{R}^{4\times 5}$, on peut effectuer $A \cdot B$ mais pas $B \cdot A$.
2. Si $A \in \mathbb{R}^{m\times n}$ et $B \in \mathbb{R}^{n\times m}$, on peut effectuer tant $A \cdot B$ que $B \cdot A$, mais en général $A \cdot B \neq B \cdot A$ car $A \cdot B \in \mathbb{R}^{m\times m}$ et $B \cdot A \in \mathbb{R}^{n\times n}$.
3. Attention : même pour des matrices *carrées* $A, B \in \mathbb{R}^{n\times n}$ en général $A \cdot B \neq B \cdot A$.
4. Le produit scalaire de deux vecteurs (voir 5.3) est un cas particulier de produit matriciel. Pour cela il faut regarder le premier vecteur comme une ligne et le deuxième comme une colonne :

$$\text{si } \vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^{1\times n} \quad \text{et} \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^{n\times 1}$$

$$\text{alors} \quad \vec{x} \cdot \vec{y} = \sum_{i=1,\ldots,n} x_i y_i \in \mathbb{R}^{1\times 1} = \mathbb{R}$$

5.7 Propriété. *Pour toutes matrices A, B, C et pour tout scalaire α, en supposant vérifiées les conditions de compatibilité nécessaires, on a :*

1. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$
2. $A \cdot I = A = I \cdot A$, *où* $I \in \mathbb{R}^{n\times n}$ *est la matrice identité : la i-ème ligne de* I *est le i-ème vecteur de la base canonique de* $\mathbb{R}^n$ *(voir 3.9)*
3. $A \cdot (B + C) = A \cdot B + A \cdot C$, $(A + B) \cdot C = A \cdot C + B \cdot C$
4. $A \cdot (\alpha \cdot B) = \alpha \cdot (A \cdot B) = (\alpha \cdot A) \cdot B$

Preuve. Voir plus loin (Corollaire 7.27). □

5.8 Remarque. Si on se restreint aux matrices carrées, on peut résumer la Propriété 5.7 en disant que $\mathbb{R}^{n\times n}$ est une $\mathbb{R}$-*algèbre* par rapport aux opérations de somme, produit par un scalaire et produit matriciel définies en 5.1 et 5.5. On reviendra au Chapitre 7 sur la définition générale de $\mathbb{R}$-algèbre.

5.9 Remarque. Attention, le produit matriciel n'est pas simplifiable : si A, B, C sont des matrices (et les conditions de compatibilité sont vérifiées), alors :

1. la condition $A \cdot C = B \cdot C$ n'implique pas $A = B$,
2. la condition $C \cdot A = C \cdot B$ n'implique pas $A = B$,
3. en particulier, on peut bien avoir $A \cdot B = 0$ avec $A \neq 0 \neq B$.

Voici un exemple de deux matrices non nulles qui donnent comme produit la matrice nulle :

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

5.3 Expression matricielle d'un système

Considérons la matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

Chaque ligne de A est un vecteur de $\mathbb{R}^n$ et nous les notons

$$\vec{a}_{1*} = (a_{11}, \dots, a_{1n}) \quad \dots \quad \vec{a}_{m*} = (a_{m1}, \dots, a_{mn})$$

Chaque colonne de A est un vecteur de $\mathbb{R}^m$ et nous les notons

$$\vec{a}_{*1} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \quad \dots \quad \vec{a}_{*n} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

5.10 Remarque. On peut exprimer le produit matriciel en utilisant la notation précédente :

$$A \cdot B = \begin{pmatrix} \vec{a}_{1*} \cdot \vec{b}_{*1} & \vec{a}_{1*} \cdot \vec{b}_{*2} & \dots & \vec{a}_{1*} \cdot \vec{b}_{*p} \\ \vec{a}_{2*} \cdot \vec{b}_{*1} & \vec{a}_{2*} \cdot \vec{b}_{*2} & \dots & \vec{a}_{2*} \cdot \vec{b}_{*p} \\ \vdots & & & \\ \vec{a}_{m*} \cdot \vec{b}_{*1} & \vec{a}_{m*} \cdot \vec{b}_{*2} & \dots & \vec{a}_{m*} \cdot \vec{b}_{*p} \end{pmatrix}$$

Cela permet de voir que

1. k-ème colonne de $A \cdot B = A \cdot$ (k-ème colonne de B); autrement dit

$$A \cdot (\vec{b}_{*1} \mid \ldots \mid \vec{b}_{*p}) = (A \cdot \vec{b}_{*1} \mid \ldots \mid A \cdot \vec{b}_{*p})$$

2. i-ème ligne de $A \cdot B =$ (i-ème ligne de A) $\cdot B$; autrement dit

$$\begin{pmatrix} \vec{a}_{1*} \\ \vdots \\ \vec{a}_{m*} \end{pmatrix} \cdot B = \begin{pmatrix} \vec{a}_{1*} \cdot B \\ \vdots \\ \vec{a}_{m*} \cdot B \end{pmatrix}$$

En particulier :

1. k-ème colonne de $A = A \cdot \vec{e}_k$
2. i-ème ligne de $A = \vec{e}_i \cdot A$

où $\vec{e}_k$ et $\vec{e}_i$ sont les vecteurs de la base canonique.

5.11 Remarque. Considérons un système

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

à m équations et n variables. En utilisant les opérations sur les matrices, on peut exprimer $\mathcal{S}$ de façon plus compacte :

1. $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$, avec

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}\ ,\ \ \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n\ ,\ \ \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

2. $\mathcal{S}\colon x_1 \cdot \vec{a}_{*1} + x_2 \cdot \vec{a}_{*2} + \ldots + x_n \cdot \vec{a}_{*n} = \vec{b}$, avec

$$\vec{a}_{*1} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}\ ,\ \ldots\ ,\ \vec{a}_{*n} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \in \mathbb{R}^m$$

Autrement dit : $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est solution de $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ si et seulement si $x_1 \cdot \vec{a}_{*1} + x_2 \cdot \vec{a}_{*2} + \ldots + x_n \cdot \vec{a}_{*n} = \vec{b}$.

Preuve.

1.
$$A \cdot \vec{x} = \begin{pmatrix} \vec{a}_{1*} \cdot \vec{x} \\ \vec{a}_{2*} \cdot \vec{x} \\ \vdots \\ \vec{a}_{m*} \cdot \vec{x} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}$$

2.
$$\begin{aligned} A \cdot \vec{x} &= A \cdot (x_1 \cdot \vec{e}_1 + x_2 \cdot \vec{e}_2 + \ldots + x_n \cdot \vec{e}_n) \\ &= x_1 \cdot A \cdot \vec{e}_1 + x_2 \cdot A \cdot \vec{e}_2 + \ldots + x_n \cdot A \cdot \vec{e}_n \\ &= x_1 \cdot \vec{a}_{*1} + x_2 \cdot \vec{a}_{*2} + \ldots + x_n \cdot \vec{a}_{*n} \end{aligned}$$

□

5.12 Exercice. A titre d'exercice, voyons une première preuve du Théorème 4.23 (pour une deuxième preuve, voir le Corollaire 7.16).
Soit $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ un système, $\mathcal{S}_0$ le système homogène associé, $\vec{x}_0 \in \mathrm{Sol}(\mathcal{S})$ et $\vec{x}_1, \ldots, \vec{x}_k$ un base de $\mathrm{Sol}(\mathcal{S}_0)$.
Montrons d'abord que pour tout $\alpha_1, \ldots, \alpha_k$, le vecteur $\vec{x}_0 + \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_k \cdot \vec{x}_k$ est solution de $\mathcal{S}$:

$$A \cdot (\vec{x}_0 + \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_k \cdot \vec{x}_k) = A \cdot \vec{x}_0 + \alpha_1 \cdot A \cdot \vec{x}_1 + \ldots \alpha_k \cdot A \cdot \vec{x}_k = \vec{b} + \vec{0} + \ldots + \vec{0} = \vec{b}$$

Réciproquement, montrons que si $\vec{y}$ est solution de $\mathcal{S}$, alors il peut s'écrire sous la forme $\vec{y} = \vec{x}_0 + \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_k \cdot \vec{x}_k$: nous avons $A \cdot \vec{x}_0 = \vec{b}$ et $A \cdot \vec{y} = \vec{b}$, d'où $A \cdot \vec{x}_0 - A \cdot \vec{y} = \vec{0}$, ou encore $A \cdot (\vec{x}_0 - \vec{y}) = \vec{0}$. Cela signifie que $\vec{x}_0 - \vec{y} \in \mathrm{Sol}(\mathcal{S}_0)$ et donc $\vec{x}_0 - \vec{y}$ est combinaison linéaire de $\vec{x}_1, \ldots, \vec{x}_k$. Par conséquent, on peut trouver $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tels que $\vec{x}_0 - \vec{y} = \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_k \cdot \vec{x}_k$, d'où la thèse.

5.4 Opérations élémentaires, changement de base

Les opérations élémentaires (voir 4.6 et 4.10) peuvent s'interpréter en termes de produit matriciel :

5.13 Définition. Une *matrice élémentaire* $E \in \mathbb{R}^{n\times n}$ est une matrice obtenue en appliquant une opération élémentaire aux lignes ou aux colonnes de la matrice identité $I \in \mathbb{R}^{n\times n}$.

5.14 Propriété. *Soit* $A \in \mathbb{R}^{m\times n}$.

1. *Soit* $A' \in \mathbb{R}^{m\times n}$ *la matrice obtenue en appliquant une certaine opération élémentaire aux colonnes de* A, *et soit* $E \in \mathbb{R}^{n\times n}$ *la matrice élémentaire obtenue en appliquant la même opération élémentaire aux colonnes de* $I \in \mathbb{R}^{n\times n}$. *On a :*
$$A' = A \cdot E$$
2. *Soit* $A' \in \mathbb{R}^{m\times n}$ *la matrice obtenue en appliquant une certaine opération élémentaire aux lignes de* A, *et soit* $E \in \mathbb{R}^{m\times m}$ *la matrice élémentaire obtenue en appliquant la même opération élémentaire aux lignes de* $I \in \mathbb{R}^{m\times m}$. *On a :*
$$A' = E \cdot A$$

5.15 Exercice. Considérons la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}^{2\times 3}$$

et l'opération élémentaire "échanger la première et la deuxième ligne". Si on applique cette opération élémentaire à la matrice identité

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

on obtient la matrice élémentaire

$$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

Calculons maintenant

$$E \cdot A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix} = A'$$

La nouvelle matrice $A' \in \mathbb{R}^{2\times 3}$ est bien la matrice que l'on obtient de A en échangeant la première et la deuxième ligne.

Pour terminer ce chapitre, voyons comment changent les coordonnées d'un vecteur si on change la base sur laquelle on décompose le vecteur.

5.16 Remarque. Soit $e = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$, $f = \{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_n\}$ deux bases de $\mathbb{R}^n$. Nous pouvons exprimer chaque vecteur de la base e par rapport à la base f :

$$\begin{array}{rcl} \vec{e}_1 & = & a_{11} \cdot \vec{f}_1 + a_{21} \cdot \vec{f}_2 + \ldots + a_{n1} \cdot \vec{f}_n \\ \vec{e}_2 & = & a_{12} \cdot \vec{f}_1 + a_{22} \cdot \vec{f}_2 + \ldots + a_{n2} \cdot \vec{f}_n \\ \ldots & & \\ \vec{e}_n & = & a_{1n} \cdot \vec{f}_1 + a_{2n} \cdot \vec{f}_2 + \ldots + a_{nn} \cdot \vec{f}_n \end{array}$$

On peut en extraire la *matrice changement de base*

$${}_fI_e = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = ({}_f\vec{e}_1 \mid {}_f\vec{e}_2 \mid \ldots \mid {}_f\vec{e}_n)$$

5.17 Propriété. *Soit $\vec{x} \in \mathbb{R}^n$. Avec la notation de 5.16 :*

$${}_f\vec{x} = {}_fI_e \cdot {}_e\vec{x}$$

Preuve. On a

$$\vec{x} = \alpha_1 \cdot \vec{e}_1 + \alpha_2 \cdot \vec{e}_2 + \ldots + \alpha_n \cdot \vec{e}_n =$$

$$= \alpha_1 \cdot (a_{11} \cdot \vec{f_1} + a_{21} \cdot \vec{f_2} + \ldots + a_{n1} \cdot \vec{f_n}) + \ldots + \alpha_n \cdot (a_{1n} \cdot \vec{f_1} + a_{2n} \cdot \vec{f_2} + \ldots + a_{nn} \cdot \vec{f_n}) =$$
$$= (\alpha_1 a_{11} + \ldots + \alpha_n a_{1n}) \cdot \vec{f_1} + \ldots + (\alpha_1 a_{n1} + \ldots + \alpha_n a_{nn}) \cdot \vec{f_n}$$

et donc

$${}_f\vec{x} = \begin{pmatrix} \alpha_1 a_{11} + \ldots + \alpha_n a_{1n} \\ \vdots \\ \alpha_1 a_{n1} + \ldots + \alpha_n a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = {}_fI_e \cdot {}_e\vec{x}$$

$\square$

5.18 Exercice. Considérons les bases e et f de $\mathbb{R}^2$ ci-dessous

$$e = \left\{ \vec{e}_1 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\} \qquad f = \left\{ \vec{f_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{f_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

et le vecteur $\vec{x} = \begin{pmatrix} 6 \\ -4 \end{pmatrix} \in \mathbb{R}^2$. On cherche à déterminer ${}_e\vec{x}$ et ${}_f\vec{x}$.

Puisque $\vec{x} = 2 \cdot \vec{e}_1 - 1 \cdot \vec{e}_2$, nous avons que ${}_e\vec{x} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Pour déterminer ${}_f\vec{x}$ nous pouvons passer par la matrice changement de base : puisque $\vec{e}_1 = 2 \cdot \vec{f_1} + 2 \cdot \vec{f_2}$ et $\vec{e}_2 = 3 \cdot \vec{f_1} - 1 \cdot \vec{f_2}$, la matrice changement de base est donnée par

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

Finalement, nous obtenons

$${}_f\vec{x} = {}_fI_e \cdot {}_e\vec{x} = \begin{pmatrix} 2 & 3 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

Pour vérifier la réponse trouvée : $1 \cdot \vec{f_1} + 5 \cdot \vec{f_2} = \vec{x}$.

Chapitre 6

Espaces vectoriels réels

Le chapitre 6 section par section

A partir de l'espace des vecteurs $\mathbb{R}^n$ et de l'espace des matrices $\mathbb{R}^{m\times n}$, nous pouvons dégager la notion générale d'espace vectoriel.

1. $\mathbb{R}^n$ et $\mathbb{R}^{m\times n}$ sont des espaces vectoriels, mais beaucoup d'autres exemples importants rentrent dans cette définition.
2. Tout espace vectoriel admet une base : si la base est finie, tout résultat mentionné dans les chapitres précédents pour $\mathbb{R}^n$ reste valable, mais il y a aussi des espaces vectoriels qui n'ont pas de bases finies.

6.1 Définition et exemples

Nous avons déjà rencontré deux exemples d'espaces vectoriels : l'espace $\mathbb{R}^n$ des vecteurs à n composantes (voir 2.11) et l'espace $\mathbb{R}^{m\times n}$ des matrices à m lignes et n colonnes (voir 5.2). Voyons maintenant la définition générale d'espace vectoriel réel :

6.1 Définition. Un *espace vectoriel réel* est donné par :

1. un ensemble V dont les éléments sont appelés vecteurs et notés $\vec{x}, \vec{y}, \ldots$
2. un élément $\vec{0} \in V$ appelé vecteur zéro ou vecteur nul
3. une opération binaire appelée somme

$$V \times V \to V \ , \ \ (\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}$$

4. une opération binaire appelée action

$$\mathbb{R} \times V \to V \ , \ \ (\alpha, \vec{x}) \mapsto \alpha \cdot \vec{x}$$

Ces données doivent satisfaire aux équations suivantes : pour tout $\vec{x}, \vec{y}, \vec{z} \in V$ et pour tout $\alpha, \beta \in \mathbb{R}$:

1. $(\vec{x}+\vec{y})+\vec{z}=\vec{x}+(\vec{y}+\vec{z})$
2. $\vec{x}+\vec{0}=\vec{x}=\vec{0}+\vec{x}$,
3. $\vec{x}+\vec{y}=\vec{y}+\vec{x}$
4. $\vec{x}+(-\vec{x})=\vec{0}$, où $-\vec{x}=-1\cdot\vec{x}$
5. $(\alpha+\beta)\cdot\vec{x}=\alpha\cdot\vec{x}+\beta\cdot\vec{x}$
6. $\alpha\cdot(\vec{x}+\vec{y})=\alpha\cdot x+\alpha\cdot\vec{y}$
7. $\alpha\cdot(\beta\cdot\vec{x})=(\alpha\beta)\cdot\vec{x}$
8. $1\cdot\vec{x}=\vec{x}$.

Un espace vectoriel réel est donc un quadruplet

$$(V,\ \vec{0}\in V,\ V\times V\to V,\ \mathbb{R}\times V\to V)$$

Pour simplifier les notations nous écrivons simplement V.

6.2 Exercice. Soit V un espace vectoriel réel.

1. Si $\vec{x}+\vec{z}=\vec{y}+\vec{z}$ alors $\vec{x}=\vec{y}$
2. $0\cdot\vec{x}=\vec{0}$ pour tout $\vec{x}\in V$.
3. $\alpha\cdot\vec{0}=\vec{0}$ pour tout $\alpha\in\mathbb{R}$.

Preuve.

1. $\vec{x}+\vec{z}=\vec{y}+\vec{z}\Rightarrow\vec{x}+\vec{z}-\vec{z}=\vec{y}+\vec{z}-\vec{z}\Rightarrow\vec{x}+\vec{0}=\vec{y}+\vec{0}\Rightarrow\vec{x}=\vec{y}$
2. $\vec{0}+0\cdot\vec{x}=0\cdot\vec{x}=(0+0)\cdot\vec{x}=0\cdot\vec{x}+0\cdot\vec{x}\Rightarrow\vec{0}=0\cdot\vec{x}$
3. $\vec{0}+\alpha\cdot\vec{0}=\alpha\cdot\vec{0}=\alpha\cdot(\vec{0}+\vec{0})=\alpha\cdot\vec{0}+\alpha\cdot\vec{0}\Rightarrow\vec{0}=\alpha\cdot\vec{0}$

$\square$

6.3 Exercice. Dans un espace vectoriel, le produit entre un scalaire et un vecteur est simplifiable :

1. Si $\alpha\cdot\vec{x}=\alpha\cdot\vec{y}$ avec $\alpha\neq 0$, alors $\vec{x}=\vec{y}$.
2. Si $\alpha\cdot\vec{x}=\beta\cdot\vec{x}$ avec $\vec{x}\neq\vec{0}$, alors $\alpha=\beta$.
3. En particulier, si $\alpha\cdot\vec{x}=\vec{0}$, alors $\alpha=0$ ou $\vec{x}=\vec{0}$.

La définition de sous-espace vectoriel de $\mathbb{R}^n$ (voir 3.3) garde son sens si on remplace $\mathbb{R}^n$ par un espace vectoriel arbitraire :

6.4 Définition. Soit V un espace vectoriel réel et W un sous-ensemble de V. On dit que W est un *sous-espace vectoriel de* V s'il remplit les conditions suivantes :

1. $\vec{0}\in W$
2. Si $\vec{x},\vec{y}\in W$ alors $\vec{x}+\vec{y}\in W$
3. Si $\vec{x}\in W$ et $\alpha\in\mathbb{R}$ alors $\alpha\cdot\vec{x}\in W$.

6.5 Exemple.

1. $\mathbb{R}^n$ avec les opérations définies en 2.10 est un espace vectoriel réel.

2. $\mathbb{R}^{m\times n}$ avec les opérations définies en 5.1 est un espace vectoriel réel.
3. L'ensemble des polynômes à une indéterminée
$$\mathbb{R}[x] = \{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \mid n \in \mathbb{N}, a_i \in \mathbb{R}\}$$
est un espace vectoriel réel par rapport à la somme usuelle de polynômes et à l'action définie par
$$\alpha \cdot (a_0 + a_1x + a_2x^2 + \ldots + a_nx^n) = \alpha a_0 + \alpha a_1x + \alpha a_2x^2 + \ldots + \alpha a_nx^n$$
4. De façon analogue, l'ensemble des polynômes à plusieurs indéterminées
$$\mathbb{R}[x_1, \ldots, x_k]$$
est un espace vectoriel réel.
5. Si X est un ensemble quelconque, l'ensemble des fonctions définies sur X et à valeurs réelles
$$\mathcal{F}(X, \mathbb{R}) = \{ \text{ fonctions } X \to \mathbb{R}\}$$
avec somme et action définies par
$$(f + g)(x) = f(x) + g(x)\ , \quad (\alpha \cdot f)(x) = \alpha f(x)$$
est un espace vectoriel réel.
6. Si V est un espace vectoriel et W est un s.e.v. de V, alors W est lui-même un espace vectoriel par rapport à la restriction à W de la somme et de l'action de V.
7. Pour chaque $n \in \mathbb{N}$ fixé, l'ensemble
$$\mathbb{R}[x]_n = \{ \text{ polynômes de degré } \leq n\}$$
est un s.e.v. de $\mathbb{R}[x]$. Et si $m \leq n$, alors $\mathbb{R}[x]_m$ est un s.e.v. de $\mathbb{R}[x]_n$.
8. Si X est un sous-ensemble convenable (voir cours d'analyse) de $\mathbb{R}$, alors
   (a) $Cont(X, \mathbb{R}) = \{ \text{ fonctions continues } X \to \mathbb{R}\}$
   (b) $Der(X, \mathbb{R}) = \{ \text{ fonctions dérivables } X \to \mathbb{R}\}$
   (c) $Int(X, \mathbb{R}) = \{ \text{ fonctions intégrables } X \to \mathbb{R}\}$

   sont des s.e.v. de $\mathcal{F}(X, \mathbb{R})$.

6.2 Espaces finiment engendrés

Il y a une différence essentielle entre les exemples 6.4.1, 6.4.2, 6.4.7 d'une part, et les exemples 6.4.3, 6.4.4, 6.4.5, 6.4.8 de l'autre part : les premiers admettent une base finie, les autres n'admettent pas de base finie. Nous devons donc adapter la terminologie utilisée pour $\mathbb{R}^n$ pour traiter ces nouveaux exemples.

6.6 Définition. Soit V un espace vectoriel réel et X un sous-ensemble de V.

1. X est une *famille génératrice* de V si tout vecteur de V est combinaison linéaire de vecteurs de X.
2. X est une *base* de V si X est une famille génératrice de V et si toute famille finie de vecteurs de X est une famille libre.
3. V est *finiment engendré* (ou *de type fini*) si V admet une famille génératrice finie.

Tout espace vectoriel admet une famille génératrice (il suffit de prendre $X = V$) mais certains espaces ne sont pas finiment engendrés :

6.7 Exemple.

1. La base canonique (voir 3.9) est une base finie de $\mathbb{R}^n$, qui est donc finiment engendré.
2. Soit $E_{ij} \in \mathbb{R}^{m\times n}$ la matrice dont tous les éléments sont nuls sauf l'élément de place (ij) qui vaut 1. Les $m \cdot n$ matrices E_{ij} forment une base finie de $\mathbb{R}^{m\times n}$, qui est donc finiment engendré.
3. Les $n+1$ polynômes
$$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2, \ldots, P_n(x) = x^n$$
forment une base finie de $\mathbb{R}[x]_n$, qui est donc finiment engendré.
4. $\mathbb{R}[x]$ n'est pas finiment engendré. En effet, soit X une famille finie de polynômes et soit
$$p = Max\,\{degP \mid P \in X\}$$
Toute combinaison linéaire de polynômes tirés de X aura degré plus petit ou égal à p. Par conséquent, un polynôme de degré plus grand que p ne peut pas être engendré à partir de X.
5. $\mathbb{R}[x]$ admet une base infinie : les polynômes
$$P_0(x) = 1, P_1(x) = x, P_2(x) = x^2, \ldots, P_n(x) = x^n, P_{n+1}(x) = x^{n+1}, \ldots$$
forment une base de $\mathbb{R}[x]$.

Les résultats concernants les notions de famille génératrice, famille libre, base et dimension énoncés dans le Chapitre 3 relativement aux s.e.v. de $\mathbb{R}^n$ restent valables pour tout espace vectoriel finiment engendré. Les voici en synthèse :

6.8 Propriété. *Soit V un espace vectoriel.*

1. *Si $\vec{x}_1, \ldots, \vec{x}_k$ est une famille génératrice de V et $\vec{y}_1, \ldots, \vec{y}_l$ est une famille libre, alors $l \leq k$.*
2. *Si $\vec{x}_1, \ldots, \vec{x}_k$ est une base de V, alors toute base de V contient k vecteurs. On pose $k = \dim V$.*
3. *Si V est finiment engendré, alors il admet une base finie.*

4. *Si* $\dim V = k$ *et si* $\vec{x}_1, \ldots, \vec{x}_l$ *est une famille génératrice de* V, *alors* $k \leq l$ *et on peut extraire de* $\vec{x}_1, \ldots, \vec{x}_l$ *une base de* V.
5. *Si* $\dim V = k$ *et si* $\vec{x}_1, \ldots, \vec{x}_l$ *est une famille libre, alors* $k \geq l$ *et on peut compléter la famille* $\vec{x}_1, \ldots, \vec{x}_l$ *pour obtenir une base de* V.
6. *Si* V *est finiment engendré, tout s.e.v. de* V *est finiment engendré*

Preuve. Les preuves vues dans le cas de $\mathbb{R}^n$ s'adaptent au cas général. En détail :

1. Voir Lemme 3.14.
2. Voir Propriété 3.15.
3. Voir Propriété 3.21.
4. Voir Exercice 3.18.1.
5. Voir Exercice 3.18.2.
6. Voir Propriété 3.21.

□

6.9 Théorème. *Tout espace vectoriel réel admet une base (éventuellement infinie).*

Preuve. La preuve de ce théorème est une version transfinie de la preuve de la Propriété 3.21. Elle dépasse le cadre de ce cours. □

6.10 Remarque. Si V est un espace vectoriel et $\vec{x}_1, \ldots, \vec{x}_k$ est une base de V, nous savons que toute autre base de V contient k vecteurs, ce qui nous permet de parler de dimension d'un espace de type fini (voir 6.8). Dans le cas des espaces non nécessairement de type fini, ce qu'on peut montrer est que si un espace admet deux bases X et Y, alors les éléments de X sont en *bijection* avec les éléments de Y. La notion de bijection entre ensembles et entre espaces vectoriels sera étudiée au Chapitre 8 (pour deux ensembles finis, être en bijection revient à avoir le même nombre d'éléments), mais la théorie de la base pour les espace non finiment engendrés ne sera pas développée dans ce cours.

Chapitre 7

Applications linéaires

Le chapitre 7 section par section

Dans le Chapitre 6 nous avons étudié les espaces vectoriels. Les applications linéaires sont les fonctions entre espaces vectoriels qui préservent la structure d'espace vectoriel.

1. A toute matrice $A \in \mathbb{R}^{m \times n}$ on peut associer une fonction $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$. Ceci est l'exemple guide d'application linéaire entre espaces vectoriels.
2. Pour chaque application linéaire, on peut introduire la notion de fibre, noyau et image. Le noyau et l'image sont des sous-espaces vectoriels.
3. L'exemple L_A permet d'appliquer les résultats généraux sur les applications linéaires aux systèmes d'équations linéaires.
4. Les applications linéaires entre deux espaces vectoriels fixés forment un espace vectoriel. De plus, les applications linéaires peuvent se composer en donnant des nouvelles applications linéaires.
5. Dans le passage $A \mapsto L_A$ entre matrices et applications linéaires, le produit matriciel correspond à la composition d'applications linéaires. Ceci permet entre autre de montrer que le produit matriciel est associatif.

7.1 Définition et exemples

7.1 Définition. Soit V, W deux espaces vectoriels réels et soit

$$L \colon V \to W$$

une fonction. La fonction L est une *application linéaire* si pour tout $\vec{x}, \vec{y} \in V$ et pour tout $\alpha \in \mathbb{R}$ on a :

1. $L(\alpha \cdot \vec{x}) = \alpha \cdot L(\vec{x})$
2. $L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y})$

7.2 Exercice. Soit $L\colon V \to W$ une application linéaire. Montrer que si $\vec{x}_1, \ldots, \vec{x}_k \in V$ sont linéairement dépendants, alors $L(\vec{x}_1), \ldots, L(\vec{x}_k) \in W$ sont linéairement dépendants.

7.3 Exemple. Soit $A \in \mathbb{R}^{m\times n}$. L'application

$$L_A\colon \mathbb{R}^n \to \mathbb{R}^m\ ,\ \ L_A(\vec{x}) = A\cdot\vec{x}$$

est linéaire. En effet, par la Propriété 5.7 on a :

$$L_A(\alpha\cdot\vec{x}) = A\cdot(\alpha\cdot\vec{x}) = \alpha\cdot(A\cdot\vec{x}) = \alpha\cdot L_A(\vec{x})$$
$$L_A(\vec{x}+\vec{y}) = A\cdot(\vec{x}+\vec{y}) = A\cdot\vec{x} + A\cdot\vec{y} = L_A(\vec{x}) + L_A(\vec{y})$$

7.4 Exemple.

1. Exemples d'applications linéaires $L\colon \mathbb{R}^2 \to \mathbb{R}^2$:
   (a) Rotation autour de l'origine.
   (b) Projection orthogonale sur une droite passant par l'origine.
   (c) Symétrie orthogonale par rapport à une droite passant par l'origine.
2. Exemples d'applications linéaires $L\colon \mathbb{R}^3 \to \mathbb{R}^3$:
   (a) Rotation autour d'une droite passant par l'origine.
   (b) Projection orthogonale sur une droite passant par l'origine ou sur un plan passant par l'origine.
   (c) Symétrie orthogonale par rapport à une droite passant par l'origine ou par rapport à un plan passant par l'origine.
3. Exemples d'applications linéaires $L\colon \mathbb{R} \to \mathbb{R}$: les seules fonctions $\mathbb{R} \to \mathbb{R}$ qui sont linéaires sont les applications du type

   $$L(x) = \alpha\cdot x$$

   avec $\alpha \in \mathbb{R}$ fixé. En effet, si $L\colon \mathbb{R} \to \mathbb{R}$ est linéaire et si l'on pose $\alpha = L(1)$, on a

   $$L(x) = L(x\cdot 1) = x\cdot L(1) = x\cdot\alpha$$
4. Les translations

   $$T\colon \mathbb{R}^n \to \mathbb{R}^n\ ,\ \ T(\vec{x}) = \vec{x} + \vec{x}_0$$

   avec $\vec{x}_0 \in \mathbb{R}^n$ fixé ne sont pas des applications linéaires (sauf si $\vec{x} = \vec{0}$).
5. L'application

   $$L\colon \mathbb{R}[x]_n \to \mathbb{R}^{n+1}\ ,\ \ L(a_0 + a_1x + a_2x^2 \ldots + a_nx^n) = (a_0, a_1, \ldots, a_n)$$

   est une application linéaire.
6. (a) La primitivation $\int\colon Cont(\,[a,b],\mathbb{R}) \to \mathcal{F}([a,b],\mathbb{R})$ est une application linéaire.
   (b) L'intégration $\int_a^b\colon Int(\,[a,b],\mathbb{R}) \to \mathbb{R}$ est une application linéaire.
   (c) La dérivation $\partial\colon Der(\,]a,b[,\mathbb{R}) \to \mathcal{F}(]a,b[,\mathbb{R})$ est une application linéaire.
7. Si W est un s.e.v. de V, l'inclusion $W \to V$ est une aplication linéaire.

7.5 Exercice. Soit $L\colon V \to W$ une application linéaire. Montrer que $L(\vec{0}\,) = \vec{0}$.

Preuve. $\vec{0} + L(\vec{0}\,) = L(\vec{0}\,) = L(\vec{0} + \vec{0}\,) = L(\vec{0}\,) + L(\vec{0}\,)$ $\square$

7.2 Fibre, noyau et image

7.6 Définition. Soit $L\colon V \to W$ une application linéaire.

1. Soit $\vec{b} \in W$ fixé. La *fibre* de L au point $\vec{b}$ est le sous-ensemble de V défini par
$$\mathcal{F}_{\vec{b}} = \{\vec{x} \in V \ \text{ tels que } \ L(\vec{x}) = \vec{b}\,\}$$
2. Le *noyau* de L est le sous-ensemble de V défini par
$$\operatorname{Ker} L = \mathcal{F}_{\vec{0}} = \{\vec{x} \in V \ \text{ tels que } \ L(\vec{x}) = \vec{0}\,\}$$
3. L'*image* de L est le sous-ensemble de W défini par
$$\operatorname{Im} L = \{L(\vec{x}) \ \text{ tels que } \ \vec{x} \in V\}$$

7.7 Remarque. On peut décrire l'image $\operatorname{Im} L$ de deux autres façons :
$$\operatorname{Im} L = \{\vec{y} \in W \ \text{ tels que il existe } \ \vec{x} \in V \ \text{ tel que } \ L(\vec{x}) = \vec{y}\}$$
$$\operatorname{Im} L = \{\vec{y} \in W \ \text{ tels que } \ \mathcal{F}_{\vec{y}} \neq \emptyset\}$$

7.8 Exemple.

1. Soit
$$L\colon \mathbb{R}^2 \to \mathbb{R}\ , \ \ L(x_1, x_2) = x_1$$
Pour tout $b \in \mathbb{R}$, la fibre de L au point b est $\mathcal{F}_b = \{(b, x_2) \mid x_2 \in \mathbb{R}\}$
2. Soit
$$L\colon \mathbb{R} \to \mathbb{R}^2\ , \ \ L(x) = (x, 0)$$
et soit $\vec{b} = (a, b) \in \mathbb{R}^2$ fixé. Si $b \neq 0$, alors la fibre $\mathcal{F}_{\vec{b}}$ est vide. Si $b = 0$ alors $\mathcal{F}_{\vec{b}} = \{a\}$.
3. Soit $D\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_{n-1}$ la dérivation
$$D(a_0 + a_1x + a_2x^2 + \ldots + a_nx^n) = a_1 + 2a_2x + \ldots + na_nx^{n-1}$$
   - Soit $b_0 + b_1x + b_2x^2 + \ldots + b_{n-1}x^{n-1} = \vec{b} \in \mathbb{R}[x]_{n-1}$ fixé. La fibre de L en $\vec{b}$ est l'ensemble des primitives du polynôme $\vec{b}$:
$$\mathcal{F}_{\vec{b}} = \{b_0x + \frac{1}{2}b_1x^2 + \frac{1}{3}b_2x^3 + \ldots + \frac{1}{n}b_{n-1}x^n + k \mid k \in \mathbb{R}\}$$
   - Le noyau de L est $\operatorname{Ker} L = \mathbb{R}[x]_0 = \mathbb{R}$.

7.9 Propriété. *Soit $L\colon V \to W$ une application linéaire.*

1. $\operatorname{Ker} L$ *est un s.e.v. de* V.

2. $\operatorname{Im} L$ *est un s.e.v. de* W.

Preuve. Voyons la deuxième condition de la définition de s.e.v. (6.4) :

1. Si $\vec{x}, \vec{y} \in \operatorname{Ker} L$, alors $L(\vec{x}+\vec{y}) = L(\vec{x}) + L(\vec{y}) = \vec{0} + \vec{0} = \vec{0}$ et donc $\vec{x}+\vec{y} \in \operatorname{Ker} L$.
2. Si $\vec{a}, \vec{b} \in \operatorname{Im} L$, alors il existe $\vec{x}, \vec{y} \in V$ tels que $L(\vec{x}) = \vec{a}$ et $L(\vec{y}) = \vec{b}$. Par conséquent, $L(\vec{x}+\vec{y}) = L(\vec{x}) + L(\vec{y}) = \vec{a}+\vec{b}$ et donc $\vec{a}+\vec{b} \in \operatorname{Im} L$.

□

7.10 Notation. Soit V un espace vectoriel, U un s.e.v. de V et $\vec{x}_0 \in V$ un vecteur fixé. On pose

$$\vec{x}_0 + U = \{\vec{x}_0 + \vec{u} \mid \vec{u} \in U\}$$

Attention : en général $\vec{x}_0 + U$ n'est pas un s.e.v.

7.11 Propriété. *Soit $L\colon V \to W$ une application linéaire et soit $\vec{b} \in W$.*

1. *$\mathcal{F}_{\vec{b}} \neq \emptyset$ si et seulement si $\vec{b} \in \operatorname{Im} L$.*
2. *Soit $\vec{b} \in \operatorname{Im} L$ et soit $\vec{x}_0 \in \mathcal{F}_{\vec{b}}$ fixé. On a : $\mathcal{F}_{\vec{b}} = \vec{x}_0 + \operatorname{Ker} L$.*

Preuve. $\vec{b} \in \operatorname{Im} L \;\Leftrightarrow\; \exists \vec{x} \in V \mid L(\vec{x}) = \vec{b} \;\Leftrightarrow\; \exists \vec{x} \in \mathcal{F}_{\vec{b}} \;\Leftrightarrow\; \mathcal{F}_{\vec{b}} \neq \emptyset$
Pour montrer $\mathcal{F}_{\vec{b}} = \vec{x}_0 + \operatorname{Ker} L$ montrons la double inclusion :

$\vec{x}_0 + \operatorname{Ker} L \subseteq \mathcal{F}_{\vec{b}}$: soit $\vec{y} \in \vec{x}_0 + \operatorname{Ker} L$, alors $\vec{y} = \vec{x}_0 + \vec{x}$ pour un certain $\vec{x} \in \operatorname{Ker} L$. On a que $L(\vec{y}) = L(\vec{x}_0 + \vec{x}) = L(\vec{x}_0) + L(\vec{x}) = \vec{b} + \vec{0} = \vec{b}$ et donc $\vec{y} \in \mathcal{F}_{\vec{b}}$.

$\mathcal{F}_{\vec{b}} \subseteq \vec{x}_0 + \operatorname{Ker} L$: soit $\vec{y} \in \mathcal{F}_{\vec{b}}$. On peut écrire $\vec{y} = \vec{x}_0 + (\vec{y} - \vec{x}_0)$ d'où $\vec{y} \in \vec{x}_0 + \operatorname{Ker} L$ car $\vec{y} - \vec{x}_0 \in \operatorname{Ker} L$. En effet, $L(\vec{y} - \vec{x}_0) = L(\vec{y}) - L(\vec{x}_0) = \vec{b} - \vec{b} = \vec{0}$.

□

7.12 Exercice. Soit $L\colon V \to W$ une application linéaire et soit $X \subseteq V$ une famille génératrice de V. Montrer que $\{L(\vec{x}) \mid \vec{x} \in X\}$ est une famille génératrice de $\operatorname{Im} L$. Autrement dit : $\operatorname{Im} L$ est le plus petit s.e.v. de W qui contient $L(\vec{x})$ pour tout $\vec{x} \in X$.

Preuve. Un vecteur $\vec{b}$ est dans $\operatorname{Im} L$ ssi il existe $\vec{v} \in V$ tel que $L(\vec{v}) = \vec{b}$. Puisque X est une famille génératrice de V, nous pouvons écrire $\vec{v}$ comme combinaison linéaire de $\vec{x}_1, \ldots, \vec{x}_n$ avec $\vec{x}_1, \ldots, \vec{x}_n \in X$:

$$\vec{v} = \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_n \cdot \vec{x}_n$$

Par linéarité, nous avons

$$\vec{b} = L(\vec{v}) = L(\alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_n \cdot \vec{x}_n) = \alpha_1 \cdot L(\vec{x}_1) + \ldots + \alpha_n \cdot L(\vec{x}_n)$$

Cela montre que $\{L(\vec{x}) \mid \vec{x} \in X\}$ est une famille génératrice de $\operatorname{Im} L$. □

7.3 Application aux systèmes linéaires

7.13 Exemple. Cet exemple explique pourquoi les applications linéaires sont l'outil approprié pour étudier les systèmes d'équations linéaires.
Soit $A \in \mathbb{R}^{m\times n}$ une matrice fixée et soit

$$L_A \colon \mathbb{R}^n \to \mathbb{R}^m \ , \quad L_A(\vec{x}) = A \cdot \vec{x}$$

l'application linéaire associée à la matrice A.

1. Pour tout $\vec{b} \in \mathbb{R}^m$, la fibre $\mathcal{F}_{\vec{b}}$ est précisement l'ensemble des solutions du système $\mathcal{S} \colon A \cdot \vec{x} = \vec{b}$
$$\mathcal{F}_{\vec{b}} = \mathrm{Sol}(\mathcal{S})$$
En effet, $\vec{x} \in \mathrm{Sol}(\mathcal{S}) \Leftrightarrow A \cdot \vec{x} = \vec{b} \Leftrightarrow L_A(\vec{x}) = \vec{b} \Leftrightarrow \vec{x} \in \mathcal{F}_{\vec{b}}$.
2. Le noyau $\mathrm{Ker}\, L_A$ est précisement l'ensemble des solutions du système homogène associé $\mathcal{S}_0 \colon A \cdot \vec{x} = \vec{0}$
$$\mathrm{Ker}\, L_A = \mathrm{Sol}(\mathcal{S}_0)$$
En effet, $\vec{x} \in \mathrm{Sol}(\mathcal{S}_0) \Leftrightarrow A \cdot \vec{x} = \vec{0} \Leftrightarrow L_A(\vec{x}) = \vec{0} \Leftrightarrow \vec{x} \in \mathrm{Ker}\, L_A$.

7.14 Corollaire. *Soit $\mathcal{S}_0 \colon A \cdot \vec{x} = \vec{0}$ un système homogène à n variables. L'ensemble* $\mathrm{Sol}(\mathcal{S}_0)$ *de ses solutions est un s.e.v. de* $\mathbb{R}^n$.

Preuve. On a $\mathrm{Sol}(\mathcal{S}_0) = \mathrm{Ker}\, L_A$ par 7.13, et $\mathrm{Ker}\, L_A$ est un s.e.v. par 7.9. □

7.15 Corollaire. *Soit $A \in \mathbb{R}^{m\times n}$ une matrice fixée et soit $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ l'application linéaire associée à A. Alors :*

$$\mathrm{Col}(A) = \mathrm{Im}\, L_A$$

Preuve. Soit $\vec{b} \in \mathbb{R}^m$. Nous savons que $\vec{b} \in \mathrm{Im}\, L_A$ ssi $\mathcal{F}_{\vec{b}} \neq \emptyset$ (voir 7.11) et que $\vec{b} \in \mathrm{Col}(A)$ ssi $\mathrm{Sol}(\mathcal{S}) \neq \emptyset$, où on considère le système $\mathcal{S} \colon A \cdot \vec{x} = \vec{b}$ (voir 4.22). Puisque $\mathcal{F}_{\vec{b}} = \mathrm{Sol}(\mathcal{S})$, (voir 7.13), nous avons que $\vec{b} \in \mathrm{Im}\, L_A$ ssi $\vec{b} \in \mathrm{Col}(A)$, c'est-à-dire $\mathrm{Im}\, L_A = \mathrm{Col}(A)$. □

7.16 Corollaire. *Soit $\mathcal{S} \colon A \cdot \vec{x} = \vec{b}$ un système d'équations linéaires et soit $\vec{x}_0$ une solution de $\mathcal{S}$. Alors :*

$$\mathrm{Sol}(\mathcal{S}) = \vec{x}_0 + \mathrm{Sol}(\mathcal{S}_0)$$

où $\mathcal{S}_0 \colon A \cdot \vec{x} = \vec{0}$ est le système homogène associé.

Preuve. Par 7.13 et 7.11 on a $\mathrm{Sol}(\mathcal{S}) = \mathcal{F}_{\vec{b}} = \vec{x}_0 + \mathrm{Ker}\, L_A = \vec{x}_0 + \mathrm{Sol}(\mathcal{S}_0)$ □

7.17 Exercice. Soit $A \in \mathbb{R}^{m\times n}$ une matrice fixée et soit $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ l'application linéaire associée à A. En utilisant l'Exercice 7.12, montrer à nouveau que $\mathrm{Col}(A) = \mathrm{Im}\, L_A$ (voir 7.15).

Preuve. Considérons la base canonique $\vec{e}_1, \dots, \vec{e}_n$ de $\mathbb{R}^n$. Par l'Exercice 7.12, $L_A(\vec{e}_1), \dots, L_A(\vec{e}_n)$ est une famille génératrice de $\operatorname{Im} L_A$:

$$\operatorname{Im} L_A = \langle L_A(\vec{e}_1), \dots, L_A(\vec{e}_n)\rangle$$

Par la Remarque 5.10 on a pour tout $i = 1, \dots, n$

$$L_A(\vec{e}_i) = A \cdot \vec{e}_i = i\text{-ème colonne de } A$$

et donc $\operatorname{Im} L_A$ est l'espace engendré par les colonnes de A. $\square$

7.18 Corollaire. *Un système $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ admet au moins une solution si et seulement si $\vec{b} \in \operatorname{Col}(A)$.*

Preuve. $\mathcal{S}$ admet au moins une solution ssi il existe $\vec{x} \in \mathbb{R}^n$ tel que $A \cdot \vec{x} = \vec{b}$ ssi il existe $\vec{x} \in \mathbb{R}^n$ tel que $L_A(\vec{x}) = \vec{b}$ ssi $\vec{b} \in \operatorname{Im} L_A = \operatorname{Col}(A)$. $\square$

7.4 Opérations sur les applications linéaires

Pour le restant de ce chapitre on va étudier les opérations qu'on peut effectuer sur les applications linéaires.

7.19 Propriété. *Soient V et W deux espaces vectoriels réels. L'ensemble*

$$\operatorname{Lin}(V, W)$$

des applications linéaires de V à W est un espace vectoriel réel par rapport aux opérations

$$\operatorname{Lin}(V, W) \times \operatorname{Lin}(V, W) \to \operatorname{Lin}(V, W)\ ,\quad (L, T) \mapsto L + T$$

$$\mathbb{R} \times \operatorname{Lin}(V, W) \to \operatorname{Lin}(V, W)\ ,\quad (\alpha, L) \mapsto \alpha \cdot L$$

définies par

$$(L + T)(\vec{x}) = L(\vec{x}) + T(\vec{x}) \qquad (\alpha \cdot L)(\vec{x}) = \alpha \cdot L(\vec{x})$$

Preuve. Il faut montrer que $L + T$ et $\alpha \cdot L$ sont des applications linéaires et que les deux opérations ci-dessus satisfont aux conditions de la définition d'espace vectoriel (voir 6.1). Nos montrons ici seulement que $L + T$ satisfait à une des conditions de la définition d'application linéaire (voir 7.1). Le restant de la preuve est laissé en exercice.

$$(L + T)(\vec{x} + \vec{y}) = L(\vec{x} + \vec{y}) + T(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y}) + T(\vec{x}) + T(\vec{y}) =$$

$$= L(\vec{x}) + T(\vec{x}) + L(\vec{y}) + T(\vec{y}) = (L + T)(\vec{x}) + (L + T)(\vec{y})$$

$\square$

7.20 Propriété. *Soit $L\colon U \to V, T\colon V \to W$ et $R\colon W \to Z$ des applications linéaires entre espaces vectoriels réels.*

1. *L'identité*
$$id_V : V \to V \ , \quad id_V(\vec{x}) = \vec{x}$$
est une application linéaire.
2. *La composée*
$$T \circ L : U \to W \ , \quad (T \circ L)(\vec{x}) = T(L(\vec{x}))$$
est une application linéaire.
3. *On a* $\quad L \circ id_U = L = id_V \circ L \quad$ *et* $\quad R \circ (T \circ L) = (R \circ T) \circ L.$

Preuve.

1. et 3. Evidents.
2. $(T \circ L)(\vec{x} + \vec{y}) = T(L(\vec{x} + \vec{y})) = T(L(\vec{x}) + L(\vec{y})) = T(L(\vec{x})) + T(L(\vec{y})) = (T \circ L)(\vec{x}) + (T \circ L)(\vec{y})$.
$(T \circ L)(\alpha \cdot \vec{x}) = T(L(\alpha \cdot \vec{x})) = T(\alpha \cdot L(\vec{x})) = \alpha \cdot T(L(\vec{x}) = \alpha \cdot (T \circ L)(\vec{x})$.

$\square$

7.21 Définition.

1. Une $\mathbb{R}$-*algèbre* est un espace vectoriel réel $V = (V, \vec{0}, +, \cdot)$ muni d'une opération de produit
$$V \times V \to V \ , \quad (\vec{x}, \vec{y}) \mapsto \vec{x} \circ \vec{y}$$
et d'un élément unité $\vec{1} \in V$ tels que pour tout $\vec{x}, \vec{y}, \vec{z} \in V$ et pour tout $\alpha \in \mathbb{R}$
   (a) $\vec{x} \circ (\vec{y} \circ \vec{z}) = (\vec{x} \circ \vec{y}) \circ \vec{z}$
   (b) $\vec{1} \circ \vec{x} = \vec{x} = \vec{x} \circ \vec{1}$
   (c) $\vec{x} \circ (\vec{y} + \vec{z}) = \vec{x} \circ \vec{y} + \vec{x} \circ \vec{z}$
   (d) $(\vec{x} + \vec{y}) \circ \vec{z} = \vec{x} \circ \vec{z} + \vec{y} \circ \vec{z}$
   (e) $\vec{x} \circ (\alpha \cdot \vec{y}) = \alpha \cdot (\vec{x} \circ \vec{y}) = (\alpha \cdot \vec{x}) \circ \vec{y}$
2. Si V et W sont deux $\mathbb{R}$-algèbres, un *homomorphisme* de $\mathbb{R}$-algèbres $F : V \to W$ est une application linéaire telle que pour tout $\vec{x}, \vec{y} \in V$
   (a) $F(\vec{x} \circ \vec{y}) = F(\vec{x}) \circ F(\vec{y})$
   (b) $F(\vec{1}_V) = \vec{1}_W$

7.22 Exemple. $\mathbb{R}^{n \times n}$ est une $\mathbb{R}$-algèbre, voir 5.7.

7.23 Propriété. *Soit V un espace vectoriel réel. L'espace vectoriel* $\mathrm{Lin}(V, V)$ *des application linéaires de V vers V est une $\mathbb{R}$-algèbre. Le produit est la composition d'applications linéaires, l'élément unité est l'application linéaire identité (voir 7.20).*

Preuve. Voyons par exemple la condition (c) de la Definition 7.21 :
$$(L \circ (R + S))(\vec{x}) = L((R + S)(\vec{x})) = L(R(\vec{x}) + S(\vec{x})) =$$
$$= L(R(\vec{x})) + L(S(\vec{x})) = (L \circ R)(\vec{x}) + (L \circ S)(\vec{x})$$
Ceci étant vrai pour tout $\vec{x} \in V$, on a $L \circ (R + S) = L \circ R + L \circ S$. $\square$

7.5 Produit matriciel et composition

Nous voulons montrer ici que le produit matriciel correspond, dans le passage

$$A \in \mathbb{R}^{m\times n} \mapsto L_A \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$$

à la composition d'applications linéaires. Commençons par observer qu'une application linéaire est complètement déterminée par sa restriction à une famille génératrice :

7.24 Lemme. *Soit X une famille génératrice d'un espace vectoriel V et soit $L, T\colon V \to W$ deux applications linéaires. Si $L(\vec{x}) = T(\vec{x})$ pour tout $\vec{x} \in X$, alors $L = T$.*

Preuve. Tout $\vec{v} \in V$ est combinaison linéaire de vecteurs de X :

$$\vec{v} = \alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_n \cdot \vec{x}_n \qquad \text{avec } \vec{x}_i \in X \ (i = 1, \ldots, n)$$

Par conséquent, grâce à la linéarité de L et T nous obtenons

$$L(\vec{v}) = L(\alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_n \cdot \vec{x}_n) = \alpha_1 \cdot L(\vec{x}_1) + \ldots + \alpha_n \cdot L(\vec{x}_n) =$$

$$= \alpha_1 \cdot T(\vec{x}_1) + \ldots + \alpha_n \cdot T(\vec{x}_n) = T(\alpha_1 \cdot \vec{x}_1 + \ldots + \alpha_n \cdot \vec{x}_n) = T(\vec{v})$$

□

La prochaine propriété justifie la définition du produit matriciel : il correspond à la composition entre applications linéaires.

7.25 Propriété. *Soit $A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{n\times p}$ deux matrices et soit $A{\cdot}B \in \mathbb{R}^{m\times p}$ la matrice produit (5.5). Considérons les applications linéaires associées (7.3)*

$$L_A\colon \mathbb{R}^n \to \mathbb{R}^m \qquad L_B\colon \mathbb{R}^p \to \mathbb{R}^n \qquad L_{A\cdot B}\colon \mathbb{R}^p \to \mathbb{R}^m$$

On a que $L_A \circ L_B = L_{A\cdot B}$.

Preuve. En utilisant le Lemme 7.24, il suffit de montrer que $(L_A \circ L_B)(\vec{e}_i) = L_{A\cdot B}(\vec{e}_i)$ pour $\vec{e}_1, \ldots, \vec{e}_p$ la base canonique de $\mathbb{R}^p$. La Remarque 5.10 nous donne

$$(L_A \circ L_B)(\vec{e}_i) = L_A(L_B(\vec{e}_i)) = A \cdot (B \cdot \vec{e}_i) = A \cdot (i\text{-ème colonne de } B) =$$

$$= i\text{-ème colonne de } A \cdot B = (A \cdot B) \cdot \vec{e}_i = L_{A\cdot B}(\vec{e}_i)$$

□

7.26 Lemme. *Soit $A, B \in \mathbb{R}^{m\times n}$ et considérons les applications linéaires associées :*

$$L_A\colon \mathbb{R}^n \to \mathbb{R}^m\ ,\ \ L_B\colon \mathbb{R}^n \to \mathbb{R}^m\ .$$

Si $L_A = L_B$ alors $A = B$.

Preuve. Si $L_A(\vec{x}) = L_B(\vec{x})$ pour tout $\vec{x} \in \mathbb{R}^n$, nous pouvons choisir comme $\vec{x}$ les n vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ de la base canonique de $\mathbb{R}^n$. On a alors

$$i\text{-ème colonne de } A = A \cdot \vec{e}_i = L_A(\vec{e}_i) = L_B(\vec{e}_i) = B \cdot \vec{e}_i = i\text{-ème colonne de } B$$

d'où $A = B$. □

7.27 Corollaire. *Le produit matriciel est associatif.*

Preuve. Soit $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}$ et $C \in \mathbb{R}^{p\times q}$. Par la Propriété 7.25 et le caractère associatif de la compositions de fonctions (7.20), nous avons

$$L_{A\cdot(B\cdot C)} = L_A \circ L_{B\cdot C} = L_A \circ (L_B \circ L_C) = (L_A \circ L_B) \circ L_C = L_{A\cdot B} \circ L_C = L_{(A\cdot B)\cdot C}$$

et par le Lemme 7.26 nous pouvons conclure que $A \cdot (B \cdot C) = (A \cdot B) \cdot C$. □

A chaque matrice $A \in \mathbb{R}^{m\times n}$ nous pouvons associer une application linéaire $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ définie par $L_A(\vec{x}) = A \cdot \vec{x}$ (voir 7.3). On a ainsi une fonction

$$\varepsilon \colon \mathbb{R}^{m\times n} \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)\ , \quad \varepsilon(A) = L_A$$

7.28 Propriété. *Soit* $\varepsilon \colon \mathbb{R}^{m\times n} \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ *définie par* $\varepsilon(A) = L_A$.

1. $\varepsilon \colon \mathbb{R}^{m\times n} \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ *est une application linéaire.*
2. $\varepsilon \colon \mathbb{R}^{n\times n} \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ *est un homomorphisme de* $\mathbb{R}$*-algèbres.*

Preuve. Il faut montrer quatre conditions (voir 7.1 et 7.21) :

1. $\varepsilon(A + B) = \varepsilon(A) + \varepsilon(B)$
2. $\varepsilon(\alpha \cdot A) = \alpha \cdot \varepsilon(A)$
3. $\varepsilon(A \cdot B) = \varepsilon(A) \circ \varepsilon(B)$
4. $\varepsilon(I) = id_{\mathbb{R}^n}$

La condition 3. est donnée par la Propriété 7.25. Voyons par exemple la condition 1.

$$\begin{aligned} \varepsilon(A + B)(\vec{x}) &= L_{A+B}(\vec{x}) = (A + B) \cdot \vec{x} = A \cdot \vec{x} + B \cdot \vec{x} = \\ &= L_A(\vec{x}) + L_B(\vec{x}) = (L_A + L_B)(\vec{x}) = (\varepsilon(A) + \varepsilon(B))(\vec{x}) \end{aligned}$$

Ceci étant vrai pour tout $\vec{x} \in \mathbb{R}^n$, on a $\varepsilon(A + B) = \varepsilon(A) + \varepsilon(B)$. □

Chapitre 8

Théorème de représentation

Le chapitre 8 section par section

L'objectif de ce chapitre est de démontrer le théorème de représentation suivant : Les espaces vectoriels $\mathbb{R}^{m\times n}$ et $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ sont *isomorphes.*

1. Un isomorphisme est un cas particulier de fonction bijective. Commençons donc par étudier les fonctions injectives, surjectives et bijectives.
2. Un isomorphisme est une application linéaire bijective. Deux espaces vectoriels sont isomorphes ssi ont même dimension.
3. Si $\dim V = n$ et $\dim W = m$, le passage $A \mapsto L_A$ entre matrices et applications linéaires (déjà introduit dans le Chapitre 7) est en effet un isomorphisme $\mathbb{R}^{m\times n} \simeq \mathrm{Lin}(V, W)$. Pour expliciter cet isomorphisme, il faut fixer une base de V et une base de W.
4. Le passage inverse $L \mapsto {}_fL_e$ donne la matrice associée à une application linéaire $L\colon V \to W$ par rapport à une base e de V et à une base f de W.
5. Si on change les bases de V et W, le lien entre les deux matrices associées à une même application linéaire s'exprime par un produit matriciel.
6. Pour terminer ce chapitre, un clin d'oeil à la théorie des catégories.

8.1 Functions injectives, surjectives, bijectives

Pour arriver au théorème de représentation qui est l'objectif de ce chapitre, nous devons tout d'abord fixer la terminologie.

8.1 Définition. Soit $f\colon X \to Y$ une fonction entre deux ensembles.

1. f est *injective* si pour tout $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implique $x_1 = x_2$.
2. f est *surjective* si pour tout $y \in Y$ il existe un $x \in X$ tel que $f(x) = y$.
3. f est *bijective* (ou f est une *bijection*) si elle est injective et surjective.

8.2 Remarque. Toute fonction $f\colon X \to Y$ peut se décomposer comme une fonction surjective suivie par une fonction injective : $f = f_i \circ f_s$

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "f_s"'] & & Y \\ & \operatorname{Im} f \arrow[ur, "f_i"'] & \end{tikzcd}$$

Il suffit pour cela de poser $f_s(x) = f(x)$ pour tout $x \in X$ et $f_i(y) = y$ pour tout $y \in \operatorname{Im} f$.
Par exemple, la fonction $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas surjective, mais elle devient surjective si on considère comme ensemble d'arrive l'ensemble $\mathbb{R}^+$ des réels positifs. Autrement dit, $f = f_i \circ f_s$ où $f_s\colon \mathbb{R} \to \mathbb{R}^+$ est définie par $f_s(x) = x^2$ et $f_i\colon \mathbb{R}^+ \to \mathbb{R}$ est simplement l'inclusion de $\mathbb{R}^+$ dans $\mathbb{R}$.

8.3 Exercice. Considérons deux fonctions

$$X \xrightarrow{f} Y \xrightarrow{g} Z\,.$$

1. Si f et g sont injectives, alors $g \circ f$ est injective.
2. Si f et g sont surjectives, alors $g \circ f$ est surjective.
3. Si $g \circ f$ est injective, alors f est injective.
4. Si $g \circ f$ est surjective, alors g est surjective.

8.4 Exercice. Soit $f\colon X \to Y$ une fonction.

1. Si $X \neq \emptyset$, les conditions suivantes sont équivalentes :
   (a) f est injective,
   (b) pour tout $h, k\colon Z \to X$, si $f \circ h = f \circ k$ alors $h = k$,
   (c) il existe une fonction $g\colon Y \to X$ telle que $g \circ f = id_X$

$$\begin{tikzcd} X \arrow[rr, "id_X"] \arrow[dr, "f"'] & & X \\ & Y \arrow[ur, "g"'] & \end{tikzcd}$$

2. Les conditions suivantes sont équivalentes :
   (a) f est surjective,
   (b) pour tout $h, k\colon Y \to Z$, si $h \circ f = k \circ f$ alors $h = k$,
   (c) il existe une fonction $g\colon Y \to X$ telle que $f \circ g = id_Y$

$$\begin{tikzcd} Y \arrow[rr, "id_Y"] \arrow[dr, "g"'] & & Y \\ & X \arrow[ur, "f"'] & \end{tikzcd}$$

8.5 Lemme. Considérons trois fonctions

$$Y \xrightarrow{g_1} X \xrightarrow{f} Y \xrightarrow{g_2} X$$

Si $f \circ g_1 = id_Y$ et $g_2 \circ f = id_X$, alors $g_1 = g_2$.

Preuve. Puisque $f \circ g_1 = id_Y$, f est surjective (8.4.2) ; donc pour montrer que $g_1 = g_2$ il suffit de montrer que $g_1 \circ f = g_2 \circ f$ (8.4.2). Puisque $g_2 \circ f = id_X$, f est injective (8.4.1) ; donc pour montrer que $g_1 \circ f = g_2 \circ f$ il suffit de montrer que $f \circ g_1 \circ f = f \circ g_2 \circ f$ (8.4.1). Or,

$$f \circ g_1 \circ f = id_Y \circ f = f = f \circ id_X = f \circ g_2 \circ f\,.$$

□

8.6 Corollaire. *Une fonction $f\colon X \to Y$ est bijective si et seulement si il existe une fonction $g\colon Y \to X$ telle que $g \circ f = id_X$ et $f \circ g = id_Y$.*

Preuve. Par l'Exercice 8.4, f est injective et surjective si et seulement si il existe $g_1, g_2\colon Y \to X$ telles que $f \circ g_1 = id_Y$ et $g_2 \circ f = id_X$. Par le Lemme 8.5 on a que $g_1 = g_2$ et le corollaire est démontré. □

8.7 Définition. Deux ensembles X et Y sont en *bijection* (notation : $X \simeq Y$) s'il existe une bijection $f\colon X \to Y$ où, de façon équivalente, s'il existe un bijection $g\colon Y \to X$.

8.8 Exercice. Montrer que deux ensembles finis sont en bijection si et seulement si ils contiennent le même nombre d'éléments.

La Définition 8.1 et les résultats 8.3, 8.4, 8.5 et 8.6 sont valables en particulier pour les applications linéaires. Dans le cas des applications linéaires on a aussi la propriété suivante :

8.9 Propriété. *Soit $L\colon V \to W$ une application linéaire.*

1. *L est injective si et seulement si* $\operatorname{Ker} L = \{\vec{0}\}$.
2. *L est surjective si et seulement si* $\operatorname{Im} L = W$.
3. *L est bijective si et seulement si il existe une application linéaire $T\colon W \to V$ telle que $L \circ T = id_W$ et $T \circ L = id_V$.*

Preuve.

1. Si L est injective et $\vec{x} \in \operatorname{Ker} L$, alors $L(\vec{x}) = \vec{0} = L(\vec{0})$ et donc $\vec{x} = \vec{0}$.
2. Evident par la définition de $\operatorname{Im} L$.
3. D'après le Corollaire 8.6, il reste à montrer que T est linéaire : en utilisant la linéarité de L et le fait que $L \circ T = id_W$, on a

   $$L(T(\vec{y}_1 + \vec{y}_2)) = \vec{y}_1 + \vec{y}_2 = L(T(\vec{y}_1)) + L(T(\vec{y}_2)) = L(T(\vec{y}_1) + T(\vec{y}_2))$$

   d'où $T(\vec{y}_1 + \vec{y}_2) = T(\vec{y}_1) + T(\vec{y}_2)$ car L est injective. De même pour montrer que $T(\alpha \cdot \vec{y}) = \alpha \cdot T(\vec{y})$.

□

8.10 Exercice. Soit V, W, U trois espaces vectoriels et $L: V \to W, T: W \to U$ deux fonctions telles que la composée $T \circ L$ est linéaire. Montrer que :

1. Si T est linéaire et injective alors L est linéaire.
2. Si L est linéaire et surjective alors T est linéaire.

8.2 Isomorphismes entre espaces vectoriels

8.11 Définition.

1. Un *isomorphisme* entre deux espaces vectoriels V et W est une application linéaire bijective $L: V \to W$.
2. Deux espaces vectoriels V et W sont dits *isomorphes* s'il existe un isomorphisme $V \to W$ (ou, de façon équivalente, $W \to V$). Notation : $V \simeq W$.
3. Si $L: V \to W$ est un isomorphisme, l'application linéaire $T: W \to V$ telle que $L \circ T = id_W$ et $T \circ L = id_V$ est appelée *application inverse* et notée L^{-1}.

8.12 Exemple. Soit $A \in \mathbb{R}^{m \times n}$ et soit $L_A: \mathbb{R}^n \to \mathbb{R}^m$ l'application linéaire associée à la matrice A.

1. L_A est surjective ssi pour tout $\vec{b} \in \mathbb{R}^m$ le système $\mathcal{S}: A \cdot \vec{x} = \vec{b}$ admet au moins une solution.
2. L_A est injective ssi pour tout $\vec{b} \in \mathbb{R}^m$ le système $\mathcal{S}: A \cdot \vec{x} = \vec{b}$ admet au plus une solution ssi le système homogène $\mathcal{S}_0: A \cdot \vec{x} = \vec{0}$ admet une solution unique.

8.13 Exemple. Les espaces vectoriels $\mathbb{R}[x]_n$ et $\mathbb{R}^{n+1}$ sont isomorphes :

$$\mathbb{R}[x]_n \simeq \mathbb{R}^{n+1} .$$

Preuve. Nous pouvons considérer l'application linéaire

$$L: \mathbb{R}[x]_n \to \mathbb{R}^{n+1} , \quad L(a_0 + a_1 x + \ldots + a_n x^n) = (a_0, a_1, \ldots, a_n) .$$

Il suffit de montrer que L est injective et surjective, ou bien de trouver

$$L^{-1}: \mathbb{R}^{n+1} \to \mathbb{R}[x]_n$$

telle que $L \circ L^{-1} = id_{\mathbb{R}^{n+1}}$ et $L^{-1} \circ L = id_{\mathbb{R}[x]_n}$. Il faut bien évidemment choisir $L^{-1}(a_0, a_1, \ldots, a_n) = a_0 + a_1 x + \ldots + a_n x^n$. □

On peut remarquer que les deux espaces vectoriels isomorphes de l'Exemple 8.13 ont même dimension. En effet, un lien étroit existe entre la dimension et l'isomorphisme. Pour préciser cela nous avons besoin de deux lemmes.

8.14 Lemme. *Soit $L: V \to W$ une application linéaire.*

1. *Si* $\vec{x}_1, \dots, \vec{x}_n$ *est une famille libre et* L *est injective, alors* $L(\vec{x}_1), \dots, L(\vec{x}_n)$ *est une famille libre.*
2. *Soit* $\vec{x}_1, \dots, \vec{x}_n$ *une famille génératrice de* V. *L'application* L *est surjective si et seulement si* $L(\vec{x}_1), \dots, L(\vec{x}_n)$ *est une famille génératrice de* W.

Preuve. 1. Si $a_1 \cdot L(\vec{x}_1) + \ldots + a_n \cdot L(\vec{x}_n) = \vec{0}$, par linéarité on a $L(a_1 \cdot \vec{x}_1 + \ldots + a_n \cdot \vec{x}_n) = \vec{0}$. Mais puisque L est injective cela implique que $a_1 \cdot \vec{x}_1 + \ldots + a_n \cdot \vec{x}_n = \vec{0}$. Et puisque $\vec{x}_1, \dots, \vec{x}_n$ est une famille libre, on en tire $a_1 = \ldots = a_n = 0$.
2. Si $\vec{x}_1, \dots, \vec{x}_n$ est une famille génératrice de V, par l'Exercice 7.12 nous savons que $\operatorname{Im} L$ est le s.e.v. de W engendré par $L(\vec{x}_1), \dots, L(\vec{x}_n)$. Par conséquent : L est surjective ssi $\operatorname{Im} L = W$ ssi $L(\vec{x}_1), \dots, L(\vec{x}_n)$ est une famille génératrice de W. □

8.15 Lemme. *(Propriété universelle de la base) Soit* V, W *deux espaces vectoriels,* $\vec{e}_1, \dots, \vec{e}_n$ *une base de* V *et* $\vec{y}_1, \dots, \vec{y}_n$ *des vecteurs arbitraires de* W. *Il existe une et une seule application linéaire*

$$L \colon V \to W$$

telle que $L(\vec{e}_i) = \vec{y}_i$ *pour* $i = 1, \dots, n$.

Preuve. L'unicité d'une telle application linéaire est garantie par le Lemme 7.24. Voyons l'existence : si $\vec{x}$ est un vecteur de V, on peut l'écrire comme

$$\vec{x} = a_1 \cdot \vec{e}_1 + \ldots + a_n \cdot \vec{e}_n$$

et cette écriture est unique. Si L est linéaire et $L(\vec{e}_i) = \vec{y}_i$, il faut définir

$$L(\vec{x}) = a_1 \cdot \vec{y}_1 + \ldots + a_n \cdot \vec{y}_n \,.$$

Il est facile de vérifier que l'application L ainsi définie est linéaire. □

8.16 Remarque. La propriété universelle de la base reste valable même si l'espace V n'est pas de type fini : soit V, W deux espaces vectoriels, $E \subseteq V$ une base (finie ou infinie) de V et $\{\vec{y}_{\vec{e}} \mid \vec{e} \in E\}$ des vecteurs arbitraires de W. Il existe une et une seule application linéaire

$$L \colon V \to W$$

telle que $L(\vec{e}) = \vec{y}_{\vec{e}}$ pour tout $\vec{e} \in E$.
La preuve est la même que dans le cas de dimension finie.

8.17 Propriété. *Soit* V, W *deux espaces vectoriels, avec* V *de type fini. Les espaces* V *et* W *sont isomorphes si et seulement si* W *est de type fini et* $\dim V = \dim W$.

Preuve.
$\Rightarrow$: Si $L \colon V \to W$ est un isomorphisme et $\vec{e}_1, \dots, \vec{e}_n$ est une base de V, alors par le Lemme 8.14 $L(\vec{e}_1), \dots, L(\vec{e}_n)$ est une base de W. Par conséquent, $\dim W =$

$n = \dim V$.
$\Leftarrow$: Soit $\vec{e}_1, \ldots, \vec{e}_n$ une base de V et soit $\vec{f}_1, \ldots, \vec{f}_n$ une base de W. Par le Lemme 8.15 il existe une (unique) application linéaire $L\colon V \to W$ telle que $L(\vec{e}_i) = \vec{f}_i$ pour $i = 1, \ldots, n$. Par le Lemme 8.14 L est surjective. Elle est aussi injective : si $\vec{x} = a_1 \cdot \vec{e}_1 + \ldots + a_n \cdot \vec{e}_n \in \operatorname{Ker} L$, alors

$$\vec{0} = L(\vec{x}) = L(a_1 \cdot \vec{e}_1 + \ldots + a_n \cdot \vec{e}_n) = a_1 \cdot L(\vec{e}_1) + \ldots + a_n \cdot L(\vec{e}_n) = a_1 \cdot \vec{f}_1 + \ldots + a_n \cdot \vec{f}_n$$

Puisque $\vec{f}_1, \ldots, \vec{f}_n$ est une famille libre, on a forcement $a_1 = \ldots = a_n = 0$ et donc $\vec{x} = \vec{0}$. □

Comme corollaire de la propriété précédente nous obtenons le théorème de représentation pour les espaces vectoriels finiment engendrés :

8.18 Corollaire. *Un espace vectoriel réel W est isomorphe à $\mathbb{R}^n$ si et seulement si W est de type fini et* $\dim W = n$.

Preuve. Il suffit de poser $V = \mathbb{R}^n$ dans la Propriété 8.17. □

8.3 Le théorème de représentation pour les applications linéaires

Nous pouvons à présent revenir à l'application linéaire

$$\varepsilon\colon \mathbb{R}^{m\times n} \to \operatorname{Lin}(\mathbb{R}^n, \mathbb{R}^m)\ ,\ \ \varepsilon(A) = L_A$$

pour completer l'énoncé de la Propriété 7.28 et obtenir ainsi le premier théorème de répresentation pour les applications linéaires.

8.19 Propriété. *L'application linéaire*

$$\varepsilon\colon \mathbb{R}^{m\times n} \to \operatorname{Lin}(\mathbb{R}^n, \mathbb{R}^m)\ ,\ \ \varepsilon(A) = L_A$$

est un isomorphisme.

Preuve. Nous avons déjà démontré dans le Lemme 7.26 que, pour $A, B \in \mathbb{R}^{m\times n}$, si $L_A = L_B$ alors $A = B$. Autrement dit :

$$\varepsilon(A) = \varepsilon(B) \ \Rightarrow\ A = B\,,$$

c'est-à-dire que ε est injective. Il reste donc à montrer qu'elle est aussi surjective. Soit $L\colon \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire. Nous cherchons une matrice $A \in \mathbb{R}^{m\times n}$ telle que $\varepsilon(A) = L$, c'est-à-dire telle que $L_A = L$. Pour découvrir comment il faut définir A, supposons pour un moment qu'une telle matrice existe. La condition $L_A = L$ signifie que pour tout $\vec{x} \in \mathbb{R}^n$ on a $L_A(\vec{x}) = L(\vec{x})$. En particulier nous pouvons choisir comme vecteurs $\vec{x}$ les n vecteurs $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ de la base canonique de $\mathbb{R}^n$ et la condition précédente devient $L_A(\vec{e}_i) = L(\vec{e}_i)$, c'est-à-dire

$$A \cdot \vec{e}_i = L(\vec{e}_i) \quad \forall\ i = 1, \ldots, n$$

Mais, d'après la Remarque 5.10,

$$A \cdot \vec{e}_i = i\text{-ème colonne de } A$$

ce qui montre que le seul candidat possible est la matrice A dont les colonnes sont l'image par L des vecteurs de la base canonique de $\mathbb{R}^n$:

$$A = (L(\vec{e}_1) \mid L(\vec{e}_2) \mid \ldots \mid L(\vec{e}_n))$$

Il reste à vérifier qu'effectivement $L_A = L$ et, grâce au Lemme 7.24, pour cela il suffit de montrer que $L_A(\vec{e}_i) = L(\vec{e}_i)$ pour $i = 1, \ldots, n$. Cela revient à

$$L_A(\vec{e}_i) = A \cdot \vec{e}_i = i\text{-ème colonne de } A = L(\vec{e}_i)$$

où la dernière égalité est la définition même de A. □

8.20 Corollaire. *L'application linéaire*

$$\varepsilon \colon \mathbb{R}^{n\times n} \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)\ , \quad \varepsilon(A) = L_A$$

est un isomorphisme de $\mathbb{R}$*-algèbres.*

Preuve. Il suffit de choisir $m = n$ dans 8.19. □

8.21 Lemme. *Soit* $L \colon V \to V'$ *et* $T \colon W \to W'$ *deux applications linéaires.*

1. *La fonction*

$$- \circ L \colon \ \mathrm{Lin}(V', W) \to \mathrm{Lin}(V, W)\ , \quad S \mapsto S \circ L$$

est une application linéaire. De plus, si L *est un isomorphisme avec inverse* L^{-1}*, alors* $- \circ L$ *aussi est un isomorphisme avec inverse* $- \circ L^{-1}$.

2. *La fonction*

$$T \circ - \colon \ \mathrm{Lin}(V, W) \to \mathrm{Lin}(V, W')\ , \quad S \mapsto T \circ S$$

est une application linéaire. De plus, si T *est un isomorphisme avec inverse* T^{-1}*, alors* $T \circ -$ *aussi est un isomorphisme avec inverse* $T^{-1} \circ -$.

Preuve. 1. Il faut vérifier que pour $S, R \in \mathrm{Lin}(V, W)$ et $\alpha \in \mathbb{R}$ on a

$$(S + R) \circ L = S \circ L + R \circ L \ \text{ et } \ (\alpha \cdot S) \circ L = \alpha \cdot (S \circ L)\,.$$

C'est un exercice. Pour la deuxième partie, on a

$$(S \circ L) \circ L^{-1} = S \circ (L \circ L^{-1}) = S \circ id = S$$

et analoguement $(S \circ L^{-1}) \circ L = S$.
2. Semblable au point 1. Voyons par exemple que $T \circ (S + R) = T \circ S + T \circ R$:

$$(T \circ (S + R))(\vec{x}) = T((S + R)(\vec{x})) = T(S(\vec{x}) + R(\vec{x})) =$$

$$= T(S(\vec{x})) + T(R(\vec{x})) = (T \circ S)(\vec{x}) + (T \circ R)(\vec{x}) = (T \circ S + T \circ R)(\vec{x})$$

□

8.22 Corollaire. *(Théorème de représentation pour les applications linéaires) Soit V, W deux espaces vectoriels avec $\dim V = n$ et $\dim W = m$. Alors*

1. $\mathrm{Lin}(V, W) \simeq \mathbb{R}^{m\times n}$
2. $\dim \mathrm{Lin}(V, W) = m \times n$.

Preuve. 1. Puisque $\dim V = n$ et $\dim W = m$, d'après le Corollaire 8.18 nous avons des isomorphimes $V \simeq \mathbb{R}^n$ et $W \simeq \mathbb{R}^m$. Par le Lemme 8.21 on a alors

$$\mathrm{Lin}(V, W) \simeq \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)\,.$$

Si on compose cet isomorphisme avec l'isomorphisme

$$\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{m\times n}$$

de la Propriété 8.19, on obtient l'isomorphisme cherché.
2. Il suffit d'appliquer la Propriété 8.17 à l'isomorphisme obtenu dans 1. □

8.4 La matrice associée à une application linéaire

Cette longue section est consacrée à analyser le Corollaire 8.22 pour bien apprendre son usage. La difficulté principale vient du fait que l'isomorphisme

$$\mathrm{Lin}(V, W) \simeq \mathbb{R}^{m\times n}$$

n'est pas *canonique*. Cela signifie que pour l'expliciter il faut effectuer des choix. Il faut notamment choisir une base de l'espace V et une base de l'espace W, et si on change les bases choisies on obtient un isomorphisme différent. Voyons cela en détail.

8.23 Remarque. Soit $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de V et $f = \{\vec{f}_1, \ldots, \vec{f}_m\}$ une base de W. Rappelons la notation 4.18

$$_e\vec{x} = (x_1, \ldots, x_n) \text{ si } \vec{x} = x_1 \cdot \vec{e}_1 + \ldots + x_n \cdot \vec{e}_n$$

$$_f\vec{y} = (y_1, \ldots, y_m) \text{ si } \vec{y} = y_1 \cdot \vec{f}_1 + \ldots + y_m \cdot \vec{f}_m$$

Nous pouvons construire les isomorphismes

$$\begin{array}{ll} E\colon V \to \mathbb{R}^n & E(\vec{x}) = {}_e\vec{x} \\ E^{-1}\colon \mathbb{R}^n \to V & E^{-1}(x_1, \ldots, x_n) = x_1 \cdot \vec{e}_1 + \ldots + x_n \cdot \vec{e}_n \end{array}$$

et

$$\begin{array}{ll} F\colon W \to \mathbb{R}^m & F(\vec{y}) = {}_f\vec{y} \\ F^{-1}\colon \mathbb{R}^m \to W & F^{-1}(y_1, \ldots, y_m) = y_1 \cdot \vec{f}_1 + \ldots + y_m \cdot \vec{f}_m \end{array}$$

Autrement dit :

$$\begin{array}{rcl} E(\vec{e}_i) & = & i\text{-ème vecteur de la base canonique de } \mathbb{R}^n \\ F(\vec{f}_i) & = & i\text{-ème vecteur de la base canonique de } \mathbb{R}^m \end{array}$$

A partir d'une application linéaire $L\colon V \to W$ nous pouvons maintenant

1. construire une nouvelle application linéaire
$$\mathbb{R}^n \xrightarrow{E^{-1}} V \xrightarrow{L} W \xrightarrow{F} \mathbb{R}^m$$
2. en vertu de la Propriété 8.19, nous obtenons une unique matrice $A \in \mathbb{R}^{m\times n}$ telle que $L_A = F \circ L \circ E^{-1}$, c'est-à-dire telle que le diagramme suivant commute
$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "E"'] & W \arrow[d, "F"] \\ \mathbb{R}^n \arrow[r, "L_A"'] & \mathbb{R}^m \end{tikzcd}$$
3. si nous adoptons la notation $A = {}_fL_e$, la commutativité du diagramme précédent nous dit que la matrice ${}_fL_e$ est l'unique matrice à m lignes et n colonnes telle que pour tout $\vec{x} \in \mathbb{R}^n$
$$_fL(\vec{x}) = {}_fL_e \cdot {}_e\vec{x} \qquad [8.1]$$

Les points 1., 2. et 3. ci-dessus décrivent explicitement l'isomorphisme
$$\mathrm{Lin}(V,W) \xrightarrow{F\circ - \circ E^{-1}} \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \xrightarrow{\varepsilon^{-1}} \mathbb{R}^{m\times n} \qquad L \mapsto {}_fL_e$$
On peut décrire explicitement l'isomorphisme dans la direction opposée :
$$\mathbb{R}^{m\times n} \xrightarrow{\varepsilon} \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \xrightarrow{F^{-1}\circ - \circ E} \mathrm{Lin}(V,W) \qquad A \mapsto A_e^f$$
où $A_e^f : V \to W$ est donc l'unique application linéaire telle que pour tout $\vec{x} \in V$
$$_f(A_e^f(\vec{x})) = A \cdot {}_e\vec{x} \qquad [8.2]$$

La formule [8.1] permet de reconstruire l'application linéaire L à partir de la matrice ${}_fL_e$ et aussi de construire la matrice ${}_fL_e$ elle-même. En effet, en choisissant comme vecteur $\vec{x}$ les vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ de la base e de V on a :
$$\begin{aligned} {}_fL(\vec{e}_i) &= {}_fL_e \cdot {}_e(\vec{e}_i) \\ &= {}_fL_e \cdot i\text{-ème vecteur de la base canonique de } \mathbb{R}^n \\ &= i\text{-ème colonne de } {}_fL_e \end{aligned}$$

8.24 Exemple.

1. Considérons l'application linéaire
$$D : \mathbb{R}[x]_3 \to \mathbb{R}[x]_2 \ , \quad D(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2$$
et les bases $e = \{1, x, x^2, x^3\}$ de $\mathbb{R}[x]_3$ et $f = \{1, x, x^2\}$ de $\mathbb{R}[x]_2$.
Pour déterminer la matrice ${}_fD_e$ il faut calculer
$$\begin{array}{rcccl} D(1) & = & 0 & = & 0\cdot 1 + 0x + 0x^2 \\ D(x) & = & 1 & = & 1\cdot 1 + 0x + 0x^2 \\ D(x^2) & = & 2x & = & 0\cdot 1 + 2x + 0x^2 \\ D(x^3) & = & 3x^2 & = & 0\cdot 1 + 0x + 3x^2 \end{array}$$

d'où

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

Une fois la matrice ${}_fD_e$ trouvée, nous pouvons l'utiliser pour calculer $D(P)$ sans plus utiliser la définition explicite de D. Calculons par exemple $D(P)$ pour $P(x) = 3 + 4x + 5x^2 + 6x^3 \in \mathbb{R}[x]_3$: puisque

$$ {}_eP = \begin{pmatrix} 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} $$

on a d'après la formule [8.2]

$$ {}_fD(P) = {}_fD_e \cdot {}_eP = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 18 \end{pmatrix} $$

Finalement, $D(P) = 4 + 10x + 18x^2$.

2. Considérons comme application linéaire

$$ L\colon \mathbb{R}^2 \to \mathbb{R}^2 $$

la rotation d'angle θ autour de l'origine, et cherchons ${}_eL_e$ où e est la base canonique de $\mathbb{R}^2$. Puisque

$$ L\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \quad \text{et} \quad L\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} $$

on a

$$ {}_eL_e = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} $$

Par exemple

$$ L\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \cos\theta - 2\sin\theta \\ \sin\theta + 2\cos\theta \end{pmatrix} $$

3. Considérons comme application linéaire

$$ L\colon \mathbb{R}^2 \to \mathbb{R}^2 $$

la projection orthogonale sur la première bissectrice et cherchons ${}_eL_e$ où e est la base canonique de $\mathbb{R}^2$. Puisque

$$ L\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad \text{et} \quad L\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} $$

on a

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

Cherchons maintenant ${}_{e'}L_{e'}$ où

$$ e' = \{\vec{e}_1' = (1,1), \vec{e}_2' = (1,-1)\} $$

Puisque $\vec{e}_1'$ est sur la droite de projection et $\vec{e}_2'$ est orthogonale à la droite de projection, on a

$$ L(\vec{e}_1') = \vec{e}_1' \ \text{ et } \ L(\vec{e}_2') = \vec{0} $$

et donc

$$ {}_{e'}L_{e'} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} $$

Pour calculer L d'un vecteur quelconque de $\mathbb{R}^2$ on peut utiliser soit ${}_eL_e$ soit ${}_{e'}L_{e'}$. Cherchons par exemple $L(1,2)$:

(a) $L \begin{pmatrix} 1 \\ 2 \end{pmatrix} = {}_eL_e \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{3}{2} \end{pmatrix}$

(b) Puisque $(1,2) = \frac{3}{2}\vec{e}_1 - \frac{1}{2}\vec{e}_2'$, on a

$$ {}_{e'}L \begin{pmatrix} 1 \\ 2 \end{pmatrix} = {}_{e'}L_{e'} \cdot \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} $$

et donc $L(1,2) = \frac{3}{2}\vec{e}_1' + 0\vec{e}_2' = (\frac{3}{2}, \frac{3}{2})$.

4. Soit V un espace vectoriel. Considérons l'application identité

$$ id_V : V \to V $$

et cherchons la matrice ${}_e(id_V)_e$ où $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ est une base de V. Puisque ${}_e(\vec{e}_i)$ est le i-ème vecteur de la base canonique de $\mathbb{R}^n$, la matrice ${}_e(id_V)_e$ est la matrice identité $I \in \mathbb{R}^{n\times n}$.

5. Considérons à nouveau l'application identité

$$ id_V : V \to V $$

et cherchons la matrice ${}_{e'}(id_V)_e$ où $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ et $e' = \{\vec{e}_1', \ldots, \vec{e}_n'\}$ sont deux bases de V. Puisque

$$ {}_{e'}(\vec{e}_i) = {}_{e'}id(\vec{e}_i) = {}_{e'}id_e \cdot {}_e(\vec{e}_i) = i\text{-ème colonne de } {}_{e'}id_e $$

on retrouve la matrice changement de base de la Remarque 5.16

$$ {}_{e'}(id_V)_e = {}_{e'}I_e $$

8.25 Exercice. Déterminer tous les polynômes P de degré plus petit ou égal à 4 tels que $\int_0^1 P = 0 = \int_{-1}^0 P$.

Preuve. L'ensemble des polynômes P de degré plus petit ou égal à 4 tels que $\int_0^1 P = 0 = \int_{-1}^0 P$ est le noyau de l'application linéaire

$$L\colon \mathbb{R}[x]_4 \to \mathbb{R}^2 \ , \ \ L(P) = \left(\int_0^1 P, \int_{-1}^0 P \right)$$

Fixons $e = \{P_0(x) = 1, P_1(x) = x, P_2(x) = x^2, P_3(x) = x^3, P_4(x) = x^4\}$ et $f = \{(1,0),(0,1)\}$ comme bases de l'espace $\mathbb{R}[x]_4$ et de l'espace $\mathbb{R}^2$. La matrice associée à L par rapport à ces bases est

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

Puisque rang ${}_fL_E = 2$ (les lignes sot linéairement indépendantes), $\dim \operatorname{Ker} L = 5 - 2 = 3$. En regardant les colonnes de la matrice, nous pouvons observer que $L(\vec{e}_2) = 2L(\vec{e}_4)$, d'où $\vec{e}_2 - 2\vec{e}_4 \in \operatorname{Ker} L$, c'est-à-dire $x - 2x^3 \in \operatorname{Ker} L$. De même, $L(\vec{e}_1) = 3L(\vec{e}_3)$ et $L(\vec{e}_1) = 5L(\vec{e}_5)$, d'où $1 - 3x^2, 1 - 5x^4 \in \operatorname{Ker} L$. Finalement,

$$\operatorname{Ker} L = \langle x - 2x^3, 1 - 3x^2, 1 - 5x^4 \rangle$$

□

8.5 Changement des bases

Les deux dernières points de l'Example 8.24 suggèrent le problème suivant : si $L\colon V \to W$ est une application linéaire et si on fixe deux bases e et e' de V et deux bases f et f' de W, quel lien existe entre la matrice ${}_fL_e$ et la matrice ${}_{f'}L_{e'}$? Pur y répondre nous utilisons la pripriété suivante :

8.26 Propriété. *Soit $L\colon V \to W$ et $T\colon W \to U$ deux applications linéaires entre espaces de type fini, e une base de V, f une base de W et g une base de U. Alors*

$${}_g(T \circ L)_e = {}_gT_f \cdot {}_fL_e$$

Preuve. D'après la Remarque 8.23 nous savons que les matrices satisfont et sont déterminées par les conditions

$$\begin{array}{rcll} {}_g(T \circ L)(\vec{x}) & = & {}_g(T \circ L)_e \cdot {}_e\vec{x} & \forall\, \vec{x} \in V \\ {}_fL(\vec{x}) & = & {}_fL_e \cdot {}_e\vec{x} & \forall\, \vec{x} \in V \\ {}_gT(\vec{y}) & = & {}_gT_f \cdot {}_f\vec{y} & \forall\, \vec{y} \in W \end{array}$$

Cela donne :

$${}_g(T \circ L)(\vec{x}) = {}_gT(L(\vec{x})) = {}_gT_f \cdot {}_fL(\vec{x}) = {}_gT_f \cdot {}_fL_e \cdot {}_e\vec{x}$$

et donc ${}_g(T \circ L)_e = {}_gT_f \cdot {}_fL_e$. □

8.27 Corollaire. *Soit $L\colon V \to W$ une application linéaire entre espaces de type fini et soient e, e' deux bases de V et f, f' deux bases de W. Alors :*

$$ {}_{f'}L_{e'} = {}_{f'}I_f \cdot {}_fL_e \cdot {}_eI_{e'} $$

où ${}_{f'}I_f$ et ${}_eI_{e'}$ sont les matrices changement de base.

Preuve. Il suffit d'appliquer la Propriété 8.26 à la situation

$$ V \xrightarrow{id_V} V \xrightarrow{L} W \xrightarrow{id_W} W $$

en prénant comme base de V la base e' au départ et e à l'arrivée, et comme base de W la base f au départ et f' à l'arrivée. $\square$

8.28 Exercice. Soit e, e' deux bases d'un espace vectoriel V. En utilisant la Propriété 8.26 et les exercices 8.24.4 et 8.24.5 montrer que

$$ {}_eI_{e'} \cdot {}_{e'}I_e = I \quad \text{et} \quad {}_{e'}I_e \cdot {}_eI_{e'} = I\,. $$

8.6 Divagation

Pour préciser davantage le théorème de représentation des applications linéaires (Corollaire 8.22) il faudrait utilisier le language de la *théorie des catégories* : la catégorie dont les objets sont les espaces vectoriels de type fini avec une base choisie et les flèches sont les applications linéaires, est équivalente à la catégorie dont les objets sont les nombres naturels et les flèches sont les matrices. D'un côté on associe à $n \in \mathbb{N}$ l'espace vectoriel $\mathbb{R}^n$ muni de sa base canonique et à une matrice $A \in \mathbb{R}^{m\times n}$ l'application linéaire $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ (voir 7.3). De l'autre côté, si $L\colon V \to W$ est une application linéaire, $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ est une base de V et $f = \{\vec{f}_1, \ldots, \vec{f}_m\}$ est une base de W, on associe à L la matrice ${}_fL_e \in \mathbb{R}^{m\times n}$ (voir 8.23). Les Propriétés 7.25 et 8.26 expriment la fonctorialité de ces constructions, et le diagramme commutatif

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "E"'] & W \arrow[d, "F"] \\ \mathbb{R}^n \arrow[r, "L_A"'] & \mathbb{R}^m \end{tikzcd}$$

de la Remarque 8.23.2 exprime la naturalité de l'isomorphisme $E\colon V \to \mathbb{R}^n$ induit par la base e de l'espace V. Tout ceci, bien que élémentaire, dépasse le cadre de ce cours.

Chapitre 9

Produit d'espaces vectoriels

Le chapitre 9 section par section

Dans ce chapitre nous étudions le produit d'espaces vectoriels. Cela nous permettra d'établir le théorème du rang, qui donne un lien entre le noyau et l'image d'une application linéaire.

1. Le produit cartésien d'espaces vectoriels est muni d'une structure canonique d'espace vectoriel. Dans le cas de dimension finie, sa dimension est la somme des dimensions des différents espaces.
2. Une application facile est le calcul de la dimension d'une somme directe de sous-espaces vectoriels.
3. Le théorème du rang nous dit que pour toute application linéaire $L\colon V \to W$, on a une décomposition $V \simeq \operatorname{Ker} L \times \operatorname{Im} L$.
4. Le théorème du rang, combiné avec le théorème de représentation des applications linéaires, permet d'étudier l'injectivité et la surjectivité des applications linéaires ...
5. ... et aussi de calculer la dimension d'une somme quelconque de sous-espaces vectoriels.
6. Une autre application du théorème de représentation est que l'espace lignes et l'espace colonnes d'une matrice ont même dimension.

9.1 Produit cartésien d'espaces vectoriels

9.1 Définition. Soit X, Y deux ensembles. Le produit (cartésien) de X et Y est l'ensemble

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

muni des projections

$$\pi_X\colon X \times Y \to X\ ,\ \pi_X(x, y) = x \quad \text{et} \quad \pi_Y\colon X \times Y \to Y\ ,\ \pi_Y(x, y) = y\,.$$

9.2 Remarque.

1. Si Z est un troisième ensemble, se donner une fonction $f\colon Z \to X \times Y$ revient exactement à se donner deux fonctions
$$g\colon Z \to X \quad \text{et} \quad h\colon Z \to Y\,.$$
En effet :
   (a) si on a f on pose $g = \pi_X \circ f$ et $h = \pi_Y \circ f$,
   (b) si on a g et h on pose $f(z) = (g(z), h(z))$.
2. On peut de la même manière faire le produit d'une famille finie d'ensembles
$$X_1 \times X_2 \times \ldots \times X_k$$
ou d'une famille infinie d'ensembles
$$\prod_{i\in I} X_i\,.$$

9.3 Propriété. *Soit V, W deux espaces vectoriels réels. Le produit cartésien*
$$V \times W$$
est un espace vectoriel par rapport aux opérations "composante par composante" :

$\vec{0}_{V\times W} = (\vec{0}_V, \vec{0}_W)$
$\mathbb{R} \times (V \times W) \to V \times W\ ,\quad \alpha \cdot (\vec{v}, \vec{w}) = (\alpha \cdot \vec{v}, \alpha \cdot \vec{w})$
$(V \times W) \times (V \times W) \to V \times W\ ,\quad (\vec{v}_1, \vec{w}_1) + (\vec{v}_2, \vec{w}_2) = (\vec{v}_1 + \vec{v}_2, \vec{w}_1 + \vec{w}_2)$

De plus, les projections
$$\pi_V\colon V \times W \to V\ \ ,\ \ \pi_W\colon V \times W \to W$$
sont des applications linéaires.

Preuve. La preuve est laissée en exercice. □

9.4 Remarque. La Propriété 9.3 explique pourquoi, si V et W sont deux espaces vectoriels, la structure d'espace vectoriel de $V \times W$ est définie composante par composante : c'est en effet la seule structure d'espace vectoriel qu'on puisse mettre sur l'ensemble $V \times W$ pour que les projections $\pi_V\colon V \times W \to V$ et $\pi_W\colon V \times W \to W$ soient deux applications linéaires. De plus, avec une telle structure, si U est un troisième espace vectoriel et $G\colon U \to V$ et $H\colon U \to W$ sont deux fonctions, alors la fonction
$$F\colon U \to V \times W \quad F(\vec{u}) = (G(\vec{u}), H(\vec{u}))$$
(voir Remarque 9.2) sera linéaire si et seulement si G et H sont linéaires.

9.5 Exemple.

1. L'espace $\mathbb{R}^n$ est le produit de n copies de $\mathbb{R}$.
2. La fonction $F\colon \mathbb{R}[x]_3 \to \mathbb{R}[x]_2 \times \mathbb{R}^4$ définie par

$$F(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_1 + 2a_2x + 3a_3x^2, (a_0, a_1, a_2, a_3))$$

est linéaire car les deux composantes

$$\mathbb{R}[x]_3 \to \mathbb{R}[x]_2\ ,\quad a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_1 + 2a_2x + 3a_3x^2$$

$$\mathbb{R}[x]_3 \to \mathbb{R}^4\ ,\quad a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto (a_0, a_1, a_2, a_3)$$

sont linéaires (voir Exemple 7.4).

9.6 Propriété. *Soit V, W deux espaces vectoriels et soit $\vec{e}_1, \ldots, \vec{e}_n$ une base de V et $\vec{f}_1, \ldots, \vec{f}_m$ une base de W. Alors :*

1. $(\vec{e}_1, \vec{0}_W), \ldots, (\vec{e}_n, \vec{0}_W), (\vec{0}_V, \vec{f}_1), \ldots, (\vec{0}_V, \vec{f}_m)$ *est une base de $V \times W$.*

2. $\dim(V \times W) = \dim V + \dim W$.

Preuve. 1. Si

$$a_1(\vec{e}_1, \vec{0}) + \ldots + a_n(\vec{e}_n, \vec{0}) + b_1(\vec{0}, \vec{f}_1) + \ldots + b_m(\vec{0}, \vec{f}_m) = (\vec{0}, \vec{0})$$

alors $a_1\vec{e}_1 + \ldots + a_n\vec{e}_n = \vec{0}$ et $b_1\vec{f}_1 + \ldots + b_m\vec{f}_m = \vec{0}$. Ceci implique $a_1 = \ldots = a_n = 0$ et $b_1 = \ldots = b_m = 0$ car les familles $\{\vec{e}_1, \ldots, \vec{e}_n\}$ et $\{\vec{f}_1, \ldots, \vec{f}_m\}$ sont libres.
Soit maintenant $(\vec{v}, \vec{w}) \in V \times W$. Puisque $\vec{v} \in V$ et $\{\vec{e}_1, \ldots, \vec{e}_n\}$ est une famille génératrice de V, on peut trouver $a_1, \ldots, a_n \in \mathbb{R}$ tels que $\vec{v} = a_1\vec{e}_1 + \ldots + a_n\vec{e}_n$. Puisque $\vec{w} \in W$ et $\{\vec{f}_1, \ldots, \vec{f}_m\}$ est une famille génératrice de W, on peut trouver $b_1, \ldots, b_m \in \mathbb{R}$ tels que $\vec{w} = b_1\vec{f}_1 + \ldots + b_m\vec{f}_m$. Finalement,

$$(\vec{v}, \vec{w}) = a_1(\vec{e}_1, \vec{0}) + \ldots + a_n(\vec{e}_n, \vec{0}) + b_1(\vec{0}, \vec{f}_1) + \ldots + b_m(\vec{0}, \vec{f}_m)$$

2. C'est une conséquence immédiate de 1. □

9.7 Remarque. La propriété précédente reste valable pour le produit de plusieurs espaces : si $V_1, \ldots, V_k$ sont des espaces vectoriels finiment engendrés, alors l'espace produit $V_1 \times \ldots \times V_k$ est finiment engendré et

$$\dim(V_1 \times \ldots \times V_k) = \dim V_1 + \ldots + \dim V_k\,.$$

9.2 Somme directe de sous-espaces vectoriels

9.8 Définition. Soit V un espace vectoriel et soit $V_1, \ldots, V_k$ des s.e.v. de V. La somme

$$V_1 + \ldots + V_k = \{\vec{v}_1 + \ldots + \vec{v}_k \mid \vec{v}_i \in V_i\}$$

(c'est-à-dire le plus petit s.e.v. de V qui contient les V_i) est dite *somme directe*, notée

$$V_1 \oplus \ldots \oplus V_k\,,$$

si pour tout famille $\vec{v}_1 \in V_1, \ldots, \vec{v}_k \in V_k$ on a :

$$\vec{v}_1 + \ldots + \vec{v}_k = \vec{0} \ \Rightarrow\ \vec{v}_1 = \ldots = \vec{v}_k = \vec{0}\,.$$

De façon équivalente, la somme $V_1 + \ldots + V_k$ est directe si pour tout $i = 1, \ldots, k-1$ on a :

$$(V_1 + \ldots + V_i) \cap V_{i+1} = \{\vec{0}\}\,.$$

9.9 Lemme. *Soit V un espace vectoriel et soit $V_1, \ldots, V_k$ des s.e.v. de V. Considérons les V_i comme des espaces vectoriels et considérons la fonction*

$$S\colon V_1 \times \ldots \times V_k \to V\ ,\ \ S(\vec{v}_1, \ldots, \vec{v}_k) = \vec{v}_1 + \ldots + \vec{v}_k\,.$$

1. *S est linéaire.*
2. *S est injective ssi la somme $V_1 + \ldots + V_k$ est directe.*
3. *S est surjective ssi $V = V_1 + \ldots + V_k$ (somme).*
4. *S est bijective ssi $V = V_1 \oplus \ldots \oplus V_k$ (somme directe).*

Preuve. Par la Propriété 8.9 nous savons que S est injective ssi $\operatorname{Ker} S = \{\vec{0}\,\}$, c'est-à-dire ssi

$$S(\vec{v}_1, \ldots, \vec{v}_n) = \vec{0} \ \Rightarrow\ (\vec{v}_1, \ldots, \vec{v}_n) = \vec{0}$$

c'est-à-dire ssi

$$\vec{v}_1 + \ldots + \vec{v}_k = \vec{0} \ \Rightarrow\ \vec{v}_1 = \ldots = \vec{v}_k = \vec{0}$$

ce qui nous donne exactement la condition de somme directe dans la Définition 9.8. Le reste de la preuve est laissé en exercice. □

9.10 Corollaire. *Soit V un espace vectoriel et soit $V_1, \ldots, V_k$ des s.e.v. de V. Si chaque V_i est finiment engendré et si $V = V_1 \oplus \ldots \oplus V_k$, alors V est finiment engendré et*

$$\dim V = \dim V_1 + \ldots + \dim V_k\,.$$

Preuve. Par le Lemme 9.9, si $V = V_1 \oplus \ldots \oplus V_k$ alors l'application

$$S\colon V_1 \times \ldots \times V_k \to V\ ,\ \ S(\vec{v}_1, \ldots, \vec{v}_k) = \vec{v}_1 + \ldots + \vec{v}_k$$

est un isomorphisme. Puisque chaque s.e.v. V_i est finiment engendré, le produit $V_1 \times \ldots \times V_k$ est finiment engendré (Remarque 9.7) et donc V, étant isomorphe à $V_1 \times \ldots \times V_k$, est lui aussi finiment engendré (Propriété 8.17). De plus,

$$\dim V = \dim(V_1 \times \ldots \times V_k) = \dim V_1 + \ldots + \dim V_k$$

où la première égalité utilise la Propriété 8.17 et la deuxième égalité utilise la Remarque 9.7. □

9.11 Exercice. Soit $V_1, \ldots, V_k$ des espaces vectoriels et soit $V = V_1 \times \ldots \times V_k$. Considérons les sous-ensembles $\widetilde{V}_i$ de V définis par

$$\widetilde{V}_i = \{(\vec{0}_{V_1}, \ldots, \vec{0}_{V_{i-1}}, \vec{v}_i, \vec{0}_{V_{i+1}}, \ldots, \vec{0}_{V_k}) \mid \vec{v}_i \in V_i\}\,.$$

Montrer que :

1. $\widetilde{V}_i$ est un s.e.v. de V et $\widetilde{V}_i \simeq V_i$.
2. $V = \widetilde{V}_1 \oplus \ldots \oplus \widetilde{V}_k$.

9.3 Le théorème du rang

9.12 Lemme. *Soit $L\colon V \to W$ une application linéaire surjective, avec W de type fini. Il existe une application linéaire $T\colon W \to V$ telle que $L \circ T = id_W$.*

Preuve. Soit $\{\vec{w}_1, \ldots, \vec{w}_n\}$ une base de W et choisissons $\vec{v}_1, \ldots, \vec{v}_n \in V$ tels que $L(\vec{v}_i) = \vec{w}_i$ pour tout i (ceci est possible car L est surjective). On peut maintenant poser $T(\vec{w}_i) = \vec{v}_i$ et utiliser le Lemme 8.15 pour obtenir l'application linéaire $T\colon W \to V$ cherchée. En effet, pour vérifier que $L \circ T = id_W$ il suffit de remarquer que $L(T(\vec{w}_i)) = \vec{w}_i = id_W(\vec{w}_i)$ pour tout i et utiliser le Lemme 7.24. □

9.13 Propriété. *Soit $L\colon V \to W$ une application linéaire surjective, avec W de type fini. Alors*

$$V \simeq \operatorname{Ker} L \times W\,.$$

Preuve. Soit $T\colon W \to V$ une application linéaire telle que $L \circ T = id_W$ (voir Lemme 9.12). Pour construire l'isomorphisme $V \simeq \operatorname{Ker} L \times W$ posons

$$S\colon \operatorname{Ker} L \times W \to V\ ,\quad S(\vec{x}, \vec{w}) = \vec{x} + T(\vec{w})$$

$$S^{-1}\colon V \to \operatorname{Ker} L \times W\ ,\quad S^{-1}(\vec{v}) = (\vec{v} - T(L(\vec{v})), L(\vec{v}))$$

– S^{-1} est bien définie, c'est-à-dire que $\vec{v} - T(L(\vec{v})) \in \operatorname{Ker} L$ pour tout $\vec{v} \in V$:

$$L(\vec{v} - T(L(\vec{v}))) = L(\vec{v}) - L(T(L(\vec{v}))) = L(\vec{v}) - L(\vec{v}) = \vec{0}$$

– S est linéaire : exercice.
– $S \circ S^{-1} = id_V$:

$$S(S^{-1}(\vec{v})) = S(\vec{v} - T(L(\vec{v})), L(\vec{v})) = \vec{v} - T(L(\vec{v})) + T(L(\vec{v})) = \vec{v}$$

– $S^{-1} \circ S = id_{\operatorname{Ker} L \times W}$:

$$\begin{aligned} S^{-1}(S(\vec{x}, \vec{w})) &= S^{-1}(\vec{x} + T(\vec{w})) \\ &= (\vec{x} + T(\vec{w}) - T(L(\vec{x} + T(\vec{w}))), L(\vec{x} + T(\vec{w}))) \\ &= (\vec{x} + T(\vec{w}) - T(L(\vec{x}) + L(T(\vec{w}))), L(\vec{x}) + L(T(\vec{w}))) \\ &= (\vec{x} + T(\vec{w}) - T(L(T(\vec{w}))), L(T(\vec{w}))) \\ &= (\vec{x} + T(\vec{w}) - T(\vec{w}), \vec{w}) \\ &= (\vec{x}, \vec{w}) \end{aligned}$$

□

9.14 Remarque. Afin d'éviter l'hypothèse que L soit surjective dans la propriété précédente, observons que toute application linéaire $L\colon V \to W$ peut se décomposer comme une application linéaire surjective suivie par une application linéaire injective

$$\begin{tikzcd} V \arrow[rr, "L"] \arrow[dr, "L_s"'] & & W \\ & \operatorname{Im} L \arrow[ur, "L_i"'] & \end{tikzcd}$$

avec $L_s(\vec{v}) = L(\vec{v})$ pour tout $\vec{v} \in V$ et $L_i(\vec{y}) = \vec{y}$ pour tout $\vec{y} \in \operatorname{Im} L$ (voir Remarque 8.2). De plus :

1. $\operatorname{Im} L_s = \operatorname{Im} L$ (car L_s est surjective) et $\operatorname{Ker} L_s = \operatorname{Ker} L$ (car $L = L_i \circ L_s$ et L_i est injective, en utilisant le point 5 de l'Exercice 9.17).
2. Si V est de type fini, alors $\operatorname{Im} L$ est de type fini aussi.

9.15 Corollaire. *(Théorème du Rang) Soit $L \colon V \to W$ une application linéaire, avec V de type fini. Alors :*

1. $V \simeq \operatorname{Ker} L \times \operatorname{Im} L$.

2. $\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$.

Preuve. Avec les notations de la Remarque 9.14, nous pouvons appliquer la Propriété 9.13 à l'application linéaire surjective

$$L_s \colon V \to \operatorname{Im} L$$

Nous obtenons $V \simeq \operatorname{Ker} L_s \times \operatorname{Im} L = \operatorname{Ker} L \times \operatorname{Im} L$. En utilisant 8.17 et 9.7 cet isomorphisme donne aussi $\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$. □

9.16 Remarque. Le théorème du rang 9.15 utilise le Lemme 9.12, qui est basé sur la propriété universelle de la base (8.15). Puisque la propriété universelle de la base reste valable dans le cas des espaces ayant une base infinie (voir Remarque 8.16) et puisque tout espace vectoriel admet une base finie ou infinie (voir Théorème 6.9), on peut se passer de l'hypothèse que W soit de type fini dans le Lemme 9.12 et, par conséquent, le théorème du rang peut s'énoncer dans la forme suivante : *soit $L \colon V \to W$ une application linéaire, alors*

$$V \simeq \operatorname{Ker} L \times \operatorname{Im} L$$

Par contre, le point 2. du 9.15 n'a plus de sens dans le cas d'espaces de dimension infinie.

9.17 Exercice. Considérons deux applications linéaires

$$V \xrightarrow{L} W \xrightarrow{T} U \ .$$

1. $\operatorname{Im}(T \circ L) \subseteq \operatorname{Im} T$.
2. Si L est surjective alors $\operatorname{Im}(T \circ L) = \operatorname{Im} T$.
3. Si L est bijective alors $\operatorname{Ker}(T \circ L) \simeq \operatorname{Ker} T$.
4. $\operatorname{Ker} L \subseteq \operatorname{Ker}(T \circ L)$.
5. Si T est injective alors $\operatorname{Ker} L = \operatorname{Ker}(T \circ L)$.
6. Si T est injective alors $\operatorname{Im} L \simeq \operatorname{Im}(T \circ L)$.

Preuve. 1. Exercice.
2. Soit $\vec{u} \in \operatorname{Im} T$. Par définition de $\operatorname{Im} T$, il existe $\vec{w} \in W$ tel que $T(\vec{w}) = \vec{u}$. Puisque L est surjective, il existe aussi $\vec{v} \in V$ tel que $L(\vec{v}) = \vec{w}$. Par conséquent,

$T(L(\vec{v})) = T(\vec{w}) = \vec{u}$, ce qui montre que $\vec{u} \in \mathrm{Im}(T \circ L)$, et donc $\mathrm{Im}\, T \subseteq \mathrm{Im}(T \circ L)$.
3. Soit $L' \colon \mathrm{Ker}(T \circ L) \to \mathrm{Ker}\, T$ la restriction de $L \colon V \to W$ aux noyaux (L' est bien définie car si $\vec{v} \in \mathrm{Ker}(T \circ L)$, alors $L(\vec{v}) \in \mathrm{Ker}\, T$). L'application L' est injective en tant que restriction de L qui est injective. Soit maintenant $\vec{w} \in \mathrm{Ker}\, T \subseteq W$. Puisque L est surjective, on peut trouver $\vec{v} \in V$ tel que $L(\vec{v}) = \vec{w}$, et ce $\vec{v}$ est dans $\mathrm{Ker}(T \circ L)$ car $T(L(\vec{v})) = T(\vec{w}) = \vec{0}$. Finalement, $L'(\vec{v}) = L(\vec{v}) = \vec{w}$, ce qui montre que L' est surjective.
4. Exercice.
5. Semblable à 2.
6. Semblable à 3. □

9.18 Exercice. Pour terminer cette section, voyons un exercice qui généralise le fait que la fonction

$$S \colon \mathrm{Ker}\, L \times W \to V \quad S(\vec{x}, \vec{w}) = \vec{x} + T(\vec{w})$$

utilisée dans la preuve de la Propriété 9.13, est linéaire.
Soient U et W deux espaces vectoriels.

1. Les fonctions

$$i_U \colon U \to U \times W \quad i_U(\vec{u}) = (\vec{u}, \vec{0}) \;\text{ et }\; i_W \colon W \to U \times W \quad i_W(\vec{w}) = (\vec{0}, \vec{w})$$

sont linéaires.

2. Si V est un troisième espace vectoriel et $R \colon U \to V$ et $T \colon W \to V$ sont deux applications linéaires, alors la fonction

$$S \colon U \times W \to V \quad S(\vec{u}, \vec{w}) = R(\vec{u}) + T(\vec{w})$$

est linéaire.

3. De plus, la fonction S définie au point précédent est l'unique application linéaire telle que $S \circ i_U = R$ et $S \circ i_W = T$

$$\begin{tikzcd} U \arrow[r, "i_U"] \arrow[dr, "R"'] & U \times W \arrow[d, "S"] & W \arrow[l, "i_W"'] \arrow[dl, "T"] \\ & V & \end{tikzcd}$$

Notons que ce même argument s'applique pour montrer que la fonction

$$S \colon V_1 \times \ldots \times V_k \to V \;, \quad S(\vec{v}_1, \ldots, \vec{v}_k) = \vec{v}_1 + \ldots + \vec{v}_k \,.$$

utilisée dans le Lemme 9.9 est linéaire.

9.4 Application : injectivité et surjectivité

Le prochain corollaire, qui est un cas particulier du Corollaire 9.15, justifie le nom "Théorème du Rang" :

9.19 Corollaire. *Soit* $A \in \mathbb{R}^{m\times n}$ *une matrice et soit* $\mathcal{S}_0\colon A \cdot \vec{x} = \vec{0}$ *le système homogène qui lui correspond. On a :*

$$\dim \operatorname{Sol}(\mathcal{S}_0) = n - \operatorname{rang} A\,.$$

Preuve. Considérons l'application linéaire $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ associée à la matrice A. Par le Corollaire 9.15 nous avons

$$n = \dim \mathbb{R}^n = \dim \operatorname{Ker} L_A + \dim \operatorname{Im} L_A = \dim \operatorname{Sol}(\mathcal{S}_0) + \operatorname{rang} A$$

où pour la troisième égalité on utilise l'Exemple 7.13 et le Corollaire 7.15. □

9.20 Corollaire. *Soit* $A \in \mathbb{R}^{m\times n}$ *une matrice et soit* $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ *l'application linéaire associée. Alors :*

1. L_A *est surjective ssi* $\operatorname{rang} A = m$.
2. L_A *est injective ssi* $\operatorname{rang} A = n$.

Preuve. 1. En utilisant que $\operatorname{Im} L_A = \operatorname{Col}(A)$ et donc que $\dim \operatorname{Im} L_A = \dim \operatorname{Col}(A) = \operatorname{rang} A$, nous avons

$$\begin{aligned} L_A \text{ surjective} \quad &\Leftrightarrow \quad \operatorname{Im} L_A = \mathbb{R}^m \\ &\Leftrightarrow \quad \dim \operatorname{Im} L_A = m \\ &\Leftrightarrow \quad \operatorname{rang} A = m \end{aligned}$$

2. En utilisant la formule $\dim \mathbb{R}^n = \dim \operatorname{Ker} L_A + \dim \operatorname{Im} L_A$ du Corollaire 9.15, nous avons

$$\begin{aligned} L_A \text{ injective} \quad &\Leftrightarrow \quad \operatorname{Ker} L_A = \{\vec{0}\,\} \\ &\Leftrightarrow \quad \dim \operatorname{Ker} L_A = 0 \\ &\Leftrightarrow \quad \dim \operatorname{Im} L_A = n \\ &\Leftrightarrow \quad \operatorname{rang} A = n \end{aligned}$$

□

Le Corollaire 9.20 peut aussi se formuler à partir d'une application linéaire entre espaces vectoriels de type fini.

9.21 Corollaire. *Soit* $L\colon V \to W$ *une application linéaire entre espaces vectoriels de type fini, soit* $n = \dim V, m = \dim W$ *et soit* $A = {}_fL_e \in \mathbb{R}^{m\times n}$ *la matrice associée à* L *par rapport à une base* e *de* V *et à une base* f *de* W*. Alors :*

1. L *est surjective ssi* $\operatorname{rang} A = m$.
2. L *est injective ssi* $\operatorname{rang} A = n$.

Preuve. Considérons le diagramma commutatif de la Remarque 8.23

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "E"'] & W \arrow[d, "F"] \\ \mathbb{R}^n \arrow[r, "L_A"'] & \mathbb{R}^m \end{tikzcd}$$

où E et F sont deux isomorphismes. En utilisant l'Exercice 8.3 et le Corollaire 9.20, nous avons :

1. L est surjective ssi L_A est surjective ssi rang $A = m$.
2. L est injective ssi L_A est injective ssi rang $A = n$.

$\square$

Comme cas particulier du Corollaire 9.15 nous avons aussi la propriété suivante.

9.22 Propriété. *Soit V, W deux espaces vectoriels ayant même dimension et soit $L\colon V \to W$ une applcation linéaire. Alors L est injective si et seulement si elle est surjective.*

Preuve. En utilisant la formule $\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L$ du 9.15 nous avons :

$$\begin{aligned} L \text{ est injective} \quad &\Leftrightarrow \quad \operatorname{Ker} L = \{\vec{0}\} \\ &\Leftrightarrow \quad \dim \operatorname{Ker} L = 0 \\ &\Leftrightarrow \quad \dim \operatorname{Im} L = \dim V \\ &\Leftrightarrow \quad \operatorname{Im} L = W \\ &\Leftrightarrow \quad L \text{ est surjective.} \end{aligned}$$

$\square$

9.23 Remarque. On pourrait espérer que la Propriété 9.22 reste valable pour les espaces de dimension infinie en remplaçant l'hypothèse $\dim V = \dim W$ par l'hypothèse $V \simeq W$. Ceci n'est pas vrai, comme l'exemple suivant le montre. L'application linéaire $D\colon \mathbb{R}[x] \to \mathbb{R}[x]$ qui envoie un polynôme sur son polynôme dérivé est surjective mais elle n'est pas injective.

9.5 Application : somme de sous-espaces

La prochaine propriété généralise le Corollaire 9.10 au cas d'une somme quelconque (pas nécessairement directe) de sous-espaces vectoriels de type fini.

9.24 Propriété. *Soit V un espace vectoriel et soit $V_1, \ldots, V_k$ des s.e.v. finiment engendrés de V. On a :*

$$\dim(V_1 + \ldots + V_k) = \sum_{i=1}^{k} \dim V_i - \sum_{i=1}^{k-1} \dim((V_1 + \ldots + V_i) \cap V_{i+1})\,.$$

En particulier, pour $k = 2$, on a :

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)\,.$$

Preuve. Par induction sur k:
$k = 2$: Considérons l'application linéaire

$$S\colon V_1 \times V_2 \to V\ , \quad S(\vec{v}_1, \vec{v}_2) = \vec{v}_1 + \vec{v}_2\,.$$

On a $\operatorname{Im} S = V_1 + V_2$ et $\operatorname{Ker} S \simeq V_1 \cap V_2$, car

$$\operatorname{Ker} S = \{(\vec{v}_1, \vec{v}_2) \mid \vec{v}_1 + \vec{v}_2 = \vec{0}\,\} = \{(-\vec{v}_2, \vec{v}_2) \mid \vec{v}_2 \in V_1 \cap V_2\}\,.$$

La formule $\dim V = \dim \operatorname{Ker} S + \dim \operatorname{Im} S$ du 9.15 nous donne :

$$\dim V_1 + \dim V_2 = \dim(V_1 \times V_2) = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)\,.$$

$k > 2$: $\dim(V_1 + \ldots + V_{k-1} + V_k) = \dim((V_1 + \ldots + V_{k-1}) + V_k) =$

$$= \dim(V_1 + \ldots + V_{k-1}) + \dim V_k - \dim((V_1 + \ldots + V_{k-1}) \cap V_k) =$$

$$= \sum_{i=1}^{k-1} \dim V_i - \sum_{i=1}^{k-2} \dim((V_1 + \ldots + V_i) \cap V_{i+1}) + \dim V_k - \dim((V_1 + \ldots + V_{k-1}) \cap V_k) =$$

$$= \sum_{i=1}^{k} \dim V_i - \sum_{i=1}^{k-1} \dim((V_1 + \ldots + V_i) \cap V_{i+1})\,.$$

□

9.25 Remarque. Si la somme $V_1 + \ldots + V_k$ de la Propriété 9.24 est une somme directe, pour tout i le sous-espace vectoriel $(V_1 + \ldots + V_i) \cap V_{i+1}$ est réduit à zéro (voir Définition 9.8) et donc sa dimension vaut zéro. Par conséquent, la formule

$$\dim(V_1 + \ldots + V_k) = \sum_{i=1}^{k} \dim V_i - \sum_{i=1}^{k-1} \dim((V_1 + \ldots + V_i) \cap V_{i+1})$$

de la Propriété 9.24 se reconduit à la formule

$$\dim(V_1 + \ldots + V_k) = \dim V_1 + \ldots + \dim V_k$$

du Corollaire 9.10.

9.6 Application : espace lignes et espace colonnes

Pour terminer ce chapitre, nous allons démontrer la Propriété 4.14 en utilisant la décomposition $L = L_i \circ L_s$ (voir 9.14) et la transposée d'une matrice :

9.26 Définition. Soit $A \in \mathbb{R}^{m \times n}$. La *transposée* de A est la matrice

$$A^t \in \mathbb{R}^{n \times m}$$

dont les lignes sont les colonnes de A (et dont les colonnes sont les lignes de A). Autrement dit, si $A = (a_{ij})$ alors $A^t = (b_{ij})$ avec $b_{ij} = a_{ji}$.

9.27 Exercice. Soit $B \in \mathbb{R}^{m \times n}$ et $C \in \mathbb{R}^{n \times p}$. Montrer que $(B \cdot C)^t = C^t \cdot B^t$.

9.28 Lemme. *Considérons deux matrices $B \in \mathbb{R}^{m \times n}$ et $C \in \mathbb{R}^{n \times p}$.*

1. $\mathrm{Col}(B \cdot C) \subseteq \mathrm{Col}(B)$.
2. $\mathrm{Lign}(B \cdot C) \subseteq \mathrm{Lign}(C)$.

Preuve. 1. Passons aux applications linéaires associées :

$$\mathbb{R}^p \xrightarrow{L_C} \mathbb{R}^n \xrightarrow{L_B} \mathbb{R}^m$$

En utilisant l'Exercice 9.17 nous avons

$$\mathrm{Col}(B \cdot C) = \mathrm{Im}\, L_{B \cdot C} = \mathrm{Im}(L_B \circ L_C) \subseteq \mathrm{Im}\, L_B = \mathrm{Col}(B)$$

2. En utilisant le point 1. nous avons

$$\mathrm{Lign}(B \cdot C) = \mathrm{Col}((B \cdot C)^t) = \mathrm{Col}(C^t \cdot B^t) \subseteq \mathrm{Col}(C^t) = \mathrm{Lign}(C)$$

□

9.29 Propriété. *Soit* $A \in \mathbb{R}^{m \times n}$. *Les espaces vectoriels* $\mathrm{Col}(A)$ *et* $\mathrm{Lign}(A)$ *ont même dimension.*

Preuve. Soit $c = \dim \mathrm{Col}(A)$ et $l = \dim \mathrm{Lign}(A) = \dim \mathrm{Col}(A^t)$, nous allons démontrer que $l \leq c$ et $c \leq l$.

$l \leq c$: considérons l'application linéaire $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ et sa décomposition en partie surjective et partie injective (Remarque 9.14)

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "L_A"] \arrow[dr, "L_s"'] & & \mathbb{R}^m \\ & \mathrm{Im}\, L_A \arrow[ur, "L_i"'] & \end{tikzcd}$$

Puisque $\mathrm{Im}\, L_A = \mathrm{Col}(A)$ et $\dim \mathrm{Col}(A) = c$, l'espace $\mathrm{Im}\, L_A$ est isomorphe à $\mathbb{R}^c$. Soit donc $E \colon \mathrm{Im}\, L_A \to \mathbb{R}^c$ un isomorphisme, nous avons deux applications linéaires

$$\mathbb{R}^n \xrightarrow{L_s} \mathrm{Im}\, L_A \xrightarrow{E} \mathbb{R}^c \qquad \mathbb{R}^c \xrightarrow{E^{-1}} \mathrm{Im}\, L_A \xrightarrow{L_i} \mathbb{R}^m$$

Par le théorème de représentation des applications linéaires, la première est du type L_C pour une certaine matrice $C \in \mathbb{R}^{c \times n}$ et la deuxième est du type L_B pour une certaine matrice $B \in \mathbb{R}^{m \times c}$. De plus, on a

$$L_A = L_i \circ L_s = L_i \circ E^{-1} \circ E \circ L_s = L_B \circ L_C = L_{B \cdot C}$$

et donc $A = B \cdot C$. Par le Lemme 9.28 cela implique $\mathrm{Lign}(A) \subseteq \mathrm{Lign}(C)$ et donc

$$l = \dim \mathrm{Lign}(A) \leq \dim \mathrm{Lign}(C) \leq c$$

où la deuxième inégalité est due au fait que c est égale au nombre de lignes de la matrice C.

$c \leq l$: considérons la matrice $A^t \in \mathbb{R}^{n\times m}$, l'application linéaire $L_{A^t}\colon \mathbb{R}^m \to \mathbb{R}^n$ et sa décomposition en partie surjective et partie injective (Remarque 9.14)

$$\begin{tikzcd} \mathbb{R}^m \arrow[rr, "L_{A^t}"] \arrow[dr, "T_s"'] & & \mathbb{R}^n \\ & \operatorname{Im} L_{A^t} \arrow[ur, "T_i"'] & \end{tikzcd}$$

Puisque $\operatorname{Im} L_{A^t} = \operatorname{Col}(A^t) = \operatorname{Lign}(A)$ et $\dim \operatorname{Lign}(A) = l$, l'espace $\operatorname{Im} L_{A^t}$ est isomorphe à $\mathbb{R}^l$. Nous avons donc deux applications linéaires

$$\mathbb{R}^m \xrightarrow{T_s} \operatorname{Im} L_{A^t} \xrightarrow{\simeq} \mathbb{R}^l$$

$$\mathbb{R}^l \xrightarrow{\simeq} \operatorname{Im} L_{A^t} \xrightarrow{T_i} \mathbb{R}^n$$

Par le théorème de représentation des applications linéaires, la première est du type L_Y pour une certaine matrice $Y \in \mathbb{R}^{l\times m}$ et la deuxième est du type L_X pour une certaine matrice $X \in \mathbb{R}^{n\times l}$. De plus, on a $L_{A^t} = L_X \circ L_Y$ et donc $A^t = X \cdot Y$. Par le Lemme 9.28 cela implique $\operatorname{Lign}(A^t) \subseteq \operatorname{Lign}(Y)$ et donc

$$c = \dim \operatorname{Col}(A) = \dim \operatorname{Lign}(A^t) \leq \dim \operatorname{Lign}(Y) \leq l$$

où la deuxième inégalité est due au fait que l est égale au nombre de lignes de la matrice Y. $\square$

Chapitre 10

Espace quotient

Le chapitre 10 section par section

Dans le chapitre précédent, nous avons étudié le produit d'espaces vectoriels. Cela nous a permis d'établir le théorème du rang pour une application linéaire $L\colon V \to W$ (Corollaire 9.15) : $V \simeq \operatorname{Ker} L \times \operatorname{Im} L$
Dans ce chapitre nous allons étudier le quotient d'un espace vectoriel par un sous-espace, ce qui nous permettra de reformuler le théorème du rang sous la forme : $\operatorname{Im} L \simeq V/\operatorname{Ker} L$

1. Nous commençons par introduire la notion de relation d'équivalence sur un ensemble. Quotienter un ensemble par une relation d'équivalence signifie construire un nouvel ensemble où les éléments qui etaint équivalents deviennent égaux.
2. Tout sous-espace vectoriel induit une relation d'équivalence sur l'espace vectoriel. L'ensemble quotient hérite de l'espace de départ une structure d'espace vectoriel : c'est l'espace quotient.

10.1 Relations d'équivalence

Pour arriver à la notion d'espace quotient, nous introduisons d'abord la notion de relation d'équivalence sur un ensemble.

10.1 Définition.

1. Une *relation* $\mathcal{R}$ entre deux ensembles X et Y est un sous-ensemble du produit cartesien : $\mathcal{R} \subseteq X \times Y$.
   Si $(x, y) \in \mathcal{R}$, nous écrivons $x\mathcal{R}y$ et nous lisons "x est en relation $\mathcal{R}$ avec y".
2. Une *relation d'équivalence* sur un ensemble X est une relation $\mathcal{R} \subseteq X \times X$ qui satisfait aux trois conditions suivantes :
   (a) $\mathcal{R}$ est réflexive : $\forall\, x \in X$, $x\mathcal{R}x$.
   (b) $\mathcal{R}$ est transitive : $\forall\, x, y, z \in X$, si $x\mathcal{R}y$ et $y\mathcal{R}z$ alors $x\mathcal{R}z$.

(c) $\mathcal{R}$ est symétrique : $\forall\, x, y \in X$, si $x\mathcal{R}y$ alors $y\mathcal{R}x$.

10.2 Exemple.

1. Soit $X = \mathbb{Z}$ (l'ensemble des nombres entiers) et soit $p \in \mathbb{N}$ un nombre naturel fixé. On obtient une relation d'équivalence sur $\mathbb{Z}$ en posant : $x\mathcal{R}y$ s'il existe $n \in \mathbb{Z}$ tel que $x - y = n \cdot p$.
2. Soit X l'ensemble des droites du plan. On obtient une relation d'équivalence sur X en posant : $\alpha\mathcal{R}\beta$ si les droites α et β sont parallèles.
3. Soit V un espace vectoriel et soit X l'ensemble des vecteurs non nuls de V. On obtient une relation d'équivalence sur X en posant : $\vec{x}\mathcal{R}\vec{y}$ s'il existe $\lambda \in \mathbb{R}, \lambda \neq 0$, tel que $\vec{x} = \lambda \cdot \vec{y}$.
4. Soit $X = \mathbb{R}^2$. On obtient une relation d'équivalence sur X en posant : $\vec{x}\mathcal{R}\vec{y}$ si $\mathrm{dist}(\vec{x}, \vec{0}\,) = \mathrm{dist}(\vec{y}, \vec{0}\,)$ où, pour $\vec{x} = (x_1, x_2)$,
$$\mathrm{dist}(\vec{x}, \vec{0}\,) = \sqrt{x_1^2 + x_2^2}$$
est la distance euclidienne entre le point $\vec{x}$ et l'origine.
5. Soit X l'ensemble de tous les espaces vectoriels réels. On obtient une relation d'équivalence sur X en posant : $V\mathcal{R}W$ si V et W sont isomorphes.
6. Soit $X = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. On obtient une relation d'équivalence sur X en posant : $(a, b)\mathcal{R}(c, d)$ si $a \cdot d = b \cdot c$.
7. Soit X un ensemble quelconque. La relation
$$x\mathcal{R}y \ \text{ ssi } \ x = y$$
est la plus petite relation d'équivalence sur X. Elle est appelée la relation discrète et notée Δ. La relation
$$x\mathcal{R}y \ \text{ pour tout } \ x, y \in X$$
est la plus grande relation d'équivalence sur X. Elle est appelée la relation indiscrète ou grossière et notée ∇.
8. Soit V un espace vectoriel et soit U un sous-espace vectoriel de V. On obtient une relation d'équivalence $\mathcal{R}[U]$ sur V en posant :
$$\vec{x}\mathcal{R}[U]\vec{y} \ \text{ si } \ \vec{x} - \vec{y} \in U$$
La relation $\mathcal{R}[U]$ est appelée relation su V induite par U. Cet exemple est le plus important pour la suite du chapitre.

10.3 Définition. Soit X un ensemble et $\mathcal{R}$ une relation d'équivalence sur X.

1. Soit $x \in X$. La *classe d'équivalence* de x est l'ensemble des éléments qui sont en relation $\mathcal{R}$ avec x :
$$[x] = \{y \in X \mid x\mathcal{R}y\}$$
Si $y \in [x]$, on dit que y est un *représentant* de la classe $[x]$.

2. L'*ensemble quotient* de X par $\mathcal{R}$ est l'ensemble des classes d'équivalence :

$$X/\mathcal{R} = \{[x] \mid x \in X\}$$

3. La *projection sur le quotient* est la fonction

$$\pi \colon X \to X/\mathcal{R} \quad \pi(x) = [x]$$

10.4 Exercice. Soit X un ensemble et $\mathcal{R}$ une relation d'équivalence sur X. Montrer que

1. $[x] = [y]$ ssi $x\mathcal{R}y$.
2. Les classes d'équivalence sont non vides, deux à deux disjointes, et leur réunion est égale à X tout entier (c'est-à-dire, l'ensemble quotient $X/\mathcal{R}$ est une partition de X).

10.5 Exemple. Décrivons l'ensemble quotient pour les relations d'équivalence introduites dans l'Exemple 10.2 :

1. L'ensemble quotient $\mathbb{Z}/\mathcal{R}$ est l'ensemble fini qui contient p éléments $\mathbb{Z}_p = \{[0], [1], \ldots, [p-1]\}$.
2. L'ensemble quotient est en bijection avec l'ensemble des droites par l'origine.
4. L'ensemble quotient est en bijection avec l'ensemble des cercles centrés à l'origine.
6. L'ensemble quotient est en bijection avec l'ensemble $\mathbb{Q}$ des nombres rationnels.
7. $X/\Delta = X$ et X/∇ contient un seul élément.
8. Dans cet exemple, la classe d'équivalence $[\vec{x}]$ est parfois notée $\vec{x} + U$

$$[\vec{x}] = \vec{x} + U = \{\vec{y} \in V \mid \vec{x} - \vec{y} \in U\}$$

en accord avec la Notation 7.10. En particulier, $[\vec{0}\,] = U$. On reviedra sur cette exemple dans la Proposition 10.9.

10.6 Exemple. Soit $f \colon X \to Y$ une fonction. La relation $\mathcal{R}[f]$ sur X définie par

$$x\mathcal{R}[f]y \text{ ssi } f(x) = f(y)$$

est une relation d'équivalence, dit *relation nucléaire* de f.
Cet exemple est générique, au sens que si $\mathcal{R}$ est une relation d'équivalence sur un ensemble X et $\pi \colon X \to X/\mathcal{R}$ est la projection sur le quotient, alors $\mathcal{R}$ coincide avec la relation nucléaire $\mathcal{R}[\pi]$. En effet, en utilisant la définition de π et le premier point de l'exercice 10.4, nous avons

$$x\mathcal{R}[\pi]y \text{ ssi } \pi(x) = \pi(y) \text{ ssi } [x] = [y] \text{ ssi } x\mathcal{R}y$$

10.7 Propriété. *Soit X un ensemble et $\mathcal{R}$ une relation d'équivalence sur X.*

1. *La projection* $\pi\colon X \to X/\mathcal{R}$ *est surjective, et* $\pi(x) = \pi(y)$ *ssi* $x\mathcal{R}y$.
2. *Si* $f\colon X \to Y$ *est une fonction telle que* $f(x) = f(y)$ *à chaque fois que* $x\mathcal{R}y$*, alors il existe une unique fonction* $\widehat{f}\colon X/\mathcal{R} \to Y$ *telle que* $\widehat{f} \circ \pi = f$

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "f"'] & & X/\mathcal{R} \arrow[dl, "\widehat{f}"] \\ & Y & \end{tikzcd}$$

Preuve.

1. Un élément de $X/\mathcal{R}$ est une classe d'équivalence s. Puisque s est non vide, on peut choisir un représentant $x \in s$ et on a $\pi(x) = [x] = s$, ce qui montre la surjectivité de π.
2. Pour que $\widehat{f} \circ \pi = f$, il faut poser $\widehat{f}([x]) = f(x)$. Il reste à montrer que cette définition est bien donnée, c'est-à-dire que si x et y sont deux représentants de la même classe, alors $f(x) = f(y)$. Mais ceci est exactement l'hypothèse sur f.

$\square$

10.8 Corollaire. *Soit* $f\colon X \to Y$ *une fonction et* $\mathcal{R}[f]$ *sa relation nucléaire. Considérons le diagramme de la Propriété 10.7*

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "f"'] & & X/\mathcal{R}[f] \arrow[dl, "\widehat{f}"] \\ & Y & \end{tikzcd}$$

1. $\widehat{f}$ *est injective,*
2. $\widehat{f}$ *est surjective ssi* f *est surjective.*

Autrement dit, toute surjection est, à isomorphisme près, la projection sur un ensemble quotient.

Preuve.

1. $\widehat{f}([x]) = \widehat{f}([y])$ ssi $f(x) = f(y)$ ssi $x\mathcal{R}[f]y$ ssi $[x] = [y]$ dans $X/\mathcal{R}[f]$.
2. C'est une conséquence facile de la condition $\widehat{f} \circ \pi = f$.

$\square$

10.2 Espace vectoriel quotient

10.9 Propriété. *Soit* V *un espace vectoriel et* U *un sous-espace vectoriel de* V*. Considérons la relation d'équivalence* $\mathcal{R}[U]$ *sur* V *définie par*

$$\vec{x}\mathcal{R}[U]\vec{y} \;\; ssi \;\; \vec{x} - \vec{y} \in U$$

(voir Exemple 10.2). Notons l'ensemble quotient $V/\mathcal{R}[U]$ *par* V/U.

1. *V/U est un espace vectoriel par rapport aux opérations définies sur les représentants. En particulier, le zéro de V/U est la classe de $\vec{0}$, c'est-à-dire U. L'espace V/U est dit* espace quotient.
2. *La projection $\pi\colon V \to V/U$ est une application linéaire surjective et son noyau est U.*

Preuve.

1. Posons $[\vec{x}]+[\vec{y}] = [\vec{x}+\vec{y}]$ et $\alpha\cdot[\vec{x}] = [\alpha\cdot\vec{x}]$. Il faut monter que ces opérations sont bien définies. Pour la somme, cela revient à montrer que si $\vec{x}\mathcal{R}[U]\vec{x}'$ et $\vec{y}\mathcal{R}[U]\vec{y}'$, alors $(\vec{x}+\vec{y})\mathcal{R}[U](\vec{x}'+\vec{y}')$. En effet, si $\vec{x}-\vec{x}'$ et $\vec{y}-\vec{y}'$ sont dans U, alors $(\vec{x}+\vec{y})-(\vec{x}'+\vec{y}') = (\vec{x}-\vec{x}')+(\vec{y}-\vec{y}')$ est dans U car U est un s.e.v. De même, pour le produit par un scalaire, cela revient à montrer que si $\vec{x}\mathcal{R}[U]\vec{x}'$, alors $(\alpha\cdot\vec{x})\mathcal{R}[U](\alpha\cdot\vec{x}')$. En effet, si $\vec{x}-\vec{x}'$ est dans U, alors $\alpha\cdot\vec{x}-\alpha\cdot\vec{x}' = \alpha\cdot(\vec{x}-\vec{x}')$ est dans U car U est un s.e.v.
2. La surjectivité et la linéarité de π sont évidentes. Calculons son noyau : $\pi(\vec{x}) = \pi(\vec{0})$ ssi $[\vec{x}] = [\vec{0}]$ ssi $\vec{x}\mathcal{R}[U]\vec{0}$ ssi $\vec{x}-\vec{0}\in U$ ssi $\vec{x}\in U$.

$\square$

10.10 Remarque. Nous venons de voir que si U est un sous-espace vectoriel d'un espace vectoriel V, alors $\mathcal{R}[U]$ n'est pas seulement une relation d'équivalence sur V, mais elle est une congruence, c'est-à-dire

1. si $\vec{x}\mathcal{R}[U]\vec{x}'$ et $\vec{y}\mathcal{R}[U]\vec{y}'$, alors $(\vec{x}+\vec{y})\mathcal{R}[U](\vec{x}'+\vec{y}')$
2. si $\vec{x}\mathcal{R}[U]\vec{x}'$, alors $(\alpha\cdot\vec{x})\mathcal{R}[U](\alpha\cdot\vec{x}')$.

Réciproquement, si $\mathcal{R}$ est une congruence sur V, alors la classe d'équivalence $[\vec{0}\,]$ est un sous-espace vectoriel de V et $\mathcal{R}[\vec{0}\,] = \mathcal{R}$. Ceci établie une bijection entre sous-espaces vectoriels de V et congruences sur V.

10.11 Propriété. *Soit V un espace vectoriel, U un sous-espace vectoriel de V, et $L\colon V \to W$ une application linéaire telle que $U \subseteq \operatorname{Ker} L$.*

1. *Il existe une unique application linéaire $\widehat{L}\colon V/U \to W$ telle que $\widehat{L}\circ\pi = L$*

$$\begin{tikzcd} V \arrow[rr, "\pi"] \arrow[dr, "L"'] & & V/U \arrow[dl, "\widehat{L}"] \\ & W & \end{tikzcd}$$

2. *$\widehat{L}$ est injective ssi* $\operatorname{Ker} L = U$.

Preuve.

1. Si $\vec{x}\mathcal{R}[U]\vec{y}$, alors $\vec{x}-\vec{y}\in U \subseteq \operatorname{Ker} L$ et donc $L(\vec{x}) = L(\vec{y})$. Nous pouvons maintenant appliquer la Proposition 10.7.
2. Si $\widehat{L}$ est injective, alors en utilisant l'Exercice 9.17 et la Propriété 10.9 nous avons : $\operatorname{Ker} L = \operatorname{Ker}(\widehat{L}\circ\pi) = \operatorname{Ker}\pi = U$. Réciproquement, si $\widehat{L}([\vec{x}]) = \vec{0}$, alors $L(\vec{x}) = \vec{0}$ et donc $\vec{x}\in\operatorname{Ker} L = U$. Cela nous dit que $[\vec{x}] = U$, le zéro de V/U, et donc $\widehat{L}$ est injective.

□

Nous arrivons à la version du théorème du rang annoncée au début du chapitre.

10.12 Corollaire. *Soit $L\colon V \to W$ une application linéaire. On a un isomorphisme canonique*

$$\operatorname{Im} L \simeq V/\operatorname{Ker} L$$

Preuve. Considérons la factorisation de L comme composée d'une application linéaire surjective L_s suivie par une application linéaire injective L_i (Remarque 9.14)

$$\begin{tikzcd} V \arrow[rr, "L"] \arrow[dr, "L_s"'] & & W \\ & \operatorname{Im} L \arrow[ur, "L_i"'] & \end{tikzcd}$$

Puisque $\operatorname{Ker} L_s = \operatorname{Ker} L$, nous pouvons appliquer la Propriété 10.11 à la situation

$$\begin{tikzcd} V \arrow[rr, "\pi"] \arrow[dr, "L_s"'] & & V/\operatorname{Ker} L \arrow[dl, "\widehat{L_s}"] \\ & \operatorname{Im} L & \end{tikzcd}$$

et $\widehat{L_s}$ est injective. De plus, $\widehat{L_s}$ est aussi surjective car L_s est surjective. □

10.13 Remarque. Nous savons qu'une application linéaire $L\colon V \to W$ est injective ssi son noyau est réduit à zéro (voir Propriété 8.9). L'espace quotient permet d'introduire la notion de *conoyau* qui mesure la surjectivité de l'application L. Pour cela, considérons le sous-espace vectoriel $\operatorname{Im} L$ de W et posons

$$\operatorname{Coker} L = W/\operatorname{Im} L$$

On a alors que :

1. Pour tout $\vec{x} \in V, \pi(L(\vec{x})) = \vec{0}$.
2. Si $T\colon W \to Z$ est une application linéaire telle que pour tout $\vec{x} \in V, T(L(\vec{x})) = \vec{0}$, alors il existe une unique application linéaire $\widehat{T}\colon \operatorname{Coker} L \to Z$ telle que $\widehat{T} \circ \pi = T$

$$\begin{tikzcd} V \arrow[r, "L"] & W \arrow[r, "\pi"] \arrow[d, "T"'] & \operatorname{Coker} L \arrow[dl, "\widehat{T}"] \\ & Z & \end{tikzcd}$$

3. L est surjective ssi $\operatorname{Coker} L$ est réduit à zéro.

Preuve. La preuve est un exercice laissé au lecteur. □

Chapitre 11

Déterminant

Le chapitre 11 section par section

D'après l'Exemple 8.12, nous savons que pour tout vecteur $\vec{b} \in \mathbb{R}^m$ le système $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ admet une solution unique si et seulement si l'application linéaire $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ est bijective. Dans ce chapitre nous allons étudier une méthode qui permet de déterminer si L_A est bijective.

1. D'abord on peut traduire le fait que L_A est bijective par une condition sur A : l'inversibilité.
2. Dans le cas d'un système de deux équations en deux variables, le fait d'avoir une unique solution peut s'exprimer en utilisant la pente des droites. Ceci nous permet de découvrir la définition de déterminant pour une matrice 2×2.
3. Dans le cas $n \geq 2$, la définition du déterminant est axiomatique : le déterminant est une fonction $\det\colon \mathbb{R}^{n\times n} \to \mathbb{R}$ qui satisfait à trois conditions.
4. On peut alors démontrer l'unicité de la fonction déterminant, en utilisant la forme réduite de Gauss-Jordan d'une matrice, et ...
5. ... démontrer par induction son existence (formule de Laplace).
6. Deux propriétés qui permettent parfois de simplifier le calcul du déterminant sont le fait que le déterminant préserve le produit matriciel, c'est-à-dire $\det(A \cdot B) = \det A \cdot \det B$, et ...
7. ... le fait que le déterminant ne change pas si on transpose la matrice, c'est-à-dire $\det A^t = \det A$.
8. Pour terminer, le déterminant peut être utilisé pour calculer le rang d'une matrice quelconque et la solution d'un système à n équations et n variables dans le cas où une telle solution soit unique (règle de Cramer).

11.1 Matrices inversibles

Commençons par traduire la bijectivité de L_A par une condition sur la matrice A:

11.1 Lemme. *Soit $A \in \mathbb{R}^{m\times n}$.*

1. *L_A est surjective ssi il existe une matrice $B \in \mathbb{R}^{n\times m}$ telle que $A \cdot B = I_m$. Dans ce cas, on dit que A est inversible à droite.*
2. *L_A est injective ssi il existe une matrice $C \in \mathbb{R}^{n\times m}$ telle que $C \cdot A = I_n$. Dans ce cas, on dit que A est inversible à gauche.*

Preuve.

1. Si $A \cdot B = I$ pour une matrice $B \in \mathbb{R}^{n\times m}$, nous pouvons considérer les applications linéaires
$$L_A \colon \mathbb{R}^n \to \mathbb{R}^m \quad L_B \colon \mathbb{R}^m \to \mathbb{R}^n$$
et nous avons $L_A \circ L_B = L_{A\cdot B} = L_I = id_{\mathbb{R}^m}$, ce qui implique que L_A est surjective.
Vice-versa, si L_A est surjective le Lemme 9.12 nous garantit l'existence d'une application linéaire $T \colon \mathbb{R}^m \to \mathbb{R}^n$ telle que $L_A \circ T = id_{\mathbb{R}^m}$. Par le théorème de représentation des applications linéaire (8.19), $T = L_B$ pour une certaine matrice $B \in \mathbb{R}^{n\times m}$, et nous avons
$$L_{A\cdot B} = L_A \circ L_B = L_A \circ T = id_{\mathbb{R}^m} = L_I$$
d'où $A \cdot B = I$ (voir 7.26).
2. La preuve du point 2 est semblable à celle du point 1 et elle est laissée en exercice.

□

11.2 Corollaire. *Soit $A \in \mathbb{R}^{m\times n}$. Si A est inversible à droite et inversible à gauche alors $m = n$.*

Preuve. Si A est inversible à droite et à gauche, par le Lemme 11.1 l'application linéaire L_A est bijective et donc $\mathbb{R}^n$ et $\mathbb{R}^m$ sont isomorphes. Mais deux espaces isomorphes ont même dimension. □

11.3 Corollaire. *Soit $A \in \mathbb{R}^{n\times n}$. Les conditions suivantes sont équivalentes :*

1. *$L_A \colon \mathbb{R}^n \to \mathbb{R}^n$ est bijective,*
2. *A est inversible (= inversible à droite et inversible à gauche),*
3. *A est inversible à droite,*
4. *A est inversible à gauche,*
5. *il existe une matrice A^{-1} telle que $A \cdot A^{-1} = I_n = A^{-1} \cdot A$ (une telle matrice A^{-1} est nécessairement unique),*
6. $\operatorname{rang} A = n$,

7. *pour tout $\vec{b} \in \mathbb{R}^n$, le système $S\colon A \cdot \vec{x} = \vec{b}$ admet une solution unique,*
8. *le système homogène $S_0\colon A \cdot \vec{x} = \vec{0}$ admet une solution unique (qui est nécessairement la solution $\vec{0}$),*
9. *les colonnes de A sont linéairement indépendantes,*
10. *les lignes de A sont linéairement indépendantes.*

Preuve. L'équivalence entre les conditions 1. 2. 3. et 4. est une conséquence du Lemme 11.1 et de la Propriété 9.22. Pour avoir l'équivalence avec la condition 5. il reste à montrer que si $B, C \in \mathbb{R}^{n\times n}$ sont telles que

$$A \cdot B = I_n \ \text{ et } \ C \cdot A = I_n$$

alors $B = C$. En effet :

$$C = C \cdot I = C \cdot (A \cdot B) = (C \cdot A) \cdot B = I \cdot B = B$$

L'équivalence entre la condition 1. et la condition 6. est une conséquence du Corollaire 9.20. Pour avoir l'équivalence entre la condition 1. et la condition 7., il suffit d'écrire la condition 7. en remplaçant $A \cdot \vec{x} = \vec{b}$ par $L_A(\vec{x}) = \vec{b}$. Pour l'équivalence entre la condition 1. et la condition 8., il suffit d'observer que la bijectivité de L_A équivaut à son injectivité (Propriété 9.22) et donc au fait que le noyau de L_A est réduit au vecteur nul (Propriété 8.9). Or, le noyau de L_A est exactement l'ensemble des solutions du système homogène S_0 (Exemple 7.13). L'équivalence entre les conditions 6., 9. et 10. est une conséquence de la Propriété 9.29. $\square$

11.4 Remarque.

1. Le fait qu'une matrice est inversible à gauche ssi elle est inversible à droite est valable pour les matrices carrées (11.3) mais pas en général : la matrice $A = (1, 0) \in \mathbb{R}^{1\times 2}$ est inversible à droite mais pas à gauche.
2. Les matrices inversibles sont appelées aussi *régulières*, les matrices non inversibles sont appelées aussi *singulières*.

11.5 Exercice. Montrer que :

1. La matrice identité I est inversible.
2. Si $A, B \in \mathbb{R}^{n\times n}$ sont inversibles, alors $A \cdot B$ est inversible et $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.
3. Toute matrice élémentaire est inversible.

11.6 Remarque. Si une matrice $A \in \mathbb{R}^{n\times n}$ est inversible, on peut tirer du Corollaire 11.3 que

1. pour tout $\vec{b} \in \mathbb{R}^n$, le système $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ admet une unique solution, qui est donnée par le vecteur $\vec{x} = A^{-1} \cdot \vec{b}$,
2. la i-ème colonne de A^{-1} est l'unique solution du système $A \cdot \vec{x} = \vec{e}_i$, où $\vec{e}_i$ est le i-ème vecteur de la base canonique de $\mathbb{R}^n$. En effet :

$$\begin{aligned} \vec{e}_i &= i\text{-ème colonne de } I \\ &= i\text{-ème colonne de } A \cdot A^{-1} \\ &= A \cdot (i\text{-ème colonne de } A^{-1}) \end{aligned}$$

11.2 Définition géométrique pour $n = 2$

Nous avons donc établi que pour une matrice $A \in \mathbb{R}^{n\times n}$ les trois conditions

- A est inversible,
- $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ est bijective,
- pour tout $\vec{b} \in \mathbb{R}^n$ le système $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$ admet une solution unique

sont équivalentes. Déterminer si une matrice est inversible en utilisant la définition est assez laborieux. Nous cherchons une méthode plus simple. Examinons d'abord le cas $n = 2$:

la matrice

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

est inversible

$\Leftrightarrow$ pour tout $(b_1, b_2) \in \mathbb{R}^2$ le système

$$\mathcal{S}\colon \begin{cases} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{cases}$$

admet une solution unique

$\Leftrightarrow$ les droites d'équation

$$\alpha\colon a_{11}x_1 + a_{12}x_2 = b_1 \ \text{ et } \ \beta\colon a_{21}x_1 + a_{22}x_2 = b_2$$

ont un seul point d'intersection

$\Leftrightarrow$ les droites α et β ne sont pas parallèles

$\Leftrightarrow$ la pente de α est différente de la pente de β: $-\frac{a_{11}}{a_{12}} \neq -\frac{a_{21}}{a_{22}}$

$\Leftrightarrow$ $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

En conclusion, si on pose

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

nous avons que A est inversible ssi $\det A \neq 0$.
Dans le restant de ce chapitre nous cherchons à généraliser cela au cas d'une matrice carrée à n lignes et n colonnes. Plus précisement, nous cherchons à définir le déterminant $\det A$ d'une matrice carrée $A \in \mathbb{R}^{n\times n}$ de façon que A soit inversible ssi $\det A \neq 0$.

11.3 Définition axiomatique pour $n \geq 2$

11.7 Définition. Soit $n \geq 2$. Le *déterminant* est une fonction

$$\det\colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

1. multilinéaire, c'est-à-dire linéaire par rapport à chaque ligne,
2. alternée, c'est-à-dire que si dans A il y a deux lignes égales, alors $\det A = 0$,
3. et telle que $\det I = 1$ (où I est la matrice identité).

11.8 Propriété. *Pour $n = 2$ si on pose*

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

les trois conditions de la Définition 11.7 sont vérifiées.

Preuve. Explicitons les trois conditions ; la preuve qu'elles sont vérifiées est un calcul facile laissé au lecteur.

$$\det \begin{pmatrix} \alpha a_{11} + \beta a'_{11} & \alpha a_{12} + \beta a'_{12} \\ a_{21} & a_{22} \end{pmatrix} = \alpha \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \beta \det \begin{pmatrix} a'_{11} & a'_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{21} + \beta a'_{21} & \alpha a_{22} + \beta a'_{22} \end{pmatrix} = \alpha \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \beta \det \begin{pmatrix} a_{11} & a_{12} \\ a'_{21} & a'_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{pmatrix} = 0$$

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

□

Comment varie le déterminant d'une matrice si l'on transforme la matrice par des opérations élémentaires sur les lignes ? La réponse est donnée par la propriété suivante :

11.9 Propriété.

1. *Si A' est obtenue de A en permutant deux lignes, alors* $\det A' = -\det A$.
2. *Si A' est obtenue de A en multipliant une ligne par un coefficient λ, alors* $\det A' = \lambda \det A$.
3. *Si A' est obtenue de A en ajoutant à une ligne un multiple d'une autre ligne, alors* $\det A' = \det A$. *(De même, si on ajoute à une ligne une combinaison linéaire des autres lignes, le déterminant de la matrice ne change pas.)*
4. *Si dans A il y a une ligne nulle, alors* $\det A = 0$.

Preuve.

1. En écrivant L_i pour la i-ème ligne de la matrice, nous avons

$$0 = \det \begin{pmatrix} \vdots \\ L_i + L_j \\ \vdots \\ L_i + L_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} =$$

$$= 0 + \det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \end{pmatrix} + 0$$

2. Par linéarité par rapport à la ligne multipliée par λ.
3. $$\det \begin{pmatrix} \vdots \\ L_i + \alpha L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + \alpha \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + 0$$
4. Par linéarité par rapport à la ligne nulle.

□

11.4 Unicité du déterminant

Pour montrer l'unicité de la fonction déterminant, nous allons utiliser la forme réduite de Gauss-Jordan : soit $A \in \mathbb{R}^{m\times n}$ et soit B une matrice échelonnée obtenue de A par une suite d'opérations élémentaires sur les lignes (voir 4.7) ; soit $\vec{b}_{i*}$ la i-ème ligne de B et, si $\vec{b}_{i*} \neq \vec{0}$, soit b_{ij} le pivot, c'est-à-dire le premier élément non nul de $\vec{b}_{i*}$ (notons que pour un pivot b_{ij} on a forcement $i < j$). Par opérations élémentaires sur les lignes on peut encore :

- rendre égale à 1 chaque pivot (pour cela, il suffit de diviser chaque ligne non nulle par son pivot), et ensuite
- rendre nul chaque élément qui se trouve au-dessus d'un pivot (si le pivot de la ligne $\vec{b}_{i*}$ est $b_{ij} = 1$, il suffit de remplacer chaque ligne $\vec{b}_{k*}$ avec $k < i$ par $\vec{b}_{k*} - b_{kj} \cdot \vec{b}_{i*}$).

La matrice $U \in \mathbb{R}^{m\times n}$ ainsi obtenue est appelée *matrice en forme réduite de Gauss-Jordan*. (On peut démontrer que la forme réduite d'une matrice A est unique, nous ne le façons pas car nous n'avons pas besoin de ce résultat.)

11.10 Remarque.

1. Puisque U est obtenue à partir de A par une suite d'opérations élémentaires sur les lignes et que les opérations élémentaires ne changent pas le rang d'une matrice, nous avons
$$\operatorname{rang} A = \operatorname{rang} U.$$

2. Si la matrice A de départ est carrée, il n'y a que deux possibilités :
   i. La dernière ligne de la matrice échelonnée B est non nulle ; dans ce cas toutes les lignes de B sont non nulles, les n pivots sont les éléments b_{ii} de la diagonale et on aura $U = I$, la matrice identité.
   ii. La dernière ligne de B est nulle ; dans ce cas la dernière ligne de U aussi sera nulle.
3. De plus, puisque d'un coté $\det I = 1$ et $\operatorname{rang} I = n$ et de l'autre coté le déterminant d'une matrice dont la dernière ligne est nulle vaut 0 et son rang est strictement inférieure à n, on a
   i. $U = I$ ssi $\det U \neq 0$ ssi $\operatorname{rang} U = n$,
   ii. la dernière ligne de U est nulle ssi $\det U = 0$ ssi $\operatorname{rang} U < n$.
4. Toujours dans le cas d'une matrice carrée, puisque l'on passe de A à U par opérations élémentaires sur les lignes, on peut exprimer le déterminant de U en fonction du déterminant de A : si pour obtenir la matrice U on a effectué sur la matrice A
   - r opérations élémentaires de type "échanger deux lignes",
   - s opérations élémentaires de type "multiplier une ligne par un scalaire λ non nul",
   - t opérations élémentaires de type "ajouter à une ligne un multiple d'une autre ligne"

   la Propriété 11.9 nous donne

$$\det U = (-1)^r \, \lambda_1 \cdot \ldots \cdot \lambda_s \det A \qquad [11.10]$$

11.11 Propriété. *(Unicité du déterminant) Soit*

$$\det \colon \mathbb{R}^{n\times n} \to \mathbb{R} \quad \textit{et} \quad \operatorname{Det} \colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

deux fonctions qui satisfont aux trois conditions de la Définition 11.7. Alors pour toute $A \in \mathbb{R}^{n\times n}$ on a

$$\det A = \operatorname{Det} A$$

Preuve. Soit $A \in \mathbb{R}^{n\times n}$ et soit U sa forme réduite. Puisque la formule [11.10] a été établie en utilisant seulement les conditions de la Définition 11.7, elle est valable tant pour la fonction det que pour la fonction Det . Par conséquent :

$$\det U = (-1)^r \, \lambda_1 \cdot \ldots \cdot \lambda_s \det A \quad \text{et} \quad \operatorname{Det} U = (-1)^r \, \lambda_1 \cdot \ldots \cdot \lambda_s \operatorname{Det} A$$

avec tous les coefficients λ_i non nuls. Examinons les deux cas possibles :

i. Si $U = I$, alors $\det U = 1 = \operatorname{Det} U$ et donc $\det A = (-1)^r \, \lambda_1^{-1} \cdot \ldots \cdot \lambda_s^{-1} = \operatorname{Det} A$.
ii. Si la dernière ligne de U est nulle, alors $\det U = 0 = \operatorname{Det} U$ et donc $\det A = 0 = \operatorname{Det} A$.

□

11.12 Propriété. *Une matrice* $A \in \mathbb{R}^{n\times n}$ *est inversible ssi* $\det A \neq 0$.

Preuve. En utilisant la formule [11.10] et le fait qu'une matrice $n \times n$ est inversible ssi son rang vaut n (voir 11.3), nous avons

$$\begin{aligned} \det A \neq 0 &\Leftrightarrow \det U \neq 0 \\ &\Leftrightarrow U = I \\ &\Leftrightarrow \operatorname{rang} U = n \\ &\Leftrightarrow \operatorname{rang} A = n \\ &\Leftrightarrow A \text{ est inversible.} \end{aligned}$$

□

11.13 Corollaire. *Soit* $A \in \mathbb{R}^{n\times n}$.

$$\begin{aligned} \det A \neq 0 &\Leftrightarrow \textit{les lignes de } A \textit{ sont linéairement indépendantes} \\ &\Leftrightarrow \textit{les colonnes de } A \textit{ sont linéairement indépendantes} \end{aligned}$$

Preuve. On sait que $\det A \neq 0 \Leftrightarrow A$ est inversible $\Leftrightarrow \operatorname{rang} A = n$, et le rang de A est par définition la dimension de l'espace $\operatorname{Lign}(A)$ et aussi la dimension de l'espace $\operatorname{Col}(A)$. □

11.5 Construction inductive du déterminant

Dans la preuve de la prochaine propriété nous présentons une formule (dite formule de Laplace) qui permet de calculer explicitement le déterminant d'une matrice $n \times n$.

11.14 Propriété. *Soit* $n \geq 2$. *Il existe une (unique) fonction*

$$\det\colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

qui satisfait aux conditions de la Définition 11.7.

Preuve. Par induction sur n:

$n = 2$: On pose

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

voir Propriété 11.8.

$n > 2$: Soit $A \in \mathbb{R}^{n\times n}$ et fixons un $j \in \{1, \dots, n\}$. Soit $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ la matrice obtenue de A en retirant la i-ème ligne et la j-ème colonne. Posons :

$$\det_j A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \qquad [11.14]$$

et vérifions que la fonction $\det_j$ ainsi définie satisfait aux conditions de la Définition 11.7. (La vérification des conditions 1 et 2 étant assez technique, le lecteur est invité à comparer chaque étape avec le calcul explicite effectué dans l'Exemple 11.15.)

1. Regardons le terme
$$a_{ij} \det A_{ij}$$
pour voir comment il varie en fonction de la k-ième ligne de A:
   - Si $k = i$, alors a_{ki} dépend de façon linéaire de la k-ième ligne de A et A_{ik} ne dépend pas de la k-ième ligne.
   - Si $k \neq i$, alors a_{ij} ne dépend pas de la k-ième ligne de A et $\det A_{ij}$ dépend de façon linéaire de la k-ième ligne (par hypothèse inductive).

   Pour tout $i = 1, \ldots, n$, le terme $a_{ij} \det A_{ij}$ dépend donc de façon linéaire de la k-ème ligne et, par conséquence, la somme $\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ dépend elle aussi de façon linéaire de la k-ième ligne de A.
2. Supposons que dans A la ligne de place i_1 soit égale à la ligne de place i_2, avec $i_1 \neq i_2$. Si $i \neq i_1$ et $i \neq i_2$, dans A_{ij} la i_1-ème ligne est égale à la i_2-ème ligne et donc, par hypothèse inductive, $\det A_{ij} = 0$. Par conséquent, dans l'expression de $\det_j A$ il ne reste que
$$\det_j A = (-1)^{i_1+j} a_{i_1 j} \det A_{i_1 j} + (-1)^{i_2+j} a_{i_2 j} \det A_{i_2 j}$$
avec $a_{i_1 j} = a_{i_2 j}$.
   i. Si i_1 et i_2 sont consécutifs, alors les deux termes sont de signe opposé et $A_{i_1 j} = A_{i_2 j}$. Par conséquent, $\det_j A = 0$.
   ii. Si on peut ramener les deux lignes égales à deux lignes consécutives par un nombre impair de permutations de lignes consécutives, les deux termes sont de même signe et on peut ramener $A_{i_2 j}$ à $A_{i_1 j}$ par un nombre impair de permutations de ligne consécutives. Par hypothèse inductive on a alors que $\det A_{i_1 j} = -\det A_{i_2 j}$ et à nouveau $\det_j A = 0$.
   iii. Si on peut ramener les deux lignes égales à deux lignes consécutives par un nombre pair de permutations de lignes consécutives, alors les deux termes sont de signe opposé et on peut ramener $A_{i_2 j}$ à $A_{i_1 j}$ par un nombre pair de permutations de ligne consécutives. Par hypothèse inductive on a alors que $\det A_{i_1 j} = \det A_{i_2 j}$ et à nouveau $\det_j A = 0$.
3. Si $A = I_n$, alors $a_{ij} = 0$ si $i \neq j$ et $a_{jj} = 1$, et $A_{jj} = I_{n-1}$. Par conséquent, $\det_j I_n = \det_j I_{n-1} = 1$ par hypotèse inductive.

Par l'unicité de la fonction déterminant (11.11), on a $\det_1 A = \ldots = \det_n A$ et nous écrivons simplement $\det A$. □

11.15 Exemple. Considérons une matrice à trois lignes et trois colonnes
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$\det_1 A = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}$$

Vérifions explicitement la linéarité de $\det_1$ par rapport à la première ligne (en écrivant $|A|$ à la place de $\det_1 A$) :

$$\begin{vmatrix} \alpha x_1 + \beta y_1 & \alpha x_2 + \beta y_2 & \alpha x_3 + \beta y_3 \\ a & b & c \\ d & e & f \end{vmatrix} =$$

$$(\alpha x_1 + \beta y_1) \begin{vmatrix} b & c \\ e & f \end{vmatrix} - a \begin{vmatrix} \alpha x_2 + \beta y_2 & \alpha x_3 + \beta y_3 \\ e & f \end{vmatrix} + d \begin{vmatrix} \alpha x_2 + \beta y_2 & \alpha x_3 + \beta y_3 \\ b & c \end{vmatrix}$$

et l'on voit ici que

i. $(\alpha x_1 + \beta y_1)$ dépend de façon linéaire de la première ligne et $\begin{vmatrix} b & c \\ e & f \end{vmatrix}$ ne dépand pas de la première ligne ;

ii. a ne dépend pas de la première ligne et $\begin{vmatrix} \alpha x_2 + \beta y_2 & \alpha x_3 + \beta y_3 \\ e & f \end{vmatrix}$ dépend de façon linéaire de la première ligne (par hypothèse inductive) ;

iii. d ne dépend pas de la première ligne et $\begin{vmatrix} \alpha x_2 + \beta y_2 & \alpha x_3 + \beta y_3 \\ b & c \end{vmatrix}$ dépend de façon linéaire de la première ligne (par hypothèse inductive).

Chacun des trois termes de la somme dépend donc de façon linéaire de la première ligne et, par conséquent, il en est ainsi de $\det_1$.

Vérifion maintenant le caractère alterné de $\det_1$. Pour examiner tous les cas significatifs, travaillons sur une matrice à quatre lignes et quatre colonnes :

i. Les deux lignes égales sont consécutives :

$$\begin{vmatrix} a & b & c & d \\ a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{vmatrix} =$$

$$= a \begin{vmatrix} b & c & d \\ f & g & h \\ j & k & l \end{vmatrix} - a \begin{vmatrix} b & c & d \\ f & g & h \\ j & k & l \end{vmatrix} + e \begin{vmatrix} b & c & d \\ b & c & d \\ j & k & l \end{vmatrix} - i \begin{vmatrix} b & c & d \\ b & c & d \\ f & g & h \end{vmatrix}$$

les deux premiers termes se simplifient entre eux, les troisième et le quatrième s'annullent (par induction) car dans les deux matrices à trois lignes et trois colonnes il y a deux lignes égales.

ii. On peut ramener les deux lignes égales à deux lignes consécutives par un nombre impair de permutations de lignes consécutives :

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ a & b & c & d \\ i & j & k & l \end{vmatrix} =$$

$$= a \begin{vmatrix} f & g & h \\ b & c & d \\ j & k & l \end{vmatrix} - e \begin{vmatrix} b & c & d \\ b & c & d \\ j & k & l \end{vmatrix} + a \begin{vmatrix} b & c & d \\ f & g & h \\ j & k & l \end{vmatrix} - i \begin{vmatrix} b & c & d \\ f & g & h \\ b & c & d \end{vmatrix}$$

(à nouveau par induction) le premier et le troisième terme se simplifient entre eux car le déterminant change de signe, le deuxième et le quatrième terme s'annullent car il y a deux lignes égales.

iii. On peut ramener les deux lignes égales à deux lignes consécutives par un nombre pair de permutations de lignes consécutives :

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ a & b & c & d \end{vmatrix} =$$

$$= a \begin{vmatrix} f & g & h \\ j & k & l \\ b & c & d \end{vmatrix} - e \begin{vmatrix} b & c & d \\ j & k & l \\ b & c & d \end{vmatrix} + i \begin{vmatrix} b & c & d \\ f & g & h \\ b & c & d \end{vmatrix} - a \begin{vmatrix} b & c & d \\ f & g & h \\ j & k & l \end{vmatrix}$$

(toujours par induction) le premier et le quatrième terme se simplifient entre eux car le déterminant change de signe deux fois, le deuxième et le troisième terme s'annullent car il y a deux lignes égales.

11.16 Remarque.

1. Une matrice $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ est *triangulaire supérieure* si tous les éléments a_{ij} avec $i > j$ sont nuls.
2. Si A est triangulaire supérieure, le déterminant de A est simplement le produit des éléments de la diagonale descendante :

   $$\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

   Ceci est facile à montrer par induction en utilisant la formule [11.14] avec $j = 1$.
3. Le point 2. ci-dessus et la Propriété 11.9 donnent une méthode pour calculer le déterminant qui est souvent plus rapide que la formule [11.14].

11.6 Déterminant et produit matriciel

La prochaine propriété permet parfois de simplifier le calcul du déterminant :

11.17 Propriété. *Soit $A, B \in \mathbb{R}^{n\times n}$. On a :*

$$\det(A \cdot B) = \det A \cdot \det B$$

Preuve. Premier cas : si $\det B = 0$, alors B n'est pas inversible (11.12) et donc $A \cdot B$ n'est pas inversible non plus (si il existe C telle que $C \cdot (A \cdot B) = I$, alors par associativité du produit matriciel $(C \cdot A) \cdot B = I$, ce qui montrerait que

B est inversible). Par conséquent, $\det(A \cdot B) = 0$ (11.12) et donc $\det(A \cdot B) = \det A \cdot \det B$.
Deuxième cas : si $\det B \neq 0$, posons

$$\delta \colon \mathbb{R}^{n \times n} \to \mathbb{R} \qquad \delta(A) = \frac{\det(A \cdot B)}{\det B}$$

et montrons que la fonction δ satisfait aux conditions de la Définition 11.7 (d'où, par l'unicité de la fonction déterminant, $\delta(A) = \det A$ et donc $\det A \cdot \det B = \det(A \cdot B)$).

1. Puisque $(-) \cdot B$ est linéaire, la formule

   $$i\text{-ème ligne de } A \cdot B = (i\text{-ème ligne de } A) \cdot B$$

   (voir Remarque 5.10) montre que $\det(A \cdot B)$ dépend de façon linéaire des lignes de A, et donc $\delta(A)$ aussi en dépend de façon linéaire. Montrons de façon plus détaillée que δ dépend de façon linéaire de la première ligne : considérons la matrice suivante (où L_i est la i-ème ligne)

   $$A = \begin{pmatrix} \alpha L_1 + \beta L_1' \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$$

   En utilisant le fait que la fonction det dépend de façon linéaire de la première ligne, nous avons :

$$\delta \begin{pmatrix} \alpha L_1 + \beta L_1' \\ L_2 \\ \vdots \\ L_n \end{pmatrix} = \frac{\det \left(\begin{pmatrix} \alpha L_1 + \beta L_1' \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \cdot B \right)}{\det B} = \frac{\det \begin{pmatrix} (\alpha L_1 + \beta L_1') \cdot B \\ L_2 \cdot B \\ \vdots \\ L_n \cdot B \end{pmatrix}}{\det B} =$$

$$= \frac{\det \begin{pmatrix} \alpha L_1 \cdot B + \beta L_1' \cdot B \\ L_2 \cdot B \\ \vdots \\ L_n \cdot B \end{pmatrix}}{\det B} = \frac{\alpha \det \begin{pmatrix} L_1 \cdot B \\ L_2 \cdot B \\ \vdots \\ L_n \cdot B \end{pmatrix} + \beta \det \begin{pmatrix} L_1' \cdot B \\ L_2 \cdot B \\ \vdots \\ L_n \cdot B \end{pmatrix}}{\det B} =$$

$$\frac{\alpha \det \left(\begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \cdot B \right) + \beta \det \left(\begin{pmatrix} L_1' \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \cdot B \right)}{\det B} =$$

$$= \alpha \frac{\det\left(\begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \cdot B\right)}{\det B} + \beta \frac{\det\left(\begin{pmatrix} L'_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \cdot B\right)}{\det B} = \alpha\delta \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} + \beta\delta \begin{pmatrix} L'_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$$

2. Si dans A il y a deux lignes égales, la formule ci-dessus tirée de la Remarque 5.10 montre que dans $A \cdot B$ aussi il y a deux lignes égales. Par conséquent, $\det(A \cdot B) = 0$ et donc $\delta(A) = 0$.
3. $\delta(I) = \frac{\det(I \cdot B)}{\det B} = \frac{\det B}{\det B} = 1$.

□

11.18 Exercice. Soit $A \in \mathbb{R}^{n \times n}$ une matrice inversible. Montrer que

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

11.19 Remarque. Attention, en ce qui concerne la somme de deux matrices on a pas une propriété semblable à la propriété précédente : en générale

$$\det(A + B) \neq \det A + \det B$$

Comme exemple on peut considérer les matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

En effet, $\det A = 0 = \det B$ mais $\det(A + B) = 1$.

11.20 Remarque. La Propriété 11.17 permet de définir le déterminant d'une application linéaire $L\colon V \to V$, où V est un espace vectoriel de type fini. Si e est une base de V et ${}_eL_e$ est la matrice associée à L par rapport à la base e, on pose

$$\det L = \det({}_eL_e)$$

Cette définition ne dépend pas de la base de V que nous avons chosie : si f est une deuxième base de V, alors

$${}_fL_f = {}_fI_e \cdot {}_eL_e \cdot {}_eI_f$$

où ${}_fI_e$ et ${}_eI_f$ sont les matrices changement de base et ${}_fI_e = ({}_eI_f)^{-1}$ (voir 8.27 et 8.28). En utilisant 11.17 et 11.18, nous avons

$$\det({}_fL_f) = \frac{1}{\det({}_eI_f)} \cdot \det({}_eL_e) \cdot \det({}_eI_f) = \det({}_eL_e)$$

Dans la théorie du déterminant comme nous l'avons présentée jusqu'à ici (11.7 – 11.19) les lignes et les colonnes d'une matrice ont un rôle fort différent. Nous allons remédier à cette asymétrie en utilisant la transposée d'une matrice (voir 9.26).

11.7 Déterminant et matrice transposée

Pour échelonner une matrice nous avons toujours travaillé sur les lignes, il faudrait donc parler de matrice à lignes échelonnées. On peut travailler sur les colonnes et obtenir ainsi une matrice à colonnes échelonnées : à ce propos il n'y a rien de nouveau à apprendre car échelonner une matrice par rapport à ses colonnes revient exactement à échelonner la matrice transposée par rapport aux lignes. Nous pouvons donc répéter la théorie du déterminant en échangeant le rôle des lignes avec celui des colonnes. Voici en résumé :

11.21 Propriété.

A. *Il existe une unique fonction*

$$\delta\colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

1. *multilinéaire, c'est-à-dire linéaire par rapport à chaque colonne,*
2. *alternée, c'est-à-dire que si dans A il y a deux colonnes égales, alors $\delta(A) = 0$,*
3. *et telle que $\delta(I) = 1$.*

B. *Pour $n = 2$ on pose*

$$\delta\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

C. *Pour $n > 2$ et $i \in \{1, \ldots, n\}$ fixé, posons*

$$\delta_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\delta(A_{ij})$$

Par unicité, on a $\delta_1(A) = \ldots = \delta_n(A)$ et nous écrivons simplement $\delta(A)$.

11.22 Lemme. *Pour toute matrice $A \in \mathbb{R}^{n\times n}$ on a $\delta(A) = \det A^t$.*

Preuve. Pour $n = 2$ c'est évident. Pour $n > 2$ remarquons d'abord que

$$(A^t)_{ij} = (A_{ji})^t$$

Si on applique maintenant la formule [11.14] à la matrice A^t avec, pour fixer les idées, $j = 1$, nous avons

$$\begin{aligned} \det\nolimits_1 A^t &= \textstyle\sum_{i=1}^{n} (-1)^{i+1} a_{1i} \det(A^t)_{i1} \\ &= \textstyle\sum_{i=1}^{n} (-1)^{i+1} a_{1i} \det(A_{1i})^t \\ &= \textstyle\sum_{i=1}^{n} (-1)^{i+1} a_{1i} \delta(A_{1i}) \\ &= \delta_1(A) \end{aligned}$$

□

11.23 Corollaire. *Pour toute matrice* $A \in \mathbb{R}^{n\times n}$ *on a* $\det A = \det A^t$.

Preuve. Puisque on sait déjà que $\delta(A) = \det A^t$, il suffit de montrer que $\delta(A) = \det A$, et pour cela il suffit de montrer que la fonction

$$\delta \colon \mathbb{R}^{n\times n} \to \mathbb{R}$$

satisfait aux conditions de la Définition 11.7. Si $n = 2$ cela est évident car $\delta(A) = \det A$ par définition. Si $n > 2$, voyons par exemple que δ est une fonction alternée par rapport aux lignes : supposons que les deux lignes égales dans A soient la ligne i_1 et la ligne i_2 et calculons $\delta_i(A)$ avec $i \neq i_1, i_2$: puisque pour tout $j = 1, \ldots, n$ la ligne i_1 et la ligne i_2 de la matrice A_{ij} sont égales, par hypothèse inductive nous avons

$$\delta_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \delta(A_{ij}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} 0 = 0$$

□

11.8 Applications : calcul du rang et règle de Cramer

Pour terminer ce chapitre, voyons deux applications. Commençons par voir comment le déterminant peut aider pour calculer le rang d'une matrice non nécessairement carrée.

11.24 Définition. Une matrice $B \in \mathbb{R}^{k\times l}$ est une *sous-matrice* d'une matrice $A \in \mathbb{R}^{m\times n}$ (avec $k \leq m$ et $l \leq n$) si B est obtenue à partir de A en retirant m-k lignes et n-l colonnes.

11.25 Exemple. La matrice

$$B = \begin{pmatrix} a & c \\ i & k \end{pmatrix}$$

est une sous-matrice de la matrice

$$A = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{pmatrix}$$

car elle est obtenue de A en retirant la deuxième et la quatrième colonne et la deuxième ligne.

11.26 Propriété. *Soit* $A \in \mathbb{R}^{m\times n}$ *et soit* $k \leq \min(m, n)$. *Alors* $\operatorname{rang} A \geq k$ *si et seulement si il existe une sous-matrice* B *de* A *avec* $B \in \mathbb{R}^{k\times k}$ *et* B *inversible.*

Preuve. Soit $B \in \mathbb{R}^{k\times k}$ une sous-matrice inversible de A et soient $L_{i_1}, \ldots, L_{i_k}$ et $C_{j_1}, \ldots, C_{j_k}$ les lignes et les colonnes de A qui interviennent dans B. Si on pose

$$\pi\colon \mathbb{R}^m \to \mathbb{R}^k \quad \pi(x_1, \ldots, x_m) = (x_{j_1}, \ldots, x_{j_k})$$

les lignes de B sont exactement $\pi(L_{i_1}), \ldots, \pi(L_{i_k})$. Puisque B est inversible, les lignes $\pi(L_{i_1}), \ldots, \pi(L_{i_k})$ sont linéairement indépendantes et donc, par l'Exercice 7.2, les lignes $L_{i_1}, \ldots, L_{i_k}$ aussi sont linéairement indépendantes. Ceci montre que $\operatorname{rang} A = \dim \operatorname{Lign}(A) \geq k$.
Vice-versa, supposons $\operatorname{rang} A \geq k$ et soient $L_{i_1}, \ldots, L_{i_k}$ des colonnes linéairement indépendantes de A. La matrice $C \in \mathbb{R}^{k\times n}$ formée par ces lignes a rang égal à k et elle contient donc (au moins) k colonnes $C_{j_1}, \ldots, C_{j_k}$ linéairement indépendantes. La sous-matrice $B \in \mathbb{R}^{k\times k}$ formée par ces colonnes a rang k et elle est donc inversible. $\square$

Comme deuxième application, voyons la règle de Cramer. Si $A \in \mathbb{R}^{n\times n}$ est une matrice inversible, le système

$$\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$$

admet une solution unique, donnée par $\vec{x} = A^{-1} \cdot \vec{b}$ (pour voir cela, il suffit de multiplier les deux termes de l'équation $A \cdot \vec{x} = \vec{b}$ à gauche par A^{-1}). Une formule plus explicite pour calculer la solution $\vec{x}$ est donnée par la pripriété suivante.

11.27 Propriété. *Considérons le système $\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$, avec $A \in \mathbb{R}^{n\times n}$ une matrice inversible. Les composantes $x_1, \ldots, x_n$ de l'unique solution $\vec{x}$ de $\mathcal{S}$ sont données par*

$$x_j = \frac{\det(C_1 \mid \ldots \mid C_{j-1} \mid \vec{b} \mid C_{j+1} \mid \ldots \mid C_n)}{\det A} \quad \textit{pour tout } \; j = 1, \ldots, n$$

où $C_1, \ldots, C_n$ sont les colonnes de A.

Preuve. Le fait que $\vec{x}$ soit solution de $\mathcal{S}$ revient à dire que

$$\vec{b} = x_1 \cdot C_1 + \ldots + x_n \cdot C_n$$

(voir Remarque 5.11). Par la linéarité du déterminant par rapport à la j-ème colonne et en utilisant le caractère alterné du déterminant par rapport aux colonnes, nous obtenons

$$\begin{gathered}
\det(C_1 \mid \ldots \mid C_{j-1} \mid \vec{b} \mid C_{j+1} \mid \ldots \mid C_n) = \\
= \det(C_1 \mid \ldots \mid C_{j-1} \mid x_1 \cdot C_1 + \ldots + x_n \cdot C_n \mid C_{j+1} \mid \ldots \mid C_n) = \\
= x_1 \cdot \det(C_1 \mid \ldots \mid C_{j-1} \mid C_1 \mid C_{j+1} \mid \ldots \mid C_n) + \ldots \\
\ldots + x_j \cdot \det(C_1 \mid \ldots \mid C_{j-1} \mid C_j \mid C_{j+1} \mid \ldots \mid C_n) + \ldots \\
\ldots + x_n \cdot \det(C_1 \mid \ldots \mid C_{j-1} \mid C_n \mid C_{j+1} \mid \ldots \mid C_n) = \\
= x_j \cdot \det(C_1 \mid \ldots \mid C_{j-1} \mid C_j \mid C_{j+1} \mid \ldots \mid C_n) = x_j \cdot \det A
\end{gathered}$$

Si A est inversible, son déterminant est différent de zéro et, en divisant par $\det A$, on obtient la thèse. $\square$

Chapitre 12

Espaces vectoriels sur un corps

Le chapitre 12 section par section

Jusqu'à ici nous avons travaillé avec des matrices à coefficients réels, avec des systèmes à coefficients réels, et avec des espaces vectoriels réels. Peut-on remplacer $\mathbb{R}$ par $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{C}, \ldots$? Ce qu'il nous faut est de choisir les coefficients dans un ensemble de “nombres” qui satisfait à certaines propriétés. Un tel ensemble est appelé un *corps commutatif.*

1. La notion de corps commutatif est introduite par étapes : en ordre de complexité croissante, nous avons les groupes, les anneaux, les corps. Parmi les exemples, on a le corps $\mathbb{R}$ des réels et le corps $\mathbb{C}$ des complexes.
2. La théorie des espaces vectoriels et des applications linéaires, développée dans les chapitres précédents pour le cas réel, se généralise sans difficulté aux cas des espaces vectoriels sur un corps commutatif quelconque.

Les notions de groupe et de anneau ne seront pas développées dans ce cours, mais elles feront l'objet d'une étude approfondie en deuxième et troisième année tant dans le programme de physique que dans le programme de mathématique.

12.1 Groupes, anneaux, corps

12.1 Définition. Un *groupe* est donné par :

1. un ensemble G
2. un élément $0 \in G$ appelé zéro
3. une opération binaire appelée somme

$$G \times G \to G \ , \ \ (x, y) \mapsto x + y$$

4. une opération unaire appelée inverse

$$G \to G\ ,\quad x \mapsto -x$$

Ces données doivent satisfaire aux équations suivantes (pour tout $x, y, z \in G$):

1. $(x + y) + z = x + (y + z)$ (associtivité de la somme)
2. $x + 0 = x = 0 + x$ (0 est l'élément neutre par rapport à la somme)
3. $x + (-x) = 0 = (-x) + x$ (l'élément $-x$ est l'élément inverse de x par rapport à la somme.

Le groupe G est dit *commutatif* (ou *abélien*) si de plus

4. $x + y = y + x$ (commutativité de la somme).

12.2 Exemple.

1. $(\mathbb{Z}, 0, +, -), (\mathbb{Q}, 0, +, -), (\mathbb{R}, 0, +, -), (\mathbb{Q}\backslash\{0\}, 1, \cdot, (\text{-})^{-1}), (\mathbb{R}\backslash\{0\}, 1, \cdot, (\text{-})^{-1})$ sont des groupes abéliens.
2. $(\mathbb{N}, 0, +), (\mathbb{Z}, 1, \cdot), (\mathbb{Q}, 1, \cdot, (\text{-})^{-1}), (\mathbb{R}, 1, \cdot, (\text{-})^{-1})$ ne sont pas de groupes.

12.3 Définition. Un *anneau* est donné par :

1. un ensemble A
2. un élément $0 \in A$ appelé zéro
3. un élément $1 \in A$ appelé unité
4. une opération binaire appelée somme

$$A \times A \to A\ ,\quad (x, y) \mapsto x + y$$

5. une opération unaire appelée inverse

$$A \to A\ ,\quad x \mapsto -x$$

6. une opération binaire appelée produit

$$A \times A \to A\ ,\quad (x, y) \mapsto x \cdot y$$

Ces données doivent satisfaire aux conditions et équations suivantes (pour tout $x, y, z \in A$):

1. $(A, 0, +, -)$ est un groupe abélien
2. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ (associativité du produit)
3. $x \cdot 1 = x = 1 \cdot x$ (1 est l'élément neutre par rapport au produit)
4. $x \cdot (y + z) = x \cdot y + x \cdot z$ (distributivité à gauche du produit par rapport à la somme)
5. $(x + y) \cdot z = x \cdot z + y \cdot z$ (distributivité à droite du produit par rapport à la somme).

L'anneau A est dit *commutatif* si de plus

6. $x \cdot y = y \cdot x$ (commutativité du produit).

12.4 Exemple. $(\mathbb{Z}, 0, 1, +, -, \cdot), (\mathbb{Q}, 0, 1, +, -, \cdot), (\mathbb{R}, 0, 1, +, -, \cdot)$ sont des anneaux commutatifs.

12.5 Exemple. Si $(A, 0, 1, +, -, \cdot)$ est un anneau, nous pouvons considérer l'ensemble des éléments inversibles

$$A^* = \{a \in A \mid \text{ il existe } a^* \in A \text{ tel que } a \cdot a^* = 1 = a^* \cdot a\}$$

et nous avons que $(A^*, 1, \cdot, (-)^*)$ est un groupe. Par contre, on ne peut pas restreindre la somme $+ \colon A \times A \to A$ à une somme $+ \colon A^* \times A^* \to A^*$ car la somme de deux éléments inversibles pourrait être un élément non inversible.

12.6 Définition. Un *corps* est un anneau

$$(K, 0, 1, +, -, \cdot)$$

qui satisfait en plus la condition

1. pour tout $x \in K, x \neq 0$, il existe $x^* \in K$ tel que $x \cdot x^* = 1 = x^* \cdot x$ (l'élément x^* est dit élément inverse de x par rapport au produit).

Le corps K est dit *commutatif* si de plus

2. $x \cdot y = y \cdot x$ (commutativité du produit).

12.7 Exemple.

1. $(\mathbb{Q}, 0, 1, +, -, \cdot), (\mathbb{R}, 0, 1, +, -, \cdot)$ sont des corps commutatifs, avec élément inverse $x^* = x^{-1}$.
2. $(\mathbb{Z}, 0, 1, +, -, \cdot)$ n'est pas un corps.

12.8 Exemple. L'ensemble $\mathbb{R}^2$ avec les éléments

$$0 = (0, 0) \ , \quad 1 = (1, 0)$$

et les opérations

$$(a, b)+(c, d) = (a+c, b+d) \ , \quad -(a, b) = (-a, -b) \ , \quad (a, b) \cdot (c, d) = (ac-bd, ad+bc)$$

est un corps commutatifs noté $\mathbb{C}$ et appelé *corps des nombres complexes.*
Comme exercice, le lecteur peut calculer l'élément inverse $(a, b)^*$ d'un nombre complexe $(a, b) \neq (0, 0)$.

Si nous écrivons a à la place de $(a, 0)$ et i (l'unité imaginaire) pour $(0, 1)$, nous obtenons

1. $(a, b) = a + ib$ (forme algébrique d'un nombre complexe)
2. $i^2 = -1$.

En utilisant la forme algébrique, on peut développer les calculs avec les nombres complexes comme si on avait à faire à des nombres réels, quitte à utiliser la relation $i^2 = -1$.

La particularité de $\mathbb{C}$ par rapport à $\mathbb{R}$ est donnée par le résultat suivant (l'exemple $P(x) = x^2 + 1$ montre qu'un tel résultat n'est pas valable si on cherche les racines réelles d'un polynôme à coefficients réels).

12.9 Théorème. *(Théorème Fondamental de l'Algèbre) Soit un polynôme de degré $n > 0$ à coefficients complexes*

$$P(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \in \mathbb{C}[x]_n$$

Alors :

1. *$P(x)$ admet n racines complexes (non nécessairement distinctes)*

$$\exists\ \alpha_1, \ldots, \alpha_n \in \mathbb{C} \text{ tels que } P(\alpha_1) = \ldots = P(\alpha_n) = 0$$

2. *et, par conséquent, $P(x)$ se décompose en produit de n facteurs de premier degré*

$$P(x) = a_n (x - \alpha_1) \cdot \ldots \cdot (x - \alpha_n)$$

Preuve. Ce théorème sera admis sans preuve, car elle est trop longue par rapport à son intérêt dans ce cours. Le lecteur peut trouver une preuve complète (rédigée par Francis Borceux) dans le document "Théorème Fondamental" disponible sur le site iCampus du cours. □

12.2 Espaces sur un corps

12.10 Définition. Soit $K = (K, 0, 1, +, -, \cdot)$ un corps commutatif. Un *espace vectoriel sur K* est donné par :

1. un ensemble V dont les éléments sont appelés vecteurs et notés $\vec{x}, \vec{y}, \ldots$
2. un élément $\vec{0} \in V$ appelé vecteur zéro ou vecteur nul
3. une opération binaire appelée somme

$$V \times V \to V\ ,\quad (\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}$$

4. une opération unaire appelée inverse

$$V \to V\ ,\quad \vec{x} \mapsto -\vec{x}$$

5. une opération binaire appelée action

$$K \times V \to V\ ,\quad (\alpha, \vec{x}) \mapsto \alpha \cdot \vec{x}$$

Ces données doivent satisfaire aux conditions et équations suivantes (pour tout $\vec{x}, \vec{y}, \vec{z} \in V$ et pour tout $\alpha, \beta \in K$):

1. $(V, 0, +, -)$ est un groupe abélien
2. $\alpha \cdot (\beta \cdot \vec{x}) = (\alpha\beta) \cdot \vec{x}$
3. $1 \cdot \vec{x} = \vec{x}$
4. $(\alpha + \beta) \cdot \vec{x} = \alpha \cdot \vec{x} + \beta \cdot \vec{x}$
5. $\alpha \cdot (\vec{x} + \vec{y}) = \alpha \cdot x + \alpha \cdot \vec{y}$.

12.11 Remarque. Pour retrouver, dans le cas $K = \mathbb{R}$, la définition d'espace vectoriel réel donnée au Chapitre 6 (Définition 6.1) il suffit de remarquer que dans tous espace vectoriel sur un corps K on a $(-1) \cdot \vec{x} = -\vec{x}$. En effet, en utilisant l'Exercice 6.2 qui reste valable dans tout espace vectoriel sur un corps K arbitraire, nous avons :

$$\vec{x} + (-1) \cdot \vec{x} = 1 \cdot \vec{x} + (-1) \cdot \vec{x} = (1 - 1) \cdot \vec{x} = 0 \cdot \vec{x} = \vec{0}$$

12.12 Définition. Soient V et W deux espaces vectoriels sur un même corps commutatifs K et soit

$$L \colon V \to W$$

une fonction. La fonction L est une *application linéaire* si pour tout $\vec{x}, \vec{y} \in V$ et pour tout $\alpha, \beta \in K$ on a :

1. $L(\alpha \cdot \vec{x}) = \alpha \cdot L(\vec{x})$
2. $L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y})$

ou, de façon équivalente, si

3. $L(\alpha \cdot \vec{x} + \beta \cdot \vec{y}) = \alpha \cdot L(\vec{x}) + \beta \cdot L(\vec{y})$

12.13 Exemple. Soit K un corps commutatif et V, W deux espaces vectoriels sur K.

1. L'ensemble K^n des n-uples d'éléments dans K, l'ensemble $K^{m\times n}$ des matrices à m lignes et n colonnes à coefficients dans K, l'ensemble $K[x]$ des polynômes à coefficients dans K, l'ensemble $K[x]_n$ des polynômes de degré plus petit ou égale à n à coefficients dans K, l'ensemble $\mathcal{F}(X, V)$ des fonctions de X vers V avec X un ensemble quelconque, l'ensemble $\mathrm{Lin}(V, W)$ des applications linéaires de V vers W, sont des espaces vectoriels sur K par rapport aux mêmes opérations utilisées dans le cas réel.
2. Les ensembles $K^{n\times n}$ et $\mathrm{Lin}(V, V)$ sont des anneaux non commutatifs si on considère comme produit le produit matriciel dans $K^{n\times n}$ et la composition d'applications linéaires dans $\mathrm{Lin}(V, V)$.

12.14 Remarque. L'entièreté de la théorie développée dans tous les chapitres précédents pour les systèmes d'équations linéaires à coefficients réels, pour les espaces vectoriels réels et pour les applications linéaires entre espaces vectoriels réels reste valable si on remplace le corps $\mathbb{R}$ des nombres réels par un corps commutatif K quelconque. La seule différence est la Remarque 2.4 :
Trois cas peuvent se présenter pour un système $\mathcal{S}$ d'équations linéaires :

1. *Le système $\mathcal{S}$ n'a pas de solutions. On dit que $\mathcal{S}$ est un système* impossible *(ou sur-déterminé).*
2. *Le système $\mathcal{S}$ admet une solution unique. On dit que $\mathcal{S}$ est un système* déterminé.
3. *Le système $\mathcal{S}$ admet une infinité de solutions. On dit que $\mathcal{S}$ est un système* indéterminé *(ou sous-déterminé).*

Cette remarque reste valable si K est $\mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, mais pas en général. Dans certains cas un système d'équations linéaires à coefficients dans K peut admettre un nombre fini strictement plus grand que 1 de solutions. Voici un exemple simple.

12.15 Exemple. L'ensemble $\{0, 1, 2, \ldots, p-1\}$ est un anneau commutatif par rapport aux opérations $+_p$ et $\cdot_p$ définies de la façon suivante :

$$a +_p b \text{ est le reste de la division de } a+b \text{ par } p$$

$$a \cdot_p b \text{ est le reste de la division de } a \cdot b \text{ par } p$$

Par exemple, si $p = 7$, alors $5 +_p 6 = 4, 1 +_p 2 = 3, 6 \cdot_p 5 = 2$.
Cet anneau est noté $\mathbb{Z}_p$ et appelé l'anneau des entiers modulo p. Si de plus p est un nombre premier, alors $\mathbb{Z}_p$ est un corps commutatif.
Le système d'une équation à une variable

$$\mathcal{S}_0 \colon 0 \cdot_p x = 0$$

interprété comme système à coefficients dans $\mathbb{Z}_p$ admet p solutions (tous les éléments de $\mathbb{Z}_p$).

12.16 Remarque. Un même groupe abélien peut être muni d'une structure d'espace vectoriel sur des corps différents. Par exemple, tout espace vectoriel sur $\mathbb{C}$ est aussi un espace vectoriel sur $\mathbb{R}$ (et sa dimension comme $\mathbb{R}$-espace est le double de sa dimension comme $\mathbb{C}$-espace).

Ceci est un cas particulier d'un phénomène plus général : si H et K sont deux corps commutatifs et

$$f \colon H \to K$$

est un *morphisme de corps*, c'est-à-dire une fonction telle que

$$f(a+b) = f(a) + f(b)\ , \quad f(a \cdot b) = f(a) \cdot f(b)\ , \quad f(1) = 1$$

alors tout groupe abélien $(V, 0, +, -)$ muni d'une structure de K-espace vectoriel

$$K \times V \xrightarrow{\cdot} V$$

peut être muni d'une structure de H-espace vectoriel grâce à l'action

$$H \times V \xrightarrow{f \times id} K \times V \xrightarrow{\cdot} V$$

Dans le cas de $\mathbb{R}$ et $\mathbb{C}$, le morphisme de corps est l'inclusion

$$\mathbb{R} \to \mathbb{C} \quad a \mapsto (a, 0)$$

12.17 Exemple. Evidemment $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel de dimension 1. Mais le groupe abélien $(\mathbb{R}, 0, +, -)$ est aussi un $\mathbb{Q}$-espace vectoriel et, en tant que $\mathbb{Q}$-espace vectoriel, sa dimension est infinie. Pour voir cela, nous utilisons les notions de nombre algébrique et de nombre transcendant : un nombre réel est *algébrique* s'il est racine d'un polynôme à coefficients rationnels non tous nuls (ou, ce qui revient au même, à coefficients entiers non tous nuls). Un nombre réel qui n'est pas algébrique est dit *transcendant.* Tout nombre rationnel est algébrique, mais il y a des nombres irrationnels qui sont algébriques, comme par exemple $\sqrt{2}$ qui est racine du polynôme $x^2 - 2$. On peut montrer que, dans un sens que nous ne préciserons pas ici, il y a beaucoup plus de nombres transcendants que de nombres algébriques. Pour cet exemple nous nous limiterons à accepter, sans preuve, que le nombre e, la base du logarithme népérien, est un nombre transcendant. Nous pouvons maintenant construire dans le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$ des suites libres de longueur arbitraire. En effet, pour tout $n \in \mathbb{N}$, la suite

$$1, e, e^2, e^3, \ldots, e^n$$

est libre car si une combinaison linéaire à coefficients rationnels

$$a_0 + a_1 e + a_2 e^2 + a_3 e^3 + \ldots + a_n e^n$$

est nulle, alors tous les coefficients sont nuls car e est transcendant.

12.18 Remarque. La définition d'espace vectoriel sur un corps K garde tout son sens si K est seulement un anneau commutatif, mais dans ce cas on préfère parler de *module* sur l'anneau K plutôt que d'espace vectoriel. Par exemple, les modules sur l'anneau $\mathbb{Z}$ sont exactement les groupes abéliens.
On peut aussi parler de modules sur un anneau A qui n'est pas commutatif, mais dans ce cas il faut distinguer entre

- module à gauche : l'action est de type $A \times V \to V$
- module à droite : l'action est de type $V \times A \to V$

(dans le cas commutatif il n'y a pas de différence : tout module à gauche est aussi un module à droite et vice-versa).

Attention : la théorie des modules sur un anneau (commutatif ou pas) est fort différente de la théorie des espaces vectoriels sur un corps commutatif. Par exemple, dans le cas d'un anneau, il n'est plus vrai que tout module admet une base (finie ou infinie).

12.19 Remarque. Dans les chapitres consacrés aux espaces euclidiens et aux formes quadratiques, nous nous limitons aux espaces vectoriels réels. La raison principale est qu'il faudra utiliser des inégalités. Or, $\mathbb{R}$ est un *corps ordonné*, c'est-à-dire muni d'une relation d'ordre totale

$$\leq \ \subset \ \mathbb{R} \times \mathbb{R}$$

compatible avec la structure de corps

1. $x \leq x$ (réflexivité)
2. si $x \leq y$ et $y \leq z$ alors $x \leq z$ (transitivité)
3. si $x \leq y$ et $y \leq x$ alors $x = y$ (anti-symétrie)

4. pour tout x, y on a $x \leq y$ ou $y \leq x$ (ordre total)
5. si $x \leq y$ alors $x + z \leq y + z$ (compatibilité avec la somme)
6. si $x \leq y$ et $z \leq 0$, alors $y \cdot z \leq x \cdot z$ (compatibilité avec le produit).

Par contre, $\mathbb{C}$ n'est pas un corps ordonnés car on peut montrer que dans tout corps ordonné on a

$$-1 < 0 \text{ et } 0 < x^2 \text{ (si } x \neq 0)$$

en contradiction avec la situation de $\mathbb{C}$ où $-1 = i^2$.

Chapitre 13

Espaces euclidiens

Le chapitre 13 section par section

A partir du problème de la recherche des solutions approchées d'un système qui n'admet pas de solutions exactes, nous introduisons la notion d'espace euclidien. Grâce aux espaces euclidiens on peut étudier de nombreux problèmes d'approximation dans les espaces vectoriels réels.

1. Un espace euclidien est un espace vectoriel réel muni d'un produit scalaire. $\mathbb{R}^n$ est l'exemple guide, mais il y en a d'autres comme l'espace des polynômes et l'espace des fonctions intégrables.
2. En utilisant le produit scalaire, dans un espace euclidien on peut définir la distance entre deux vecteurs et l'orthogonalité entre vecteurs.
3. La notion la plus importante qu'on peut introduire dans un espace euclidien est celle de projection orthogonale sur un sous-espace : la projection orthogonale minimise la distance.
4. Pour garantir l'existence de la projection orthogonale sur un sous-espace V, il faut disposer d'une base orthonormée de V. C'est toujours le cas si V est finiment engendré.
5. Quelques problèmes d'approximation qu'on peut traiter grâce aux espaces euclidiens : les solutions approchées d'un système, la droite d'interpolation linéaire, la droite tangente au graphe d'une fonction dérivable, le polynôme de Taylor.

13.1 Définition et exemples

Ce chapitre est consacré à la recherche des solutions approchées d'un système qui n'a pas de solutions exactes : nous savons qu'un système

$$\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$$

admet des solutions si et seulement si le vecteur des termes indépendants est combinaison linéaire des colonnes de la matrice des coefficients

$$\vec{b} \in \mathrm{Col}(A)$$

Dans le cas où cette condition n'est pas vérifiée, on peut chercher les solutions approchées : on remplace le vecteur $\vec{b}$ par un vecteur $\vec{b}'$ qui remplit les conditions suivantes

1. $\vec{b}' \in \mathrm{Col}(A)$
2. parmi les vecteurs de $\mathrm{Col}(A)$, $\vec{b}'$ est le plus proche de $\vec{b}$

$$\forall\, \vec{y} \in \mathrm{Col}(A),\ \vec{y} \neq \vec{b}' \colon \quad \mathrm{dist}(\vec{b}, \vec{b}') < \mathrm{dist}(\vec{b}, \vec{y})$$

Le nouveau système $\mathcal{S}' \colon A \cdot \vec{x} = \vec{b}'$ admet des solutions, qui sont appelées *solutions approchées* du système $\mathcal{S}$ de départ.

Pour que tout cela puisse fonctionner, il faut préciser ce que distance

$$\mathrm{dist}(\vec{x}, \vec{y})$$

entre vecteurs de $\mathbb{R}^n$ signifie. Les espace euclidiens sont des espaces vectoriels où l'on peut parler de distance : l'espace $\mathbb{R}^n$ en est un exemple parmi beaucoup d'autres.

13.1 Définition. Un *espace euclidien* est donné par un espace vectoriel réel E et un *produit scalaire*, c'est-à-dire une fonction

$$(- \mid -) \colon E \times E \to \mathbb{R}\,, \quad \vec{x}, \vec{y} \mapsto (\vec{x} \mid \vec{y})$$

1. bilinéaire, c'est-à-dire linéaire par rapport à chaque variable :

$$(\alpha \cdot \vec{x} + \beta \cdot \vec{y} \mid \vec{z}) = \alpha(\vec{x} \mid \vec{z}) + \beta(\vec{y} \mid \vec{z})$$

$$(\vec{x} \mid \alpha \cdot \vec{y} + \beta \cdot \vec{z}) = \alpha(\vec{x} \mid \vec{y}) + \beta(\vec{x} \mid \vec{z})$$

2. symétrique : $(\vec{x} \mid \vec{y}) = (\vec{y} \mid \vec{x})$
3. définie positive : pour tout $\vec{x} \in E, \vec{x} \neq 0$, on a $(\vec{x} \mid \vec{x}) > 0$.

13.2 Remarque. A cause de la linéarité par rapport à chaque variable, on a $(\vec{x} \mid \vec{0}\,) = 0 = (\vec{0} \mid \vec{x})$ pour tout vecteur $\vec{x}$.

13.3 Exemple.

1. L'espace $\mathbb{R}^n$ muni du *produit scalaire canonique*

$$(\vec{x} \mid \vec{y}) = \sum_{i=1}^{n} x_i y_i \quad \text{pour} \quad \vec{x} = (x_1, \ldots, x_n) \quad \text{et} \quad \vec{y} = (y_1, \ldots, y_n)$$

   est un espace euclidien. (Nous avons déjà utilisé le produit scalaire canonique dans la définition du produit matriciel, voir 5.3.)

2. Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique (c'est-à-dire que $A^t = A$). Si on pose
$$(\vec{x} \mid \vec{y}) = (x_1, \ldots, x_n) \cdot A \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$
on a une fonction bilinéaire et symétrique. Ci cette fonction est aussi définie positive, on dira que la matrice A est définie positive. Les notions de matrice définie positive ou négative et semi-définie positive ou négative seront étudiées plus en profondeur au Chapitre 17 (voir Remarque 17.15). Notons que pour $A = I$ on retrouve le produit scalaire canonique de $\mathbb{R}^n$.
3. Soit $a_0, a_1, \ldots, a_n \in \mathbb{R}$ des réels fixés, avec $a_i \neq a_j$ si $i \neq j$. L'espace $\mathbb{R}[x]_n$ muni du produit scalaire
$$(P \mid Q) = \sum_{i=0}^{n} P(a_i)Q(a_i)$$
est un espace euclidien.
4. Soit $a, b \in \mathbb{R}$ deux réels fixés, avec $a < b$. L'espace des fonctions intégrables $Int([a, b], \mathbb{R})$ muni du produit scalaire
$$(f \mid g) = \int_a^b f(x)g(x)dx$$
est un espace euclidien.

13.2 Norme, distance, orthogonalité

13.4 Notation. Soit E un espace euclidien et soit $\vec{x}, \vec{y} \in E$.
1. Norme de $\vec{x}$: $\| \vec{x} \| = \sqrt{(\vec{x} \mid \vec{x})}$.
2. Distance entre $\vec{x}$ et $\vec{y}$: $\mathrm{dist}(\vec{x}, \vec{y}) = \| \vec{x} - \vec{y} \|$.

13.5 Propriété. *Soit E un espace euclidien.*
1. *Pour tout $\vec{x}, \vec{y} \in E$ on a $\| \vec{x} \| \geq 0$ et $\mathrm{dist}(\vec{x}, \vec{y}) \geq 0$.*
2. *Si $\vec{x} \neq 0$, alors $\| \vec{x} \| > 0$; si $\vec{x} \neq \vec{y}$, alors $\mathrm{dist}(\vec{x}, \vec{y}) > 0$.*
3. *Pour tout $\vec{x} \in E$ et pour tout $\alpha \in \mathbb{R}$ on a $\| \alpha \cdot \vec{x} \| = | \alpha | \cdot \| \vec{x} \|$.*
4. *Inégalité de Cauchy : $| (\vec{x} \mid \vec{y}) | \leq \| \vec{x} \| \cdot \| \vec{y} \|$.*
5. *Inégalité triangulaire : $\| \vec{x} + \vec{y} \| \leq \| \vec{x} \| + \| \vec{y} \|$.*

Preuve. 1. 2. 3. : Exercice.
4. Pour tout $\alpha \in \mathbb{R}$, $0 \leq (\vec{x} + \alpha\vec{y} \mid \vec{x} + \alpha\vec{y}) = \alpha^2 \| \vec{y} \|^2 + 2\alpha(\vec{x} \mid \vec{y}) + \| \vec{x} \|^2$. Si on regarde cette expression comme un polynôme de deuxième degré par rapport à la variable α, il faut que son discriminant soit negatif ou nul :
$$\Delta = (\vec{x} \mid \vec{y})^2 - \| \vec{x} \|^2 \cdot \| \vec{y} \|^2 \leq 0$$

d'où $(\vec{x} \mid \vec{y})^2 \leq \| \vec{x} \|^2 \cdot \| \vec{y} \|^2$ et donc $| (\vec{x} \mid \vec{y}) | \leq \| \vec{x} \| \cdot \| \vec{y} \|$.
5. En utilisant le point 4. nous avons

$$\| \vec{x}+\vec{y} \|^2 = (\vec{x}+\vec{y} \mid \vec{x}+\vec{y}) = (\vec{x} \mid \vec{x}) + 2(\vec{x} \mid \vec{y}) + (\vec{y} \mid \vec{y}) \leq$$

$$\leq \| \vec{x} \|^2 + 2 \| \vec{x} \| \cdot \| \vec{y} \| + \| \vec{y} \|^2 = (\| \vec{x} \| + \| \vec{y} \|)^2$$

d'où $\| \vec{x}+\vec{y} \| \leq \| \vec{x} \| + \| \vec{y} \|$. □

13.6 Exemple. Dans $\mathbb{R}^2$ avec le produit scalaire canonique :

1. $\| \vec{x} \| = \sqrt{x_1^2 + x_2^2}$,
2. $\mathrm{dist}(\vec{x}, \vec{y}) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$,
3. $(\vec{x} \mid \vec{y}) = \| \vec{x} \| \cdot \| \vec{y} \| \cdot \cos(\vec{x}, \vec{y})$.

Preuve. 3. Soit $\vec{x} = (x_1, x_2), \vec{y} = (y_1, y_2)$. Il faut montrer que

$$x_1y_1 + x_2y_2 = \sqrt{x_1^2 + x_2^2} \cdot \sqrt{y_1^2 + y_2^2} \cdot \cos(\vec{x}, \vec{y})$$

Puisque la norme d'un vecteur et l'angle entre deux vecteurs ne changent pas par rotation autour de l'origine, pour simplifier les calculs nous pouvons supposer que le vecteur $\vec{x}$ se trouve sur le demi-axe horizontal positif : $x_1 \geq 0$ et $x_2 = 0$. La formule à montrer devient alors

$$x_1y_1 = x_1\sqrt{y_1^2 + y_2^2} \cdot \cos(\vec{x}, \vec{y})$$

Après simplification, cela donne

$$\frac{y_1}{\sqrt{y_1^2 + y_2^2}} = \cos(\vec{x}, \vec{y})$$

qui est vrai par définition de cosinus. □

13.7 Définition. Soit E un espace euclidien.

1. Deux vecteurs $\vec{x}, \vec{y} \in E$ sont *orthogonaux* si $(\vec{x} \mid \vec{y}) = 0$. Notation : $\vec{x} \perp \vec{y}$
2. Soit V un sous-ensemble de E. L'*espace orthogonale* de V est

$$V^\perp = \{\vec{x} \in E \mid \vec{x} \perp \vec{v} \text{ pour tout } \vec{v} \in V\}$$

13.8 Exemple. Dans $\mathbb{R}^2$ muni du produit scalaire canonique, soit $\vec{y} = (a, b)$ et soit $V = \langle \vec{y} \rangle$. Alors

$$V^\perp = \{(x_1, x_2) \in \mathbb{R}^2 \mid ax_1 + bx_2 = 0\}$$

est la droite par l'origine orthogonale à la droite V de direction $\vec{y}$.

13.9 Exercice. Dans un espace euclidien E montrer que :

1. $V^\perp$ est un s.e.v. de E.
2. Si $\vec{x} \in V \cap V^\perp$ alors $\vec{x} = \vec{0}$.

3. $E^\perp = \{\vec{0}\}$, $\{\vec{0}\}^\perp = E$.
4. Si $V_1 \subseteq V_2$ alors $V_1^\perp \supseteq V_2^\perp$.
5. $V \subseteq (V^\perp)^\perp$, $V^\perp = ((V^\perp)^\perp)^\perp$.
6. $V_1 \subseteq V_2^\perp \;\Leftrightarrow\; V_1^\perp \supseteq V_2$.
7. Si V_1 et V_2 sont deux s.e.v., alors $(V_1 + V_2)^\perp = V_1^\perp \cap V_2^\perp$.

13.10 Remarque. L'égalité $V = (V^\perp)^\perp$ n'est pas toujours vérifiée.

1. Si V est un sous-ensemble de E mais pas un s.e.v., on peut trouver un contre-exemple en dimension finie : $E = \mathbb{R}^2$ avec le produit scalaire canonique et V le demi-axe horizontal positif; $(V^\perp)^\perp$ est alors l'axe horizontal.
2. Si V est un s.e.v., il faut chercher un contre-exemple en dimension infinie (car en dimension finie on a bien l'égalité $V = (V^\perp)^\perp$, voir 13.14 et 13.23). Nous pouvons prendre $E = \mathbb{R}[x]$ avec le produit scalaire

$$(P \mid Q) = \int_0^1 P(x) \cdot Q(x)dx$$

et comme V le s.e.v. des polynômes qui s'annulent en zéro. Dans ce cas on peut montrer que $V^\perp = \{\vec{0}\}$ et donc $(V^\perp)^\perp = \mathbb{R}[x]$.

13.3 Projection orthogonale

13.11 Définition. Soit E un espace euclidien et V un s.e.v. de E. La *projection orthogonale* de E sur V est une application

$$P_V \colon E \to E$$

telle que pour tout $\vec{x} \in E$:

1. $P_V(\vec{x}) \in V$,
2. $\vec{x} - P_V(\vec{x}) \in V^\perp$.

13.12 Propriété. *(Unicité de la projection orthogonale) Soit E un espace euclidien et soit V un s.e.v. de E. Si*

$$f \colon E \to E \quad \text{et} \quad g \colon E \to E$$

sont deux fonctions telles que pour tout $\vec{x} \in E$

1. $f(\vec{x}) \in V$ et $g(\vec{x}) \in V$

2. $\vec{x} - f(\vec{x}) \in V^\perp$ et $\vec{x} - g(\vec{x}) \in V^\perp$

alors $f = g$.

Preuve. On a $f(\vec{x}) - g(\vec{x}) = f(\vec{x}) - \vec{x} + \vec{x} - g(\vec{x})$. Or, le terme de gauche est dans V et celui de droite est dans $V^\perp$. Donc $f(\vec{x}) - g(\vec{x}) \in V \cap V^\perp$. Par 13.9 cela donne $f(\vec{x}) - g(\vec{x}) = \vec{0}$, c'est-à-dire $f(\vec{x}) = g(\vec{x})$. □

13.13 Propriété. *Soit E un espace euclidien et soit V un s.e.v. de E. Soit $P_V : E \to E$ la projection orthogonale de E sur V.*

1. *P_V est une application linéaire,*
2. $\operatorname{Ker} P_V = V^\perp$,
3. $\operatorname{Im} P_V = V = \{\vec{x} \in E \mid P_V(\vec{x}) = \vec{x}\}$,
4. $P_V \circ P_V = P_V$.

Preuve. A titre d'exercice montrons que $P_V(\vec{x} + \vec{y}) = P_V(\vec{x}) + P_V(\vec{y})$. Pour cela, grâce à la Propriété 13.12 il suffit de montrer que $P_V(\vec{x}) + P_V(\vec{y}) \in V$ et que $\vec{x} + \vec{y} - (P_V(\vec{x}) + P_V(\vec{y})) \in V^\perp$. La première condition est satisfaite car $P_V(\vec{x}), P_V(\vec{y}) \in V$ et V est un s.e.v. par hypothèse. La deuxième condition est satisfaite car $\vec{x}+\vec{y}-(P_V(\vec{x})+P_V(\vec{y})) = \vec{x}-P_V(\vec{x})+\vec{y}-P_V(\vec{y})$ et $\vec{x}-P_V(\vec{x}), \vec{y}-P_V(\vec{y}) \in V^\perp$, qui est un s.e.v. par 13.9. □

13.14 Propriété. *Soit E un espace euclidien et V un s.e.v. de E. Si la projection orthogonale P_V de E sur V existe, alors :*

1. $E = V \oplus V^\perp$,
2. *la projection orthogonale $P_{V^\perp}$ de E sur $V^\perp$ existe,*
3. $V = (V^\perp)^\perp$.

Preuve. 1. Tout vecteur de E se décompose comme somme d'un vecteur de V et d'un vecteur de $V^\perp$:

$$\vec{x} = P_V(\vec{x}) + \vec{x} - P_V(\vec{x})$$

Donc $E = V + V^\perp$ et cette somme est directe car $V \cap V^\perp = \{\vec{0}\}$ (voir 13.9).
2. Montrons qu'on peut poser $P_{V^\perp}(\vec{x}) = \vec{x} - P_V(\vec{x})$. Pour cela, vérifions les deux conditions de la Définition 13.11 avec $V^\perp$ à la place de V. Evidemment $\vec{x} - P_V(\vec{x}) \in V^\perp$. En autre, $\vec{x} - (\vec{x} - P_V(\vec{x})) = P_V(\vec{x}) \in V \subseteq (V^\perp)^\perp$ (voir 13.9).
3. Il faut montrer que $(V^\perp)^\perp \subseteq V$ (l'inclusion $V \subseteq (V^\perp)^\perp$ est toujours valable, voir 13.9). Soit $\vec{x} \in (V^\perp)^\perp$. Par le point 1 nous pouvons écrire

$$\vec{x} = \vec{a} + \vec{b}$$

avec $\vec{a} \in V \subseteq (V^\perp)^\perp$ et $\vec{b} \in V^\perp$. On a donc $\vec{x} - \vec{a} = \vec{b}$ et, puisque $\vec{x} - \vec{a} \in (V^\perp)^\perp$ et $\vec{b} \in V^\perp$, $\vec{x} - \vec{a} = \vec{0}$ (voir à nouveau 13.9). Cela donne $\vec{x} = \vec{a}$ et donc $\vec{x} \in V$. □

La propriété la plus importante de la projection orthogonale est qu'elle minimise la distance :

13.15 Propriété. *Soit E un espace euclidien et soit V un s.e.v. de E. Soit $P_V : E \to E$ la projection orthogonale de E sur V. Soit $\vec{x} \in E$ et $\vec{y} \in V$. Si $\vec{y} \neq P_V(\vec{x})$, alors*

$$\operatorname{dist}(\vec{x}, P_V(\vec{x})) < \operatorname{dist}(\vec{x}, \vec{y})$$

Preuve. Puisque la racine carrée est une fonction monotone croissante, il suffit de montrer que $\mathrm{dist}^2(\vec{x}, P_V(\vec{x})) < \mathrm{dist}^2(\vec{x}, \vec{y})$:

$$\mathrm{dist}^2(\vec{x}, \vec{y}) = \| \vec{x} - \vec{y} \|^2 = (\vec{x} - \vec{y} \mid \vec{x} - \vec{y}) =$$

$$= (\vec{x} - P_V(\vec{x}) + P_V(\vec{x}) - \vec{y} \mid \vec{x} - P_V(\vec{x}) + P_V(\vec{x}) - \vec{y}) =$$

$$= (\vec{x} - P_V(\vec{x}) \mid \vec{x} - P_V(\vec{x})) + 2(\vec{x} - P_V(\vec{x}) \mid P_V(\vec{x}) - \vec{y}) + (P_V(\vec{x}) - \vec{y} \mid P_V(\vec{x}) - \vec{y}) =$$

$$= (\vec{x} - P_V(\vec{x}) \mid \vec{x} - P_V(\vec{x})) + (P_V(\vec{x}) - \vec{y} \mid P_V(\vec{x}) - \vec{y}) >$$

$$> (\vec{x} - P_V(\vec{x}) \mid \vec{x} - P_V(\vec{x})) = \| \vec{x} - P_V(\vec{x}) \|^2 = \mathrm{dist}^2(\vec{x}, P_V(\vec{x}))$$

où nous avons utilisé que

- $(\vec{x} - P_V(\vec{x}) \mid P_V(\vec{x}) - \vec{y}) = 0$ car $\vec{x} - P_V(\vec{x}) \in V^{\perp}$ et $P_V(\vec{x}) - \vec{y} \in V$
- $(P_V(\vec{x}) - \vec{y} \mid P_V(\vec{x}) - \vec{y}) > 0$ car $\vec{y} \neq P_V(\vec{x})$.

□

13.4 Bases orthonormées

Il reste à regler le problème de l'existence de la projection orthogonale. Pour cela nous allons utiliser la notion de base orthonormée :

13.16 Définition. Soit E un espace euclidien et soit $\vec{e}_1, \ldots, \vec{e}_n \in E$.

1. $\vec{e}_1, \ldots, \vec{e}_n$ est une *famille orthogonale* si $\vec{e}_i \neq \vec{0}$ et si $\vec{e}_i \perp \vec{e}_j$ pour tout $i, j = 1, \ldots, n, i \neq j$.
2. $\vec{e}_1, \ldots, \vec{e}_n$ est une *famille orthonormée* si $\| \vec{e}_i \| = 1$ et si $\vec{e}_i \perp \vec{e}_j$ pour tout $i, j = 1, \ldots, n, i \neq j$.

13.17 Exemple. La base canonique de $\mathbb{R}^n$ est une base orthonormée par rapport au produit scalaire canonique.

13.18 Remarque.

1. Toute famille orthonormée est orthogonale (car si $\| \vec{e}_i \| = 1$ alors $\vec{e}_i \neq \vec{0}$).
2. (Cette remarque sera utilisée au Chapitre 16 pour établir l'existence de l'application adjointe en dimension finie.) Si un espace euclidien E admet une base orthonormée $\vec{e}_1, \ldots, \vec{e}_n$ alors le produit scalaire de E peut s'exprimer à travers le produit scalaire canonique de $\mathbb{R}^n$: si $\vec{x} = x_1 \cdot \vec{e}_1 + \ldots + x_n \cdot \vec{e}_n$ et $\vec{y} = y_1 \cdot \vec{e}_1 + \ldots + y_n \cdot \vec{e}_n$ alors

$$(\vec{x} \mid \vec{y}) = x_1 y_1 + \ldots + x_n y_n$$

En particulier :

$$(\vec{x} \mid \vec{e}_i) = x_i$$

En effet, par la bilinéarité du produit scalaire, nous avons :

$$\left(\sum_{i=1}^{n} x_i \cdot \vec{e}_i \mid \sum_{i=1}^{n} y_i \cdot \vec{e}_i\right) = \sum_{i,j=1}^{n} x_i y_j (\vec{e}_i \mid \vec{e}_j) = \sum_{i=1}^{n} x_i y_i$$

car $(\vec{e}_i \mid \vec{e}_j) = 1$ si $i = j$ et $(\vec{e}_i \mid \vec{e}_j) = 0$ si $i \neq j$.

13.19 Propriété. *Soit E un espace euclidien, V un s.e.v. de E, et $\vec{e}_1, \ldots, \vec{e}_n$ une base orthonormée de V. Alors la projection orthogonale de E sur V existe et elle est donnée par*

$$P_V(\vec{x}) = (\vec{x} \mid \vec{e}_1) \cdot \vec{e}_1 + \ldots + (\vec{x} \mid \vec{e}_n) \cdot \vec{e}_n \qquad [13.19]$$

Preuve. Il faut montrer que $P_V(\vec{x})$ défini dans [13.19] satisfait aux conditions de la Définition 13.11 :
1. $P_V(\vec{x}) \in V$ car il est une combinaison linéaire de $\vec{e}_1, \ldots, \vec{e}_n \in V$ et V est un s.e.v.
2. Pour monter que $\vec{x} - P_V(\vec{x}) \in V^\perp$ il faut monter que $(\vec{x} - P_V(\vec{x}) \mid \vec{v}) = 0$ pour tout $\vec{v} \in V$, et pour cela il suffit de montrer que $(\vec{x} - P_V(\vec{x}) \mid \vec{e}_i) = 0$ pour tout $i = 1, \ldots, n$:

$$\begin{aligned} (\vec{x} - P_V(\vec{x}) \mid \vec{e}_i) &= (\vec{x} - \textstyle\sum_{j=1}^n (\vec{x} \mid \vec{e}_j) \cdot \vec{e}_j \mid \vec{e}_i) \\ &= (\vec{x} \mid \vec{e}_i) - \textstyle\sum_{j=1}^n (\vec{x} \mid \vec{e}_j)(\vec{e}_j \mid \vec{e}_i) \\ &= (\vec{x} \mid \vec{e}_i) - (\vec{x} \mid \vec{e}_i) = 0 \end{aligned}$$

car $(\vec{e}_i \mid \vec{e}_j) = 1$ si $i = j$ et $(\vec{e}_i \mid \vec{e}_j) = 0$ si $i \neq j$. □

La propriété précédente montre que l'existence de la projection orthogonale sur V dépend de l'existence d'une base orthonormée de V.

13.20 Lemme. *Toute famille orthogonale est une famille libre.*

Preuve. Soit $\vec{e}_1, \ldots, \vec{e}_n$ une famille orthogonale et supposons que

$$\sum_{i=1}^n \alpha_i \cdot \vec{e}_i = \vec{0}$$

Alors pour tout $j = 1, \ldots, n$ nous avons

$$0 = (\vec{0} \mid \vec{e}_j) = (\sum_{i=1}^n \alpha_i \cdot \vec{e}_i \mid \vec{e}_j) = \sum_{i=1}^n \alpha_i (\vec{e}_i \mid \vec{e}_j) = \alpha_j (\vec{e}_j \mid \vec{e}_j)$$

et donc $\alpha_j = 0$ car $(\vec{e}_j \mid \vec{e}_j) > 0$.

13.21 Propriété. *Soit E un espace euclidien et V un s.e.v. de E. Si V est finiment engendré et $V \neq \{\vec{0}\}$, alors V admet une base orthonormée.*

Preuve. Par induction sur la dimension n de V.
Si $n = 1$, soit $\vec{v}_1 \in V, \vec{v}_1 \neq \vec{0}$. Alors

$$\frac{\vec{v}_1}{\| \vec{v}_1 \|}$$

est une base orthonormée de V.
Si $n > 1$, soit $\vec{v}_1, \ldots, \vec{v}_n$ une base de V. Par hypothèse inductive le s.e.v.

$W = \langle \vec{v}_1, \ldots, \vec{v}_{n-1} \rangle$ admet une base orthonormée $\vec{e}_1, \ldots, \vec{e}_{n-1}$ et donc, par la Propriété 13.19, il existe la projection orthogonale P_W de E sur W. Posons

$$\vec{e}_n = \vec{v}_n - P_W(\vec{v}_n)$$

et montrons que $\vec{e}_1, \ldots, \vec{e}_{n-1}, \vec{e}_n$ est une famille orthogonale de vecteurs de V (et donc

$$\vec{e}_1, \ldots, \vec{e}_{n-1}, \frac{\vec{e}_n}{\| \vec{e}_n \|}$$

est une base orthonormée de V).

- $\vec{e}_n \in V$ car $\vec{e}_n = \vec{v}_n - P_W(\vec{v}_n)$ avec $\vec{v}_n \in V$ et $P_W(\vec{v}_n) \in W \subset V$.
- $\vec{e}_n \neq \vec{0}$ car $P_W(\vec{v}_n)$ appartient à W mais $\vec{v}_n$ n'appartient pas à W (si $\vec{v}_n \in W$, alors $\vec{v}_n$ est combinaison linéaire de $\vec{v}_1, \ldots, \vec{v}_{n-1}$, ce qui est impossible car $\vec{v}_1, \ldots, \vec{v}_{n-1}, \vec{v}_n$ est une famille libre). Donc $\vec{v}_n \neq P_W(\vec{v}_n)$, c'est-à-dire $\vec{v}_n - P_W(\vec{v}_n) \neq \vec{0}$.
- $\vec{e}_i \perp \vec{e}_n$ pour tout $i = 1, \ldots, n-1$ car $\vec{e}_i \in W$ et $\vec{e}_n = \vec{v}_n - P_W(\vec{v}_n) \in W^{\perp}$.

$\square$

13.22 Remarque. La Propriété 13.21 est valable même si $V = \{\vec{0}\}$. Dans ce cas l'unique base de V est consituée par la famille vide, qui est une famille orthonormée.

13.23 Corollaire. *Soit E un espace euclidien et V un s.e.v. finiment engendré de E. La projection orthogonale $P_V \colon E \to E$ de E sur V existe et, par conséquent, $E = V \oplus V^{\perp}$.*

Preuve. C'est une conséquence de 13.19, 13.21 et 13.14. $\square$

13.24 Remarque. D'après la Propriété 13.19 pour calculer la projection orthogonale P_V il nous faut une base orthonormée de V. Comme démontré dans 13.21 une telle base existe si V est de dimension finie. Explicitons la construction inductive utilisée dans la preuve de la Propriété 13.21.
Soit $\vec{v}_1, \ldots, \vec{v}_n$ une base quelconque de V ; on pose :

$$\begin{aligned}
\vec{e}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} \\
\vec{e}_2 &= \frac{\vec{v}_2 - (\vec{v}_2|\vec{e}_1)\cdot\vec{e}_1}{\|\vec{v}_2 - (\vec{v}_2|\vec{e}_1)\cdot\vec{e}_1\|} \\
\vec{e}_3 &= \frac{\vec{v}_3 - (\vec{v}_3|\vec{e}_1)\cdot\vec{e}_1 - (\vec{v}_3|\vec{e}_2)\cdot\vec{e}_2}{\|\vec{v}_3 - (\vec{v}_3|\vec{e}_1)\cdot\vec{e}_1 - (\vec{v}_3|\vec{e}_2)\cdot\vec{e}_2\|} \\
\ldots \\
\vec{e}_n &= \frac{\vec{v}_n - \sum_{i=1}^{n-1}(\vec{v}_n|\vec{e}_i)\cdot\vec{e}_i}{\|\vec{v}_n - \sum_{i=1}^{n-1}(\vec{v}_n|\vec{e}_i)\cdot\vec{e}_i\|}
\end{aligned}$$

Les vecteurs $\vec{e}_1, \ldots, \vec{e}_n$ sont une base orthonormée de V. Cette méthode pour construire une base orthonormée est la méthode de Gram-Schmidt.

13.5 Problèmes d'approximation

Nous pouvons revenir au problème de départ pour voir comment la projection orthogonale permet de trouver les solutions approchées d'un système qui n'a pas de solutions exactes.

13.25 Exercice. Considérons le système

$$\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$$

avec $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ et $\vec{b} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

1. $\mathcal{S}$ n'a pas de solutions exactes : en effet $\operatorname{rang} A = 2 < 3 = \operatorname{rang}(A \mid \vec{b}\,)$, ce qui signifie que $\vec{b}$ n'est pas combinaison linéaire des colonnes de A.
2. Puisque $\operatorname{rang} A = 2$, les colonnes de A sont linéairement indépendantes, elle sont donc une base de $\mathrm{Col}(A)$. Nous pouvons appliquer la méthode de G.-S. pour obtenir une base orthonormée de $\mathrm{Col}(A)$:
$$\vec{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \qquad \vec{e}_2 = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$
3. La projection orthogonale de $\vec{b}$ sur $\mathrm{Col}(A)$ est donc
$$\vec{b}' = (\vec{b} \mid \vec{e}_1) \cdot \vec{e}_1 + (\vec{b} \mid \vec{e}_2) \cdot \vec{e}_2 = \begin{pmatrix} \frac{4}{3} \\ \frac{4}{3} \\ \frac{4}{3} \end{pmatrix}$$
4. On obtient ainsi un nouveau système
$$\mathcal{S}'\colon A \cdot \vec{x} = \vec{b}'$$
Les solutions approchées du système de départ $\mathcal{S}$ sont exactement les solutions du nouveau système $\mathcal{S}'$. Dans ce cas le nouveau système $\mathcal{S}'$ admet une solution unique (car $\operatorname{rang} A = 2$) qui est
$$\vec{x} = \begin{pmatrix} \frac{4}{3} \\ 0 \end{pmatrix}$$

13.26 Remarque. Pour calculer les solutions approchées d'un système

$$\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$$

on peut éviter de calculer explicitement la projection orthogonale $\vec{b}'$ de $\vec{b}$ sur $\mathrm{Col}(A)$, et même de calculer une base orthonormée de $\mathrm{Col}(A)$. En effet, $\vec{x}$ est solution approchées de $\mathcal{S}$ ssi $A \cdot \vec{x} = \vec{b}'$, c'est-à-dire ssi le vecteur $A \cdot \vec{x}$ est la projection orthogonale de $\vec{b}$ sur $\mathrm{Col}(A)$. D'après la Définition 13.11 cela revient à satisfaire deux conditions :

1. $A \cdot \vec{x} \in \mathrm{Col}(A)$
2. $\vec{b} - A \cdot \vec{x} \in \mathrm{Col}(A)^{\perp}$

La première condition est toujours vérifiée et la deuxième revient à

$$(\vec{b} - A \cdot \vec{x}) \perp \vec{c} \quad \text{pour tout} \quad \vec{c} \in \mathrm{Col}(A)$$

c'est-à-dire à

$$(\vec{b} - A \cdot \vec{x} \mid \vec{a}_{*j}) = 0 \quad \text{pour tout} \quad j = 1, \dots, n$$

ou encore à

$$(\vec{b} - \sum_{i=1}^{n} x_i \cdot \vec{a}_{*i} \mid \vec{a}_{*j}) = 0 \quad \text{pour tout} \quad j = 1, \dots, n$$

et finalement à

$$\sum_{i=1}^{n} x_i (\vec{a}_{*i} \mid \vec{a}_{*j}) = (\vec{b} \mid \vec{a}_{*j}) \quad \text{pour tout} \quad j = 1, \dots, n$$

13.27 Exercice. Reprenons l'Exercice 13.25 en utilisant cette deuxième méthode. Les solutions approchées de $\mathcal{S}$ sont les solutions du système

$$\left\{ \begin{array}{rcl} x_1(\vec{a}_{*1} \mid \vec{a}_{*1}) + x_2(\vec{a}_{*2} \mid \vec{a}_{*1}) & = & (\vec{b} \mid \vec{a}_{*1}) \\ x_1(\vec{a}_{*1} \mid \vec{a}_{*2}) + x_2(\vec{a}_{*2} \mid \vec{a}_{*2}) & = & (\vec{b} \mid \vec{a}_{*2}) \end{array} \right.$$

avec $\vec{a}_{*1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{a}_{*2} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. Cela donne

$$\left\{ \begin{array}{rcl} 3x_1 + 6x_2 & = & 4 \\ 6x_1 + 14x_2 & = & 8 \end{array} \right.$$

et l'unique solution est $x_1 = \frac{4}{3}$, $x_2 = 0$.

Une autre application des espaces euclidiens est la recherche du polynôme de degré n qui approxime au mieux une famille de m points du plan, avec $n \leq m$. Nous nous limitons ici à un exercice de recherche de la droite qui approxime au mieux trois points non alignés du plan (une telle droite est appelée *droite d'interpolation linéaire.*

13.28 Exercice. Déterminer la droite $d(x) = ax + b$ qui approxime au mieux les trois points

$$P_1 = (-1, 0)\,,\ P_2 = (0, -1)\,,\ P_3 = (1, 0)$$

Deux approches sont possibles :
Première méthode : soit $q(x) = \alpha x^2 + \beta x + \gamma$ l'unique polynôme de degré 2 tel que

$$q(-1) = 0\ ,\quad q(0) = -1\ ,\quad q(1) = 0$$

La droite $d(x)$ que nous cherchons est la projection orthogonale de q sur l'espace $\mathbb{R}[x]_1$ des polynômes de degré plus petit ou égal à 1 muni du produit scalaire

$$(f(x) \mid g(x)) = f(-1)g(-1) + f(0)g(0) + f(1)g(1)$$

La base usuelle de $\mathbb{R}[x]_1$ est donnée par $\{p_0(x) = 1, p_1(x) = x\}$. Il s'agit en effet d'une base orthogonale par rapport au produit scalaire choisi, et donc pour avoir une base orthonormée il suffit de normer $p_0(x)$ et $p_1(x)$. Nous obtenons

$$\left\{u_0(x) = \frac{1}{\sqrt{3}}\ ,\ u_1(x) = \frac{x}{\sqrt{2}}\right\}$$

ce qui nous permet de calculer la projection orthogonale, qui resulte être une droite horizontale :

$$d(x) = (q(x) \mid u_0(x)) \cdot u_0(x) + (q(x) \mid u_1(x)) \cdot u_1(x) = -\frac{1}{3}$$

Deuxième méthode : si nous imposons à la droite $d(x) = ax + b$ de passer par les trois points P_1, P_2, P_3, nous obtenons le système $S\colon A \cdot \vec{x} = \vec{b}$ suivant

$$S\colon \begin{pmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

qui est évidemment impossible. Nous devons donc remplacer $\vec{b}$ par sa projection orthogonale $\vec{b}'$ sur l'espace $\mathrm{Col}(A)$. Or, les colonnes de A sont orthogonales par rapport au produit scalaire canonique de $\mathbb{R}^3$, il suffit donc de les normer pour obtenir la base orthonormée suivante

$$\{\vec{e}_1 = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})^t\ ,\ \vec{e}_2 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^t\}$$

Cela nous donne comme projection orthogonale

$$\vec{b}' = (\vec{b} \mid \vec{e}_1) \cdot \vec{e}_1 + (\vec{b} \mid \vec{e}_2) \cdot \vec{e}_2 = (-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3})^t$$

et le nouveau système $S'\colon A \cdot \vec{x} = \vec{b}'$ admet comme unique solution $a = 0, b = -\frac{1}{3}$, ce qui finalement nous donne comme droite d'interplation la droite horizontale d'équation $d(x) = -\frac{1}{3}$.

13.29 Remarque.

1. Remarquons que dans la première méthode utilisée dans l'Exercice 13.28 nous n'avons pas besoin de calculer explicitement le polynôme $q(x)$, mais seulement de savoir qu'un tel polynôme existe (et est unique) et de connaître les valeurs $q(-1), q(0), q(1)$. Ceci est intéressant car, bien que dans notre exemple la détermination de $q(x)$ est immédiate (c'est le polynôme $q(x) = x^2 - 1$) dans le cas d'un nuage de m points avec m très grand l'expression de $q(x)$ est assez laborieuse.

2. On peut justifier le fait que les deux méthodes utilisées dans l'Exercice 13.28 donnent le même résultat par un argument similaire à celui de la Remarque 13.26. Voyons-le pour la droite $d(x)$ qui approxime au mieux trois points arbitraires $P_1 = (x_1, y_1), P_2 = (x_2, y_2), P_3 = (x_3, y_3)$:
Première méthode : le fait que la droite $d(x) = ax + b$ soit la projection orthogonale de $q(x)$ sur $\mathbb{R}[x]_1$ revient au fait que

$$q(x) - d(x) \in \mathbb{R}[x]_1^\perp$$

et cela s'exprime par les équations

$$(q(x) - d(x) \mid p_0(x)) = 0 \qquad (q(x) - d(x) \mid p_1(x)) = 0$$

c'est-à-dire, en vertu du produit scalaire choisi dans $\mathbb{R}[x]$,

$$\sum_{i=1}^{3}(y_i - (ax_i + b)) = 0 \qquad \sum_{i=1}^{3}(y_i x_i - (ax_i + b)x_i) = 0$$

Deuxième méthode : le fait que $\vec{x} = (a, b)^t$ soit solution approchée du système $S\colon A \cdot \vec{x} = \vec{b}$, avec $\vec{b} = (y_1, y_2, y_3)^t$, revient au fait que $A \cdot \vec{x}$ soit la projection orthogonale de $\vec{b}$ sur $\mathrm{Col}(A)$, c'est-à-dire

$$\vec{b} - A \cdot \vec{x} \in \mathrm{Col}(A)^\perp$$

et cela s'exprime par les équations

$$(\vec{b} - A \cdot \vec{x} \mid (x_1, x_2, x_3)^t) = 0 \qquad (\vec{b} - A \cdot \vec{x} \mid (1, 1, 1)^t) = 0$$

car $(x_1, x_2, x_3)^t$ et $(1, 1, 1)^t$ sont les colonnes de la matrice A. En utilisant le produit scalaire canonique de $\mathbb{R}^3$, cela donne exactement les équations

$$\sum_{i=1}^{3}(y_i - (ax_i + b))x_i = 0 \qquad \sum_{i=1}^{3}(y_i - (ax_i + b)) = 0$$

13.30 Remarque. Il est intéressant de remarquer que, dans les Propriétés 13.12, 13.15, 13.19 et 13.21, le caractère défini positif du produit scalaire

$$(\vec{x} \mid \vec{x}) > 0 \ \text{ si } \ \vec{x} \neq \vec{0}$$

(voir Définition 13.1) n'intervient que pour les vecteurs non nuls du sous-espace vectoriel V sur lequel on effectue la projection orthogonale. Par conséquent, on peut répéter la théorie développée dans ce chapitre pour un espace vectoriel réel E muni d'une application

$$(- \mid -)\colon E \times E \to \mathbb{R}$$

bilinéaire, symétrique et semi-définie positive, où semi-définie positive signifie

$$(\vec{x} \mid \vec{x}) \geq 0 \ \text{ pour tout } \ \vec{x} \in E,$$

et d'un sous-espace vectoriel V de E tel que la restriction de la fonction $(- \mid -)$ à $V \times V$ soit définie positive :

$$(\vec{x} \mid \vec{x}) > 0 \;\text{ pour tout }\; \vec{x} \in V,\; \vec{x} \neq \vec{0}.$$

La différence entre cette situation nouvelle et la situation traitée dans le chapitre est que, dans la situation nouvelle, on peut bien avoir deux vecteurs différents $\vec{x}$ et $\vec{y}$ de E qui n'appartiennent pas à V et qui sont tells que $\operatorname{dist}(\vec{x}, \vec{y}) = 0$.

La remarque précédente permet d'appliquer la théorie des espaces euclidiens à des nouveaux exemples intéressants. En voici deux (très proches l'un à l'autre).

13.31 Exemple. Soit $a, b, c \in \mathbb{R}$ trois nombres réels fixés, avec $a < b < c$, et soit

$$E = \mathcal{D}_b(]a, c[, \mathbb{R}) = \{f\colon]a, c[\to \mathbb{R} \mid f \;\text{ est dérivable au point }\; b\}$$

l'espace vectoriel réel des fonctions dérivables au point b.
L'espace E est muni d'une forme bilinéaire, symétrique et semi-définie positive

$$(- \mid -)\colon E \times E \to \mathbb{R}\,,\;\; (f \mid g) = f(b) \cdot g(b) + f'(b) \cdot g'(b)$$

Cette forme n'est pas définie positive : par exemple, la fonction $f(x) = (x - b)^2$ est un vecteur non nul de E et $(f \mid f) = 0$.
Si on considère comme sous-espace vectoriel de E l'espace

$$V = \mathbb{R}[x]_1$$

des polynômes de degré plus petit ou égal à 1, la restriction de $(- \mid -)$ à $V \times V$ est définie positive : en effet pour un polynôme de premier degré $p(x) = a_0 + a_1 x$, la condition $(p \mid p) = 0$ implique tout de suite $a_0 = a_1 = 0$, c'est-à-dire que p est le polynôme nul.
Nous pouvons donc calculer la projection orthogonale d'une fonction $f \in E$ sur V, et il est facile de vérifier que $P_V(f)$ est exactement la droite tangente au graphe de f au point de coordonnées $(b, f(b))$.
Le polynôme de Taylor de degré n d'une fonction n fois dérivable peut être traité de façon similaire.

13.32 Exemple. Soit E l'espace vectoriel réel des fonctions $f\colon \mathbb{R}^2 \to \mathbb{R}$ différentiables et soit V son sous-espace vectoriel des polynômes de degré au plus 1, en deux variables et à coefficients réels

$$V = \{p\colon \mathbb{R}^2 \to \mathbb{R}\,,\; p(x, y) = ax + by + c \mid a, b, c \in \mathbb{R}\}.$$

L'espace E est muni d'une forme bilinéaire, symétrique et semi-définie positive

$$(f \mid g) = f(0) \cdot g(0) + \frac{\partial f}{\partial x}(0) \cdot \frac{\partial g}{\partial x}(0) + \frac{\partial f}{\partial y}(0) \cdot \frac{\partial g}{\partial y}(0)$$

où 0 est l'origine de $\mathbb{R}^2$. Cette forme n'est pas définie positive : par exemple, la fonction $f(x, y) = x \cdot y$ est un vecteur non nul de E et $(f \mid f) = 0$.

La restriction de $(- \mid -)$ à $V \times V$ est définie positive : si on considère un polynôme du type $p(x, y) = ax + by + c$, alors $(p \mid p) = c^2 + a^2 + b^2$. La seule possibilité pour avoir $(p \mid p) = 0$ est donc $a = b = c = 0$, c'est-à-dire que p doit être le polynôme nul.
Nous pouvons donc calculer la projection orthogonale d'une fonction $f \in E$ sur V : la base canonique $\{1, x, y\}$ de V est orthonormée par rapport au produit scalaire choisi, et nous obtenons

$$P_V(f) = (f \mid 1) \cdot 1 + (f \mid x) \cdot x + (f \mid y) \cdot y = f(0) + \frac{\partial f}{\partial x}(0) \cdot x + \frac{\partial f}{\partial y}(0) \cdot y$$

c'est-à-dire le plan tangent en $(0, f(0))$ au graphe de la fonction f.

Chapitre 14

Opérateurs linéaires

Le chapitre 14 section par section

Les systèmes d'équations linéaires sont suvent utilisés pour modéliser des systèmes physiques, biologiques, exceptera, qui évoluent dans le temps. Ceci conduit à se poser des nouvelles questions qui mènent aux notions de valeur propre et vecteur propre.

1. Si une population évolue de façon linéaire, on peut prédire son évolution à partir de la population initiale, mais pour cela il faut utiliser les puissances d'une matrice carrée.
2. Les valeurs propres et les vecteurs propres d'un opérateur linéaire L sont liés par la condition $L(\vec{x}) = \lambda \cdot \vec{x}$. Par itération, cela donne $L^k(\vec{x}) = \lambda^k \cdot \vec{x}$.
3. Un opérateur linéaire est diagonalisable si tout vecteur peut s'écrire comme combinaison linéaire de vecteurs propres. Autrement dit, sa matrice associée est, à un changement de base près, diagonale.
4. Le polynôme caractéristique permet de calculer les valeurs propres d'un opérateur linéaire ou d'une matrice carrée.
5. Pour déterminer si un opérateur linéaire est diagonalisable, il faut comparer la dimension de chaque espace propre avec la multiplicité algébrique de la valeur propre correspondante.
6. Pour terminer, retour au problème de l'évolution linéaire d'une population et, en particulier, au cas où l'opérateur ou la matrice qui décrivent l'évolution sont diagonalisables.

14.1 Evolution linéaire d'une population

Commençons par un exemple.

14.1 Exemple. Soit P une population animale dont la durée de vie maximale est de n années. La population P est donc répartie en n tranches d'age :

$$P = (p_1, p_2, \dots, p_n)$$

Soit r_i le taux de reproduction des individus de la i-ème tranche d'age et s_i leur taux de survie. Si on dénote par

$$P_0 = (p_{01}, p_{02}, \ldots, p_{0n})$$

la population au temps 0, la population après un an

$$P_1 = (p_{11}, p_{12}, \ldots, p_{1n})$$

est donnée par le système

$$\mathcal{S}\colon \left\{ \begin{array}{lcl} p_{11} & = & r_1 p_{01} + r_2 p_{02} + \ldots + r_n p_{0n} \\ p_{12} & = & s_1 p_{01} \\ \vdots & & \\ p_{1n} & = & s_{n-1} p_{0,n-1} \end{array} \right.$$

On peut alors se poser deux questions :

1. Comment peut-on déterminer la population après plusieurs années (k années)
$$P_k = (p_{k1}, p_{k2}, \ldots, p_{kn})$$
en fonction de la population initiale P_0 et des paramètres r_i et s_i ($i = 1, \ldots, n$) ?
2. Est-il possible de déterminer si la population restera stable, ou si elle évoluera vers l'extinction ou vers une surpopulation ?

Ces questions conduisent au calcul des puissances d'une matrice carrée. En effet, si on met le système $\mathcal{S}$ sous forme matricielle

$$\mathcal{S}\colon A \cdot \vec{x} = \vec{b}$$

on a :

$$\begin{array}{rcccccccc} \text{population après 1 an} & : & P_1 & = & A \cdot P_0 & & & & \\ \text{population après 2 ans} & : & P_2 & = & A \cdot P_1 & = & A \cdot A \cdot P_0 & = & A^2 \cdot P_0 \\ \vdots & & & & & & & & \\ \text{population après } k \text{ ans} & : & P_k & = & A^k \cdot P_0 & & & & \end{array}$$

et même

$$\lim_{k \to +\infty} P_k = \left(\lim_{k \to +\infty} A^k \right) \cdot P_0$$

14.2 Remarque. Le calcul des puissances d'une matrice carrée est en général assez fastidieux. Il y a néanmoins deux cas où ce calcul devient facile :

1. Si la matrice est une matrice *diagonale*
$$D = (d_{ij})_{i,j=1,\ldots,n} \quad d_{ij} = 0 \;\text{ si }\; i \neq j$$
alors
$$D^k = (d_{ij}^k)_{i,j=1,\ldots,n}$$

2. Si la matrice $A \in \mathbb{K}^{n\times n}$ peut s'écrire comme produit

$$A = P \cdot D \cdot P^{-1}$$

avec $D \in \mathbb{K}^{n\times n}$ diagonale et $P \in \mathbb{K}^{n\times n}$ inversible, alors

$$\begin{aligned} A^k &= (P \cdot D \cdot P^{-1})^k \\ &= P \cdot D \cdot P^{-1} \cdot P \cdot D \cdot P^{-1} \cdot \ldots \cdot P \cdot D \cdot P^{-1} \\ &= P \cdot D \cdot D \cdot \ldots \cdot D \cdot P^{-1} \\ &= P \cdot D^k \cdot P^{-1} \end{aligned}$$

et on se reconduit essentiellement au premier cas.

L'objectif de ce chapitre est de trouver des conditions nécessaires et suffisantes pour qu'une matrice $A \in \mathbb{K}^{n\times n}$ puisse s'écrire comme

$$A = P \cdot D \cdot P^{-1}$$

avec P inversible et D diagonale. Pour cela nous utilisons les valeurs propres, qui interviendront aussi dans le Chapitre 17 consacré aux formes quadratiques. Dans ce chapitre $\mathbb{K}$ est un corps commutatif quelconque.

14.2 Valeurs propres et vecteurs propres

14.3 Définition. Soit V un espace vectoriel sur $\mathbb{K}$. Un *opérateur linéaire* sur V est une application linéaire

$$L\colon V \to V$$

Pour mieux comprendre la définition de valeur propre et de vecteur propre voyons d'abord un exemple.

14.4 Exemple. Considérons les matrices

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ inversible, et } D = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \text{ diagonale}$$

Posons

$$A = P \cdot D \cdot P^{-1} = \begin{pmatrix} 0 & 5 \\ -10 & 15 \end{pmatrix}$$

Il est facile de vérifier que la matrice A, les colonnes de P et les éléments de la diagonale de D sont liés par les conditions

$$A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad A \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 10 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

c'est-à-dire $L_A(\vec{a}_{*1}) = 5 \cdot \vec{a}_{*1}$, $L_A(\vec{a}_{2*}) = 10 \cdot a_{2*}$.

14.5 Définition. Soit $L\colon V \to V$ un opérateur linéaire.

1. Un scalaire $\lambda \in \mathbb{K}$ est dit *valeur propre* de L s'il existe $\vec{x} \in V, \vec{x} \neq 0$ tel que
$$L(\vec{x}) = \lambda \cdot \vec{x} \qquad [14.5]$$
Un tel $\vec{x}$ est dit *vecteur propre* de L de valeur propre λ (ou associé à la valeur propre λ).
2. Soit $\lambda \in \mathbb{K}$, l'*espace propre* associé à λ est définie par
$$E(\lambda) = \{\vec{x} \in V \mid L(\vec{x}) = \lambda \cdot \vec{x}\}$$
Autrement dit,
$$E(\lambda) = \{ \text{ vecteurs propres associés à } \lambda \ \} \cup \{\vec{0}\,\}$$

14.6 Exercice. Soit $L\colon V \to V$ un opérateur linéaire et considérons $\vec{x} \in V$ et $\lambda \in \mathbb{K}$ tels que $L(\vec{x}) = \lambda \cdot \vec{x}$. Montrer que pour tout $n \in \mathbb{N}$ on a $L^n(\vec{x}) = \lambda^n \cdot \vec{x}$. Autrement dit, si la population initiale est un vecteur propre, l'évolution de la population est completement déterminée par la valeur propre correspondante.

14.7 Remarque. Soit $L\colon V \to V$ un opérateur linéaire et soit $\lambda \in \mathbb{K}$.

1. $E(\lambda)$ est un s.e.v. de V. En effet, $E(\lambda)$ est le noyau de l'opérateur linéaire
$$L - \lambda \cdot id_V \colon V \to V \ , \quad \vec{x} \mapsto L(\vec{x}) - \lambda \cdot \vec{x}$$
En particulier, $E(0) = \operatorname{Ker} L$.
2. λ est valeur propre de L ssi $\dim E(\lambda) \geq 1$.

14.3 Opérateurs diagonalisables

14.8 Définition. Un opérateur linéaire $L\colon V \to V$ est *diagonalisable* si V admet une famille génératrice formée par des vecteurs propres de L.
(Si V est de dimension finie, ceci revient à dire que V admet une base formée par des vecteurs propres de V.)

14.9 Exemple.

1. Soit $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ la projection sur la droite $\alpha\colon y = ax$ parallèlement à la droite $\beta\colon y = bx$ (avec $a \neq b$).
   (a) Tout vecteur non nul $\vec{a}$ parallèle à la droite α est vecteur propre de L associé à la valeur propre 1 car $L(\vec{a}) = \vec{a}$.
   (b) Tout vecteur non nul $\vec{b}$ parallèle à la droite β est vecteur propre de L associé à la valeur propre 0 car $L(\vec{b}) = \vec{0}$.

   L est donc diagonalisable. De plus, la matrice associée à L par rapport à une base formée par un vecteur parallèle à α et un vecteur parallèle à β est la matrice diagonale
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$
qui a les valeurs propres de L sur la diagonale.

2. Soit $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle α autour de l'origine, avec $\alpha \in [0, 2\pi[$.
   (a) Si $\alpha \neq 0$ et $\alpha \neq \pi$ alors L n'a pas de vecteurs propres (car si $\vec{x} \neq \vec{0}$ alors $L(\vec{x}) \neq \lambda \cdot \vec{x}$ pour tout $\lambda \in \mathbb{R}$). L n'est donc pas diagonalisable.
   (b) Si $\alpha = 0$ alors L est l'identité. Elle est diagonalisable car tout vecteur non nul est vecteur propre associé à la valeur propre 1. De plus, la matrice associée à L par rapport à une base quelconque est la matrice identité, qui est diagonale et dont les éléments sur la diagonale sont les valeurs propres.
   (c) Si $\alpha = \pi$ alors tout vecteur non nul est vecteur propre associé à la valeur propre -1. L est donc diagonalisable et la matrice associée à L par rapport à une base quelconque est la matrice diagonale

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

   Encore une fois on retrouve sur la diagonale les valeurs propres de L.

14.10 Définition. Une matrice carrée $A \in \mathbb{K}^{n\times n}$ est *diagonalisable* s'il existe $P \in \mathbb{K}^{n\times n}$ inversible et $D \in \mathbb{K}^{n\times n}$ diagonale telles que

$$A \cdot P = P \cdot D \quad (\text{ou bien } \; A = P \cdot D \cdot P^{-1}) \qquad [14.10]$$

14.11 Propriété.

1. *Soit $L\colon V \to V$ un opérateur linéaire et soit $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de V. L est diagonalisable au sens de la Définition 14.8 ssi ${}_eL_e$ est diagonalisable au sens de la Définition 14.10.*
2. *Soit $A \in \mathbb{K}^{n\times n}$. A est diagonalisable au sens de la Définition 14.10 ssi $L_A\colon \mathbb{K}^n \to \mathbb{K}^n$ est diagonalisable au sens de la Définition 14.8.*

Preuve. 1. Si $L\colon V \to V$ est diagonalisable, il existe une base $f = \{\vec{f}_1, \ldots, \vec{f}_n\}$ formée par des vecteurs propres de L:

$$L(\vec{f}_i) = \lambda_i \cdot \vec{f}_i \quad (i = 1, \ldots, n)$$

On obtient donc une matrice diagonale

$${}_fL_f = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} \in \mathbb{K}^{n\times n}$$

Par conséquent

$${}_eL_e = {}_eI_f \cdot {}_f L_f \cdot {}_f I_e = P \cdot D \cdot P^{-1}$$

avec $D = {}_fL_f$ et $P = {}_eI_f$, ce qui montre que ${}_eL_e$ est diagonalisable.
Réciproquement, si ${}_eL_e$ est diagonalisable on peut la décomposer comme

$${}_eL_e = P \cdot D \cdot P^{-1}$$

avec P inversible et D diagonale. Considérons l'isomorphisme

$$E\colon V \to \mathbb{K}^n \quad \vec{x} \mapsto {}_e\vec{x} \quad , \quad E^{-1}\colon \mathbb{K}^n \to V \quad (x_1, \ldots, x_n) \mapsto x_1 \cdot \vec{e}_1 + \ldots x_n \cdot \vec{e}_n$$

déterminé par la base e de V (voir 8.23). Posons

$$\vec{f_i} = E^{-1}(i\text{-ème colonne de } P)$$

Puisque P est inversible, ses colonnes sont linéairement indépendantes (voir 11.12 et 11.13). Par conséquent, les vecteurs $\vec{f_1}, \ldots, \vec{f_n}$ sont eux aussi linéairement indépendants (8.14) et forment donc une base de V (3.19). De plus, puisque

$${}_e(\vec{f_i}) = E(E^{-1}(i\text{-ème colonne de } P)) = i\text{-ème colonne de } P$$

la matrice changement de base ${}_eI_f$ est exactement la matrice P. Il reste à montrer que les vecteurs $\vec{f_i}$ sont des vecteurs propres de L :

$$D = P^{-1} \cdot_e L_e \cdot P = {}_fI_e \cdot_e L_e \cdot_e I_f = {}_fL_f$$

et donc si

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

on a

$${}_fL(\vec{f_1}) = {}_fL_f \cdot_f (\vec{f_1}) = D \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 \cdot_f (\vec{f_1})$$

d'où $L(\vec{f_1}) = \lambda_1 \cdot \vec{f_1}$. De même $L(\vec{f_i}) = \lambda_i \cdot \vec{f_i}$ pour $i = 2, \ldots, n$.
2. C'est un cas particulier de 1. car, si $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ est la base canonique de $\mathbb{K}^n$, alors ${}_e(L_A)_e = A$. □

14.12 Remarque.

1. Grâce à la Propriété 14.11 nous pouvons désormais parler de valeurs propres et vecteurs propres d'une matrice carrée : si $A \in \mathbb{K}^{n \times n}$, les valeurs propres et les vecteurs propres de A sont par définition les valeurs propres et les vecteurs propres de l'opérateur linéaire

$$L_A\colon \mathbb{K}^n \to \mathbb{K}^n \ , \ \ \vec{x} \mapsto A \cdot \vec{x}$$

2. Explicitement, $\lambda \in \mathbb{K}$ est valeur propre de A s'il existe un vecteur non nul $\vec{x} \in \mathbb{K}^n$ tel que $A \cdot \vec{x} = \lambda \cdot \vec{x}$. Un tel vecteur $\vec{x}$ est dit vecteur propre de A associé à la valeur propre λ.

3. De plus, si $P \in \mathbb{K}^{n\times n}$ est une matrice inversible et si on pose $B = P^{-1}\cdot A\cdot P$, alors :
   (a) Les valeurs propres de B sont exactement les valeurs propres de A.
   (b) $\vec{x} \in \mathbb{K}^n$ est vecteur propre de B associé à la valeur propre λ ssi $P \cdot \vec{x}$ est vecteur propre de A associé à la valeur propre λ.

   (pour la preuve, voir l'Exercice 14.15).
4. On peut extraire de la preuve de la Propriété 14.11 une méthode pour construire, si elles existent, la matrice inversible P et la matrice diagonale D telles que
$$A = P \cdot D \cdot P^{-1}:$$
   (a) les colonnes de P sont une base de $\mathbb{K}^n$ de vecteurs propres de A,
   (b) les éléments sur la diagonale de D sont les valeurs propres de A.

   Attention : il faut respecter l'ordre entre les colonnes de P et les éléments de la diagonale de D.
5. Retenons aussi de la preuve de la Propriété 14.11 que si $L\colon V \to V$ est un opérateur linéaire et f est une base de V, alors la matrice ${}_fL_f$ est diagonale ssi f est formée par des vecteurs propres de L. Dans ce cas, ${}_fL_f$ est la matrice diagonale des valeurs propres de L.

14.4 Polynôme caractéristique

Il reste à trouver une méthode pour déterminer les valeurs propres :

14.13 Propriété. *Soit $L\colon V \to V$ un opérateur linéaire, $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de V et $A = {}_eL_e \in \mathbb{K}^{n\times n}$ la matrice associée. Les valeurs propres de L sont les solutions de l'équation*
$$\det(A - \lambda \cdot I) = 0 \qquad [14.13]$$
où $I \in \mathbb{K}^{n\times n}$ est la matrice identité. La fonction de la variable λ
$$\det(A - \lambda \cdot I)\colon \mathbb{K} \to \mathbb{K}$$
est un polynôme de degré n dit polynôme caractéristique *de L.*

Preuve.
$$\begin{aligned}
\lambda \text{ est valeur propre de } L \quad &\Leftrightarrow \quad \exists\, \vec{x} \neq \vec{0}\colon L(\vec{x}) = \lambda \cdot \vec{x} \\
&\Leftrightarrow \quad \exists\, \vec{x} \neq \vec{0}\colon \vec{x} \in \operatorname{Ker}(L - \lambda \cdot id_V) \\
&\Leftrightarrow \quad L - \lambda \cdot id_V \ \text{ n'est pas injective} \\
&\Leftrightarrow \quad L - \lambda \cdot id_V \ \text{ n'est pas bijective} \\
&\Leftrightarrow \quad A - \lambda \cdot I \ \text{ n'est pas inversible} \\
&\Leftrightarrow \quad \det(A - \lambda \cdot I) = 0
\end{aligned}$$

La preuve que $\det(A - \lambda \cdot I)$ est un polynôme de degré n est laissée en exercice (suggestion : à faire par induction sur n en utilisant la formule [11.14] avec $j = 1$). □

14.14 Exercice. Montrer que si $A \in \mathbb{K}^{n\times n}$ est une matrice triangulaire, alors les valeurs propres de A sont exactement les éléments de la diagonale principale de A.

14.15 Exercice. Soit $A, P \in \mathbb{K}^{n\times n}$ avec P inversible, et soit $B = P^{-1} \cdot A \cdot P$. Montrer que

1. Les valeurs propres de B sont exactement les valeurs propres de A.
2. $\vec{x} \in \mathbb{K}^n$ est vecteur propre de B associé à la valeur propre λ ssi $P \cdot \vec{x}$ est vecteur propre de A associé à la valeur propre λ.

Preuve. 1. C'est une conséquence de la Propriété 14.13 :

$$\begin{aligned}\det(B - \lambda \cdot I) &= \det(P^{-1} \cdot A \cdot P - \lambda \cdot P^{-1} \cdot I \cdot P)\\ &= \det(P^{-1} \cdot (A - \lambda \cdot I) \cdot P)\\ &= \det P^{-1} \cdot \det(A - \lambda \cdot I) \cdot \det P\\ &= \det(A - \lambda \cdot I)\end{aligned}$$

2. Si $B \cdot \vec{x} = \lambda \cdot \vec{x}$, alors $A \cdot P \cdot \vec{x} = P \cdot B \cdot \vec{x} = P \cdot (\lambda \cdot \vec{x}) = \lambda \cdot P \cdot \vec{x}$.
Si $A \cdot P \cdot \vec{x} = \lambda \cdot P \cdot \vec{x}$, alors $B \cdot \vec{x} = P^{-1} \cdot A \cdot P \cdot \vec{x} = P^{-1} \cdot (\lambda \cdot P \cdot \vec{x}) = \lambda \cdot \vec{x}$. □

14.16 Remarque.

1. Au point 1 de l'Exercice 14.15 nous avons en effet montré que si $B = P^{-1} \cdot A \cdot P$, alors A et B ont le même polynôme caractéristique.
2. Une autre façon de justifier le premier point de l'Exercice 14.15 est de se rappeler que les valeurs propres d'une matrice A sont par définition les valeurs propres d'un opérateur linéaire dont A est une matrice associée (14.12.1). Or, si $B = P^{-1} \cdot A \cdot P$, les matrices A et B sont associées au même opérateur linéaire par rapport à deux bases différentes.

14.17 Exercice. Revenons à l'Exemple 14.4, mais cette fois-ci prenons comme donnée de départ la matrice

$$A = \begin{pmatrix} 0 & 5 \\ -10 & 15 \end{pmatrix}$$

et essayons de déterminer une matrice P inversible et une matrice D diagonale telles que $A = P \cdot D \cdot P^{-1}$.

- Le polynôme caractéristique est

$$\det(A - \lambda \cdot I) = \lambda^2 - 15\lambda + 50$$

dont les racines sont $\lambda_1 = 5$ et $\lambda_2 = 10$. La matrice D est donc

$$D = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix}$$

- Le système $A \cdot \vec{x} = 5 \cdot \vec{x}$ admet comme solutions les vecteurs du type (α, α) avec $\alpha \in \mathbb{R}$ arbitraire. On peut donc choisir comme première colonne de P le vecteur $(1, 1)$.

- Le système $A \cdot \vec{x} = 10 \cdot \vec{x}$ admet comme solutions les vecteurs du type $(\alpha, 2\alpha)$ avec $\alpha \in \mathbb{R}$ arbitraire. On peut donc choisir comme deuxième colonne de P le vecteur $(1, 2)$.
- Finalement, comme matrice P nous pouvons choisir

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

et on vérifie facilement que $A \cdot P = P \cdot D$. (On peut inverser l'ordre des valeurs propres dans la diagonale de D. Il faut alors inverser l'ordre des colonnes dans P.)

14.5 Critères de diagonalisabilité

Pour montrer qu'un opérateur linéaire est diagonalisable il faut trouver une base de vecteurs propres ; le prochain lemme simplifie le travail à faire :

14.18 Lemme. *Soit $L\colon V \to V$ un opérateur linéaire et soit $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ des valeurs propres distinctes de L. Soit $\vec{x}_1, \ldots, \vec{x}_r \in V$ des vecteurs propres, avec $\vec{x}_i$ associé à λ_i. La famille $\vec{x}_1, \ldots, \vec{x}_r$ est libre.*
Autrement dit : la somme

$$E(\lambda_1) + \ldots + E(\lambda_r)$$

est une somme directe.

Preuve. Par induction sur r :

1. Si $r = 1$ alors $\vec{x}_1$ est une famille libre car $\vec{x}_1 \neq \vec{0}$.
2. Supposons que $\vec{x}_1, \ldots, \vec{x}_{r-1}$ soit une famille libre et que

$$a_1 \cdot \vec{x}_1 + \ldots + a_r\vec{x}_r = \vec{0}$$

il faut montrer que $a_1 = \ldots = a_r = 0$. Examinons deux cas :
   - Si $a_r = 0$ alors $a_1 \cdot \vec{x}_1 + \ldots + a_{r-1} \cdot \vec{x}_{r-1} = \vec{0}$ et par hypothèse inductive $a_1 = \ldots = a_{r-1} = 0$.
   - Si $a_r \neq 0$ alors $\vec{x}_r = -\frac{a_1}{a_r} \cdot \vec{x}_1 - \ldots - \frac{a_{r-1}}{a_r} \cdot \vec{x}_{r-1}$ et donc

$$\begin{array}{rcl} L(\vec{x}_r) & = & -\frac{a_1}{a_r} \cdot L(\vec{x}_1) - \ldots - \frac{a_{r-1}}{a_r} \cdot L(\vec{x}_{r-1}) \\ & = & -\frac{a_1}{a_r}\lambda_1 \cdot \vec{x}_1 - \ldots - \frac{a_{r-1}}{a_r}\lambda_{r-1} \cdot \vec{x}_{r-1} \\ L(\vec{x}_r) & = & \lambda_r \cdot \vec{x}_r \\ & = & -\frac{a_1}{a_r}\lambda_r \cdot \vec{x}_1 - \ldots - \frac{a_{r-1}}{a_r}\lambda_r \cdot \vec{x}_{r-1} \end{array}$$

Puisque $\vec{x}_1, \ldots, \vec{x}_{r-1}$ est une famille libre, on en tire que

$$\frac{a_i}{a_r}\lambda_i = \frac{a_i}{a_r}\lambda_r \quad \text{pour tout } i = 1, \ldots, r-1$$

Cela donne $a_i = 0$ pour $i = 1, \ldots, r-1$ car $\lambda_i \neq \lambda_r$. Il reste donc $a_r \cdot \vec{x}_r = \vec{0}$, ce qui donne $\vec{x}_r = \vec{0}$ car $a_r \neq 0$. Cela est impossible car $\vec{x}_r$ est un vecteur propre.

□

14.19 Corollaire. *Soit V un espace vectoriel avec* $\dim V = n$ *et soit* $L \colon V \to V$ *un opérateur linéaire. Si L admet n valeurs propres distinctes, alors L est diagonalisable.*

Preuve. Il suffit de choisir un vecteur propre pour chaque valeur propre. Par le Lemme 14.18 nous avons ainsi n vecteurs propres linéairement indépendants et donc une base de V formée par des vecteurs propres de L □

14.20 Remarque. La condition du Corollaire 14.19 est une condition suffisante pour la diagonalisabilité, mais elle n'est pas nécessaire. Par exemple, l'opérateur linéaire identité $id_V \colon V \to V$ est diagonalisable, mais il admet $\lambda = 1$ comme unique valeur propre.

14.21 Propriété. *(Premier critère de diagonalisabilité) Soit* $L \colon V \to V$ *un opérateur linéaire avec* $\dim V = n$, *et soit* $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ *les valeurs propres distinctes de L. Sont équivalentes :*

1. *L est diagonalisable*
2. $V = E(\lambda_1) \oplus \ldots \oplus E(\lambda_r)$
3. $n = \dim E(\lambda_1) + \ldots + \dim E(\lambda_r)$

Preuve. $1 \Leftrightarrow 2$: L est diagonalisable ssi tout vecteur de V est combinaison linéaire de vecteurs propres de L ssi $V = E(\lambda_1) + \ldots + E(\lambda_r)$, et on sait déjà que cette somme est directe (Lemme 14.18).
$2 \Leftrightarrow 3$: Puisque chaque $E(\lambda_i)$ est un s.e.v. de V, on a :

$$\begin{aligned} V = E(\lambda_1) \oplus \ldots \oplus E(\lambda_r) \quad &\Leftrightarrow \quad \dim V = \dim(E(\lambda_1) \oplus \ldots \oplus E(\lambda_r)) \\ &\Leftrightarrow \quad n = \dim E(\lambda_1) + \ldots + \dim E(\lambda_r) \end{aligned}$$

□

14.22 Remarque. Dans la Propriété 14.21, l'équivelance entre les conditions 1. et 2. reste valable même si V n'est pas finiment engendré.

14.23 Remarque. Soit $L \colon V \to V$ un opérateur linéaire avec $\dim V = n$, et soit $\lambda \in \mathbb{K}$ une valeur propre de L.

1. La *multiplicité géométrique* de λ est la dimension de $E(\lambda)$. Puisque $E(\lambda) = \operatorname{Ker}(L - \lambda \cdot id_V)$, d'après le Théorème du rang nous avons

$$\dim E(\lambda) = n - \dim \operatorname{Im}(L - \lambda \cdot id_V) = n - \operatorname{rang}(A - \lambda \cdot I)$$

   où A est la matrice associée à L par rapport à une base quelconque de V.

2. La *multiplicité algébrique* $m_a(\lambda)$ de λ est sa multiplicité en tant que racine du polynôme caractéristique $\det(A - \lambda \cdot I)$. On a :

$$1 \leq \dim E(\lambda) \leq m_a(\lambda) \leq n$$

Preuve. Montrons que $\dim E(\lambda) \leq m_a(\lambda)$: Soit $\vec{e}_1, \ldots, \vec{e}_k$ une base de $E(\lambda)$ et $\vec{e}_1, \ldots, \vec{e}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n$ une base de V. Alors $L(\vec{e}_i) = \lambda \cdot \vec{e}_i$ pour tout $i = 1, \ldots, k$ et donc

$$A = {}_eL_e = \begin{pmatrix} M & B \\ 0 & C \end{pmatrix}$$

avec $M \in \mathbb{K}^{k\times k}$ la matrice diagonale avec tous les éléments de la diagonale egaux à λ, $0 \in \mathbb{K}^{(n-k)\times k}$ la matrice nulle, $B \in \mathbb{K}^{k\times(n-k)}$ et $C \in \mathbb{K}^{(n-k)\times(n-k)}$. Par conséquent, le polynôme caractéristique est

$$\det(A - \alpha \cdot I) = (\lambda - \alpha)^k \det(C - \alpha \cdot I)$$

d'où $\dim E(\lambda) = k \leq m_a(\lambda)$. □

14.24 Propriété. *(Deuxième critère de diagonalisabilité) Soit $L\colon V \to V$ un opérateur linéaire avec $\dim V = n$, et soit $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ les valeurs propres distinctes de L. Sont équivalentes :*

1. *L est diagonalisable*
2. *$n = m_a(\lambda_1) + \ldots + m_a(\lambda_r)$ et $\dim E(\lambda_i) = m_a(\lambda_i)$ pour tout $i = 1, \ldots, r$.*

Preuve. $2 \Rightarrow 1$: Du point 2 on tire que $\dim E(\lambda_1) + \ldots + \dim E(\lambda_r) = n$, et donc L est diagonalisable grâce au premier critère (Propriété 14.21).
$1 \Rightarrow 2$: Si L est diagonalisable, par 14.21 et 14.23 on a

$$n = \dim E(\lambda_1) + \ldots + \dim E(\lambda_r) \leq m_a(\lambda_1) + \ldots + m_a(\lambda_r) \leq n$$

(la dernière inégalité est valable car les λ_i sont les racines d'un polynôme de degré n, voir 14.13). Cela donne

$$m_a(\lambda_1) + \ldots + m_a(\lambda_r) = n$$

et aussi

$$\dim E(\lambda_1) + \ldots + \dim E(\lambda_r) = m_a(\lambda_1) + \ldots + m_a(\lambda_r)$$

Cette dernière égalité avec la condition $\dim E(\lambda_i) \leq m_a(\lambda_i)$ donne $\dim E(\lambda_i) = m_a(\lambda_i)$ pour tout $i = 1, \ldots, r$. □

14.6 Applications

14.25 Remarque. Avant de revenir aux applications, voyons en synthèse comment les outils introduits dans ce chapitre permettent d'estimer l'évolution d'une population.
Premier cas : soit $L\colon V \to V$ un opérateur linéaire. Si un vecteur $\vec{x} \in V$ satisfait à la condition $L(\vec{x}) = \lambda \cdot \vec{x}$ (c'est-à-dire, si $\vec{x}$ est soit le vecteur nul, soit un vecteur propre associé à une valeur propre λ), alors pour tout $k \in \mathbb{N}$ nous avons

$$L^k(\vec{x}) = \lambda^k \cdot \vec{x}$$

(voir Exercice 14.6) et ceci sans hypothèse sur l'opérateur L. On peut donc estimer facilement l'évolution de la population, mais seulement si la population initiale est un vecteur propre.
Deuxième cas : soit $L\colon V \to V$ un opérateur linéaire diagonalisable sur un espace de dimension finie. Il existe donc une base $f = \{\vec{f_1}, \dots \vec{f_n}\}$ de V formée par des vecteurs propres de L, avec valeurs propres respectives $\lambda_1, \dots, \lambda_n$. Considérons un vecteur $\vec{x} \in V$ avec sa décomposition $\vec{x} = x_1 \cdot \vec{f_1} + \dots + x_n \cdot \vec{f_n}$. Pour tout $k \in \mathbb{N}$ nous avons :

$$L^k(\vec{x}) = x_1\lambda_1^k \cdot \vec{f_1} + \dots + x_n\lambda_n^k \cdot \vec{f_n}$$

Dans le cas d'un opérateur diagonalisable nous pouvons donc estimer facilement l'évolution de la population, et ceci quelle que soit la population initiale.

14.26 Exercice. Soit E un espace euclidien et soit V un s.e.v. de E. Si la projection orthogonale $P_V\colon E \to E$ de E sur V existe, alors elle est un opérateur linéaire diagonalisable.
En effet, on sait que $E = V \oplus V^\perp$ (voir 13.14). De plus :

$$V = \{\vec{x} \in E \mid P_V(\vec{x}) = \vec{x}\} = E(1)$$

$$V^\perp = \operatorname{Ker} P_V = \{\vec{x} \in E \mid P_V(\vec{x}) = \vec{0}\,\} = E(0)$$

et donc tout vecteur de E est combinaison linéaire de vecteurs propres de P_V.

14.27 Exercice. Soit $\mathbb{K}$ un corps commutatif, $A \in \mathbb{K}^{n\times n}$ une matrice inversible et diagonalisable, et $\vec{b} \in \mathbb{K}^n$ un vecteur fixé. On peut exprimer l'unique solution du système

$$S\colon A \cdot \vec{x} = \vec{b}$$

en utilisant les valeurs propres et les vecteurs propres de A.
En effet, puisque A est diagonalisable, $\mathbb{K}^n$ admet une base $f = \{\vec{f_1}, \dots, \vec{f_n}\}$ formée par des vecteurs propres de A. Soit $\lambda_1, \dots, \lambda_n$ les valeurs propres associées à $\vec{f_1}, \dots, \vec{f_n}$:

$$A \cdot \vec{f_i} = \lambda_i \vec{f_i}\,, \quad i = 1, \dots, n$$

Puisque A est inversible, toutes les valeurs propres sont différentes de zéro. En multipliant par A^{-1} et par λ_i^{-1}, la condition précédente donne aussi

$$\lambda_i^{-1}\vec{f_i} = A^{-1} \cdot \vec{f_i}$$

Soit

$$\vec{b} = b_1\vec{f_1} + \dots + b_n\vec{f_n}$$

la décomposition de $\vec{b}$ par rapport à la base f. Si $\vec{x}$ est l'unique solution du système S, nous avons

$$\begin{aligned}
\vec{x} &= A^{-1} \cdot A \cdot \vec{x} \\
&= A^{-1} \cdot \vec{b} \\
&= A^{-1} \cdot (b_1\vec{f_1} + \dots + b_n\vec{f_n}) \\
&= b_1 A^{-1} \cdot \vec{f_1} + \dots + b_n A^{-1} \cdot \vec{f_n} \\
&= b_1\lambda_1^{-1}\vec{f_1} + \dots + b_n\lambda_n^{-1}\vec{f_n}
\end{aligned}$$

14.28 Exercice. Pour terminer ce chapitre, reprenons l'Exemple 14.1 sur l'évolution d'une population.

1. Supposons que la population P soit divisée en trois tranches d'age
$$P = (p_1, p_2, p_3)$$
avec taux de reproduction et de survie donnés par
$$r_1 = 0\ ,\ \ r_2 = 1\ ,\ \ r_3 = 3\ ,\ \ s_1 = \frac{1}{2}\ ,\ \ s_2 = \frac{1}{3}\ ,\ \ s_3 = 0$$
Le système qui donne la population après k années est
$$P_k = A^k \cdot P_0$$
où P_0 est la population initiale et
$$A = \begin{pmatrix} 0 & 1 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$$
Le polynôme caractéristique est $-\lambda^3 + \frac{1}{2}\lambda + \frac{1}{2}$. Par conséquent, $\lambda = 1$ est valeur propre et $(6, 3, 1) \in E(1)$. Une population initiale avec 60 % des individus dans la première tranche d'age, 30 % dans la deuxième tranche et 10 % dans la troisième tranche restera stable dans le temps.
2. Supposons mantenant que la population P soit divisée en deux tranches d'age
$$P = (p_1, p_2)$$
et que les taux de reproduction et de survie, combinés à des phénomènes migratoires, donnent un système
$$P_k = A^k \cdot P_0$$
où P_0 est la population initiale et
$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$
Le polynôme caractéristique est $\lambda^2 - 6\lambda + 8$. On trouve donc que $\lambda = 4$ est valeur propre et que $(1, 1) \in E(4)$. Une population initiale P_0 ayant 50 % d'individus dans la prmeière tranche d'age et 50 % d'individus dans la deuxième tranche est destinée à s'accroitre très rapidement car (voir Exercice 14.6)
$$P_k = A^k \cdot P_0 = 4^k \cdot P_0$$

14.29 Exercice. Supposons qu'une population animale P soit divisée en trois tranches d'age
$$P = (p_1, p_2, p_3)$$

et que les taux de reproduction et de survie, combinés à des phénomènes migratoires, donnent un système

$$P_k = A^k \cdot P_0$$

où P_0 est la population initiale et

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

Montrer que, quel que soit la population initiale P_0, la population P est destinée à s'éteindre.

Chapitre 15

Triangularisabilité

Le chapitre 15 section par section

Pour résoudre certains systèmes d'équations différentielles linéaires, il faut parfois reconduire la matrice des coefficients du système à une matrice triangulaire. Pour cette raison, voyons dans ce chapitre une condition nécessaire et suffisante pour la triangularisabilité (ou trigonalisabilité) d'une matrice carrée.

1. Pour une matrice carrée, la condition de triangularisabilité est plus faible que la condition de diagonalisabilité, mais elle permet d'établir un lien intéressant entre valeurs propres, déterminant et trace.
2. Une autre aide pour calculer les valeurs propres est donnée par le théorème de Cayley-Hamilton et son réciproque.

15.1 Matrices triangularisables

15.1 Définition.

1. Une matrice $T = (t_{ij}) \in \mathbb{K}^{n\times n}$ est *triangulaire (supérieure)* si $t_{ij} = 0$ pour tout $i > j$.
2. Une matrice $A \in \mathbb{K}^{n\times n}$ est *triangularisable* s'il existe une matrice $P \in \mathbb{K}^{n\times n}$ inversible et une matrice $T \in \mathbb{K}^{n\times n}$ triangulaire telles que $A = P \cdot T \cdot P^{-1}$.

15.2 Exemple.

1. La matrice
$$T = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$
est triangulaire.
2. Toute matrice diagonale est triangulaire. Toute matrice diagonalisable est triangularisable.

15.3 Propriété. *Soit $A \in \mathbb{K}^{n\times n}$ une matrice carrée et $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ les valeurs propres distinctes de A. La matrice A est triangularisable si et seulement si $m_a(\lambda_1) + \ldots + m_a(\lambda_r) = n$.*

Preuve. $\Rightarrow$: Si $T = P^{-1} \cdot A \cdot P$, alors T et A ont les mêmes valeurs propres (Exercice 14.15), et si T est triangulaire alors les valeurs prores de T sont exactement les éléments de la diagonale de T (Exercice 14.14). Cela montre que A admet n valeurs propres (non nécessairement distinctes).
$\Leftarrow$: Cette implication se démontre par induction sur n. Elle ne sera pas abordée dans ce cours. □

15.4 Corollaire. *Toute matrice $A \in \mathbb{C}^{n\times n}$ est triangularisable.*

Preuve. Le polynôme caractéristique de A est un polynôme à coefficients complexes de degré n. En suivant le Théorème fondamental de l'algèbre (voir 12.9), un tel polynôme admet n racines complexes (non nécessairement distinctes). La condition de la Propriété 15.3 s'applique et A est donc triangularisable. □

La notion de matrice triangularisable nous permet d'établir des liens entre valeurs propres, trace et déterminant d'une matrice. Ces liens sont souvent utiles pour calculer les valeurs propres d'une matrice.

15.5 Définition. Soit $A \in \mathbb{K}^{n\times n}$ une matrice carrée. La trace de A est la somme des éléments sur la diagonale principale de A:

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$$

15.6 Exercice. Soit $A, B \in \mathbb{K}^{n\times n}$. Montrer que $\mathrm{tr}(A \cdot B) = \mathrm{tr}(B \cdot A)$.

15.7 Propriété. *Soit $A \in \mathbb{K}^{n\times n}$ une matrice triangularisable et soit $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ les valeurs propres (non nécessairement distinctes) de A*

1. $\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$
2. $\mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n$

Preuve. Si $A = P \cdot T \cdot P^{-1}$, alors A et T ont mêmes valeurs propres (14.15), même déterminant (11.17) et même trace (15.6). Il suffit donc de démontrer les deux énoncés pour une matrice triangulaire T. Or, si $T = (t_{ij})$ est triangulaire, alors $T - \lambda I$ l'est aussi et, par 11.16, nous avons

$$\det T = t_{11} \cdot t_{22} \cdot \ldots \cdot t_{nn} \quad \text{et} \quad \det(T - \lambda I) = (t_{11} - \lambda) \cdot (t_{22} - \lambda) \cdot \ldots \cdot (t_{nn} - \lambda)$$

La deuxième relation montre que les valeurs propres $\{\lambda_1, \ldots, \lambda_n\}$ de T sont exactement les éléments $\{t_{11}, \ldots, t_{nn}\}$ de la diagonale principale. Les conditions $\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$ et $\mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n$ sont maintenant évidentes. □

15.2 Théorème de Cayley-Hamilton

La notion de matrice triangularisable nous permet aussi d'énoncer le Théorème de Cayley-Hamilton.

15.8 Notation. Si $f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m \in \mathbb{K}[x]$ et si $A \in \mathbb{K}^{n\times n}$, on pose $f(A) = a_0I + a_IA + a_2A^2 + \ldots + a_mA^m \in \mathbb{K}^{n\times n}$.

15.9 Lemme. *Soit $f \in \mathbb{K}[x]$.*

1. *Si $A, P \in \mathbb{K}^{n\times n}$ avec P inversible, alors $f(P^{-1} \cdot A \cdot P) = P^{-1} \cdot f(A) \cdot P$.*
2. *Si* $A = \begin{pmatrix} a_{11} & ? & \ldots & ? \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$ *alors* $f(A) = \begin{pmatrix} f(a_{11}) & ? & \ldots & ? \\ 0 & & & \\ \vdots & & f(B) & \\ 0 & & & \end{pmatrix}$

Preuve.

1. Il suffit d'expliciter $f(P^{-1} \cdot A \cdot P)$ et de mettre en évidence P^{-1} et P en utilisant que $I = P^{-1} \cdot I \cdot P$.
2. Exercice [suggestion : démontrer d'abord le cas particulier $f(x) = x^m$].

15.10 Propriété. *(Théorème de Cayley-Hamilton) Soit $A \in \mathbb{K}^{n\times n}$ une matrice triangularisable et soit $P_A \in \mathbb{K}[x]_n$ son polynôme caractéristique. On a : $P_A(A) = 0$.*

Preuve. Par induction sur n.
1) $n = 1$: Si $A = (a_{11})$ alors $P_A(x) = a_{11} - x$ et donc $P_A(A) = a_{11} - a_{11} = 0$.
2) Supposons que la propriété soit valable pour toute matrice à n-1 lignes et n-1 colonnes.
3) Soit $A \in \mathbb{K}^{n\times n}$ et soit $T = P^{-1} \cdot A \cdot P$ une matrice triangulaire. En utilisant 14.16 et 15.9.1 nous avons $P_A(A) = P_T(A) = P \cdot P_T(T) \cdot P^{-1}$. Il suffit donc de montrer que $P_T(T) = 0$ pour T une matrice triangulaire. Pour cela : si

$$T = \begin{pmatrix} t & ? & \ldots & ? \\ 0 & & & \\ \vdots & & T' & \\ 0 & & & \end{pmatrix}$$

avec $T' \in \mathbb{K}^{(n-1)\times(n-1)}$ triangulaire, alors

$$P_T(x) = \det(T - xI) = \det \begin{pmatrix} t - x & ? & \ldots & ? \\ 0 & & & \\ \vdots & & T' - xI & \\ 0 & & & \end{pmatrix} =$$

$$= (t - x)\det(T' - xI) = (t - \lambda)P_{T'}(x)$$

d'où, en utilisant 15.9.2 et l'hypothèse inductive, nous avons

$$P_T(T) = (tI - T)P_{T'}(T) =$$

$$= \begin{pmatrix} 0 & ? & \dots & ? \\ 0 & & & \\ \vdots & & tI - T' & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} P_{T'}(t) & ? & \dots & ? \\ 0 & & & \\ \vdots & & P_{T'}(T') & \\ 0 & & & \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & ? & \dots & ? \\ 0 & & & \\ \vdots & & tI - T' & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} P_{T'}(t) & ? & \dots & ? \\ 0 & & & \\ \vdots & & O & \\ 0 & & & \end{pmatrix} = 0$$

□

Le théorème de Cayley-Hamilton admet une sorte de réciproque :

15.11 Propriété. *Soit $A \in \mathbb{K}^{n\times n}$ une matrice carrée et soit $\lambda \in \mathbb{K}$ une valeur propre de A. Si $f \in \mathbb{K}[x]$ est tel que $f(A) = 0$, alors $f(\lambda) = 0$.*

Preuve. Soit $\vec{x}$ un vecteur propre de A associé à λ. Puisque $A \cdot \vec{x} = \lambda \cdot \vec{x}$, nous avons (par induction) que pour tout $m \in \mathbb{N}$

$$A^m \cdot \vec{x} = \lambda^m \cdot \vec{x}$$

Soit maintenant $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$ un polynôme tel que $f(A) = 0$. Nous avons :

$$\begin{aligned} \vec{0} = 0 \cdot \vec{x} = f(A) \cdot \vec{x} &= \\ = (a_0I + a_1A + a_2A^2 + \dots + a_mA^m) \cdot \vec{x} &= \\ = a_0\vec{x} + a_1A\vec{x} + a_2A^2\vec{x} + \dots + a_mA^m\vec{x} &= \\ = a_0\vec{x} + a_1\lambda\vec{x} + a_2\lambda^2\vec{x} + \dots + a_m\lambda^m\vec{x} &= \\ = (a_0 + a_1\lambda + a_2\lambda^2 + \dots + a_m\lambda^m) \cdot \vec{x} &= \\ = f(\lambda) \cdot \vec{x} & \end{aligned}$$

d'où $0 = f(\lambda)$ car $\vec{x} \neq \vec{0}$. □

Chapitre 16

Application adjointe

Le chapitre 16 section par section

L'objectif de ce chapitre est d'établir le théorème spectral pour les matrices réelles symétriques. Ce théorème sera employé dans la suite pour étudier le caractère d'une forme quadratique.

1. Dans le passage entre matrices et applications linéaires, la matrice transposée correspond à l'application adjointe.
2. Les opérateurs qui sont représentés par une matrice symétrique sont les opérateurs auto-adjoints. Leurs valeurs propres sont toutes réelles.
3. Dans sa versione matricielle, le théorème spectral assure qu'une matrice réelle est symétrique ssi elle est diagonalisable par une matrice orthogonale.

16.1 Application adjointe et matrice transposée

16.1 Définition. Soit $L\colon E \to F$ une application linéaire entre espaces euclidiens. L'*application adjointe* est une application linéaire

$$L^*\colon F \to E$$

qui remplit la condition suivante : pour tout $\vec{x} \in E$ et pour tout $\vec{y} \in F$, on a

$$(L(\vec{x}) \mid \vec{y})_F = (\vec{x} \mid L^*(\vec{y}))_E$$

16.2 Propriété.

1. *L'application adjointe, si elle existe, est unique.*
2. $Id^* = Id\ ,\ \ (L^*)^* = L\ ,\ \ (T \circ L)^* = L^* \circ T^*$

Preuve. 1. Si $R\colon F \to E$ est une deuxième application linéaire telle que pour tout $\vec{x} \in E$ et pour tout $\vec{y} \in F$ on a

$$(L(\vec{x}) \mid \vec{y})_F = (\vec{x} \mid R(\vec{y}))_E$$

alors $(\vec{x} \mid L^*(\vec{z}))_E = (\vec{x} \mid R(\vec{z}))_E$ et donc $(\vec{x} \mid L^*(\vec{z}) - R(\vec{z}))_E = 0$. Puisque cela est valable pour tout $\vec{x} \in E$, nous avons que $L^*(\vec{z}) - R(\vec{z}) \in E^\perp = \{\vec{0}\}$ (voir 13.9) et donc $L^*(\vec{z}) = R(\vec{z})$.
2. Voyons par exemple que $(T \circ L)^* = L^* \circ T^*$ pour $L\colon E \to F$ et $T\colon F \to G$. $(T \circ L)^*\colon G \to E$ est l'unique application linéaire telle que

$$((T \circ L)(\vec{x}) \mid \vec{z})_G = (\vec{x} \mid (T \circ L)^*(\vec{z}))_E$$

pour tout $\vec{x} \in E$ et pour tout $\vec{z} \in G$. Par définition de L^* et T^* nous avons

$$((T \circ L)(\vec{x}) \mid \vec{z})_G = (T(L(\vec{x})) \mid \vec{z})_G = (L(\vec{x}) \mid T^*(\vec{z}))_F =$$

$$= (\vec{x} \mid L^*(T^*(\vec{z}))_E = (\vec{x} \mid (L^* \circ T^*)(\vec{z}))_E$$

d'où, par unicité de l'application adjointe, $(T \circ L)^* = L^* \circ T^*$. □

16.3 Exemple. Si dans la Définition 16.1 nous prenons $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$ munis de leur produit scalaire canonique, ainsi que $L = L_A\colon \mathbb{R}^n \to \mathbb{R}^m$, l'application linéaire associée à une matrice $A \in \mathbb{R}^{m\times n}$, nous avons $(L_A)^* = L_{A^t}\colon \mathbb{R}^m \to \mathbb{R}^n$, l'application linéaire associée à la matrice transposée. En effet,

$$(\vec{x} \mid L_{A^t}(\vec{y})) = (\vec{x} \mid A^t \cdot \vec{y}) = \vec{x}^t \cdot A^t \cdot \vec{y} = (A \cdot \vec{x})^t \cdot \vec{y} = (A \cdot \vec{x} \mid \vec{y}) = (L_A(\vec{x}) \mid \vec{y})$$

Dans la prochaine propriété nous établissons le lien entre application adjointe et matrice transposée, en généralisant l'Exemple 16.3.

16.4 Propriété. *Si $L\colon E \to F$ est une application linéaire entre espaces euclidiens de dimension finie, l'application adjointe $L^*\colon F \to E$ existe :*
si e est une base orthonormée de E et f est une base orthonormée de F, alors l'application L^* est donnée par

$${}_eL^*(\vec{y}) = ({}_fL_e)^t \cdot {}_f\vec{y}$$

Autrement dit

$${}_e(L^*)_f = ({}_fL_e)^t$$

Preuve. Il faut montrer, en utilisant la condition ${}_eL^*(\vec{y}) = ({}_fL_e)^t \cdot {}_f\vec{y}$, que

$$(L(\vec{x}) \mid \vec{y})_F = (\vec{x} \mid L^*(\vec{y}))_E$$

Puisque e et f sont des bases orthonormées, nous pouvons appliquer la Remarque 13.18.2 pour calculer le produit scalaire dans E et dans F. Nous obtenons ainsi

$$(\vec{x} \mid L^*(\vec{y}))_E = {}_e\vec{x}^t \cdot {}_eL^*(\vec{y}) = {}_e\vec{x}^t \cdot ({}_fL_e)^t \cdot {}_f\vec{y} =$$

$$= ({}_fL_e \cdot {}_e\vec{x})^t \cdot {}_f\vec{y} = {}_fL(\vec{x})^t \cdot {}_f\vec{y} = (L(\vec{x}) \mid \vec{y})_F$$

□

16.2 Opérateurs auto-adjoints et matrices symétriques

16.5 Définition. Soit E un espace euclidien et $L\colon E \to E$ un opérateur linéaire. L'opérateur L est *auto-adjoint* si $L^* = L$, c'est-à-dire si pour tout $\vec{x}, \vec{y} \in E$ on a

$$(L(\vec{x}) \mid \vec{y})_E = (\vec{x} \mid L(\vec{y}))_E$$

16.6 Corollaire. *Soit E un espace euclidien de dimension n et e une base orthonormée de E. Un opérateur linéaire $L\colon E \to E$ est auto-adjoint si et seulement si la matrice ${}_eL_e \in \mathbb{R}^{n\times n}$ est symétrique.*

Preuve. L est auto-adjoint ssi ${}_eL_e = {}_e(L^*)_e$ ssi ${}_eL_e = ({}_eL_e)^t$ ssi ${}_eL_e$ est symétrique. □

Dans la suite de ce chapitre nous démontrons le théorème spectral pour les opérateurs linéaires (16.10) et pour les matrices réelles 16.13). Ce dernier sera utilisé au Chapitre 17 pour déterminer le caractère d'une forme quadratique. Pour préparer le théorème spectral il nous faut un rappel sur le corps complexe.

16.7 Remarque.

1. Soit $z = (a, b) = a + ib \in \mathbb{C}$ un nombre complexe. Son *conjugué* est le nombre
$$\overline{z} = (a, -b) = a - ib \in \mathbb{C}$$
c'est-à-dire le symétrique de z par rapport à l'axe des réels. Nous avons :
   (a) $\overline{z} \cdot z = a^2 + b^2 = \mid z \mid^2$, et donc $\overline{z} \cdot z > 0$ si $z \neq 0$
   (b) $z = \overline{z}$ ssi $z \in \mathbb{R}$
   (c) $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$ et $\overline{z_1 \cdot z_2} = \overline{z}_1 \cdot \overline{z}_2$
2. On peut prolonger l'opération de conjugaison aux matrices complexes :
$$\text{si } A = (a_{ij}) \in \mathbb{C}^{m\times n}, \quad \text{on pose } \overline{A} = (\overline{a_{ij}}) \in \mathbb{C}^{m\times n}$$
et on a
$$\overline{A+B} = \overline{A} + \overline{B}\ ,\ \ \overline{A \cdot B} = \overline{A} \cdot \overline{B}\ ,\ \ \overline{A} = A \ \text{ ssi } \ A \in \mathbb{R}^{m\times n}$$

Soit maintenant $A \in \mathbb{R}^{n\times n}$ et soit $P_A \in \mathbb{R}[x]_n$ son polynôme caractéristique (14.13). Par le Théorème fondamental de l'algèbre (12.9), P_A admet n racines complexes $\lambda_1, \ldots, \lambda_n$ (non nécessairement distinctes). Parmi ces racines, on peut avoir des racines réelles et des racines complexes non réelles, comme l'exemple suivant le montre.

16.8 Exemple. Soit

$$A = \begin{pmatrix} 3 & 0 & -1 \\ 0 & -1 & 2 \\ 0 & -2 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

Son polynôme caractéristique est $P_A(\lambda) = (3-\lambda)\cdot(\lambda^2+3)$ et les valeurs propres sont donc

$$\lambda_1 = 3 \in \mathbb{R}\ , \quad \lambda_2 = i\sqrt{3} \in \mathbb{C}\setminus\mathbb{R}\ , \quad \lambda_3 = -i\sqrt{3} \in \mathbb{C}\setminus\mathbb{R}$$

Le problème mis en évidence par l'Exemple 16.8 ne peut pas se présenter pour une matrice symétrique, car nous avons le lemme suivant.

16.9 Lemme.

1. *Soit $A \in \mathbb{R}^{n\times n}$ et soit $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ les valeurs propres de A. Si A est symétrique, alors $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.*
2. *Soit $L\colon E \to E$ un opérateur linéaire sur un espace euclidien E de dimension n. Soit $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ les valeurs propres de L. Si L est auto-adjoint, alors $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.*

Preuve. 1. Soit $\lambda \in \mathbb{C}$ et $\vec{x}^t = (x_1, \ldots, x_n) \in \mathbb{C}^n$ tells que $A\cdot\vec{x} = \lambda\vec{x}$, avec $\vec{x} \neq \vec{0}$, et posons $\vec{y}^t = (\overline{x}_1, \ldots, \overline{x}_n)$. Puisque $\vec{y}^t\cdot A\cdot\vec{x} = \vec{y}^t\cdot\lambda\vec{x}$ et A est symétrique, nous avons

$$\vec{x}^t\cdot A\cdot\vec{y} = \vec{x}^t\cdot A^t\cdot\vec{y} = (\vec{y}^t\cdot A\cdot\vec{x})^t = (\vec{y}^t\cdot\lambda\vec{x})^t = \vec{x}^t\cdot\lambda\vec{y}$$

Puisque A est à coefficients réels, en passant aux conjugués cela donne

$$\vec{y}^t\cdot\overline{\lambda}\vec{x} = \overline{\vec{x}^t\cdot\lambda\vec{y}} = \overline{\vec{x}^t\cdot A\cdot\vec{y}} = \vec{y}^t\cdot A\cdot\vec{x} = \vec{y}^t\cdot\lambda\vec{x}$$

ce qui revient, en utilisant 16.7.1.(a), à

$$\overline{\lambda}(|\,x_1\,|^2+\ldots+|\,x_n\,|^2) = \lambda(|\,x_1\,|^2+\ldots+|\,x_n\,|^2)$$

Puisque $|\,x_1\,|^2+\ldots+|\,x_n\,|^2 > 0$ (car $\vec{x} \neq \vec{0}$) l'égalité précédente implique $\overline{\lambda} = \lambda$ et donc $\lambda \in \mathbb{R}$.
2. Soit e une base othonormée de E. Puisque L est auto-adjoint, la matrice ${}_eL_e$ est symétrique (16.6) et on peut appliquer le point 1. □

16.3 Théorème spectral

16.10 Théorème. *(Théorème Spectral pour les opérateurs linéaires) Soit E un espace euclidien de dimension finie et soit $L\colon E \to E$ un opérateur linéaire. Les conditions suivantes sont équivalentes :*

1. *E admet une base orthonormée formée par des vecteurs propres de L.*
2. *E admet une base orthonormée e telle que ${}_eL_e$ est diagonale.*
3. *L est auto-adjoint.*

Preuve. L'équivalence entre les conditions 1 et 2 est un cas particulier de la Propriété 14.11 (voir point 5 de la Remarque 14.12).
$2 \Rightarrow 3$: La matrice ${}_eL_e$ est diagonale et donc symétrique. Par le Corollaire 16.6 l'opérateur L est auto-adjoint.

$3 \Rightarrow 1$: La preuve est par induction sur $n = \dim E$.
Si $n = 1$ l'énoncé est vrai car dans ce cas tout vecteur non nul de E est vecteur propre de L. Il suffit alors d'en normer un pour avoir une base orthonormée de E de vecteurs propres.
Supposons l'énoncé vrai pour $n - 1$ et démontrons-le pour n:
Par le Lemme 16.9, L admet au moins une valeur propre réelle λ_1, et donc au moins un vecteur propre $\vec{e}_1 \in E(\lambda_1)$ que nous pouvons supposer être normé. Soit $V = \langle \vec{e}_1 \rangle$. Montrons que $L(\vec{x}) \in V^\perp$ pour tout $\vec{x} \in V^\perp$: puisque $L = L^*$, nous avons

$$(\vec{e}_1 \mid L(\vec{x})) = (\vec{e}_1 \mid L^*(\vec{x})) = (L(\vec{e}_1) \mid \vec{x}) = (\lambda_1 \vec{e}_1 \mid \vec{x}) = 0$$

Par conséquent, nous pouvons considérer la restriction

$$L' \colon V^\perp \to V^\perp$$

de L à $V^\perp$. Or, L' est un opérateur auto-adjoint (car L est auto-adjoint) et $\dim V^\perp = n - 1$ (car $E = V \oplus V^\perp$, voir 13.23). Par hypothèse inductive, $V^\perp$ admet une base orthonormée $\{\vec{e}_2, \dots, \vec{e}_n\}$ de vecteurs propres de L'. Il est clair que $\vec{e}_2, \dots, \vec{e}_n$ sont aussi vecteurs propres pour L. Finalement, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ est une base orthonormée de E formée de vecteurs propres de L. □

Pour énoncer la formulation matricielle du théorème spectral nous avons besoin de la notion de matrice orthogonale.

16.11 Définition. Une matrice $Q \in \mathbb{R}^{n \times n}$ est *orthogonale* si $Q^t \cdot Q = I$.
Autrement dit, Q est orthogonale si ses colonnes forment une base de $\mathbb{R}^n$ orthonormée par rapport au produit scalaire canonique.

16.12 Remarque. Toute matrice orthogonale Q est inversible et $Q^{-1} = Q^t$.

16.13 Corollaire. *(Théorème Spectral pour les matrices réelles) Soit $A \in \mathbb{R}^{n \times n}$. Les conditions suivantes sont équivalentes :*

1. *$\mathbb{R}^n$ admet une base orthonormée (par rapport au produit scalaire canonique) formée par des vecteurs propres de A.*
2. *Il existe $Q, D \in \mathbb{R}^{n \times n}$, avec Q orthogonale et D diagonale, telles que*

$$A \cdot Q = Q \cdot D$$

3. *A est symétrique.*

Preuve. L'équivalence entre les conditions 1 et 2 est un cas particulier de la Propriété 14.11.
$2 \Rightarrow 3$: Si $A = Q \cdot D \cdot Q^{-1}$ avec Q orthogonale, alors $A = Q \cdot D \cdot Q^t$. Par conséquent :

$$A^t = (Q \cdot D \cdot Q^t)^t = (Q^t)^t \cdot D^t \cdot Q^t = Q \cdot D \cdot Q^t = A$$

et donc A est symétrique.
$3 \Rightarrow 1$: Considérons l'opérateur linéaire $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ et soit c la base canonique de $\mathbb{R}^n$ (qui est orthonormée par rapport au produit scalaire canonique, voir 13.17). Puisque ${}_c(L_A)_c = A$ est symétrique, l'opérateur L_A est auto-adjoint (16.6) et donc, par le Théorème 16.10, $\mathbb{R}^n$ admet une base orthonormée formée par des vecteurs propres de L_A, qui sont exactement les vecteurs propres de A (voir 14.12). $\square$

16.14 Remarque. Le théorème spectral ne s'applique qu'aux matrices réelles. Par exemple la matrice

$$A = \begin{pmatrix} 3 & i \\ i & 1 \end{pmatrix} \in \mathbb{C}^{n\times n}$$

est symétrique mais elle n'est pas diagonalisable par une matrice orthogonale. En effet, elle n'est pas diagonalisable du tout, car elle admet une seule valeur propre $\lambda = 2$ avec $m_a(\lambda) = 2$ et $\dim E(\lambda) = 1$.

Nous savons que les vecteurs propres relatifs à des valeurs propres différentes sont linéairement indépendants (voir Lemme 14.18). Dans le cas d'une matrice symétrique une propriété plus forte est valable : ils sont orthogonaux. Cette propriété aide pour construire la base orthonormée de vecteurs propres et la matrice orthogonale qui apparaissent dans le Corollaire 16.13.

16.15 Propriété. *Soit $L\colon E \to E$ un opérateur linéaire auto-adjoint sur un espace euclidien E et soient $\vec{x} \in E(\lambda_1), \vec{y} \in E(\lambda_2)$, avec $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \neq \lambda_2$. Alors $\vec{x}$ et $\vec{y}$ sont orthogonaux.*

Preuve. En utilisant la définition de L^* (16.1) et que $L = L^*$, nous avons

$$0 = (L(\vec{x}) \mid \vec{y}) - (\vec{x} \mid L^*(\vec{y})) = (L(\vec{x}) \mid \vec{y}) - (\vec{x} \mid L(\vec{y})) = (\lambda_1 \cdot \vec{x} \mid \vec{y}) - (\vec{x} \mid \lambda_2 \cdot \vec{y}) =$$

$$= \lambda_1 \cdot (\vec{x} \mid \vec{y}) - \lambda_2 \cdot (\vec{x} \mid \vec{y}) = (\lambda_1 - \lambda_2) \cdot (\vec{x} \mid \vec{y})$$

d'où $(\vec{x} \mid \vec{y}) = 0$ car $\lambda_1 \neq \lambda_2$. $\square$

16.16 Corollaire. *Soit $L\colon E \to E$ un opérateur linéaire auto-adjoint sur un espace euclidien E de dimension finie et soient $\lambda_1, \ldots, \lambda_r$ les valeurs propres de L (avec $\lambda_i \neq \lambda_j$ si $i \neq j$). Fixons un $i \in \{1, \ldots, r\}$. On a :*

$$(E(\lambda_1) \oplus \ldots \oplus E(\lambda_i))^\perp = E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r)$$

Preuve. En utilisant 14.21 et 13.23, nous avons respectivement

$$E = E(\lambda_1) \oplus \ldots \oplus E(\lambda_i) \oplus E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r)$$

$$E = E(\lambda_1) \oplus \ldots \oplus E(\lambda_i) \oplus (E(\lambda_1) \oplus \ldots \oplus E(\lambda_i))^\perp$$

Par conséquent,

$$\dim(E(\lambda_1) \oplus \ldots \oplus E(\lambda_i))^\perp = \dim(E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r))$$

Pour démontrer l'égalité

$$(E(\lambda_1) \oplus \ldots \oplus E(\lambda_i))^\perp = E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r)$$

il suffit donc de montrer l'inclusion

$$E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r) \subseteq (E(\lambda_1) \oplus \ldots \oplus E(\lambda_i))^\perp$$

Soient

$$\vec{x} = \vec{v}_1 + \ldots + \vec{v}_i \in E(\lambda_1) \oplus \ldots \oplus E(\lambda_i) \ , \ \vec{y} = \vec{v}_{i+1} + \ldots + \vec{v}_r \in E(\lambda_{i+1}) \oplus \ldots \oplus E(\lambda_r)$$

En utilisant la bilinéarité du produit scalaire, nous obtenons

$$(\vec{x} \mid \vec{y}) = \sum_{j=1,\ldots,i \ ; \ k=i+1,\ldots,r} (\vec{v}_j \mid \vec{v}_k) = 0$$

car chaque $\vec{v}_j$ est orthogonale à chaque $\vec{v}_k$ (voir 16.15). $\square$

Chapitre 17

Formes quadratiques

Le chapitre 17 section par section

Pour étudier les extrema d'une fonction en plusieurs variables ou pour classifier les quadriques réelles, on utilise les formes quadratiques sur un espace vectoriel réel.

1. Les formes quadratiques $V \to \mathbb{R}$ sont en bijection avec les formes bilinéaires symétriques $V \times V \to \mathbb{R}$ et ...
2. ... une fois une base de V fixée, avec les matrices symétriques.
3. Le caractère d'une forme quadratique q est défini moyennant le signe de $q(\vec{x})$ pour $\vec{x}$ qui varie dans V, mais ...
4. ... il est plus pratique de le déterminer en regardant le signe des valeurs propres de la matrice symétrique associée à la forme quadratique.
5. La loi d'inertie garantit que les indices de positivité et de négativité sont des invariants de la forme quadratique, c'est-à-dire qu'ils ne dépendent pas de la base et donc de la matrice utilisée pour les calculer.
6. Une autre méthode pour déterminer les indices d'une forme quadratique est la complétion des carrés.

17.1 Formes quadratiques et formes bilinéaires

Si $f\colon \mathbb{R} \to \mathbb{R}$ est une fonction deux fois dérivable, pour déterminer les extrema locaux on cherche les zéros de la dérivée première et ensuite on regarde le signe de la dérivée seconde pour distinguer les maxima et les minima :

- Si $f'(a) = 0$ et $f''(a) < 0$ alors a est un maximum local.
- Si $f'(a) = 0$ et $f''(a) > 0$ alors a est un minimum local.

Si on a une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ en plusieurs variables, la stratégie est la même, sauf qu'on aura n dérivées premières et n^2 dérivées secondes. Voyons un exemple : la fonction

$$f\colon \mathbb{R}^2 \to \mathbb{R}\,, \quad f(x,y) = e^x \sin y$$

a deux dérivées premières

$$D_x f(x,y) = e^x \sin y \ , \ \ D_y f(x,y) = e^x \cos y$$

et 4 dérivées secondes que nous pouvons mettre dans une matrice

$$H(f(x,y)) = \begin{pmatrix} D_x(D_x f(x,y)) = e^x \sin y & D_y(D_x f(x,y)) = e^x \cos y \\ D_x(D_y f(x,y)) = e^x \cos y & D_y(D_y f(x,y)) = -e^x \sin y \end{pmatrix}$$

C'est la *matrice Hessienne* de f. Si la fonction f est suffisamment régulière, la matrice Hessienne est symétrique. Il faut donc donner un sens à la notion de *signe* (ou *caractère*) d'une matrice symétrique. Pour cela nous utilisons les formes quadratiques (qui seront aussi utilisées dans le cours de géométrie pour classifier les coniques et les quadriques réelles).

17.1 Notation. Soit V un espace vectoriel réel.

1. Soit $q\colon V \to \mathbb{R}$ une fonction. On pose :

$$\bar{q}\colon V \times V \to \mathbb{R} \ , \ \ \bar{q}(\vec{u},\vec{v}) = \frac{1}{2}(q(\vec{u}+\vec{v}) - q(\vec{u}) - q(\vec{v}))$$

   Remarquons que $\bar{q}$ est symétrique : $\bar{q}(\vec{u},\vec{v}) = \bar{q}(\vec{v},\vec{u})$.

2. Soit $Q\colon V \times V \to \mathbb{R}$ une fonction. On pose :

$$\widetilde{Q}\colon V \to \mathbb{R} \ , \ \ \widetilde{Q}(\vec{u}) = Q(\vec{u},\vec{u})$$

17.2 Définition. Soit V un espace vectoriel réel. Une *forme quadratique* sur V est une fonction

$$q\colon V \to \mathbb{R}$$

telle que

1. $q(\lambda \cdot \vec{x}) = \lambda^2 q(\vec{x})$ pour tout $\vec{x} \in V$ et pour tout $\lambda \in \mathbb{R}$
2. $\bar{q}\colon V \times V \to \mathbb{R}$ est *bilinéaire* (c'est-à-dire linéaire par rapport à chaque variable).

Une forme quadratique $q\colon V \to \mathbb{R}$ est complètement déterminée par la forme bilinéaire symétrique $\bar{q}\colon V \times V \to \mathbb{R}$. En effet :

17.3 Propriété. *Soit V un espace vectoriel réel.*

1. *Si $q\colon V \to \mathbb{R}$ est une forme quadratique, alors $\bar{q}\colon V \times V \to \mathbb{R}$ est bilinéaire et symétrique et $\widetilde{\bar{q}} = q$.*
2. *Si $Q\colon V \times V \to \mathbb{R}$ est bilinéaire et symétrique, alors $\widetilde{Q}\colon V \to \mathbb{R}$ est une forme quadratique et $\overline{\widetilde{Q}} = Q$.*

Preuve. 1. Nous savons déjà que $\bar{q}$ est est bilinéaire et symétrique.

$$\begin{array}{rcl} \widetilde{\bar{q}}(\vec{x}) & = & \bar{q}(\vec{x},\vec{x}) \\ & = & \frac{1}{2}(q(\vec{x}+\vec{x}) - q(\vec{x}) - q(\vec{x})) \\ & = & \frac{1}{2}(q(2\vec{x}) - 2q(\vec{x})) \\ & = & \frac{1}{2}(4q(\vec{x}) - 2q(\vec{x})) \\ & = & q(\vec{x}) \end{array}$$

2. $\widetilde{Q}(\lambda\cdot\vec{x}) = Q(\lambda\cdot\vec{x},\lambda\cdot\vec{x}) = \lambda^2 Q(\vec{x},\vec{x}) = \lambda^2\widetilde{Q}(\vec{x})$

$$\begin{array}{rcl}
\overline{\widetilde{Q}}(\vec{x},\vec{y}) & = & \frac{1}{2}(\widetilde{Q}(\vec{x}+\vec{y}) - \widetilde{Q}(\vec{x}) - \widetilde{Q}(\vec{y})) \\
 & = & \frac{1}{2}(Q(\vec{x}+\vec{y},\vec{x}+\vec{y}) - Q(\vec{x},\vec{x}) - Q(\vec{y},\vec{y})) \\
 & = & \frac{1}{2}(Q(\vec{x},\vec{x}) + Q(\vec{x},\vec{y}) + Q(\vec{y},\vec{x}) + Q(\vec{y},\vec{y}) - Q(\vec{x},\vec{x}) - Q(\vec{y},\vec{y})) \\
 & = & \frac{1}{2}(Q(\vec{x},\vec{y}) + Q(\vec{y},\vec{x})) \\
 & = & Q(\vec{x},\vec{y})
\end{array}$$

□

17.4 Exemple. L'exemple le plus facile de forme quadratique est donné par la fonction

$$q\colon \mathbb{R}\to\mathbb{R}\ ,\ \ q(x) = x^2$$

La forme bilinéaire symétrique associée est simplement le produit de nombre réels : $\bar{q}(x,y) = \frac{1}{2}((x+y)^2 - x^2 - y^2) = \frac{1}{2}(x^2+2xy+y^2-x^2-y^2) = xy$.

17.2 Formes quadratiques et matrices symétriques

Voyons à présent le lien entre formes quadratiques et matrices symétriques :

17.5 Propriété. *Soit V un espace vectoriel réel et soit $e = \{\vec{e}_1,\ldots,\vec{e}_n\}$ une base de V.*

1. *Si $A\in\mathbb{R}^{n\times n}$ est une matrice symétrique, on obtient une forme quadratique*

$$q_{A,e}\colon V\to\mathbb{R}\ ,\ \ q_{A,e}(\vec{x}) = {}_e\vec{x}^t\cdot A\cdot{}_e\vec{x}$$

2. *Si $q\colon V\to\mathbb{R}$ est une forme quadratique, on obtient une matrice symétrique*

$$A_{q,e} = (\bar{q}(\vec{e}_i,\vec{e}_j)_{i=1,\ldots,n}^{j=1,\ldots,n} \in \mathbb{R}^{n\times n}$$

3. *Ces deux constructions établissent une bijection entre formes quadratiques sur V et matrices symétriques réelles à n lignes et n colonnes. Autrement dit :*
   (a) *Si $q\colon V\to\mathbb{R}$ est une forme quadratique, alors $A_{q,e}\in\mathbb{R}^{n\times n}$ est l'unique matrice symétrique telle que $q(\vec{x}) = {}_e\vec{x}^t\cdot A_{q,e}\cdot{}_e\vec{x}$ pour tout $\vec{x}\in V$.*
   (b) *Si $A\in\mathbb{R}^{n\times n}$ est une matrice symétrique, alors $q_{A,e}\colon V\to\mathbb{R}$ est l'unique forme quadratique sur V telle que $\bar{q}_{A,e}(\vec{x},\vec{y}) = {}_e\vec{x}^t\cdot A\cdot{}_e\vec{y}$ pour tout $\vec{x},\vec{y}\in V$.*

Preuve. 1. Commençons par montrer que si $A\in\mathbb{R}^{n\times n}$ est une matrice symétrique, alors $\bar{q}_{A,e}(\vec{x},\vec{y}) = {}_e\vec{x}^t\cdot A\cdot{}_e\vec{y}$:

$$\begin{array}{rcl}
\bar{q}_{A,e}(\vec{x},\vec{y}) & = & \frac{1}{2}(q_{A,e}(\vec{x}+\vec{y}) - q_{A,e}(\vec{x}) - q_{A,e}(\vec{y})) \\
 & = & \frac{1}{2}(\,{}_e(\vec{x}+\vec{y})^t\cdot A\cdot{}_e(\vec{x}+\vec{y}) - {}_e\vec{x}^t\cdot A\cdot{}_e\vec{x} - {}_e\vec{y}^t\cdot A\cdot{}_e\vec{y}) \\
 & = & \frac{1}{2}(\,{}_e\vec{x}^t\cdot A\cdot{}_e\vec{x} + {}_e\vec{x}^t\cdot A\cdot{}_e\vec{y} + {}_e\vec{y}^t\cdot A\cdot{}_e\vec{x} + {}_e\vec{y}^t\cdot A\cdot{}_e\vec{y} \\
 & & -\,{}_e\vec{x}^t\cdot A\cdot{}_e\vec{x} - {}_e\vec{y}^t\cdot A\cdot{}_e\vec{y}) \\
 & = & \frac{1}{2}(\,{}_e\vec{x}^t\cdot A\cdot{}_e\vec{y} + {}_e\vec{y}^t\cdot A\cdot{}_e\vec{x}) \\
 & = & {}_e\vec{x}^t\cdot A\,{}_e\vec{y}
\end{array}$$

où la dernière étape est justifiée par le fait que A est symétrique et donc

$$ {}_e\vec{y}^t \cdot A \cdot {}_e\vec{x} = ({}_e\vec{y}^t \cdot A \cdot {}_e\vec{x})^t = {}_e\vec{x}^t \cdot A^t \cdot {}_e\vec{y}^{tt} = {}_e\vec{x}^t \cdot A \cdot {}_e\vec{y} $$

Puisque le produit matriciel est distributif à gauche et à droite, on déduit de la formule précédente que $\bar{q}_{A,e}$ est bilinéaire. La condition $q_{A,e}(\lambda \cdot \vec{x}) = \lambda^2 q_{A,e}(\vec{x})$ est évidente et le point 1 est démontré.
2. Evident car $\bar{q}$ est symétrique et donc en particulier $\bar{q}(\vec{e}_i, \vec{e}_j) = \bar{q}(\vec{e}_j, \vec{e}_i)$.
3. Soit $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ symétrique et soit $A_{q,e} = (\bar{q}_{A,e}(\vec{e}_i, \vec{e}_j))$. Il faut montrer que $A = A_{q,e}$, c'est-à-dire que $a_{ij} = \bar{q}_{A,e}(\vec{e}_i, \vec{e}_j)$. En utilisant la formule établie au point 1, nous avons

$$ \bar{q}_{A,e}(\vec{e}_i, \vec{e}_j) = {}_e\vec{e}_i^{\,t} \cdot A \cdot {}_e\vec{e}_j = a_{ij} $$

où la dernière égalité est justifiée par le fait que ${}_e\vec{e}_i$ est le i-ème vecteur de la base canonique de $\mathbb{R}^n$, et donc ${}_e\vec{e}_i^{\,t} \cdot A \cdot {}_e\vec{e}_j$ est l'élément de place i, j de la matrice A.
Viceversa, soit $q\colon V \to \mathbb{R}$ une forme quadratique et soit $A = A_{q,e}$. Il faut montrer que $q_{A,e} = q$, et pour cela il suffit de montrer que $\overline{q}_{A,e} = \bar{q}$. Grâce à la bilinéarité, il suffit de montrer que $\bar{q}_{A,e}(\vec{e}_i, \vec{e}_j) = \bar{q}(\vec{e}_i, \vec{e}_j)$ pour tout $i, j = 1, \ldots, n$. En utilisant à nouveau la formule établie au point 1, nous avons

$$ \bar{q}_{A,e}(\vec{e}_i, \vec{e}_j) = {}_e\vec{e}_i^{\,t} \cdot A_{q,e} \cdot {}_e\vec{e}_j = \bar{q}(\vec{e}_i, \vec{e}_j) $$

où la dernière égalité est la définition même de l'élément de place i, j dans la matrice $A_{q,e}$. □

Il faut maintenant voir comment la matrice symétrique $A_{q,e}$ associée à une forme quadratique $q\colon V \to \mathbb{R}$ varie lorsque on change la base e de V.

17.6 Lemme. *Soit $A, P \in \mathbb{R}^{n\times n}$ deux matrices, avec A symétrique. Alors la matrice $P^t \cdot A \cdot P$ est symétrique.*

Preuve. Exercice □

17.7 Propriété. *Soit V un espace vectoriel réel, $q\colon V \to \mathbb{R}$ une forme quadratique et $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$, $f = \{\vec{f}_1, \ldots, \vec{f}_n\}$ deux bases de V. Le lien entre $A_{q,e}$ et $A_{q,f}$ est donné par*

$$ A_{q,f} = ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f $$

(où ${}_eI_f$ est la matrice changement de base, voir 5.16).

Preuve. Nous savons que ${}_e\vec{x} = {}_eI_f \cdot {}_f\vec{x}$ et donc que ${}_e\vec{x}^t = {}_f\vec{x}^t \cdot ({}_eI_f)^t$. En utilisant la Propriété 17.5 nous avons

$$ q(\vec{x}) = {}_e\vec{x}^t \cdot A_{q,e} \cdot {}_e\vec{x} = {}_f\vec{x}^t \cdot ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f \cdot {}_f\vec{x} $$

et donc par l'unicité (Propriété 17.5.3, que nous pouvons utiliser car par le Lemme 17.6 la matrice $({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f$ est symétrique)

$$ A_{q,f} = ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f $$ □

17.8 Exercice. Soit V un espace vectoriel réel, $q\colon V \to \mathbb{R}$ une forme quadratique et $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$, $f = \{\vec{f}_1, \ldots, \vec{f}_n\}$ deux bases de V. Le déterminant de $A_{q,e}$ et le déterminant de $A_{q,f}$ ont le même signe.

17.9 Remarque.

1. Si $A \in \mathbb{R}^{n\times n}$ est une matrice symétrique et $q = q_{A,c}\colon \mathbb{R}^n \to \mathbb{R}$ est la forme quadratique associée à A par rapport à la base canonique c de $\mathbb{R}^n$, alors

$$q\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

   est un polynôme homogène de degré 2 par rapport aux variables $x_1, \ldots, x_n$.

2. Par exemple si

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

   alors $q\colon \mathbb{R}^3 \to \mathbb{R}$ est donnée par

$$q\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 3x_1^2 + 3x_2^2 + 3x_3^2 - 2x_1x_3$$

3. Plus en général, si $q\colon V \to \mathbb{R}$ est une forme quadratique et $e = \{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de V, la formule

$$q(\vec{x}) = {}_e\vec{x}^t \cdot A_{q,e} \cdot {}_e\vec{x}$$

   montre que $q(\vec{x})$ est un polynôme homogène de degré 2 par rapport aux coordonnées de $\vec{x}$ dans la base e.

4. Retenons aussi de la preuve de la Propriété 17.5 que si $A \in \mathbb{R}^{n\times n}$ est une matrice symétrique et e est une base de $\mathbb{R}^n$, alors $\bar{q}_{A,e}(\vec{x}, \vec{y}) = {}_e\vec{x}^t \cdot A \cdot {}_e\vec{y}$.

17.3 Le caractère d'une forme quadratique

Nous arrivons à la définition de *caractère* d'une forme quadratique et d'une matrice symétrique.

17.10 Définition. Une forme quadratique $q\colon V \to \mathbb{R}$ est

1. définie positive si $q(\vec{x}) > 0$ pour tout $\vec{x} \in V, \vec{x} \neq \vec{0}$
2. semi-définie positive si $q(\vec{x}) \geq 0$ pour tout $\vec{x} \in V$
3. définie négative si $q(\vec{x}) < 0$ pour tout $\vec{x} \in V, \vec{x} \neq \vec{0}$
4. semi-définie négative si $q(\vec{x}) \leq 0$ pour tout $\vec{x} \in V$
5. indéfinie s'il existe un $\vec{x} \in V$ tel que $q(\vec{x}) > 0$ et un $\vec{y} \in V$ tel que $q(\vec{y}) < 0$.

17.11 Remarque. La bijection entre formes quadratiques sur V et fonctions bilinéaires symétriques sur $V \times V$ de la Propriété 17.3 se restreint donc à une bijection entre formes quadratiques définies positives et produits scalaires. En effet, puisque $q(\vec{x}) = \bar{q}(\vec{x}, \vec{x})$, une forme quadratique q est définie positive au sens de la Définition 17.10 si et seulement si la fonction bilinéaire symétrique $\bar{q}$ est définie positive au sens de la Définition 13.1.

17.12 Définition. Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et soit e une base de $\mathbb{R}^n$. Le caractère de A est le caractère de la forme quadratique

$$q_{A,e} \colon \mathbb{R}^n \to \mathbb{R}\ , \quad q_{A,e}(\vec{x}) = {}_e\vec{x}^t \cdot A \cdot {}_e\vec{x}$$

Il faut montrer que la Définition 17.12 est bien donnée, c'est-à-dire que le caractère d'une matrice symétrique A ne dépend pas de la base e choisie.

17.13 Lemme. *Soit $q \colon V \to \mathbb{R}$ et $q' \colon V' \to \mathbb{R}$ deux formes quadratiques et soit $L \colon V \to V'$ une application linéaire bijective telle que le diagramme suivant commute*

$$\begin{tikzcd} V \arrow[rr, "L"] \arrow[dr, "q"'] & & V' \arrow[dl, "q'"] \\ & \mathbb{R} & \end{tikzcd}$$

(ce qui revient à dire que $q' \circ L = q$). Alors q et q' ont même caractère.

Preuve. Montrons par exemple que q est définie positive si et seulement si q' est définie positive (la preuve pour les autres cas est similaire). Supposons d'abord que q' soit définie positive et soit $\vec{x} \in V, \vec{x} \neq \vec{0}$. Nous avons

$$q(\vec{x}) = q'(L(\vec{x})) > 0$$

car q' est définie positive et $L(\vec{x}) \neq \vec{0}$ (car L est injective). L'argument pour l'implication réciproque est le même, en utilisant l'application $L^{-1} \colon V' \to V$ à la place de L (notons que L^{-1} est telle que $q \circ L^{-1} = q'$). □

17.14 Corollaire. *Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique. Le caractère de A au sens de la Définition 17.12 ne dépend pas de la base de $\mathbb{R}^n$ choisie : si e, f sont deux bases de $\mathbb{R}^n$, alors*

$$q_{A,e} \colon \mathbb{R}^n \to \mathbb{R} \quad \text{et} \quad q_{A,f} \colon \mathbb{R}^n \to \mathbb{R}$$

ont même caractère.

Preuve. Soit $e = \{\vec{e}_1, \ldots, \vec{e}_n\}, f = \{\vec{f}_1, \ldots, \vec{f}_n\}$ deux bases de $\mathbb{R}^n$. Il suffit d'appliquer le Lemme 17.13 à la situation suivante

$$\begin{tikzcd} \mathbb{R}^n \arrow[rr, "L"] \arrow[dr, "q_{A,e}"'] & & \mathbb{R}^n \arrow[dl, "q_{A,f}"] \\ & \mathbb{R} & \end{tikzcd}$$

où L est l'unique application linéaire telle que $L(\vec{e}_i) = \vec{f}_i$ pour tout $i = 1, \ldots, n$ (voir 8.15). Puisque L est bijective (voir 8.14), il reste à montrer que le diagramme commute. Pour cela, il suffit d'observer que ${}_fL(\vec{x}) = {}_e\vec{x}$ (égalité qu'on peut montrer en prenant $\vec{x} = \vec{e}_i$ pour $i = 1, \ldots, n$) et donc

$$q_{A,e}(\vec{x}) = {}_e\vec{x}^t \cdot A \cdot {}_e\vec{x} = {}_fL(\vec{x})^t \cdot A \cdot {}_fL(\vec{x}) = q_{A,f}(L(\vec{x})) \qquad \square$$

Remarquons en passant qu'on peut exprimer l'équation qui conclu la preuve du Corollaire 17.14 par la commutativité du diagramme suivant :

$$\begin{tikzcd}
\mathbb{R}^n \arrow[rr, "L"] \arrow[dr, "E"] \arrow[ddr, bend right, "q_{A,e}"'] & & \mathbb{R}^n \arrow[dl, "F"'] \arrow[ddl, bend left, "q_{A,f}"] \\
& \mathbb{R}^n \arrow[d, "q_{A,\mathrm{b.c.}}"] & \\
& \mathbb{R} &
\end{tikzcd}$$

17.15 Remarque. Grâce au Corollaire 17.14, pour déterminer le caractère d'une matrice symétrique $A \in \mathbb{R}^{n \times n}$ nous pouvons choisir comme base e de $\mathbb{R}^n$ la base canonique. Dans ce cas, la forme quadratique associée à A devient simplement (voir 17.9)

$$q_{A,e} \colon \mathbb{R}^n \to \mathbb{R}\ , \ \ q_{A,e}(\vec{x}) = \vec{x}^t \cdot A \cdot \vec{x}$$

et la Définition 17.12 peut être écrite de façon plus explicite. Par exemple :

$$A \text{ est définie positive } \Leftrightarrow \ \ \vec{x}^t \cdot A \cdot \vec{x} > 0 \text{ pour tout } \vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$$

Si on compare ceci avec l'Exemple 13.3.2, on voit que la forme bilinéaire et symétrique

$$(- \mid -) \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\ , \ \ (\vec{x} \mid \vec{y}) = \vec{x}^t \cdot A \cdot \vec{y} = \bar{q}_{A,e}(\vec{x}, \vec{y})$$

est définie positive (et donc elle est un produit scalaire) si et seulement si la matrice A est définie positive.
Attention, le fait qu'une matrice A soit, par exemple, définie positive n'implique pas que tous les éléments de A soient positifs. Pour s'en convaincre, il suffit de considérer la matrice

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

Un critère pour déterminer la caractère d'une matrice sera introduit dans la prochaine section.

17.16 Corollaire. *Soit $q \colon V \to \mathbb{R}$ une forme quadratique, e une base de V, et $A_{q,e} \in \mathbb{R}^{n \times n}$ la matrice symétrique associée à q par rapport à la base e. Alors q et $A_{q,e}$ ont le même caractère.*
En particulier, deux matrices symétriques associées à une même forme quadratique par rapport à deux bases différentes ont même caractère.

Preuve. Soit $A = A_{q,e}$. Il suffit d'appliquer le Lemme 17.13 à la situation suivante

$$\begin{tikzcd} V \arrow[rr, "E"] \arrow[dr, "q"'] & & \mathbb{R}^n \arrow[dl, "q_{A,f}"] \\ & \mathbb{R} & \end{tikzcd}$$

où E est l'isomorphisme définie par $E(\vec{x}) = {}_e\vec{x}$ (voir 8.23). Par le Corollaire 17.14, pour déterminer le caractère de la matrice A nous pouvons choisir de façon arbitraire une base f de $\mathbb{R}^n$. Or, pour que le diagramme commute, il faut choisir comme base de $\mathbb{R}^n$ la base $f = \{E(\vec{e}_1), \ldots, E(\vec{e}_n)\}$, qui n'est rien d'autre que la base canonique, de façon que ${}_fE(\vec{x}) = E(\vec{x}) = {}_e\vec{x}$. Nous avons donc

$$q_{A,f}(E(\vec{x})) = {}_fE(\vec{x})^t \cdot A \cdot {}_fE(\vec{x}) = {}_e\vec{x}^t \cdot A_{q,e} \cdot {}_e\vec{x} = q(\vec{x}) \qquad \square$$

17.17 Corollaire. *Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et $P \in \mathbb{R}^{n\times n}$ une matrice inversibles. Les matrices A et $P^t \cdot A \cdot P$ ont même caractère.*

Preuve. Car A et $P^t \cdot A \cdot P$ sont les matrices associées à une même forme quadratique $q\colon \mathbb{R}^n \to \mathbb{R}$ par rapport à deux bases différentes de $\mathbb{R}^n$: si e est une base de $\mathbb{R}^n$ et $q = q_{A,e}$ alors

$$A_{q,e} = A \;\text{ et }\; A_{q,f} = P^t \cdot A \cdot P \;\text{ avec }\; P = {}_eI_f \qquad \square$$

17.4 Caractère et valeurs propres

Il est difficile de déterminer le caractère d'une forme quadratique (ou d'une matrice symétrique) en utilisant la définition. Voyons une méthode plus efficace :

17.18 Propriété. *Soit $q\colon V \to \mathbb{R}$ une forme quadratique, e une base de V et $A = A_{q,e}$ la matrice symétrique associée.*

1. *q est définie positive ssi les valeurs propres de A sont strictement positives.*
2. *q est semi-définie positive ssi les valeurs propres de A sont positives ou nulles.*
3. *q est définie négative ssi les valeurs propres de A sont strictement négatives.*
4. *q est semi-définie négative ssi les valeurs propres de A sont négatives ou nulles.*
5. *q est indéfinie ssi A admet au moins une valeur propre strictement positive et au moins une valeur propre strictement négative.*

Preuve. Puisque A est symétrique, par le Corollaire 16.13 nous savons qu'il existe une matrice orthogonale Q formée par des vecteurs propres de A telle que

$$Q^t \cdot A \cdot Q$$

est la matrice diagonale D des valeurs propres de A. Si f est la base de V déterminée par $Q = {}_eI_f$ (voir Remarque 17.19), on aura $A_{q,f} = D$. La preuve devient maintenant facile : si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A, alors

$$A_{q,f} = D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & & \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

Par conséquent, si $\vec{x} = x_1 \cdot \vec{f_1} + \ldots x_n \cdot \vec{f_n}$, nous avons

$$q(\vec{x}) = (x_1, \ldots, x_n) \cdot D \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda_1 x_1^2 + \ldots \lambda_n x_n^2$$

Voyons par exemple le premier cas :
1. Si chaque λ_i est strictement positive et au moins un x_i est non nul, alors $q(\vec{x}) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 > 0$. Réciproquement, si $q(\vec{x}) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 > 0$ pour tout $\vec{x} \neq \vec{0}$, nous pouvons choisir $\vec{x} = \vec{f_i}$ et nous obtenons $\lambda_i = q(\vec{f_i}) > 0$. □

17.19 Remarque. La phrase “Si f est la base de V déterminée par $Q = {}_eI_f$”, utilisée dans la preuve de la Propriété 17.18, signifie que si la matrice Q est par exemple donnée par

$$Q = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

alors la base f sera donnée par

$$\vec{f_1} = a \cdot \vec{e}_1 + d \cdot \vec{e}_2 + g \cdot \vec{e}_3\ ,\ \vec{f_2} = b \cdot \vec{e}_1 + e \cdot \vec{e}_2 + h \cdot \vec{e}_3\ ,\ \vec{f_3} = c \cdot \vec{e}_1 + f \cdot \vec{e}_2 + i \cdot \vec{e}_3$$

Autrement dit, si $E \colon V \to \mathbb{R}^n$ est l'isomorphisme déterminé par la base e de $\mathbb{R}^n$ et défini par $E(\vec{x}) = {}_e\vec{x}$ (voir 8.23), alors $\vec{f_i} = E^{-1}(C_i)$, où C_i est la i-ème colonne de la matrice Q.

17.20 Remarque. On peut extraire de la preuve de la Propriété 17.18 que, si $q \colon V \to \mathbb{R}$ est une forme quadratique sur un espace V de type fini, on peut toujours trouver une base f de V telle que la matrice $A_{q,f}$ soit une matrice diagonale. Il faut d'abord choisir une base quelconque e de V, construire la matrice $A_{q,e}$ et choisir ensuite une base orthonormée c de $\mathbb{R}^n$ formée par des vecteurs propres de $A_{q,e}$. Finalement, on peut prendre comme base f de V les vecteurs

$$\{\vec{f_1} = E^{-1}(\vec{c}_1), \ldots, \vec{f_n} = E^{-1}(\vec{c}_n)\}$$

où $E \colon V \to \mathbb{R}^n$ est l'isomorphisme définie par $E(\vec{x}) = {}_e\vec{x}$. En effet, les colonnes de ${}_eI_f$ sont les vecteurs $\vec{c}_i$ et donc

$$A_{q,f} = ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f = D$$

est la matrice diagonale des valeurs propres de $A_{q,e}$. (Si en particulier $V = \mathbb{R}^n$, on peut choisir comme base e la base canonique, de façon que $E\colon \mathbb{R}^n \to \mathbb{R}^n$ est l'identité et la base f est une base orthonormée de vecteurs propres de $A_{q,e}$.)
De plus, la formule établie dans la preuve de la Propriété 17.18

$$\text{si } \ \vec{x} = x_1 \cdot \vec{f_1} + \dots x_n \cdot \vec{f_n} \ \text{ alors } \ q(\vec{x}) = \lambda_1 x_1^2 + \dots \lambda_n x_n^2$$

montre que q peut s'exprimer comme *somme pondérée de carrés.*

17.5 Loi d'inertie

Si $A_{q,e}$ et $A_{q,f}$ sont deux matrices associées à une même forme quadratique par rapport à deux bases différentes, elle sont liées par la formule de la Propriété 17.7

$$A_{q,f} = ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f$$

ce qui permet de montrer que $A_{q,e}$ et $A_{q,f}$ ont le même caractère (voir 17.16). Par contre, en général $A_{q,e}$ et $A_{q,f}$ n'ont pas les mêmes valeurs prores. (Les deux matrices $A_{q,e}$ et $A_{q,f}$ ont les mêmes valeurs propres si la matrice changement de base eI_f est orthogonale, car alors la formule précédente devient

$$A_{q,f} = ({}_eI_f)^{-1} \cdot A_{q,e} \cdot {}_eI_f$$

et on peut appliquer l'Exercice 14.15.)
Cependant, ce qui reste vrai dans le cas général, est que $A_{q,e}$ et $A_{q,f}$ ont même indice de positivité et même indice de négativité, comme expliqué dans la prochaine propriété, la loi d'inertie. La loi d'inertie précise la Propriété 17.18 qui ne donne pas d'informations sur le nombre de valeurs propres positives et négatives d'une forme quadratique indéfinie. Elle sera admise sans preuve. (Dans l'énoncé de la loi de Sylvester, par "nombre de valeurs propres" il faut entendre "nombre compté avec la multiplicité algébrique de la valeur propre". Par exemple, si les valeurs propres d'une matrice sont 1, 1, 2, 3, 3, -1, -1, on dira que la matrice a cinq valeurs propres strictement positives et deux valeurs propres strictement négatives.)

17.21 Propriété. *(Loi d'inertie de Sylvester) Soit $q\colon V \to \mathbb{R}$ une forme quadratique avec V de dimension finie, et soit e, f deux bases de V. Alors :*

1. *Le nombre de valeurs propres strictement positives de $A_{q,f}$ est égal au nombre de valeurs propres strictement positives de $A_{q,e}$.*
   Ce nombre est noté $\operatorname{ind}_+ q$ *et appelé* indice de positivité *de q.*
2. *Le nombre de valeurs propres strictement négatives de $A_{q,f}$ est égal au nombre de valeurs propres strictement négatives de $A_{q,e}$.*
   Ce nombre est noté $\operatorname{ind}_- q$ *et appelé* indice de négativité *de q.*
3. *Le nombre de valeurs propres nulles de $A_{q,f}$ est égal au nombre de valeurs propres nulles de $A_{q,e}$.*
4. *De plus :* $\operatorname{rang} A_{q,e} = \operatorname{ind}_+ q + \operatorname{ind}_- q$. *On pose* $\operatorname{rang} q = \operatorname{rang} A_{q,e}$.

Preuve. To be inserted. □

17.22 Remarque. Le dernier point de la loi de Sylvester est valable pour toute matrice diagonalisable. Plus en générale, pour une matrice $A \in \mathbb{R}^{n\times n}$, notons

$\mathrm{ind}_+ A$ le nombre de valeurs propres strictement positives de A,

$\mathrm{ind}_- A$ le nombre de valeurs propres strictement négatives de A.

Soient $\lambda_1, \ldots, \lambda_r$ les valeurs propres distinctes non nulles de A, de multiplicité algébrique $m_a(\lambda_1), \ldots, m_a(\lambda_r)$. Si les deux conditions suivantes sont satisfaite

1. $m_a(0) = \dim E(0)$,
2. $m_a(0) + m_a(\lambda_1) + \ldots + m_a(\lambda_r) = n$,

alors $\mathrm{rang}\, A = \mathrm{ind}_+ A + \mathrm{ind}_- A$.

Preuve. Considérons l'application linéaire associée à la matrice A

$$L_A : \mathbb{R}^n \to \mathbb{R}^n \ , \ \ L_A(\vec{x}) = A \cdot \vec{x}$$

dont on sait que $\mathrm{Ker}\, L_A = E(0)$ et $\mathrm{Im}\, L_A = \mathrm{Col}(A)$. Par le théorème du rang (9.15) et la première hypothèse, nous avons que

$$m_a(0) = \dim E(0) = n - \mathrm{rang}\, A$$

d'où

$$\mathrm{rang}\, A = n - m_a(0) = m_a(\lambda_1) + \ldots + m_a(\lambda_r) = \mathrm{ind}_+ A + \mathrm{ind}_- A$$

grâce à la deuxième hypothèse. □

17.23 Exemple. Voici deux exemples qui montrent que dans la remarque précédente on ne peut pas se passer des deux hypothèses.

1. Soit $R\colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\frac{\pi}{2}$ autour de l'origine, et soit $A \in \mathbb{R}^{2\times 2}$ la matrice associée à R (peu importe par rapport à quelles bases). Puisque R n'a pas de vecteurs propres, $m_a(0) = 0$ et $\mathrm{ind}_+ A = 0 = \mathrm{ind}_- A$. De plus, R est bijective et donc $E(0) = \mathrm{Ker}\, R = \{\vec{0}\}$. La condition $m_a(0) = \dim E(0)$ est donc satisfaite. Par contre, la condition $m_a(0) + m_a(\lambda_1) + \ldots + m_a(\lambda_r) = n$ n'est pas remplie et $\mathrm{ind}_+ A + \mathrm{ind}_- A = 0 \neq 2 = \mathrm{rang}\, A$.
2. La matrice
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
admet 0 comme seule valeur propre, et elle est une valeur propre double. La condition $m_a(0) + m_a(\lambda_1) + \ldots + m_a(\lambda_r) = n$ est donc satisfaite. Par contre, la condition $m_a(0) = \dim E(0)$ n'est pas remplie, car $m_a(0) = 2$ et $\dim E(0) = 1$, et $\mathrm{ind}_+ A + \mathrm{ind}_- A = 0 \neq 1 = \mathrm{rang}\, A$.

L'intérêt de la loi de Sylvester vient, entre autre, de son application à la classification des quadriques réelles (voir cours de géométrie). En effet, ce qui intervient dans la classification des quadriques est l'indice de positivité et l'indice de négativité des formes quadratiques associées aux quadriques, et pas seulement leur caractère.

17.6 Complétion des carrés

La loi de Sylvester permet aussi de justifier d'autres méthodes qui sont utilisés pour déterminer le caractère et les indices d'une forme quadratique, comme par exemple la méthode dite de complétion des carrés. Cette méthode, partiellement illustrée par les deux exemples suivants, ne rentre pas dans le programme de ce cours.

17.24 Exercice. Réduire q à une forme diagonale par la méthode de complétion des carrés, et déterminer la matrice de changement de variables correspondante ; donner, après avoir effectué la complétion des carrés, une matrice P telle $P^t \cdot A \cdot P$ soit diagonale (A étant la matrice symétrique associée à la forme q).

$$q(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 2x_3^2 + 4x_1x_2 + 8x_1x_3 + 4x_2x_3$$

Preuve. On considère la forme quadratique

$$q(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 2x_3^2 + 4x_1x_2 + 8x_1x_3 + 4x_2x_3 \tag{17.1}$$

On fixe une variable, par exemple x_1, et on cherche des coefficients a, b, c tels que $(ax_1 + bx_2 + cx_3)^2$ ait les mêmes termes en x_1^2, x_1x_2 et x_1x_3 que la forme q. Il faut choisir $a = \sqrt{2}, b = \sqrt{2}$ et $c = 2\sqrt{2}$. Si on compare la forme q avec $(\sqrt{2}x_1 + \sqrt{2}x_2 + 2\sqrt{2}x_3)^2$, on voit que

$$q(x_1, x_2, x_3) = (\sqrt{2}x_1 + \sqrt{2}x_2 + 2\sqrt{2}x_3)^2 - (3x_2^2 + 6x_3^2 + 4x_2x_3) \tag{17.2}$$

Il faut donc recommencer avec la forme quadratique (plus simple que celle de depart)

$$q_1(x_2, x_3) = 3x_2^2 + 6x_3^2 + 4x_2x_3 \tag{17.3}$$

On fixe une variable, par exemple x_2, et on cherche des coefficients a, b tels que $(ax_2 + bx_3)^2$ ait les mêmes termes en x_2^2 et x_2x_3 que la forme q_1. Il faut chosir $a = \sqrt{3}$ et $b = \frac{2}{\sqrt{3}}$. Si on compare la forme q_1 avec $(\sqrt{3}x_2 + \frac{2}{\sqrt{3}}x_3)^2$, on voit que

$$q_1(x_2, x_3) = (\sqrt{3}x_2 + \frac{2}{\sqrt{3}}x_3)^2 + \frac{14}{3}x_3^2 \tag{17.4}$$

Les équations (2) et (4) nous permettent d'écrir la forme q comme somme pondérée de carrés

$$\begin{aligned} q(x_1, x_2, x_3) &= (\sqrt{2}x_1 + \sqrt{2}x_2 + 2\sqrt{2}x_3)^2 - (\sqrt{3}x_2 + \tfrac{2}{\sqrt{3}}x_3)^2 - \tfrac{14}{3}x_3^2 \\ &= y_1^2 - y_2^2 - \tfrac{14}{3}y_3^2 \end{aligned} \tag{17.5}$$

où on a posé

$$\begin{cases} y_1 = \sqrt{2}x_1 + \sqrt{2}x_2 + 2\sqrt{2}x_3 \\ y_2 = \sqrt{3}x_2 + \frac{2}{\sqrt{3}}x_3 \\ y_3 = x_3 \end{cases}$$

Si on pose $x = (x_1, x_2, x_3)^t$ et $y = (y_1, y_2, y_3)^t$, on a la forme matricielle

$$q(x_1, x_2, x_3) = x^t \cdot A \cdot x = y^t \cdot P^t \cdot A \cdot P \cdot y$$

avec A la matrice symétrique associée à la forme q et P la matrice du changement de variables inverse

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

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{4}{3} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{cases} x_1 &= \frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{3}}y_2 - \frac{4}{3}y_3 \\ x_2 &= \frac{1}{\sqrt{3}}y_2 - \frac{2}{3}y_3 \\ x_3 &= y_3 \end{cases}$$

On vérifie en effet que

$$P^t \cdot A \cdot P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -\frac{14}{3} \end{pmatrix}$$

Si on compare cette matrice diagonale avec celle qu'on peut obtenir par diagonalisation de A via une matrice orthogonale Q de vecteurs propres

$$Q^t \cdot A \cdot Q = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

on voit que les éléments sur la diagonale principale sont différents (-2, -2 et 7 sont en effet les valeurs propres de A), mais l'indice de positivité et l'indice de négativité sont les mêmes : $\operatorname{ind}_+ q = 1$ et $\operatorname{ind}_- q = 2$. □

17.25 Exercice. Déterminer l'indice de positivité et l'indice de négativité de la forme quadratique

$$q \colon \mathbb{R}^3 \to \mathbb{R}\ , \quad q(x_1, x_2, x_3) = x_1^2 - 2x_2x_2 + 2\alpha x_1x_3 + 2x_2^2 + 8x_3^2$$

Discuter en fonction du paramètre réel α.

Preuve. La matrice associée à q par rapport à la base canonique e de $\mathbb{R}^3$ est

$$A_{q,e} = \begin{pmatrix} 1 & -1 & \alpha \\ -1 & 2 & 0 \\ \alpha & 0 & 8 \end{pmatrix}$$

Une première méthode consiste à calculer le polynôme caractéristique de $A_{q,e}$ et estimer le signe des trois valeurs propres. Ceci est détaillé dans l'Exercice 3 du TP 17. Voyons une deuxième possibilité qui se base sur la méthode de complétion des carrés.

Par la loi d'inertie, pour déterminer $\operatorname{ind}_+ q$ et $\operatorname{ind}_- q$ nous pouvons choisir une base arbitraire f de $\mathbb{R}^3$. Prenons

$$f = \{\vec{f_1} = (1, 0, 0)^t, \vec{f_2} = (1, 1, 0)^t, \vec{f_3} = (-2\alpha, -\alpha, 1)^t$$

Si ${}_eI_f$ est la matrice changement de base, nous obtenons (voir 17.7)

$$A_{q,f} = ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 8-2\alpha^2 \end{pmatrix}$$

qui admet comme valeurs propres 1, 1 et $8-2\alpha^2$ (qui sont différentes des valeurs propres de $A_{q,e}$). Il faut donc examiner trois cas :

1. si $\alpha = \pm 2$, alors $\operatorname{ind}_+ q = 2$ et $\operatorname{ind}_- q = 0$
2. si $-2 < \alpha < 2$, alors $\operatorname{ind}_+ q = 3$ et $\operatorname{ind}_- q = 0$
3. si $\alpha < -2$ ou $\alpha > 2$, alors $\operatorname{ind}_+ q = 2$ et $\operatorname{ind}_- q = 1$

Le problème qui se pose est : comment déterminer une telle base f pour que $A_{q,f}$ soit diagonale, sans connaître à priori les valeurs propres de $A_{q,e}$? C'est ici que la méthode de complétion des carrés intervient. En suivant la même démarche que dans l'Exercice 17.24, nous obtenons

$$q(x_1, x_2, x_3) = (x_1 - x_2 + \alpha x_3)^2 + (x_2 + \alpha x_3)^2 + (8 - 2\alpha^2)x_3^2$$

Par conséquent, si on pose

$$y_1 = x_1 - x_2 + \alpha x_3 \ , \ \ y_2 = x_2 + \alpha x_3 \ , \ \ y_3 = x_3$$

nous avons

$$q(x_1, x_2, x_3) = y_1^2 + y_2^2 + (8 - 2\alpha^2)y_3^2$$

qui est précisément la forme diagonale de q obtenue en utilisant la base f. En effet, le changement de variables inverse en forme matricielle donne

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -2\alpha \\ 0 & 1 & -\alpha \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

et les colonnes de la matrice des coefficients sont exactement les vecteurs de la base f. Autrement dit, cette matrice est précisément la matrice de changement de base ${}_eI_f$ et donc

$$q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot A_{q,e} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \cdot ({}_eI_f)^t \cdot A_{q,e} \cdot {}_eI_f \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$$

$$= \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \cdot A_{q,f} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = y_1^2 + y_2^2 + (8 - 2\alpha^2)y_3^2$$

$\square$

17.26 Remarque. Pour appliquer la méthode de complétion des carrés il faut que le coefficient d'au moins un terme carré soit non nul. Si tous les coefficients des termes carrés sont nuls, il faut d'abord effectuer un changement de variables du type $y_1 = x_1 + x_2, y_2 = x_1 - x_2$. On peut traiter ainsi par exemple la forme $q(x_1, x_2) = x_1 x_2$.

Pour terminer, voyons, sans entrer dans les détails, une autre méthode pour déterminer le signe des valeurs propres d'une matrice symétrique sans calculer explicitement ces valeurs propres.

17.27 Définition. Une sous-matrice B d'une matrice A est *principale* si elle est obtenue en supprimant certains lignes de A ainsi que les colonnes de même indice.

17.28 Propriété. *Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et soit B une sous-matrice principale de A. Si B admet au moins k valeurs propres strictement positives (respectivement : positives ou nulles, strictement négatives, négatives ou nulles), alors il en est de même pour A.*

17.29 Remarque. La même propriété n'est pas valable pour les valeurs propres nulles. Comme exemple on peut considérer la matrice

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

et la sous-matrice principale $B = 0$. B admet 0 comme valeur propre, mais les valeurs propres de A sont 1 et -1.

17.30 Exercice. Prenons à nouveau la matrice

$$A = \begin{pmatrix} 1 & -1 & \alpha \\ -1 & 2 & 0 \\ \alpha & 0 & 8 \end{pmatrix}$$

de l'Exercice 17.25 et essayons de déterminer le signe des trois valeurs propres $\lambda_1, \lambda_2, \lambda_3$. Considérons la sous-matrice principale

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

Elle admet 2 et 8 comme valeurs propres. Par la Propriété 17.28, A admet deux valeurs propres strictment positives. De plus

$$\det A = 8 - 2\alpha^2 = \lambda_1 \cdot \lambda_2 \cdot \lambda_3$$

On a donc trois cas possibles :

1. si $\alpha = \pm 2$, alors $\det A = 0$ et A admet donc deux valeurs propres strictement positives et une valeur propre nulle,
2. si $-2 < \alpha < 2$, alors $\det A > 0$ et les trois valeurs propres de A doivent être strictement positives,
3. si $\alpha < -2$ ou $\alpha > 2$, alors $\det A < 0$ et A admet donc deux valeurs propres strictement positives et une valeur propre strictement négative.

Index