I Les matrices

Introduction

L'objet de cette partie du cours est de vous donner des outils mathématiques qui vous seront nécessaires dans les années à venir. Les objets que nous manipulerons principalement s'appellent des matrices, et servent à coder certains problèmes, tels que par exemple certains systèmes d'équations, ou certains systèmes d'équations différentielles. Nous reviendrons sur ces questions plus loin dans le cours, mais pour l'instant il nous faut définir et manipuler les objets dont nous aurons besoin.

A Matrices, opérations sur les matrices

Définition A.1. Une *matrice* $m \times n$ est un tableau rectangulaire de nombres, à m lignes et n colonnes:

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

où les a_{ij}, pour $1 \leqslant i \leqslant m$ et $1 \leqslant j \leqslant n$ sont des réels. On la note aussi $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$
Les indices i et j de a_{ij} signifient que a_{ij} est situé sur la i-ème ligne et la j-ème colonne.

L'ensemble de toutes les matrices à m lignes et n colonnes est noté $\mathcal{M}_{m,n}(\mathbb{R})$.

Exemple A.2. $\begin{pmatrix} 2 & 4 & -1 \\ 0 & 3 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & -2 \\ 3 & -1 \end{pmatrix}$ sont des matrices (2×3 et 2×2 respectivement).

Cas particuliers
Les éléments de $\mathcal{M}_{1,n}(\mathbb{R})$ sont appelés *vecteurs ligne* (ou *matrices ligne*). Ils sont de la forme $\begin{pmatrix} a_{11} & \cdots & a_{1n} \end{pmatrix}$.

Les éléments de $\mathcal{M}_{m,1}(\mathbb{R})$ sont appelés *vecteurs colonne* ou *matrices colonne*. Ils sont de la forme $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}$.

Définition A.3. On définit

- l'*addition* de deux matrices: si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ et $B = (b_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ on pose

$$A + B = (a_{ij} + b_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$$

soit

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

- la *multiplication* d'une matrice par un réel: si $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ et λ est réel, on pose

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

soit

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

Remarque A.4. On ne peut additionner des matrices que si elles sont du *même type* (même nombre de lignes et même nombre de colonnes).

Exemple A.5.

$$\begin{pmatrix} 2 & -1 & -3 \\ 4 & -7 & 5 \end{pmatrix} + \begin{pmatrix} 6 & 8 & -9 \\ -10 & 13 & 11 \end{pmatrix} = \begin{pmatrix} 2+6 & -1+8 & -3-9 \\ 4-10 & -7+13 & 5+11 \end{pmatrix}$$
$$= \begin{pmatrix} 8 & 7 & -12 \\ -6 & 6 & 16 \end{pmatrix}$$

Nous allons maintenant définir un produit de matrices, qui existe dans certaines conditions, et qui est un peu plus complexe que les opérations ci-dessus.

Définition A.6. Etant données deux matrices A de $\mathcal{M}_{m,n}(\mathbb{R})$ et B de $\mathcal{M}_{n,p}(\mathbb{R})$ telles que le nombre n de colonnes de A est égal au nombre de lignes de B, on définit leur produit AB, qui est une matrice de $\mathcal{M}_{m,p}(\mathbb{R})$, de la façon suivante: notons $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ et $B = (b_{jk})_{\substack{1 \leqslant j \leqslant n \\ 1 \leqslant k \leqslant p}}$, alors $AB = (c_{ik})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant k \leqslant p}}$ avec $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$.

Exemple A.7.

$$\begin{array}{cc} & \begin{pmatrix} 4 & -3 & 6 \\ -2 & 5 & -4 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} & \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \end{array}$$

avec
$$\begin{cases} a = 1 \cdot 4 + 2 \cdot (-2) = 0 \\ b = 1 \cdot (-3) + 2 \cdot 5 = 7 \\ c = 1 \cdot 6 + 2 \cdot (-4) = -2 \\ d = (-1) \cdot 4 + 3 \cdot (-2) = -10 \\ e = (-1) \cdot (-3) + 3 \cdot 5 = 18 \\ f = (-1) \cdot 6 + 3 \cdot (-4) = -18 \end{cases}$$

Donc
$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 4 & -3 & 6 \\ -2 & 5 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 7 & -2 \\ -10 & 18 & -18 \end{pmatrix}.$$

On remarque que le produit

$$\begin{pmatrix} 4 & -3 & 6 \\ -2 & 5 & -4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$$

n'est pas défini.

Définition A.8. On définit des matrices particulières: les *matrices unité*, ou *identité*, $I_n \in \mathcal{M}_{n,n}(\mathbb{R})$, par

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

c'est-à-dire que les termes sur la diagonale sont des 1, et les autres sont des 0.

Exemple A.9. $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Propriétés A.10.

- $(AB)C = A(BC)$ pour toutes matrices convenables A, B, C (c'est-à-dire telles que les produits existent). On peut donc oublier les parenthèses et écrire ABC. On dit que le produit est *associatif*.
- $AI_n = A$ et $I_m A = A$ pour toute matrice A de $\mathcal{M}_{m,n}(\mathbb{R})$.
- $(A + A')B = AB + A'B$ pour toutes matrices A, A', B convenables. On dit que le produit est *distributif* (à droite) par rapport à l'addition.
- $A(B + B') = AB + AB'$ pour toutes matrices A, B, B' convenables. On dit que le produit est *distributif* (à gauche) par rapport à l'addition.
- $(\alpha A)B = \alpha(AB) = A(\alpha B)$ pour toutes matrices A, B convenables et tout réel α.

Remarque A.11. *Attention:* Ce produit n'est **pas commutatif**! C'est-à-dire qu'on peut avoir $AB \neq BA$ pour certaines matrices (on peut même avoir AB définie alors que BA n'est pas définie).

Remarque A.12. Les matrices unité I_n jouent un rôle semblable au "1" de $\mathbb{R}$.

Définition A.13. On dit qu'une matrice A de $\mathcal{M}_{n,n}(\mathbb{R})$ est *inversible* s'il existe une matrice B de $\mathcal{M}_{n,n}(\mathbb{R})$ telle que $AB = I_n$ et $BA = I_n$. On dit alors que B est *l'inverse* de A et on la note A^{-1}.

On a donc $AA^{-1} = I_n$ et $A^{-1}A = I_n$.

Attention: On ne doit *jamais* diviser par une matrice, mais on peut multiplier par son inverse si elle est inversible. (La division de matrices n'existe pas).

Remarque A.14. On peut montrer que l'une des égalités ci-dessus suffit, c'est-à-dire que s'il existe une matrice B telle que $AB = I_n$ (*ou* $BA = I_n$), alors A est inversible (on rappelle que A est dans $\mathcal{M}_{n,n}(\mathbb{R})$).

Remarque A.15. Pour qu'une matrice soit inversible, il faut qu'elle ait le même nombre de lignes et de colonnes (on dit qu'elle est *carrée*). Mais **attention**, cela ne suffit pas: toutes les matrices carrées ne sont pas inversibles.

Exemple A.16. Soit $A = \begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix}$. Si A est inversible, alors il existe $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que $AB = I_2$. Donc $\begin{pmatrix} a - 3c & b - 3d \\ 2a - 6c & 2b - 6d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. En particulier: $a - 3c = 1$ mais $2(a - 3c) = 2a - 6c = 0$: contradiction. Donc B n'existe pas et A n'est pas inversible.

Exemple A.17. Soit $A = \begin{pmatrix} 1 & -3 \\ 2 & 6 \end{pmatrix}$. Si A est inversible, alors il existe $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que $AB = I_2$. Donc $\begin{pmatrix} a-3c & b-3d \\ 2a+6c & 2b+6d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. On en déduit le système d'équations $\begin{cases} a-3c=1 \\ b-3d=0 \\ 2a+6c=0 \\ 2b+6d=1 \end{cases}$ qui a pour solution $a=\frac{1}{2}$, $b=\frac{1}{4}$, $c=-\frac{1}{6}$ et $d=\frac{1}{12}$.

On vérifie que l'on a bien $A\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{6} & \frac{1}{12} \end{pmatrix} = I_2$.

Donc A est inversible et $A^{-1} = \begin{pmatrix} \dfrac{1}{2} & \dfrac{1}{4} \\ -\dfrac{1}{6} & \dfrac{1}{12} \end{pmatrix}$.

Définition A.18. On définit la *transposée* A^t d'une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant m \\ 1\leqslant j\leqslant n}}$ de $\mathcal{M}_{m,n}(\mathbb{R})$ de la façon suivante: $A^t = \left(a'_{ij}\right)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant m}}$ est une matrice de $\mathcal{M}_{n,m}(\mathbb{R})$ avec $a'_{ij} = a_{ji}$. Elle est obtenue en écrivant en lignes les colonnes de A.

Exemple A.19. Soit $A = \begin{pmatrix} 2 & 4 & 6 \\ -1 & 3 & 5 \end{pmatrix}$. Alors $A^t = \begin{pmatrix} 2 & -1 \\ 4 & 3 \\ 6 & 5 \end{pmatrix}$

II Systèmes linéaires et matrices échelonnées

Introduction

Nous serons souvent amenés à résoudre un système d'équations linéaires. Nous allons introduire une méthode systématique de résolution, la méthode du *Pivot de Gauss*, que nous allons décrire à l'aide de matrices échelonnées. Nous verrons aussi comment appliquer cette méthode pour calculer l'inverse d'une matrice carrée.

A Matrices échelonnées; pivot de Gauss

Soit à résoudre le système suivant:

$$(S)\quad \left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right.$$

à m équations et n inconnues $x_1, x_2 \ldots, x_n$. Les a_{ij} sont des coefficients réels.

Ce système d'équations peut aussi s'écrire sous forme d'une équation matricielle

$$AX = B$$

avec:

- $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, la matrice du "second membre",
- $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, la "matrice des inconnues",
- A est la matrice $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, dite *matrice du système*. Chaque colonne de A est formée par les coefficients de l'une des inconnues (x_1 pour la première colonne, etc., x_n pour la n-me colonne). Chaque ligne de A correspond à une équation.

Nous allons faire correspondre au système (S) une matrice dite "augmentée", qui est formée de A à laquelle on rajoute une colonne qui est B:

$$(A \mid B) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Proposition A.1. On ne change pas les solutions du système (S) si:

I. On échange deux équations de (S), c'est-à-dire si on échange deux *lignes* de $(A \mid B)$.

II. On multiplie une équation de (S) par un réel *non nul*, c'est-à-dire si on multiplie une *ligne* de $(A \mid B)$ par un réel *non nul*.

III. On remplace une équation de (S) par sa somme avec une *autre* équation de (S), c'est-à-dire si on remplace une *ligne* de $(A \mid B)$ par sa somme avec une *autre* ligne de $(A \mid B)$.

Nous allons utiliser ces opérations pour donner une méthode systématique de résolution de (S), la méthode du *pivot de Gauss*. Il s'agit de transformer la matrice $(A \mid B)$ en une matrice dite "échelonnée" à l'aide des opérations ci-dessus.

Définition A.2. On appelle *élément de tête* d'une ligne non nulle d'une matrice l'élément non nul situé le plus à gauche de la ligne.

Définition A.3. Une matrice est dite *échelonnée* si elle remplit les trois conditions suivantes:

- Toutes ses lignes non nulles sont situées au-dessus de ses lignes nulles.
- Chaque élément de tête d'une ligne se trouve dans une colonne à droite de l'élément de tête de la ligne précédente.
- Tous les éléments de la colonne sous un élément de tête sont nuls.

Une matrice est dite *échelonnée réduite* si elle remplit les trois conditions suivantes:

- Elle est échelonnée.
- L'élément de tête de chaque ligne vaut 1.
- Chaque 1 de tête d'une ligne est le seul élément non nul de sa colonne.

Exemple A.4. La matrice $\begin{pmatrix} {\color{red}*} & * & * & * & * & * & * & * \\ 0 & {\color{red}*} & * & * & * & * & * & * \\ 0 & 0 & 0 & {\color{red}*} & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & {\color{red}*} & * \end{pmatrix}$ est échelonnée (les ${\color{red}*}$ représentent les éléments de tête, qui sont quelconques (non nuls), et les $*$ sont des réels quelconques) mais $\begin{pmatrix} 2 & 3 & -1 \\ 0 & {\color{green}-2} & 1 \\ 0 & {\color{blue}3} & 2 \end{pmatrix}$ n'est pas échelonnée car l'élément de tête de la troisième ligne $({\color{blue}3})$ n'est pas à droite ce celui de la deuxième ligne $({\color{green}-2})$.

Exemple A.5. La matrice $\begin{pmatrix} 1 & 0 & * & 0 & * \\ 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 1 & * \end{pmatrix}$ est échelonnée réduite, mais $\begin{pmatrix} 1 & 0 & * & {\color{blue}3} & * \\ 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & {\color{green}1} & * \end{pmatrix}$ est échelonnée mais pas réduite (à cause du ${\color{blue}3}$ situé au-dessus du ${\color{green}1}$ de tête de la troisième ligne).

Proposition A.6. Soit M une matrice quelconque. A l'aide des opérations sur les lignes décrites ci-dessus, on peut toujours transformer M en une matrice échelonnée réduite. Cette matrice échelonnée *réduite* est unique, elle ne dépend que de M et pas de la succession d'opérations effectuées pour s'y ramener.

Nous allons maintenant décrire la méthode du pivot de Gauss sur un exemple:

Exemple A.7. Soit à résoudre le système

$$(S)\quad \begin{cases} 2x+6y+1z-3t &= 2 \\ x+3y-2z+6t &= 7 \\ -x-3y+4z+4t &= -3 \\ 3x+9y-6z-22t &= -1 \end{cases}$$

On écrit la matrice augmentée associée:

$$\left(\begin{array}{cccc|c} 2 & 6 & 1 & -3 & 2 \\ 1 & 3 & -2 & 6 & 7 \\ -1 & -3 & 4 & 4 & -3 \\ 3 & 9 & -6 & -22 & -1 \end{array}\right)$$

On échange (éventuellement) des lignes pour que l'élément de tête de la première ligne soit le plus simple possible; ici $L_1 \leftrightarrow L_2$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 2 & 6 & 1 & -3 & 2 \\ -1 & -3 & 4 & 4 & -3 \\ 3 & 9 & -6 & -22 & -1 \end{array}\right)$.

On fixe maintenant la première ligne, et on s'en sert pour faire apparaître des 0 sous l'élément de tête de la première ligne; ici, $L_2 \to L_2 - 2L_1$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ -1 & -3 & 4 & 4 & -3 \\ 3 & 9 & -6 & -22 & -1 \end{array}\right)$,

puis $L_3 \to L_3 + L_1$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ 0 & 0 & 2 & 10 & 4 \\ 3 & 9 & -6 & -22 & -1 \end{array}\right)$

et enfin $L_4 \to L_4 - 3L_1$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ 0 & 0 & 2 & 10 & 4 \\ 0 & 0 & 0 & -40 & -22 \end{array}\right)$.

On recommence avec l'élément de tête de la deuxième ligne. L'opération $L_3 \to L_3 - \frac{2}{5}L_2$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ 0 & 0 & 0 & 16 & \frac{44}{5} \\ 0 & 0 & 0 & -40 & -22 \end{array}\right)$.

On recommence enfin avec la troisième ligne:

l'opération $L_4 \to L_4 + \frac{5}{2}L_3$ donne $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ 0 & 0 & 0 & 16 & \frac{44}{5} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$.

Cette matrice est échelonnée.

Maintenant, pour terminer la résolution du système, nous allons nous ramener à une matrice échelonnée *réduite*. Cette fois, nous partons du bas vers le haut.

Il n'y a rien à faire sur la dernière ligne puisqu'il n'y a pas d'élément de tête.

On divise la troisième ligne par 16 pour que l'élément de tête soit 1: $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & -15 & -12 \\ 0 & 0 & 0 & 1 & \frac{11}{20} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$.

Ensuite, il nous faut faire apparaître des 0 au-dessus du 1.

Les opérations $L_2 \to L_2 + 15L_3$ puis $L_1 \to L_1 - 6L_3$ donnent successivement $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 6 & 7 \\ 0 & 0 & 5 & 0 & -\frac{15}{4} \\ 0 & 0 & 0 & 1 & \frac{11}{20} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ puis $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 0 & \frac{37}{10} \\ 0 & 0 & 5 & 0 & -\frac{15}{4} \\ 0 & 0 & 0 & 1 & \frac{11}{20} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$.

On recommence avec la deuxième ligne. On divise L_2 par 5: $\left(\begin{array}{cccc|c} 1 & 3 & -2 & 0 & \frac{37}{10} \\ 0 & 0 & 1 & 0 & -\frac{3}{4} \\ 0 & 0 & 0 & 1 & \frac{11}{20} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$.

Puis on fait apparaître des 0 au-dessus de l'élément de tête de la deuxième ligne: $L_1 \to L_1 + 2L_2$ donne $\left(\begin{array}{cccc|c} 1 & 3 & 0 & 0 & \frac{11}{5} \\ 0 & 0 & 1 & 0 & -\frac{3}{4} \\ 0 & 0 & 0 & 1 & \frac{11}{20} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$.

La matrice obtenue est échelonnée réduite. Elle correspond au système suivant:

$$\begin{cases} x + 3y &= \frac{11}{5} \\ z &= -\frac{3}{4} \\ t &= \frac{11}{20}. \end{cases}$$

Ce système est entièrement résolu. Il a une infinité de solutions: par exemple, on peut dire que y est un réel arbitraire et alors $x = -3y + \dfrac{11}{5}$.

B Cas de figure

Trois cas de figure peuvent se présenter:

(i) le système n'a aucune solution. (C'est le cas lorsqu'on obtient une ligne de la forme $\left(\begin{array}{ccc|c} 0 & \cdots & 0 & * \end{array}\right)$ où $*$ est un réel non nul, puisque cette ligne correspond à l'équation $0 = *$.)

(ii) le système a une solution unique, que l'on peut déterminer suivant les cas soit par la méthode du pivot de Gauss, soit par la *méthode de Cramer* que nous verrons plus tard.

(iii) le système a une infinité de solutions. Dans ce cas, certaines inconnues sont arbitraires, et déterminent les autres. On les appelle les *paramètres* du système.

C Calcul de l'inverse d'une matrice carrée

La méthode précédente peut aussi être utilisée pour calculer l'inverse d'une matrice carrée de la façon suivante.

Soit A une matrice carrée à n lignes et n colonnes. On considère la matrice augmentée $(A \mid I_n)$ (à n lignes et $2n$ colonnes).

Si la matrice A est inversible, alors en appliquant la méthode ci-dessus pour se ramener à une matrice échelonnée réduite, ce que l'on obtient est $(I_n \mid A^{-1})$, la partie de droite étant alors l'inverse de A.

Si A n'est pas inversible, la partie de gauche de la matrice échelonnée réduite obtenue ne sera pas I_n.

Exemple C.1. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -4 \\ 0 & -2 & 2 \end{pmatrix}$. Nous voulons l'inverser si possible:

$$(A \mid I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ -1 & 0 & -4 & 0 & 1 & 0 \\ 0 & -2 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$\begin{array}{lc}
L_2 \to L_2 + L_1 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & -1 & 1 & 1 & 0 \\ 0 & -2 & 2 & 0 & 0 & 1 \end{array}\right) \\[2ex]
L_3 \to L_3 + L_2 & \rightsquigarrow \underset{\text{(échelonnée)}}{\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)} \\[2ex]
L_2 \to L_2 + L_3 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \\[2ex]
L_1 \to L_1 - 3L_3 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & -3 & -3 \\ 0 & 2 & 0 & 2 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \\[2ex]
L_2 \to L_2/2 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & -3 & -3 \\ 0 & 1 & 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \\[2ex]
L_1 \to L_1 - 2L_2 & \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & -5 & -4 \\ 0 & 1 & 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right).
\end{array}$$

Cette dernière matrice est échelonnée réduite, et sa partie gauche est bien I_3, donc A est inversible et $A^{-1} = \begin{pmatrix} -4 & -5 & -4 \\ 1 & 1 & \frac{1}{2} \\ 1 & 1 & 1 \end{pmatrix}$.

Exemple C.2. Soit $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. La matrice $(A \mid I_3) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$ est déjà échelonnée.

La matrice échelonnée réduite est obtenue en faisant $L_1 \to L_1 - L_2$ et est égale à $\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$.

La partie gauche de cette matrice n'est pas I_3 donc A n'est pas inversible.

III Espaces $\mathbb{R}^n$ et bases

A Espaces $\mathbb{R}^n$, applications linéaires

Définition A.1. Soit n un entier positif non nul. $\mathbb{R}^n$ est l'ensemble de tous les n-uplets $(x_1, x_2, \dots, x_n)$ de réels, muni des deux opérations suivantes:

(I) l'*addition*: $(x_1, x_2, \dots, x_n)+(y_1, y_2, \dots, y_n) = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$ (addition composante à composante),

(II) la *multiplication par les réels*: si λ est un réel, on pose $\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$.

Les éléments de $\mathbb{R}^n$ sont appelés *vecteurs*, et les réels (λ) sont appelés *scalaires*.
Si $n = 0$, on pose $\mathbb{R}^n = \mathbb{R}^0 = \{0\}$ (l'ensemble contenant un unique élément, qui est 0).

Exemple A.2. $\mathbb{R} = \mathbb{R}^1$ peut être identifié à une droite.
$\mathbb{R}^2$ peut être identifié à un plan. Les vecteurs de $\mathbb{R}^4$ peuvent par exemple représenter la position d'un point en fonction du temps (le temps étant la quatrième composante).

Remarque A.3. On peut identifier naturellement $\mathbb{R}^n$, $\mathcal{M}_{1,n}(\mathbb{R})$ (matrices ligne) et $\mathcal{M}_{n,1}(\mathbb{R})$ (matrices colonne). On passe par exemple de $\mathbb{R}^2$ à $\mathcal{M}_{1,2}(\mathbb{R})$ en remplaçant le couple (x, y) par la matrice $\begin{pmatrix} x & y \end{pmatrix}$, et on passe de $\mathbb{R}^2$ à $\mathcal{M}_{2,1}(\mathbb{R})$ en remplaçant le couple (x, y) par la matrice $\begin{pmatrix} x \\ y \end{pmatrix}$.

Définition A.4. Soient n et m des entiers. Une *application linéaire* de $\mathbb{R}^n$ dans $\mathbb{R}^m$ est une application $f : \mathbb{R}^n \to \mathbb{R}^m$ qui vérifie:

- Pour tous $x, y \in \mathbb{R}^n$, on a $f(x + y) = f(x) + f(y)$
- Pour tous $x \in \mathbb{R}^n$ et λ réel, on a $f(\lambda x) = \lambda f(x)$.

Exemple A.5. Soit A une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$ et soit $X \in \mathcal{M}_{n,1}(\mathbb{R})$ une matrice colonne à n lignes. L'application

$$\begin{array}{rcl} \mathcal{M}_{n,1}(\mathbb{R}) & \to & \mathcal{M}_{m,1}(\mathbb{R}) \\ X & \mapsto & AX \end{array}$$

est une application linéaire. (Cela découle des propriétés du produit des matrices, et est laissé en exercice). Notons que le nombre n de colonnes de la matrice A est déterminé par le fait que le produit AX doit exister, et que le nombre m de lignes de A est déterminé par le nombre de composantes du vecteur d'arrivée (l'application arrive dans $\mathcal{M}_{m,1}(\mathbb{R})$).

Exemple A.6. L'application linéaire

$$\begin{array}{rcl} f : \mathcal{M}_{3,1}(\mathbb{R}) & \to & \mathcal{M}_{2,1}(\mathbb{R}) \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto & \begin{pmatrix} 2 & -1 & 0 \\ 4 & 3 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y \\ 4x + 3y - 5z \end{pmatrix} \end{array}$$

peut aussi s'écrire

$$\begin{array}{rcl} f : \mathbb{R}^3 & \to & \mathbb{R}^2 \\ (x, y, z) & \mapsto & (2x - y, 4x + 3y - 5z). \end{array}$$

Remarque A.7. On peut montrer (mais on ne le fera pas, et on ne l'utilisera pas vraiment) que toute application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ peut être représentée par une matrice comme dans l'exemple ci-dessus.

B Bases

Définition B.1. Etant donnés k réels $\lambda_1, \dots, \lambda_k$ et k vecteurs $v_1, \dots, v_k$ de $\mathbb{R}^n$, le vecteur $u = \lambda_1 v_1 + \cdots + \lambda_k v_k$ est appelé *combinaison linéaire* de $v_1, \dots, v_k$ à coefficients $\lambda_1, \dots, \lambda_k$.

Si u appartient à $\mathbb{R}^n$, on dit que u est une *combinaison linéaire* des vecteurs $v_1, \dots, v_k$ s'il existe des réels $\lambda_1, \dots, \lambda_k$ tels que $u = \lambda_1 v_1 + \cdots + \lambda_k v_k$.

Exemple B.2. Dans $\mathbb{R}^2$, on a $(2,3) = 2(1,-1) + 5(0,1)$: le vecteur $(2,3)$ est une combinaison linéaire des vecteurs $(1,-1)$ et $(0,1)$.

Définition B.3. On dit que des vecteurs $v_1, \dots, v_k$ *engendrent* $\mathbb{R}^n$ (ou forment un *système générateur* ou une *famille génératrice* de $\mathbb{R}^n$) si *tout* vecteur de $\mathbb{R}^n$ peut s'écrire comme une combinaison linéaire des vecteurs $v_1, \dots, v_k$.

Exemples B.4. (a) Dans $\mathbb{R}^2$, soit $v = (1,0)$. Les combinaisons linéaires de v sont les vecteurs de la forme $\lambda v = (\lambda, 0)$ avec $\lambda \in \mathbb{R}$. Donc en particulier, le vecteur $(1,1)$ de $\mathbb{R}^2$ n'est pas une combinaison linéaire de v : le vecteur v n'engendre pas $\mathbb{R}^2$.

(b) Toujours dans $\mathbb{R}^2$, soient $v_1 = (1,0)$ et $v_2 = (2,1)$. Soit (x,y) un vecteur quelconque de $\mathbb{R}^2$. Alors $(x,y) = (x-2y)v_1 + y v_2$ est une combinaison linéaire de v_1 et v_2. C'est vrai pour n'importe quel vecteur de $\mathbb{R}^2$, donc v_1 et v_2 engendrent $\mathbb{R}^2$.

Remarque B.5. On se place dans $\mathbb{R}^n$. Posons $\vec{e}_1 = (1, 0, \dots, 0)$, $\vec{e}_2 = (0, 1, 0, \dots, 0)$, et ainsi de suite jusqu'à $\vec{e}_n = (0, 0, \dots, 0, 1)$, de sorte que $\vec{e}_i$ est le n-uplet dont la i-ème composante est 1 et toutes les autres composantes sont des 0.

Prenons par exemple $n = 3$. Alors $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ et $\vec{e}_3 = (0,0,1)$. Soit $(2,-3,1)$ un élément de $\mathbb{R}^3$. On peut écrire $(2,-3,1) = 2\vec{e}_1 + (-3)\vec{e}_2 + 1\vec{e}_3$.

Plus généralement, soit $x = (x_1, x_2, \dots, x_n)$ un élément de $\mathbb{R}^n$. On constate que l'on peut écrire $x = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$. Donc $\vec{e}_1, \dots, \vec{e}_n$ engendrent $\mathbb{R}^n$. De plus, c'est la seule manière dont x peut s'écrire comme combinaison linéaire de $\vec{e}_1, \dots, \vec{e}_n$.

Définition B.6. On dit que des vecteurs $v_1, \dots, v_k$ sont *linéairement indépendants* (on dit aussi que la famille $\{v_1, \dots, v_k\}$ est *libre*) si aucun de ces vecteurs n'est combinaison linéaire des autres. Dans le cas contraire, on dit qu'ils sont *linéairement dépendants* ou *liés*.

Dans la pratique

Pour déterminer si une famille $v_1, \dots, v_k$ de vecteurs de $\mathbb{R}^n$ est libre ou liée, on résout l'équation en $\lambda_1, \dots, \lambda_k$:

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = (0, 0, \dots, 0).$$

C'est en réalité un système de n équations, en regardant les vecteurs composante à composante. On constate qu'il y a une solution évidente qui est $\lambda_1 = 0, \dots, \lambda_k = 0$. Alors:

- Les vecteurs $v_1, \dots, v_k$ sont linéairement indépendants si et seulement si $\lambda_1 = 0, \dots, \lambda_k = 0$ est l'unique solution.

- Les vecteurs $v_1, \dots, v_k$ sont linéairement dépendants si et seulement s'il existe une autre solution (qui n'est pas $\lambda_1 = 0, \dots, \lambda_k = 0$).

Définition B.7. Dans le deuxième cas (linéairement dépendants), on en déduit une *relation de dépendance* entre les vecteurs (c'est-à-dire que l'on peut en exprimer certains en fonction des autres).

Exemples B.8. (a) Dans $\mathbb{R}^2$, soient $v_1 = (1, 0)$ et $v_2 = (2, 1)$. Si $\lambda_1 v_1 + \lambda_2 v_2 = (0, 0)$, alors $(\lambda_1 + 2\lambda_2, \lambda_2) = (0, 0)$ et donc $\lambda_1 + 2\lambda_2 = 0$ et $\lambda_2 = 0$, donc $\lambda_1 = 0 = \lambda_2$ est la seule solution. Donc v_1 et v_2 sont linéairement indépendants.

(b) Dans $\mathbb{R}^2$, soient $v_1 = (1, 0)$ et $v_2 = (2, 0)$. Si $\lambda_1 v_1 + \lambda_2 v_2 = (0, 0)$, alors $(\lambda_1 + 2\lambda_2, 0) = (0, 0)$ et donc $\lambda_1 + 2\lambda_2 = 0$. Ceci a une infinité de solutions, par exemple on peut prendre $\lambda_2 = 1$ et $\lambda_1 = -2$. Donc v_1 et v_2 sont liés, et $-2v_1 + v_2 = (0, 0)$ (soit $v_2 = 2v_1$).

Définition B.9. On dit qu'une famille de vecteurs $v_1, v_2, \ldots, v_k$ de $\mathbb{R}^n$ est une *base* de $\mathbb{R}^n$ si c'est une famille libre de $\mathbb{R}^n$ et si $v_1, \ldots, v_k$ engendrent $\mathbb{R}^n$.

Exemple B.10. $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ est une base de $\mathbb{R}^n$. Elle est appelée la *base canonique* de $\mathbb{R}^n$.

Théorème B.11. Toutes les bases de $\mathbb{R}^n$ ont n éléments.

Théorème B.12. Soit $\{v_1, \ldots, v_n\}$ une famille de vecteurs de $\mathbb{R}^n$ à n éléments. Si la famille est libre, alors c'est une base.

Proposition B.13. Si $\{v_1, \ldots, v_n\}$ est une base de $\mathbb{R}^n$, alors tout vecteur de $\mathbb{R}^n$ s'écrit de manière *unique* comme combinaison linéaire de $v_1, \ldots, v_n$. Les coefficients de cette combinaison linéaire sont appelés *coordonnées*.

Exemple B.14. Dans $\mathbb{R}^2$, soient $v_1 = (1, 0)$ et $v_2 = (2, 1)$. Ce sont deux vecteurs de $\mathbb{R}^2$, et on a vu qu'ils sont linéairement indépendants, donc ils forment une base de $\mathbb{R}^2$. On peut voir que si $x = (x_1, x_2)$ est un vecteur quelconque de $\mathbb{R}^2$, alors $x = (x_1 - 2x_2)v_1 + x_2 v_2$, donc les coordonnées de x dans la base $\{v_1, v_2\}$ sont $x_1 - 2x_2$ et x_2.

IV Déterminants et systèmes linéaires

A Systèmes de deux équations à deux inconnues

On considère le système
$$\begin{cases} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{cases}$$
Nous avons déjà vu qu'il est équivalent à l'équation matricielle $AX = B$ où A est la matrice dy système, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, X est le vecteur colonne $\begin{pmatrix} x \\ y \end{pmatrix}$ et B est le vecteur colonne $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Ceci est équivalent à l'équation à coefficients vectoriels
$$\vec{v}_1 x + \vec{v}_2 y = \vec{b}$$
où $\vec{v}_1$, $\vec{v}_2$ et $\vec{b}$ sont les vecteurs suivants: $\vec{v}_1 = (a_{11}, a_{21})$, $\vec{v}_2 = (a_{12}, a_{22})$ et $\vec{b} = (b_1, b_2)$; on cherche donc à écrire $\vec{b}$ comme combinaison linéaire de $\vec{v}_1$ et $\vec{v}_2$.

On peut interpréter ce système géométriquement: on considère les deux droites suivantes du plan: D_1 la droite d'équation $a_{11}x + a_{12}y = b_1$ (ou $y = -\frac{a_{11}}{a_{12}}x + \frac{b_1}{a_{12}}$) et D_2 d'équation $a_{21}x + a_{22}y = b_2$ (ou $y = -\frac{a_{21}}{a_{22}}x + \frac{b_2}{a_{22}}$). Les solutions du système sont les points d'intersection de ces deux droites. Ces droites sont parallèles (distinctes ou confondues) ou bien se coupent en un point exactement. Le système a une unique solution si et seulement si les deux droites se coupent, c'est-à-dire si et seulement si leurs pentes sont différentes (puisque si les pentes sont égales, les droites sont parallèles). La pente de D_1 est $-\frac{a_{11}}{a_{12}}$ et la pente de D_2 est $-\frac{a_{21}}{a_{22}}$.

Donc le système a une solution unique si et seulement si $a_{11}a_{22} - a_{21}a_{12} \neq 0$.

Définition A.1. **(i)** Soit A la matrice $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. On appelle *déterminant* de A le réel $\det(A) = a_{11}a_{22} - a_{21}a_{12}$.

(ii) Soient $\vec{v}_1$ et $\vec{v}_2$ les vecteurs suivants: $\vec{v}_1 = (a_{11}, a_{21})$ et $\vec{v}_2 = (a_{12}, a_{22})$ (colonnes de A). On appelle *déterminant* de $\vec{v}_1$ et $\vec{v}_2$ le réel $\det(\vec{v}_1, \vec{v}_2) = a_{11}a_{22} - a_{21}a_{12}$.

Proposition A.2. $\det(\vec{v}_1, \vec{v}_2)$ est l'aire algébrique du parallélogramme déterminé par $\vec{v}_1$ et $\vec{v}_2$ (dans cet ordre).

On peut alors résumer et compléter ce que nous avons dit plus haut:

Proposition A.3. Avec les notations précédentes, $\det(A) \neq 0$ si et seulement si pour tous b_1, b_2, il existe une unique solution au système linéaire ci-dessus ce qui est équivalent à dire que $\{\vec{v}_1, \vec{v}_2\}$ est une base de $\mathbb{R}^2$. Cette solution est alors donnée par
$$x = \frac{\det(\vec{b}, \vec{v}_2)}{\det(\vec{v}_1, \vec{v}_2)} \qquad y = \frac{\det(\vec{v}_1, \vec{b})}{\det(\vec{v}_1, \vec{v}_2)}.$$
De plus, $\det(A) \neq 0$ si et seulement si A est inversible.

Exemple A.4. Soit à résoudre le système $\begin{cases} 2x - y = 3 \\ x + 3y = 0 \end{cases}$.

La matrice associée au système est $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$, dont le déterminant est $\det(A) = 2 \cdot 3 - 1 \cdot (-1) = 7 \neq 0$. Donc le système a une solution unique:

$$\begin{cases} x = \frac{1}{7} \det \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} = \frac{1}{7}(3 \cdot 3 - 0 \cdot (-1)) = \frac{9}{7} \\ y = \frac{1}{7} \det \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} = \frac{1}{7}(2 \cdot 0 - 1 \cdot 3) = -\frac{3}{7}. \end{cases}$$

De plus, la matrice A est inversible puisque $\det(A) \neq 0$.

B Systèmes de trois équations à trois inconnues

Un système linéaire

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases} \tag{B.1}$$

peut s'écrire matriciellement sous la forme $AX = B$.

On peut aussi l'écrire sous forme d'une équation linéaire à coefficients vectoriels $x\vec{v}_1 + y\vec{v}_2 + z\vec{v}_3 = \vec{b}$, avec $\vec{v}_1 = (a_{11}, a_{21}, a_{31})$, $\vec{v}_2 = (a_{12}, a_{22}, a_{32})$, $\vec{v}_3 = (a_{13}, a_{23}, a_{33})$ et $\vec{b} = (b_1, b_2, b_3)$. Donc résoudre le système revient à écrire $\vec{b}$ comme combinaison linéaire de $\vec{v}_1$, $\vec{v}_2$ et $\vec{v}_3$.

Donc dire que pour tout $(b_1, b_2, b_3) \in \mathbb{R}^3$ il existe une unique solution au système est équivalent à dire que tout $\vec{b} \in \mathbb{R}^3$ s'écrit de façon unique comme combinaison linéaire de $\vec{v}_1$, $\vec{v}_2$ et $\vec{v}_3$, ce qui revient à dire que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ est une base de $\mathbb{R}^3$.

Définition géometrique

Soit $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ un système de trois vecteurs de $\mathbb{R}^3$.

$\det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est le volume algébrique du parallélépipède construit sur $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$. Donc $|\det(\vec{v}_1, \vec{v}_2, \vec{v}_3)|$ est le volume du parallélépipède, et on met un signe $+$ si $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est *direct*, un signe $-$ si $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est *indirect*.

Pour déterminer si un système est direct, on utilise la règle du *bonhomme d'Ampère*: on place les vecteurs de façon à ce qu'ils aient tous la même origine. On imagine une personne le long de $\vec{v}_1$ (avec les pieds à l'origine du vecteur) regardant $\vec{v}_2$. Le système $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ est alors direct si et seulement si $\vec{v}_3$ est à *gauche* de la personne.

B.1 Propriétés du déterminant

(i) $\det(\vec{e}_1, \vec{e}_2, \vec{e}_3) = 1$ (où $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ est la base canonique de $\mathbb{R}^3$). *[C'est le volume du cube de côté de longueur 1].*

(ii) $\det(\vec{v}_1, \vec{v}_2, \vec{v}_3) = 0$ si et seulement si $\vec{v}_1, \vec{v}_2, \vec{v}_3$ sont liés.

En particulier, si deux des vecteurs sont égaux, le déterminant est nul.

(iii) Le déterminant est une forme *multilinéaire*: chacune des applications $\varphi_1 : \vec{v}_1 \mapsto \det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ ($\vec{v}_2$ et $\vec{v}_3$ fixés), $\varphi_2 : \vec{v}_2 \mapsto \det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ ($\vec{v}_1$ et $\vec{v}_3$ fixés), et $\varphi_3 : \vec{v}_3 \mapsto \det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ ($\vec{v}_1$ et $\vec{v}_2$ fixés), est une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}$.

Par exemple, pour la première composante, on a :

$$\det(\vec{v}_1 + \vec{v}_1{}', \vec{v}_2, \vec{v}_3) = \det(\vec{v}_1, \vec{v}_2, \vec{v}_3) + \det(\vec{v}_1{}', \vec{v}_2, \vec{v}_3)$$
$$\det(\lambda \vec{v}_1, \vec{v}_2, \vec{v}_3) = \lambda \det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$$

(iv) Le déterminant est une forme *alternée*: si on permute deux vecteurs, on change le signe du déterminant. Par exemple, $\det(\vec{v}_2, \vec{v}_1, \vec{v}_3) = -\det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$.

Théorème B.1. Les propriétés ci-dessus déterminent $\det(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ de façon unique pour tout système de vecteurs $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Essayons de comprendre pourquoi sur un exemple. Prenons $\vec{v}_1 = (1, 1, 0)$, $\vec{v}_2 = (-2, 0, -1)$ et $\vec{v}_3 = (3, 0, 0)$.

Pour pouvoir utiliser **(i)**, on va écrire ces vecteurs dans la base canonique. On a $\vec{v}_1 = \vec{e}_1 + \vec{e}_2$, $\vec{v}_2 = -2\vec{e}_1 - \vec{e}_3$ et $\vec{v}_3 = 3\vec{e}_1$.

$$\begin{aligned}
\det(\vec{v}_1, \vec{v}_2, \vec{v}_3) &= \det(\vec{e}_1 + \vec{e}_2, -2\vec{e}_1 - \vec{e}_3, 3\vec{e}_1) \\
&\overset{\text{(iii)}}{=} \det(\vec{e}_1, -2\vec{e}_1 - \vec{e}_3, 3\vec{e}_1) + \det(\vec{e}_2, -2\vec{e}_1 - \vec{e}_3, 3\vec{e}_1) \\
&\overset{\text{(ii)}}{=} 0 + \det(\vec{e}_2, -2\vec{e}_1 - \vec{e}_3, 3\vec{e}_1) \\
&\overset{\text{(iii)}}{=} -2\det(\vec{e}_2, \vec{e}_1, 3\vec{e}_1) - \det(\vec{e}_2, \vec{e}_3, 3\vec{e}_1) \\
&\overset{\text{(ii)}}{=} -\det(\vec{e}_2, \vec{e}_3, 3\vec{e}_1) \\
&\overset{\text{(iii)}}{=} -3\det(\vec{e}_2, \vec{e}_3, \vec{e}_1) \\
&\overset{\text{(iv)}}{=} -3[-\det(\vec{e}_1, \vec{e}_3, \vec{e}_2)] \\
&= 3\det(\vec{e}_1, \vec{e}_3, \vec{e}_2) \\
&\overset{\text{(iv)}}{=} -3\det(\vec{e}_1, \vec{e}_2, \vec{e}_3) \\
&\overset{\text{(i)}}{=} -3 \cdot 1 \\
&= -3
\end{aligned}$$

Définition B.2. Le déterminant d'une matrice est le déterminant de la famille de ses vecteurs colonne.

Comment calculer un déterminant dans la pratique?

B.2 Développement du déterminant suivant une ligne ou une colonne

On se ramène à des déterminants de matrices 2×2: par exemple, le développement suivant la première colonne donne:

$$\begin{aligned}
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &\overset{\text{notation}}{=} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}
\end{aligned}$$

La règle des signes est donnée par la matrice $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$.

Exemple B.3. On développe le déterminant suivant suivant la troisième colonne, pour profiter du 0:

$$\begin{vmatrix} 1^{(+)} & -6^{(-)} & 3^{(+)} \\ 2 & 4 & 0^{(-)} \\ -3 & 5 & 7^{(+)} \end{vmatrix} = 3\begin{vmatrix} 2 & 4 \\ -3 & 5 \end{vmatrix} - 0\begin{vmatrix} 1 & -6 \\ -3 & 5 \end{vmatrix} + 7\begin{vmatrix} 1 & -6 \\ 2 & 4 \end{vmatrix}$$
$$= 3(10-(-12)) + 7(4-(-12))$$
$$= 66 + 112 = 178.$$

Proposition B.4. $\det(A) = \det(A^t)$

Remarque B.5. On peut donc remplacer dans tout ce qu'on a dit les colonnes par les lignes.

B.3 Systèmes de Cramer

Théorème B.6. Les énoncés suivants sont équivalents:

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

(ii) Pour tout $(b_1, b_2, b_3) \in \mathbb{R}^3$, le système $(B.1)$ admet une unique solution.

(iii) $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ est une base de $\mathbb{R}^3$.

(iv) A est inversible.

Si l'une des propositions équivalentes ci-dessus est vérifiée, on dit que le système $(B.1)$ est un *système de Cramer*, et l'unique solution de ce système est:

$$x = \frac{\det(\vec{b}, \vec{v}_2, \vec{v}_3)}{\det(A)}, \quad y = \frac{\det(\vec{v}_1, \vec{b}, \vec{v}_3)}{\det(A)}, \quad z = \frac{\det(\vec{v}_1, \vec{v}_2, \vec{b})}{\det(A)}.$$

Exemple B.7. Soit $A = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 4 \\ 1 & 2 & -1 \end{pmatrix}$.

Alors, en développant suivant la première ligne,

$$\det(A) = \begin{vmatrix} 2 & 1 & 3 \\ 3 & 2 & 4 \\ 1 & 2 & -1 \end{vmatrix} = 2\begin{vmatrix} 2 & 4 \\ 2 & -1 \end{vmatrix} - 1\begin{vmatrix} 3 & 4 \\ 1 & -1 \end{vmatrix} + 3\begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix}$$
$$= 2(-10) - (-7) + 3 \cdot 4 = -1 \neq 0.$$

Donc A est inversible et le système

$$\begin{cases} 2x + y + 3z = 13 \\ 3x + 2y + 4z = 4 \\ x + 2y - z = -7 \end{cases}$$

(dont la matrice est A) a une unique solution (il est de Cramer).

Cette solution est donnée par:

$$x = \frac{\begin{vmatrix} 13 & 1 & 3 \\ 4 & 2 & 4 \\ -7 & 2 & -1 \end{vmatrix}}{-1} = 88,$$

$$y = \frac{\begin{vmatrix} 2 & 13 & 3 \\ 3 & 4 & 4 \\ 1 & -7 & -1 \end{vmatrix}}{-1} = -64$$

$$z = \frac{\begin{vmatrix} 2 & 1 & 13 \\ 3 & 2 & 4 \\ 1 & 2 & -7 \end{vmatrix}}{-1} = -33$$

C Systèmes de n équations à n inconnues

C.1 Déterminants et systèmes de Cramer

On considère le système

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right. \tag{C.1}$$

Si on pose $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, le système $(C.1)$ est équivalent à l'équation matricielle $AX = B$.

Si on pose $\vec{v}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} = a_{11}\vec{e}_1 + \cdots + a_{n1}\vec{e}_n$, $\ldots$, $\vec{v}_n = a_{1n}\vec{e}_1 + \cdots + a_{nn}\vec{e}_n$ (le vecteur $\vec{v}_i$ est la i^{me} colonne de A) et $\vec{b} = b_1\vec{e}_1 + \cdots + b_n\vec{e}_n$, le système $(C.1)$ est équivalent à l'équation à coefficients vectoriels $x_1\vec{v}_1 + \cdots + x_n\vec{v}_n = \vec{b}$.

Définition C.1. Le *déterminant* d'une famille de n vecteurs $\vec{v}_1, \ldots, \vec{v}_n$ de $\mathbb{R}^n$, noté $\det(\vec{v}_1, \ldots, \vec{v}_n)$, est l'unique réel possédant les propriétés suivantes:

- $\det(\vec{e}_1, \ldots, \vec{e}_n) = 1$,
- déterminant est une forme multilinéaire (c'est-à-dire linéaire en chaque variable)
- déterminant est une forme alternée.

Conséquence C.2. (i) Si deux vecteurs sont identiques, le déterminant est nul:

$$\det(\cdots, \vec{v}, \cdots, \vec{v}, \cdots) = -\det(\cdots, \vec{v}, \cdots, \vec{v}, \cdots),$$

on obtient deux nombres réels opposés et égaux donc nuls.

(ii) Si la famille de vecteurs $\{\vec{v}_1, \ldots, \vec{v}_n\}$ est liée, alors $\det(\vec{v}_1, \ldots, \vec{v}_n) = 0$.

Proposition C.3. La famille de vecteurs $\{\vec{v}_1, \dots, \vec{v}_n\}$ est liée si et seulement si $\det(\vec{v}_1, \dots, \vec{v}_n) = 0$.

Définition C.4. On pose $\det(A) = \det(\vec{v}_1, \dots, \vec{v}_n)$ en identifiant la matrice à la famille de ses vecteurs colonne.

Théorème C.5. Les énoncés suivants sont équivalents:

(i) $\det(A) \neq 0$

(ii) $\{\vec{v}_1, \dots, \vec{v}_n\}$ est une base de $\mathbb{R}^n$

(iii) Pour tout $(b_1, \dots, b_n) \in \mathbb{R}^n$, il existe une unique solution $(x_1, \dots, x_n)$ au système $(C.1)$

(iv) A est inversible

Le système $(C.1)$ est alors appelé *système de Cramer*.

Proposition C.6. Formules de Cramer: Supposons que $(C.1)$ soit un système de Cramer. Alors l'unique solution est donnée par

$$\begin{aligned}
x_1 &= \frac{\det(\vec{b}, \vec{v}_2, \dots, \vec{v}_n)}{\det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)} = \frac{\det(\vec{b}, \vec{v}_2, \dots, \vec{v}_n)}{\det(A)} \\
x_2 &= \frac{\det(\vec{v}_1, \vec{b}, \dots, \vec{v}_n)}{\det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)} = \frac{\det(\vec{v}_1, \vec{b}, \dots, \vec{v}_n)}{\det(A)} \\
&\cdots \\
x_n &= \frac{\det(\vec{v}_1, \vec{v}_2, \dots, \vec{b})}{\det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)} = \frac{\det(\vec{v}_1, \vec{v}_2, \dots, \vec{b})}{\det(A)}
\end{aligned}$$

Proposition C.7. Soient A et B deux matrices carrées $n \times n$. Alors

$$\det(AB) = \det(A)\det(B)$$
$$\det(A^t) = \det(A)$$
$$\det(I_n) = 1.$$

C.2 Calcul pratique des déterminants

On développe suivant une ligne ou une colonne en utilisant la *règle des signes* similaire au cas 3×3: les signes affectés à deux coefficients voisins sont différents lorsqu'on développe, en commençant par le signe $+$ pour le coefficient de la première ligne et de la première colonne.

De plus on utilise la règle suivante: on ne change pas la valeur du déterminant en ajoutant à une colonne une combinaison linéaire des *autres* colonnes (de même avec les lignes).

Ceci permet de faire apparaître des zéros dans une même ligne ou colonne. Cela simplifie ensuite le calcul du déterminant:

Exemple C.8. Soit à calculer $\begin{vmatrix} 5 & -2 & 7 & 3 \\ 3 & 4 & 2 & 1 \\ -4 & -1 & 0 & 7 \\ 1 & 3 & 1 & -2 \end{vmatrix}$. On choisit un "pivot" (de préférence le plus simple possible et/ou dans une ligne ou une colonne ou il y a déjà des 0), ici le 1, et on "fait apparaître" des 0 dans la colonne ou la ligne où il se trouve:

$$\begin{vmatrix} 5 & -2 & 7 & 3 \\ 3 & 4 & 2 & 1 \\ -4 & -1 & 0 & 7 \\ 1 & 3 & 1 & -2 \end{vmatrix} = \begin{vmatrix} -2 & -23 & 0 & 17 \\ 1 & -2 & 0 & 5 \\ -4 & -1 & 0 & 7 \\ 1 & 3 & 1 & -2 \end{vmatrix}$$

en remplaçant la ligne L_1 par $L_1 - 7L_4$ et la ligne L_2 par $L_2 - 2L_4$

En développant suivant la troisième colonne, on obtient

$$\begin{aligned}
\begin{vmatrix} -2 & -23 & 0 & 17 \\ 1 & -2 & 0 & 5 \\ -4 & -1 & 0 & 7 \\ 1 & 3 & 1 & -2 \end{vmatrix} &= -1 \begin{vmatrix} -2 & -23 & 17 \\ 1 & -2 & 5 \\ -4 & -1 & 7 \end{vmatrix} \\
&= - \begin{vmatrix} -1 & -27 & 27 \\ 1 & 0 & 0 \\ -4 & -9 & -27 \end{vmatrix} \\
& \quad C_2 \to C_2 + 2C_1, \quad C_3 \to C_3 - 5C_1 \\
&= - \left(-1 \begin{vmatrix} -27 & 27 \\ -9 & -27 \end{vmatrix} \right) \\
&= (-27)^2 - (-9)(27) = -486
\end{aligned}$$

en développant suivant la deuxième ligne.

Attention!
Il y a des similarités mais aussi des *différences* entre les opérations que l'on peut faire sur les déterminants et la méthode de Gauss pour résoudre les systèmes ou inverser les matrices.

- On peut faire des opérations sur les colonnes d'un déterminant, on ne peut pas faire d'opérations sur les colonnes dans la méthode de Gauss.

- On peut multiplier une ligne par un réel non nul dans la méthode de Gauss, mais si on le fait dans un déterminant on multiplie *tout* le déterminant par ce réel. En particulier, dans le calcul d'un déterminant, il ne faut pas faire une opération du type $L_2 \to 3L_1 + 4L_4$, car on multiplie alors le déterminant par 3.

- On peut échanger des lignes dans la méthode de Gauss, mais si on échange des lignes (ou des colonnes) d'un déterminant, on change le signe du déterminant.

V Diagonalisation de matrices: valeurs propres et vecteurs propres

A Introduction

Le but de ce chapitre est de se "ramener", lorsque c'est possible, d'une matrice carrée quelconque à une matrice *diagonale*, c'est-à-dire de la forme

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & & \cdots & & 0 \\ 0 & \cdots & 0 & 0 & a_{nn} \end{pmatrix}$$

B Diagonalisation

Définition B.1. On dit qu'une matrice A est *diagonalisable* s'il existe une matrice inversible P telle que $P^{-1}AP$ soit diagonale.

Définition B.2. Soit A une matrice $n \times n$. On dit qu'un réel λ est une *valeur propre* pour A s'il existe une matrice colonne V dans $\mathcal{M}_{n,1}(\mathbb{R})$ (c'est-à-dire un vecteur de $\mathbb{R}^n$) qui soit *non nulle* et telle que $AV = \lambda V$. Le vecteur V est alors appelé *vecteur propre* associé à la valeur propre λ.

Remarque B.3. Si V est un vecteur propre de A associé à la valeur propre λ, et si α est un réel quelconque non nul, alors αV est aussi un vecteur propre de A associé à la valeur propre λ.

En effet, $A(\alpha V) = \alpha AV = \alpha(\lambda V) = \lambda(\alpha V)$.

Exemple B.4. Soit $A = \begin{pmatrix} 1 & 2 \\ -3 & -6 \end{pmatrix}$ et cherchons les valeurs propres de A et les vecteurs propres de A.

Soit λ une valeur propre de A. Alors il existe un vecteur propre associé $V = \begin{pmatrix} x \\ y \end{pmatrix}$, où x et y ne sont pas tous les deux nuls, qui vérifie donc $AV = \lambda V$. Or $AV = \begin{pmatrix} x + 2y \\ -3x - 6y \end{pmatrix}$ et $\lambda V = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$, d'où le système $\left\{ \begin{array}{l} x + 2y = \lambda x \\ -3x - 6y = \lambda y \end{array} \right.$. On a donc $\begin{cases} x + 2y = \lambda x \\ 3\lambda x + \lambda y = 0 \quad L_2 \to L_2 + 3L_1 \end{cases}$

On a alors deux cas:

- Soit $\lambda \neq 0$; alors le système devient $\begin{cases} x + 2y = \lambda x \\ 3x + y = 0 \end{cases}$ avec $x \neq 0$ (sinon $y = 0$ aussi et alors $V = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$). Alors $y = -3x$ et $-5x = \lambda x$, d'où $\lambda = -5$.

  On vérifie que -5 est bien valeur propre, avec comme vecteur propre associé par exemple $\begin{pmatrix} 1 \\ -3 \end{pmatrix}$ (obtenu en choisissant $x = 1$).

- Soit $\lambda = 0$, et alors le système devient $x = -2y$. On vérifie que 0 est bien valeur propre, avec comme vecteur propre associé par exemple $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ (obtenu en choisissant $y = 1$).

La méthode décrite dans l'exemple ci-dessus n'est pas systématique. Pour faciliter le calcul des valeurs propres, on a la proposition suivante:

Proposition B.5. (Calcul des valeurs propres) Un réel λ est valeur propre pour la matrice A si et seulement si $\det(A - \lambda I_n) = 0$.

Exemple B.6. Dans l'exemple précédent: soit λ un réel. Alors

$$\begin{aligned}\det(A - \lambda I_2) &= \begin{vmatrix} 1-\lambda & 2 \\ -3 & -6-\lambda \end{vmatrix} = (1-\lambda)(-6-\lambda) + 6 \\ &= -6 + 5\lambda + \lambda^2 + 6 = \lambda(\lambda + 5).\end{aligned}$$

Il y a donc deux valeurs propres: $\lambda = 0$ et $\lambda = -5$, les racines de $\lambda(\lambda + 5) = 0$.

Théorème B.7. Une matrice A de $\mathcal{M}_{n,n}(\mathbb{R})$ est diagonalisable si et seulement s'il existe une base de $\mathbb{R}^n$ formée de vecteurs propres pour A. Une matrice P telle que $P^{-1}AP$ soit diagonale est alors obtenue en écrivant les vecteurs de cette base en colonnes successives.

Exemple B.8. Dans l'exemple précédent, la famille $\left\{ \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ est une base de $\mathbb{R}^2$ (le déterminant formé par les deux vecteurs est égal à -5 donc est non nul), donc A est diagonalisable.

Posons $P = \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix}$ et calculons P^{-1}:

- $\det(P) = -5$ donc P est bien inversible.
- $$(P \mid I_2) = \left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ -3 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \underset{L_2 \to L_2 + 3L_1}{\left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & -5 & 3 & 1 \end{array}\right)}$$
  $$\rightsquigarrow \underset{L_2 \to L_2/(-5)}{\left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 1 & -\frac{3}{5} & -\frac{1}{5} \end{array}\right)} \rightsquigarrow \underset{L_1 \to L_1 + 2L_2}{\left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 1 & -\frac{3}{5} & -\frac{1}{5} \end{array}\right)}$$
  La dernière matrice est échelonnée réduite et sa partie gauche est I_2 donc $P^{-1} = -\frac{1}{5}\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$.

Si on calcule $P^{-1}AP$, on obtient: $\begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix}$.

On remarque que la diagonale est formée des valeurs propres de A.

Remarque B.9. Si on change l'ordre des vecteurs propres de la base, on change P et la matrice diagonale, mais pas le fait que A soit diagonalisable, et on obtient toujours les valeurs propres sur la diagonale mais peut-être dans un ordre différent.

Exemple B.10. Dans l'exemple précédent, si on échange les deux vecteurs de base, on a alors $P = \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix}$ et $P^{-1}AP = \begin{pmatrix} 0 & 0 \\ 0 & -5 \end{pmatrix}$.

C Méthode pratique générale

(1) On calcule les valeurs propres à l'aide du determinant de $A - \lambda I_n$.

(2) On cherche ensuite les vecteurs propres associés (dans l'exemple, on cherche les vecteurs V tels que $AV = 0V$ et W tels que $AW = -5W$). Pour cela on résout des systèmes linéaires (sans λ!) dont les inconnues sont les coordonnées des vecteurs propres que l'on cherche. On applique pour cela la méthode du pivot de Gauss. Le système a toujours une infinité de solutions. *Le nombre de vecteurs propres linéairement indépendants qui sont associés à une valeur propre donnée est égal au nombre de paramètres* (c'est-à-dire au nombre d'inconnues qui sont arbitraires). Ce nombre est égal à

n - nombre d'équations non triviales restant à la fin de la résolution par la méthode de Gauss

(3) Une fois qu'on a trouvé tous les vecteurs propres qui sont linéairement indépendants, on cherche à voir s'ils forment une base de $\mathbb{R}^n$ (c'est-à-dire qu'il y a n vecteurs, si A est une matrice $n \times n$). Si c'est le cas, alors A est diagonalisable, sinon elle ne l'est pas.

Remarque C.1. Les vecteurs propres indépendants associés à une valeur propre donnée sont obtenus en fixant tour à tour chacun des paramètres égal à 1 (ou une valeur non nulle quelconque) et les autres égaux à 0.

Exemple C.2. Soit $A = \begin{pmatrix} 2 & 4 & -1 \\ -4 & -8 & 2 \\ 0 & 0 & -6 \end{pmatrix}$. Nous voulons savoir si A est diagonalisable.

Cherchons d'abord ses valeurs propres:

$$\begin{aligned} \det(A - \lambda I_3) &= \begin{vmatrix} 2-\lambda & 4 & -1 \\ -4 & -8-\lambda & 2 \\ 0 & 0 & -6-\lambda \end{vmatrix} \\ &= (-6-\lambda)\begin{vmatrix} 2-\lambda & 4 \\ -4 & -8-\lambda \end{vmatrix} \\ &= (-6-\lambda)\left((2-\lambda)(-8-\lambda) - (-4)\cdot 4\right) \\ &= (-6-\lambda)\lambda(\lambda+6) = -\lambda(\lambda+6)^2 \end{aligned}$$

en développant le premier déterminant suivant la troisième ligne.

Les valeurs propres sont donc 0 et -6.

Cherchons les vecteurs propres associés à la valeur propre 0 : nous cherchons les vecteurs $V = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ tels que $AV = 0.V = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

$$AV = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{aligned} 2x + 4y - z &= 0 \\ -4x - 8y + 2z &= 0 \\ -6z &= 0 \end{aligned} \right.$$

On résout:

$$\left(\begin{array}{ccc|c} 2 & 4 & -1 & 0 \\ -4 & -8 & 2 & 0 \\ 0 & 0 & -6 & 0 \end{array}\right)$$

$$L_2 \to L_2 + 2L_1 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 2 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -6 & 0 \end{array}\right)$$

$$L_3 \leftrightarrow L_2 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 2 & 4 & -1 & 0 \\ 0 & 0 & -6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$L_2 \to L_2/-6 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 2 & 4 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$L_1 \to L_1 + L_2 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 2 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$L_1 \to L_1/2 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Cette dernière matrice est échelonnée réduite, et correspond au système

$$\left\{\begin{array}{r} x + 2y = 0 \\ z = 0 \end{array}\right.$$

Donc les solutions sont les vecteurs V tels que $z = 0$ et $x+2y = 0$ c'est-à-dire $x = -2y$. Le nombre de paramètres est $3 - 2 = 1$. Il nous faut donc un vecteur propre. On choisit par exemple y comme paramètre (on peut aussi choisir x). Posons $y = 1$: le vecteur propre obtenu est $V = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$.

Cherchons maintenant les vecteurs propres associés à la valeur propre -6 : nous cherchons les vecteurs $W = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ tels que $AW = -6.W = \begin{pmatrix} -6x \\ -6y \\ -6z \end{pmatrix}$.

$$\begin{aligned} AW = -6W &\iff \left\{\begin{array}{r} 2x + 4y - z = -6x \\ -4x - 8y + 2z = -6y \\ -6z = -6z \end{array}\right. \\ &\iff \left\{\begin{array}{r} 8x + 4y - z = 0 \\ -4x - 2y + 2z = 0 \end{array}\right. \end{aligned}$$

On résout:

$$\left(\begin{array}{ccc|c} 8 & 4 & -1 & 0 \\ -4 & -2 & 2 & 0 \end{array}\right)$$

$$L_2 \to L_2 + \tfrac{1}{2}L_1 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 8 & 4 & -1 & 0 \\ 0 & 0 & \frac{3}{2} & 0 \end{array}\right)$$

$$L_2 \to \tfrac{2}{3}L_2 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 8 & 4 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

$$L_1 \to L_1 + L_2 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 8 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

$$L_1 \to L_1/8 \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

On obtient une matrice échelonnée réduite qui correspond au système

$$\begin{cases} x + \frac{1}{2}y = 0 \\ z = 0 \end{cases}$$

Les solutions sont donc les vecteurs W tels que $z = 0$ et $x + \frac{1}{2}y = 0$. Il y a donc un paramètre (x ou y), et il nous faut donc exhiber un vecteur propre, obtenu par exemple en posant $x = 1$. Donc $W = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$.

Les vecteurs propres indépendants sont donc V et W. Il n'y en a que deux, donc ils ne peuvent pas former une base de $\mathbb{R}^3$. Donc A n'est pas diagonalisable.

Exemple C.3. On considère la matrice $A = \begin{pmatrix} 0 & 2 & -6 & -9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Est-elle diagonalisable?

Calculons ses valeurs propres:

$$\begin{aligned} \det(A - \lambda I_4) &= \begin{vmatrix} -\lambda & 2 & -6 & -9 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} -\lambda & 2 & -6 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)^2 \begin{vmatrix} -\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = -\lambda(1-\lambda)^3 \end{aligned}$$

en développant successivement suivant la 4me ligne puis la 3me ligne. Donc les valeurs propres sont 0 et 1 (les racines de $-\lambda(1-\lambda)^3 = 0$).

Cherchons les vecteurs propres associés à la valeur propre 0 : ce sont les vecteurs $V = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ vérifiant $AV = 0V$, d'où le système $\begin{cases} 2y - 6z - 9t = 0 \\ y = 0 \\ z = 0 \\ t = 0 \end{cases} \iff \begin{cases} y = 0 \\ z = 0 \\ t = 0 \end{cases}$

Il y a donc un paramètre $(4-3)$, qui est x, il nous faut un vecteur propre, obtenu en posant par exemple $x=1$: $V = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Cherchons les vecteurs propres associés à la valeur propre 1 : ce sont les vecteurs $W = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ vérifiant $AW = W$, d'où le système

$$\begin{cases} 2y - 6z - 9t = x \\ y = y \\ z = z \\ t = t \end{cases} \iff \begin{cases} -x + 2y - 6z - 9t = 0 \end{cases}$$

Il y a donc 3 paramètres $(4-1)$, par exemple y, z et t, il nous faut 3 vecteurs propres, obtenus en posant successivement chacun des paramètres égal à 1, les autres égaux à 0; on obtient

$$W_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \; W_2 = \begin{pmatrix} -6 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \; W_3 = \begin{pmatrix} -9 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Les vecteurs $\{V, W_1, W_2, W_3\}$ forment un système de 4 vecteurs linéairement indépendants de $\mathbb{R}^4$, c'est donc une base de $\mathbb{R}^4$. Donc A est diagonalisable.

Une matrice inversible P telle que $P^{-1}AP$ soit diagonale est obtenue en écrivant les vecteurs propres que l'on a obtenus en colonnes:

$$P = \begin{array}{c} \begin{pmatrix} 1 & 2 & -6 & -9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \begin{matrix} V & W_1 & W_2 & W_3 \end{matrix} \end{array}$$

Calculons P^{-1} (facultatif, à moins que cela soit demandé ou que l'on veuille calculer des puissances de A – voir plus loin). Il nous faut transformer la matrice $(P \mid I_4)$ en une matrice échelonnée réduite. Elle est déjà échelonnée. En appliquant successivement les opérations $L_1 \to L_1 + 9L_4$, $L_1 \to L_1 + 6L_3$, puis $L_1 \to L_1 - 2L_2$, on obtient

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -2 & 6 & 9 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

donc

$$P^{-1} = \begin{pmatrix} 1 & -2 & 6 & 9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

On a $P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

D Applications

Application D.1. Soit A une matrice diagonalisable. Il existe donc une matrice inversible P telle que $D = P^{-1}AP$ soit diagonale. On a donc $PDP^{-1} = PP^{-1}APP^{-1} = I_nAI_n = A$.

Notons $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Si on calcule D^2, on voit qu'elle est égale à

$$D^2 = \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix}.$$

De proche en proche, on obtient

$$D^r = \begin{pmatrix} \lambda_1^r & & 0 \\ & \ddots & \\ 0 & & \lambda_n^r \end{pmatrix}$$

pour tout entier $r \geqslant 1$.

On veut calculer A^r pour un entier $r \geqslant 1$ quelconque. On a tout d'abord

$$\begin{aligned} A^2 = AA &= (PDP^{-1})(PDP^{-1}) \\ &= PD(P^{-1}P)DP^{-1} \\ &= PDI_nDP^{-1} \\ &= PD^2P^{-1} \end{aligned}$$

et on connaît D^2. Il suffit ensuite de faire le produit.

De proche en proche, on a

$$A^r = PD^rP^{-1}.$$

Exemple D.2. Soit $A = \begin{pmatrix} -1 & 12 \\ 4 & 1 \end{pmatrix}$. Nous voulons calculer A^{132}. Voyons d'abord si A est bien diagonalisable.

Cherchons les valeurs propres de A :

$$\begin{aligned} \det(A - \lambda I_2) = \begin{vmatrix} -1-\lambda & 12 \\ 4 & 1-\lambda \end{vmatrix} &= (-1-\lambda)(1-\lambda) - 4 \cdot 12 \\ &= \lambda^2 - 49 \\ &= (\lambda - 7)(\lambda + 7) \end{aligned}$$

donc les valeurs propres sont 7 et -7.

Recherchons les vecteurs propres associés à la valeur propre 7 : on recherche les vecteurs $V = \begin{pmatrix} x \\ y \end{pmatrix}$ tels que $AV = 7V$, c'est-à-dire les solutions du système $\begin{cases} -x + 12y = 7x \\ 4x + y = 7y \end{cases}$ qui est équivalent à $2x-3y=0$. Il y a donc un paramètre, par exemple y, que l'on peut choisir égal à 2 par exemple, d'où $V=\begin{pmatrix}3\\2\end{pmatrix}$.

Recherchons les vecteurs propres associés à la valeur propre -7 : on recherche les vecteurs $W=\begin{pmatrix}x\\y\end{pmatrix}$ tels que $AW=-7W$, c'est-à-dire les solutions du système $\begin{cases}-x+12y=-7x\\4x+y=-7y\end{cases}$ qui est équivalent à $x+2y=0$. Il y a donc un paramètre, par exemple y, que l'on peut choisir égal à 1 par exemple, d'où $W=\begin{pmatrix}-2\\1\end{pmatrix}$.

Puisqu'il y a deux vecteurs propres indépendants dans $\mathbb{R}^2$, ils forment une base de $\mathbb{R}^2$, et donc A est diagonalisable. On peut prendre $P=\begin{pmatrix}3&-2\\2&1\end{pmatrix}$, qui est inversible.

On a alors $D=P^{-1}AP=\begin{pmatrix}7&0\\0&-7\end{pmatrix}$.

Puisque $P^{-1}AP=D$, on a $A=PDP^{-1}$. On obtient alors

$$\begin{aligned}A^{132}=PD^{132}P^{-1}&=P\left(7\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)^{132}P^{-1}\\&=7^{132}P\begin{pmatrix}1^{132}&0\\0&(-1)^{132}\end{pmatrix}P^{-1}\\&=7^{132}P\begin{pmatrix}1&0\\0&1\end{pmatrix}P^{-1}\\&=7^{132}PI_2P^{-1}\\&=7^{132}PP^{-1}=7^{132}I_2.\end{aligned}$$

Application D.3. On va utiliser l'application précédente pour étudier le problème suivant (*suite de Fibonacci*):

On considère qu'un couple de lapins, nés à la date 0, donne naissance, à partir du deuxième mois, à un couple de lapins tous les mois. On suppose que, pendant l'étude, aucun lapin ne meurt. On note F_n le nombre de couples de lapins au mois n. On a donc $F_0=1$, et $F_1=1$ puisque le premier couple n'a pas encore donné naissance à de nouveaux couples. De plus, pour tout $n\geqslant 1$, on a $F_{n+1}=F_n+F_{n-1}$ (les lapins présents au mois n sont encore là mais ne se sont pas reproduits, et tous ceux déjà présents au mois $n-1$ se sont reproduits). Nous voulons calculer F_n.

Pour résoudre le problème, posons $X_n=\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}$. On a alors $X_n=\begin{pmatrix}F_n+F_{n-1}\\F_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}F_n\\F_{n-1}\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}X_{n-1}$. Posons $A=\begin{pmatrix}1&1\\1&0\end{pmatrix}$. On a donc

$$X_n=AX_{n-1}=A(AX_{n-2})=A^2X_{n-2}=\cdots=A^nX_0=A^n\begin{pmatrix}1\\1\end{pmatrix}.$$

Il nous faut donc calculer A^n.

Pour cela, diagonalisons A. Recherchons les valeurs propres de A :

$$\det(A-\lambda I_2)=\begin{vmatrix}1-\lambda&1\\1&-\lambda\end{vmatrix}=(1-\lambda)(-\lambda)-1\cdot 1=\lambda^2-\lambda-1.$$

En calculant le discriminant, on obtient les racines $\frac{1+\sqrt{5}}{2}$ et $\frac{1-\sqrt{5}}{2}$. Cherchons les vecteurs propres associés:

- Pour la valeur propre $\dfrac{1+\sqrt{5}}{2}$: on cherche les vecteurs $\begin{pmatrix} x \\ y \end{pmatrix}$ tels que $A\begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{1+\sqrt{5}}{2}\begin{pmatrix} x \\ y \end{pmatrix}$ ce qui revient à $x = \dfrac{1+\sqrt{5}}{2}y$ donc il y a un paramètre, y, que l'on peut choisir égal à 1 pour fixer un vecteur propre: $V_1 = \begin{pmatrix} \dfrac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$.

- De même, pour la valeur propre $\dfrac{1-\sqrt{5}}{2}$, il nous faut un vecteur propre, que l'on peut choisir égal à $V_2 = \begin{pmatrix} \dfrac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$.

Ces deux vecteurs V_1 et V_2 forment une base de $\mathbb{R}^2$. Donc en posant $P = \begin{pmatrix} \dfrac{1+\sqrt{5}}{2} & \dfrac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$

on a $D = P^{-1}AP = \begin{pmatrix} \dfrac{1+\sqrt{5}}{2} & 0 \\ 0 & \dfrac{1-\sqrt{5}}{2} \end{pmatrix}$.

On a également $P^{-1} = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & \dfrac{\sqrt{5}-1}{2} \\ -1 & \dfrac{\sqrt{5}+1}{2} \end{pmatrix}$. On peut donc calculer

$$\begin{aligned} A^n &= PD^nP^{-1} \\ &= P\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} P^{-1} \\ &= \frac{1}{\sqrt{5}}\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} & \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{pmatrix}. \end{aligned}$$

On obtient donc

$$X_n = \frac{1}{\sqrt{5}}\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} + \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} + \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{pmatrix}$$

et donc

$$F_n = \tfrac{1}{\sqrt{5}}\left[\left(\tfrac{1+\sqrt{5}}{2}\right)^{n} - \left(\tfrac{1-\sqrt{5}}{2}\right)^{n} + \left(\tfrac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\tfrac{1-\sqrt{5}}{2}\right)^{n-1}\right]$$

(la deuxième composante de X_n).

Application $\mathbf{D.4.}$ Soit à résoudre le système d'équations différentielles

$$\begin{cases} y_1' = y_1 + 2y_2 \\ y_2' = -3y_1 - 6y_2. \end{cases}$$

Posons $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ et $Y' = \begin{pmatrix} y_1' \\ y_2' \end{pmatrix}$. Ce système s'écrit alors $Y' = AY$ avec $A = \begin{pmatrix} 1 & 2 \\ -3 & -6 \end{pmatrix}$.

On a vu que A est diagonalisable: $P^{-1}AP = \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix}$ avec $P = \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix}$. Notons $D = \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix}$. On a donc $A = PDP^{-1}$.

Réécrivons l'équation $Y' = AY$: on a $Y' = PDP^{-1}Y$ donc $P^{-1}Y' = DP^{-1}Y$ en multipliant les deux membres de l'égalité par P^{-1} à gauche. En posant $Z = P^{-1}Y = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, on a donc $Z' = P^{-1}Y' = DZ$.

Cette dernière équation se traduit par

$$\begin{cases} z_1' = -5z_1 \\ z_2' = 0. \end{cases}$$

On résout les deux équations. On trouve $z_1 = \lambda_1 e^{-5x}$ et $z_2 = \lambda_2$ où λ_1 et λ_2 sont des constantes.

Finalement,

$$Y = PZ = \begin{pmatrix} z_1 - 2z_2 \\ -3z_1 + z_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 e^{-5x} - 2\lambda_2 \\ -3\lambda_1 e^{-5x} + \lambda_2 \end{pmatrix}.$$

Donc les solutions du système d'origine sont les fonctions

$$\begin{cases} y_1 = \lambda_1 e^{-5x} - 2\lambda_2 \\ y_2 = -3\lambda_1 e^{-5x} + \lambda_2 \end{cases}$$

où λ_1 et λ_2 sont des constantes.

VI Statistiques descriptives

A Cas monodimensionnel

A.1 Présentation

Dans ce chapitre, Ω est un ensemble fini non vide, appelé *population*. Un élément ω de Ω est appelé *individu*.

Une *variable statistique* ou *série statistique* est une fonction $x : \Omega \to M$ où M est un ensemble. Si ω est un individu de Ω, $x(\omega)$ est un *caractère* de ω.

Exemple A.1. Soit Ω un ensemble de 100 personnes. On peut considérer les variables statistiques suivantes:

(1) $x : \Omega \to \mathbb{N}$ définie par $x(\omega) =$ nombre de frères et sœurs de ω.

(2) $M = \{$bleu, marron, vert, noir$\}$ et $x : \Omega \to M$ définie par $x(\omega) =$ couleur des yeux de ω.

(3) $x : \Omega \to [120, 240]$ définie par $x(\omega) =$ taille de ω en centimètres.

Remarque A.2. Puisque Ω est fini, l'ensemble $x(\Omega) = \{x(\omega); \omega \text{ dans } \Omega\}$ est fini.

Définition A.3. A chaque sous-ensemble C de M on associe l'entier

$$n(C) = \mathrm{card}(\{\omega \in \Omega \text{ tq } x(\omega) \in C\}),$$

le nombre d'individus dont le caractère est dans C. On appelle C une *classe* et $n(C)$ son *effectif* (ou sa *fréquence absolue*). La *fréquence relative* de C est $f(C) = \dfrac{n(C)}{\mathrm{card}(\Omega)}$; c'est un nombre réel entre 0 et 1.

Exemple A.4. Dans les exemples ci-dessus:

(1) $C = \{0, 1, 2\}$: alors $n(C)$ est le nombre d'individus ayant au plus 2 frères et sœurs.

(2) $C = \{$bleu$\}$: alors $n(C)$ est le nombre d'individus ayant les yeux bleus.

(3) $C = [120, 200]$: alors $n(C)$ est le nombre d'individus mesurant moins de $2m$.

Soit $\mathcal{P} = \{C_1, C_2, \ldots, C_k\}$ une partition de M en k classes, c'est-à-dire que tout élément de M appartient à une et une seule classe C_i. Souvent on représente graphiquement les effectifs ou les fréquences relatives à l'aide d'un diagramme en bâtons dont les bâtons sont de hauteur $n(C_i)$ ou $f(C_i)$ ou d'un diagramme en barres (ou circulaire) où les aires des barres (ou secteurs) sont proportionnelless aux effectifs ou fréquences des classes qu'ils représentent (cf. les exemples qui vont suivre). D'autres types de graphiques sont également utilisés.

Dans la suite, nous allons supposer que l'ensemble M est inclus dans $\mathbb{R}$ et qu'il est borné (c'est-à-dire que c'est une réunion d'intervalles dont les bornes sont des nombres réels, finis).

Soit $\mathcal{P} = \{C_1, \ldots, C_k\}$ une partition de M.

On distingue deux situations:

(a) Cas discret: pour tout i on a $C_i = \{x_i\}$ avec $x_i \in x(\Omega)$ (c'est-à-dire que x_i est un caractère qu'au moins un individu possède).

(b) Cas continu: pour tout i on a $C_i = [e_{i-1}, e_i[$ un intervalle, et on pose $x_i = \frac{1}{2}(e_{i-1}+e_i)$ (le milieu de l'intervalle, ou *centre de la classe*). Dans ce cas, il se peut qu'aucun individu de Ω ne possède le caractère x_i.

Remarque A.5. Le regroupement de données en classes (choix du cas continu plutôt que du cas discret) s'impose lorsque ces données sont en nombre important.

Exemple A.6. (Cas discret) Distribution des ménages selon le nombre d'enfants.

Nombre d'enfants (x)	Fréquence relative (f)
0	$0,42$
1	$0,18$
2	$0,33$
3	$0,05$
4	$0,02$

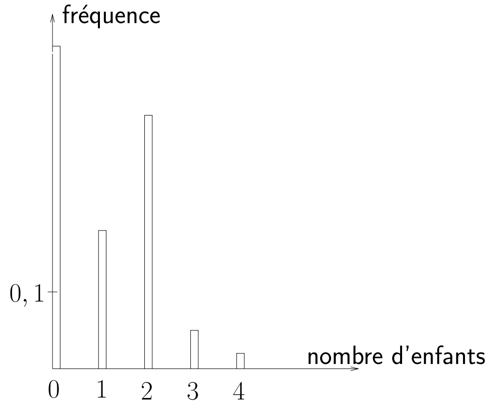


diagramme en bâtons

Exemple A.7. (Cas continu) 70 expériences de rupture en charge d'un fil de même diamètre ont donné la répartition suivante des poids limites, regroupés par classes:

Charges en kg	Effectifs n_i	Centres de classes x_i	$n_i x_i$	$n_i x_i^2$
de 800 à 850	7	825	5775	4764375
de 850 à 900	15	875	13125	11484375
de 900 à 950	32	925	29600	27380000
de 950 à 1000	16	975	15600	15210000
Total	70		64100	58838750

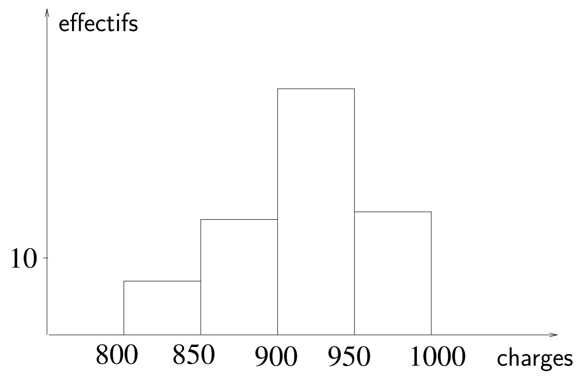


diagramme en barres

A.2 Caractéristiques d'une série statistique

Dans les cas discret et continu, notons n_i la fréquence absolue de la classe C_i (le nombre d'individus dont le caractère est dans C_i) et f_i la fréquence relative de C_i. On pose $n = \sum_{i=1}^{k} n_i$, c'est le nombre total d'individus de Ω. On a donc $f_i = \frac{1}{n} n_i$.

Définition $\mathbf{A.8.}$ On définit la *moyenne* de x par $\bar{x} = \sum_{i=1}^{k} f_i x_i = \frac{1}{n} \sum_{i=1}^{k} n_i x_i$.

Remarque $\mathbf{A.9.}$ $\bar{x}$ est un indice *central*. Il représente globalement le caractère x. Il montre une tendance.

Définition $\mathbf{A.10.}$ On définit la *variance de* x par $\sigma^2 = \sigma_x^2 = V(x) = \sum_{i=1}^{k} f_i (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - \bar{x})^2$.

Remarquons que $\sigma^2 = \frac{1}{n} \sum_{i=1}^{k} n_i x_i^2 - (\bar{x})^2 = \sum_{i=1}^{k} f_i x_i^2 - (\bar{x})^2$, c'est-à-dire que la variance est égale à la "moyenne des carrés moins le carré de la moyenne".

Le réel $\sigma = \sqrt{V(x)}$ est appelé *écart-type*.

Remarque $\mathbf{A.11.}$ L'écart-type est un indice de *dispersion*. Il mesure l'étendue du caractère x.

Définition $\mathbf{A.12.}$ La *médiane* est la valeur de la variable statistique telle qu'il y ait autant d'observations ayant une valeur supérieure que d'observations ayant une valeur inférieure. Pour la déterminer, on fait un tableau des *effectifs cumulés* ou *fréquences cumulées*.

Reprenons l'exemple $\mathbf{A.6}$ des ménages répartis suivant le nombre d'enfants.

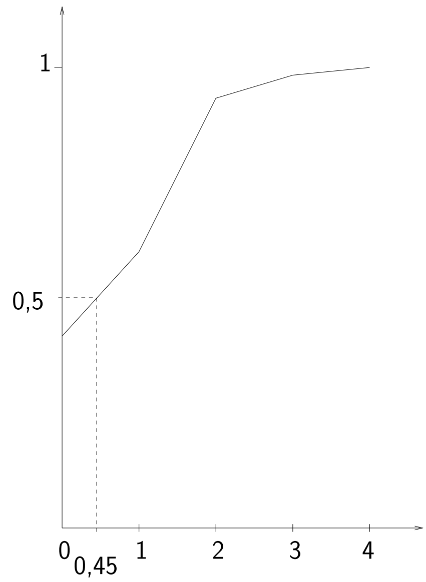
Exemple $\mathbf{A.13.}$ La moyenne est $\bar{x} = 0 \cdot 0,42 + 1 \cdot 0,18 + 2 \cdot 0,33 + 3 \cdot 0,05 + 4 \cdot 0,02 = 1,07$. La variance est $V(x) = (0^2 \cdot 0,42 + 1^2 \cdot 0,18 + 2^2 \cdot 0,33 + 3^2 \cdot 0,05 + 4^2 \cdot 0,02) - \bar{x}^2 = 1,12$ et l'écart-type est $\sigma_x = \sqrt{V(x)} = 1,06$.

Le tableau des fréquences cumulées est

nombre d'enfants	fréquences cumulées
0	$0,42$
1	$0,60$ (effectif de la classe $[0,1]$)
2	$0,93$ (effectif de la classe $[0,2]$)
3	$0,98$ (effectif de la classe $[0,3]$)
4	1 (effectif de la classe $[0,4]$)

On lit ensuite sur le diagramme des fréquences cumulées ci-dessous la médiane, qui vaut environ $0,45$.

On voit également dans le tableau que la médiane doit se situer entre 0 et 1 puisque $0,5$ est entre $0,42$ et $0,60$.



Reprenons l'exemple **A.7** de rupture en charge d'un fil.

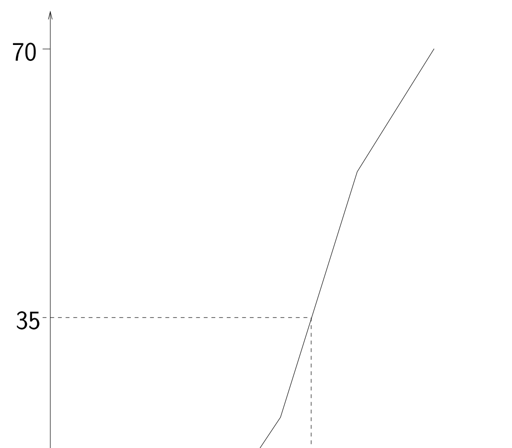
Exemple A.14. La moyenne est $\bar{x} = \dfrac{64100}{70} = 915,71\,kg$. La variance est $V(x) = \dfrac{58838750}{70} - \bar{x}^2 = 2028,77$ et l'écart-type est $\sigma_x = \sqrt{V(x)} = 45,04\,kg$.

Le tableau des effectifs cumulés est

charge en kg	effectifs cumulés
800	0
850	7
900	22
950	54
1000	70

On lit ensuite sur le diagramme des effectifs cumulés ci-dessous la médiane, qui vaut environ $920\,kg$.

En général, la médiane est différente de la moyenne.



Remarque A.15. On peut définir bien d'autres indices, qui représentent divers aspects de la population, comme par exemple sa disymétrie par rapport à la moyenne.

B Cas bidimensionnel

B.1 Description

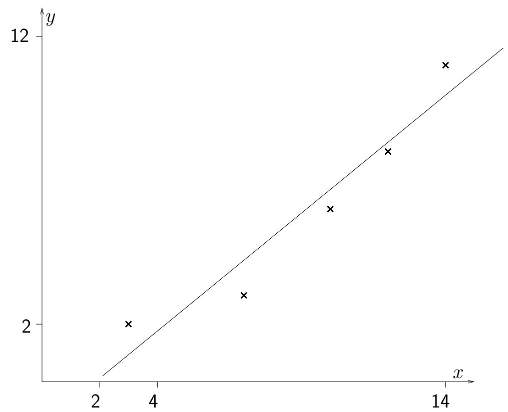
Nous considérons ici deux variables statistiques x et y définies sur une population Ω et toutes deux à valeurs dans un ensemble borné de $\mathbb{R}$. L'intérêt est de savoir si d'une certaine manière on peut *expliquer* y par x.

Exemple B.1. (Cas discret) Notes d'étudiants en mathématiques et en physique.

Elèves	Notes x en maths	Notes y en physique
1	12	8
2	14	11
3	3	2
4	10	6
5	7	3

On peut calculer $\bar{x} = 9,2$ et $\sigma_x = 3,87$. On peut aussi calculer $\bar{y} = 6$ et $\sigma_y = 3,28$.

Pour voir s'il y a un lien entre x et y, disposons les résultats graphiquement sous forme d'un *nuage de points*.



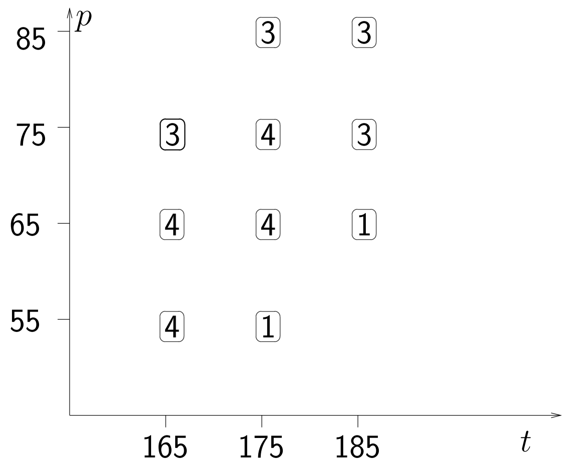
Il semble y avoir un lien linéaire. On est alors tentés d'ajuster une droite à l'ensemble de ces points.

Exemple B.2. (Cas continu) Classification de 30 adultes suivant poids et tailles.

Taille t (cm) $\rightarrow$ Poids p (kg) $\downarrow$	$160 \leqslant t < 170$	$170 \leqslant t < 180$	$180 \leqslant t < 190$	Total
$80 \leqslant p < 90$	0	3	3	6
$70 \leqslant p < 80$	3	4	3	10
$60 \leqslant p < 70$	4	4	1	9
$50 \leqslant p < 60$	4	1	0	5
Total	11	12	7	

La dernière ligne et la dernière colonne donnent les *distributions marginales*: distribution des tailles et distribution des poids (sans faire intervenir l'autre caractère). On peut en déduire $\bar{t}$, σ_t, $\bar{p}$ et σ_p.

Graphiquement, on a un "nuage de points pondérés".

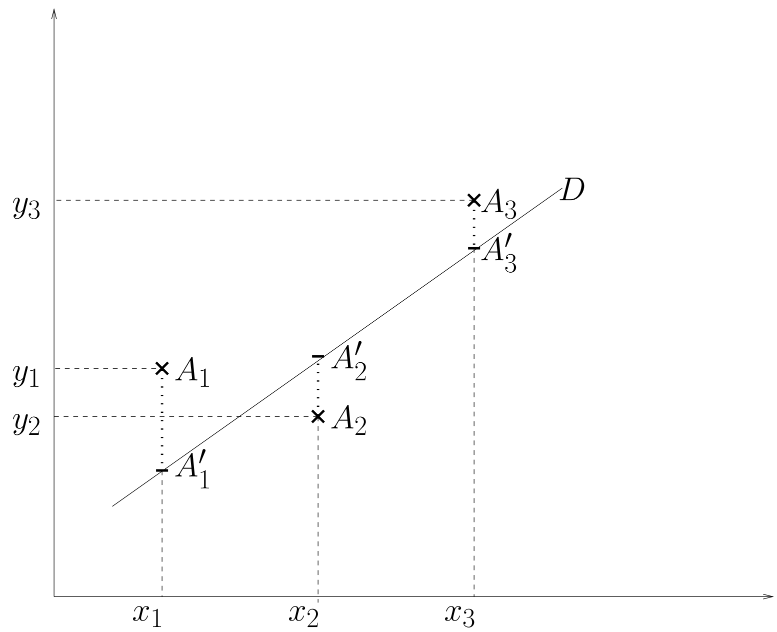


Difficile d'ajuster une droite.

B.2 Ajustement linéaire

Nous allons voir la *méthode des moindres carrés de Gauss*.

Soient $A_i(x_i, y_i)$ les n points du nuage représentant les deux variables statistiques x et y (chaque point du nuage a le poids 1 mais peut être répété). On cherche la droite D telle que la somme $S = \sum_{i=1}^{n} \operatorname{dist}(A_i, A'_i)^2$ soit minimale:



Définition B.3. On définit la *covariance* de x et y par $\operatorname{cov}(x, y) = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n} = \dfrac{\sum_{i=1}^{n} x_i y_i}{n} - \bar{x}\bar{y}$.

> **Proposition** B.4. La droite D a pour équation $y - \bar{y} = \dfrac{\operatorname{cov}(x, y)}{\sigma_x^2}(x - \bar{x})$. Elle est appelée la *droite de régression linéaire (ou d'ajustement) de y en x*.

Remarque B.5. En échangeant les rôles de y et x, on obtient la *droite de régression linéaire de x en y*, D', qui a pour équation

$$x - \bar{x} = \frac{\operatorname{cov}(x, y)}{\sigma_y^2}(y - \bar{y}).$$

Remarque $\mathbf{B.6.}$ Les droites D et D' passent par le point G de coordonnées $(\bar{x}, \bar{y})$, appelé *point moyen* ou *barycentre* du nuage. Il est utile pour placer les droites.

Définition $\mathbf{B.7.}$ Le nombre $\rho = \rho(x,y) = \dfrac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y}$ s'appelle le *coefficient de corrélation linéaire*.

Proposition $\mathbf{B.8.}$ La valeur minimale de S (qui représente en quelque sorte la distance entre les points du nuage et la droite D) est $n\sigma_y^2\left(1-\rho(x,y)^2\right)$.

Remarque $\mathbf{B.9.}$ Comme $S \geqslant 0$ on en déduit que $\rho(x,y)^2 \leqslant 1$ et que $S = 0$ si et seulement si $\rho(x,y)^2 = 1$.

Propriétés $\mathbf{B.10.}$ **(1)** $-1 \leqslant \rho \leqslant 1$.

(2) Les points du nuage sont alignés si et seulement si $\rho = -1$ ou $\rho = 1$.

(3) Les points du nuage sont presque alignés si et seulement si ρ est proche de -1 ou de 1. On dit alors que l'on a une forte corrélation linéaire (d'où le nom de ρ).

(4) D a pour coefficient directeur $\dfrac{\operatorname{cov}(x,y)}{\sigma_x^2}$ et D' a pour coefficient directeur $\dfrac{\sigma_y^2}{\operatorname{cov}(x,y)}$ donc $D = D'$ si et seulement si $\rho = -1$ ou $\rho = 1$, et D est proche de D' si et seulement si ρ est proche de -1 ou de 1.

Exemple $\mathbf{B.11.}$ Reprenons l'exemple $\mathbf{B.1}$ des notes en maths et en physique.

Elèves	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	12	8	$2,8$	2	$7,84$	4	$5,6$
2	14	11	$4,8$	5	$23,04$	25	24
3	3	2	$-6,2$	-4	$38,44$	16	$24,8$
4	10	6	$0,8$	0	$0,64$	0	0
5	7	3	$-2,2$	-3	$4,84$	9	$6,6$
Total	46	30			$74,8$	54	61
	$\bar{x} = 9,2$	$\bar{y} = 6$			$\sigma_x^2 = 14,96$	$\sigma_y^2 = 10,8$	$\operatorname{cov}(x,y) = 12,2$
					$\sigma_x = 3,87$	$\sigma_y = 3,29$	

Donc D a pour équation $y - \bar{y} = 0,815(x-\bar{x})$ soit $y = 0,815x - 1,5$ et D' a pour équation $x - \bar{x} = 1,13(y-\bar{y})$ soit $x = 1,13y + 2,42$.

Pour construire D et D' on utilise le barycentre qui a pour coordonnées ici $(9,2;6)$. On constate de plus que D passe par $(0;-1,5)$ et que D' passe par $(2,42;0)$.

On calcule $\rho = 0,96$ donc on a une très forte corrélation linéaire. On constate sur la figure que D et D' sont très proches.