Université de Bourgogne

Algèbre linéaire [1]

2017-2018

1. Ce polycopié a été rédigé par Johan Taflin et modifié à la marge par Emmanuel Wagner.

emmanuel.wagner@u-bourgogne.fr
Bureau A322

Préface

Ce texte est destiné aux étudiants de première année de licence physique/chimie à l'université de Bourgogne. Il s'agit d'une introduction à l'algèbre linéaire et au calcul matriciel. Cette version est susceptible de contenir des coquilles et des erreurs. Tout commentaire est le bienvenu.

Méthodologie : L'assimilation d'un nouveau cours est difficile et exige un travail personnel assidu. Il est important de s'approprier les notions, de trouver ses propres exemples et contre-exemples, de connaître les définitions, de s'interroger sur les hypothèses des énoncés etc.

Par ailleurs, il est illusoire de penser que ces notes de cours puissent remplacer une présence attentive en cours magistral.

Table des matières

Chapitre 1

Matrices

Ce chapitre introduit des objets très pratiques en algèbre linéaire : les matrices. Ce sont des tableaux de nombres sur lesquels il est possible de faire une série d'opérations. Bien que leur manipulation soit très simple, elles interviennent dans de nombreux problèmes. Nous les utiliserons par la suite aussi bien pour résoudre des systèmes linéaires que pour représenter des applications linéaires. Par ailleurs, les méthodes introduites dans ce chapitre sont algorithmiques, elles jouent un rôle important en informatique et sont présentes dans presque toutes les modélisations.

1.1 Calcul matriciel

Définition 1.1. Soit m et n deux entiers. Une *matrice de taille* $m \times n$ est un tableau de nombres avec m lignes et n colonnes de la forme

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

Les a_{ij} sont des réels appelés *coefficients* de A. On notera aussi

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

s'il n'y a pas d'ambiguïté sur les indices.

Exemple 1.2. La matrice

$$A = (a_{ij}) = \begin{pmatrix} 0 & 5 & 8 & -1 \\ 1 & -4 & 18 & 2 \end{pmatrix}$$

est de taille 2×4. Le coefficient a_{23} est 18 et le a_{12} est 5.

Définition 1.3. L'ensemble des matrices de taille $m \times n$ sera noté $\mathscr{M}_{m,n}(\mathbb{R})$.

Définition 1.4. Un élément de $\mathscr{M}_{m,1}(\mathbb{R})$ est un *vecteur colonne de taille* m. Un élément de $\mathscr{M}_{1,n}(\mathbb{R})$ est un *vecteur ligne de taille* n.

Exemple 1.5. La matrice

$$\begin{pmatrix} 1 & 2 & 10 & -8 & 3 \end{pmatrix}$$

est un vecteur ligne de taille 5 alors que la matrice

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

est un vecteur colonne de taille 4.

Définition 1.6. Soit $A = (a_{ij})$ une matrice de $\mathscr{M}_{m,n}(\mathbb{R})$. La *$i$-ème ligne* de A est le vecteur ligne $\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$. La *$i$-ème colonne* de A est le vecteur colonne

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Exemple 1.7. Soit A la matrice 3×4 donnée par

$$A = \begin{pmatrix} 5 & -1 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ -1 & -7 & 0 & 12 \end{pmatrix}.$$

La troisième ligne de A est $\begin{pmatrix} -1 & -7 & 0 & 12 \end{pmatrix}$ et sa deuxième colonne

$$\begin{pmatrix} -1 \\ 1 \\ -7 \end{pmatrix}.$$

Définition 1.8. Soit $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{R})$. La *transposée* de A est la matrice $B = (b_{ij})$ de $\mathscr{M}_{n,m}(\mathbb{R})$ avec $b_{ij} = a_{ji}$. On la note ${}^tA$.

Exemple 1.9. Si on reprend l'exemple précédent, la transposée de

$$A = \begin{pmatrix} 5 & -1 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ -1 & -7 & 0 & 12 \end{pmatrix}$$

est

$${}^tA = \begin{pmatrix} 5 & 0 & -1 \\ -1 & 1 & -7 \\ 0 & 2 & 0 \\ 3 & 3 & 12 \end{pmatrix}.$$

Il est possible de définir deux opérations naturelles sur les matrices : l'addition de deux matrices de même taille et la multiplication d'une matrice par un nombre réel.

Définition 1.10. Si $A = (a_{ij})$ et $B = (b_{ij})$ sont deux éléments de $\mathscr{M}_{m,n}(\mathbb{R})$, la *somme* de A et B, notée $A + B$, est la matrice $C = (c_{ij})$ avec $c_{ij} = a_{ij} + b_{ij}$.

Exemple 1.11. Par exemple dans le cas de matrice 2×4

$$\begin{pmatrix} 0 & 2 & 7 & 9 \\ 17 & -4 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -2 & 0 & -5 \\ -12 & -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 7 & 4 \\ 5 & -5 & -1 & 1 \end{pmatrix}.$$

Définition 1.12. Si $A = (a_{ij})$ est une matrice de taille $m \times n$ et λ un nombre réel alors la *multiplication* de A par λ est donnée par la matrice $\lambda A = (\lambda a_{ij})$.

Exemple 1.13. Si la matrice A est donnée par

$$A = \begin{pmatrix} 8 & 2 & -1 & 6 \\ 7 & 4 & 0 & -4 \end{pmatrix}$$

alors la matrice $(1/2)A$ est

$$\begin{pmatrix} 4 & 1 & -(1/2) & 3 \\ 7/2 & 2 & 0 & -2 \end{pmatrix}.$$

Définition 1.14. Pour $m \geq 1$ et $n \geq 1$ deux entiers fixés, on définit la matrice $0_{\mathscr{M}_{m,n}(\mathbb{R})}$ comme étant la matrice de $\mathscr{M}_{m,n}(\mathbb{R})$ dont tous les coefficients sont nuls. On l'appelle *zéro* de $\mathscr{M}_{m,n}(\mathbb{R})$. Par exemple, si $m = 2$ et $n = 4$ alors

$$0_{\mathscr{M}_{2,4}(\mathbb{R})} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il est facile de voir qu'il s'agit de l'une unique matrice A de $\mathscr{M}_{m,n}(\mathbb{R})$ telle que $A + B = B$ pour toute matrice B de $\mathscr{M}_{m,n}(\mathbb{R})$. S'il n'y a pas d'ambiguïté sur les indices, on la notera simplement par 0.

Il existe une troisième opération sur les matrices, à première vue moins intuitive mais essentielle pour la suite.

Définition 1.15. Soit m, n et p trois entiers. La *multiplication* d'une matrice $A = (a_{ij})$ de taille $m \times n$ par une matrice $B = (b_{ij})$ de taille $n \times p$ est une matrice $AB = (c_{ij})$ de taille $m \times p$ où

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Remarque 1.16. Il n'est pas possible en général de faire la multiplication de deux matrices quelconques. Le produit AB n'est possible que si le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B.

Exemple 1.17. 1) Un cas simple mais intéressant est celui de la multiplication d'un vecteur ligne par un vecteur colonne de même taille. Le résultat est un réel donné par

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

2) De manière plus générale si $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{R})$ et $B = (b_{ij}) \in \mathscr{M}_{n,p}(\mathbb{R})$, le coefficient c_{ij} de $AB = (c_{ij}) \in \mathscr{M}_{m,p}(\mathbb{R})$ est donnée par la multiplication de la i-ème ligne de A par la j-ème colonne de B.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix}$$

avec $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.
3) De manière plus concrète :

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

Attention : Même si le produit AB est défini, ce n'est pas forcément le cas pour le produit BA. Et quand bien même les deux produits seraient bien définis, en général $AB \neq BA$. Le produit matriciel n'est pas commutatif !

Exemple 1.18. Soit $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. On peut vérifier que

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

Attention : Comme on vient de le voir dans l'exemple précédent, le fait que $AB = 0$ n'implique pas que $A = 0$ ou $B = 0$.

Néanmoins, il est facile de vérifier que les opérations sur les matrices vérifient les règles suivantes.

Proposition 1.19. *1) Le produit est distributif :*

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

et

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

2) Le produit est associatif :

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

3) Si $\lambda \in \mathbb{R}$ alors

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

4) Si $A \in \mathscr{M}_{m,n}(\mathbb{R})$ alors

$$0_{\mathscr{M}_{q,m}(\mathbb{R})} \times A = 0_{\mathscr{M}_{q,n}(\mathbb{R})} \text{ et } A \times 0_{\mathscr{M}_{n,p}(\mathbb{R})} = 0_{\mathscr{M}_{m,p}(\mathbb{R})}.$$

Proposition 1.20. *Soit A et B deux matrices. Si le produit AB est bien défini alors le produit ${}^tB{}^tA$ l'est aussi et ${}^tB{}^tA = {}^t(AB)$.*

Les matrices avec autant de lignes que de colonnes interviennent souvent dans la pratique.

Définition 1.21. Une matrice de taille $n \times n$ est dite *carrée*. L'ensemble de toutes ces matrices est noté $\mathscr{M}_n(\mathbb{R})$. Pour une matrice $A = (a_{ij}) \in \mathscr{M}_n(\mathbb{R})$, les coefficients a_{ii}, $1 \leq i \leq n$, sont appelés les coefficients diagonaux de A.

Parmi les matrices carrées, une est particulièrement importante.

Définition 1.22. La matrice *identité* de taille n est la matrice de $\mathscr{M}_n(\mathbb{R})$ dont les coefficients diagonaux valent 1 et tous les autres coefficients valent 0. On la note I_n, Id_n ou simplement Id.

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

Comme son nom l'indique, la multiplication par I_n agit par identité.

Proposition 1.23. *Si $A \in \mathscr{M}_{m,n}(\mathbb{R})$ alors*

$$AI_n = I_m A = A.$$

Définition 1.24. Une matrice $A \in \mathscr{M}_n(\mathbb{R})$ est dite *inversible* s'il existe $B \in \mathscr{M}_n(\mathbb{R})$ tel que

$$AB = BA = I_n.$$

Dans ce cas, B est appelé *inverse* de A et est noté A^{-1}. L'ensemble des matrices inversibles de $\mathscr{M}_n(\mathbb{R})$ est désigné par $\mathrm{Gl}_n(\mathbb{R})$, pour *groupe linéaire*.

Attention : Si A et B sont dans $\mathrm{Gl}_n(\mathbb{R})$ alors le produit AB l'est aussi et

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

Remarque 1.25. Nous verrons dans le chapitre sur les applications linéaires que pour qu'une matrice A soit inversible, il suffit qu'elle ait un inverse à gauche (c.à.d une matrice B telle que $BA = I_n$) ou un inverse à droite (c.à.d une matrice C telle que $AC = I_n$).

Dans la suite du cours, nous rencontrerons un certain nombre d'ensembles de matrices particulières.

Définition 1.26. Soit $A = (a_{ij}) \in \mathscr{M}_n(\mathbb{R})$.
1) Si $a_{ij} = 0$ dès que $i > j$ alors la matrice A est dite *triangulaire supérieure*,
2) Si $a_{ij} = 0$ dès que $i < j$ alors la matrice A est dite *triangulaire inférieure*,
3) Si $a_{ij} = a_{ji}$ pour tout $1 \leq i, j \leq n$ (autrement dit si ${}^tA = A$) alors A est dite *symétrique*,
4) Si $a_{ij} = -a_{ji}$ pour tout $1 \leq i, j \leq n$ (autrement dit si ${}^tA = -A$) alors A est dite *anti-symétrique*. En particulier, les coefficients diagonaux de A sont nuls.

Exemple 1.27. 1) La matrice

$$\begin{pmatrix} 2 & 4 & -8 \\ 0 & 7 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

est triangulaire supérieure alors que

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

est triangulaire inférieure.
2) La matrice

$$\begin{pmatrix} 1 & 2 & -3 & -1 \\ 2 & 0 & 9 & 7 \\ -3 & 9 & 5 & 1 \\ -1 & 7 & 1 & 8 \end{pmatrix}$$

est symétrique alors que

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

est anti-symétrique.

1.2 Systèmes linéaires

Un des avantages des matrices est qu'elles permettent de réécrire les systèmes linéaire de manière concise et manipulable. Rappelons qu'un *système linéaire* est un système d'équations linéaires

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

Dans cette exemple, il y a m équations avec n inconnues $x_1, \dots, x_n$. En utilisant une écriture matricielle, le système linéaire est équivalent à

$$AX = B$$

où $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{R})$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. Résoudre un tel système revient à déterminer tous les n-uplets $(x_1, \dots, x_n) \in \mathbb{R}^n$ tels que toutes les équations sont vérifiées. Nous verrons dans la section suivante qu'en utilisant la méthode du pivot, il est possible de résoudre de manière systématique tout système linéaire.

Dans le cas particulier où il y a autant d'équations que d'inconnues, c'est-à-dire quand A est une matrice carrée, et que de plus A est inversible alors pour tout B il existe un unique X vérifiant $AX = B$, donné par $X = A^{-1}B$.

1.3 La méthode du pivot

Le but de cette section est d'expliquer une méthode, souvent dite du *pivot de Gauss*, particulièrement efficace pour résoudre des systèmes linéaires. Elle repose sur seulement trois opérations sur les matrices.

Définition 1.28. Soit A une matrice dont on note par L_i la i-ème ligne. Une *opération élémentaire* sur les lignes de A consiste à

1. remplacer une ligne L_i par $L_i + L_j$,
2. échanger deux lignes,
3. ou multiplier une ligne par un scalaire non nul.

On notera $A \to A'$ quand on passe d'une matrice A à une matrice A' par une opération élémentaire. Dans ce poly, on écrira les opérations à gauche de la nouvelle matrice.

Exemple 1.29.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \to \begin{array}{c} L_2 \\ L_1 \\ \end{array} \begin{pmatrix} 4 & 6 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \to \begin{array}{c} \\ L_3 \\ L_2 \end{array} \begin{pmatrix} 4 & 6 \\ 5 & 6 \\ 3 & 4 \end{pmatrix} \to \begin{array}{c} \\ \\ 2L_3 \end{array} \begin{pmatrix} 4 & 6 \\ 5 & 6 \\ 6 & 8 \end{pmatrix}.$$

Remarque 1.30. 1) Il est aussi possible de faire plusieurs opérations élémentaires à la fois, tant qu'elles restent réversibles. De cette façon, on peut remplacer L_i par $L_i + \lambda L_j$ avec $\lambda \in \mathbb{R}$ et $i \neq j$.
2) Par contre, il n'est pas possible de multiplier une ligne par 0 ou de faire des opérations du type

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \to \begin{matrix} L_1 + L_2 \\ L_1 + L_2 \\ \end{matrix} \begin{pmatrix} 4 & 6 \\ 4 & 6 \\ 5 & 6 \end{pmatrix}.$$

Les systèmes linéaires associés aux matrices suivantes sont particulièrement simples à résoudre.

Définition 1.31. Une matrice est dite *échelonnée* si le premier coefficient non nul de la ligne $i+1$ est à droite de celui de la ligne i. Dans ce cas, le premier coefficient non nul d'une ligne est appelé *pivot*.

Définition 1.32. Une matrice est dite *échelonnée réduite* si elle est échelonnée et vérifie les deux propriétés suivantes :
1) Les pivots valent tous 1.
2) Dans sa colonne, le pivot est le seul coefficient non nul.

Exemple 1.33. 1) Les matrices

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

ne sont pas échelonnées alors que

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 4 & 2 & 3 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & 0 & 7 \end{pmatrix}, \begin{pmatrix} 9 & 0 & 2 & 8 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

le sont et leurs pivots sont marqués en bleu. Par contre elles ne sont pas échelonnées réduites.
2) La matrice

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

est échelonnée réduite.

Avant d'introduire la méthode du pivot, il reste à fixer quelques notations.

Si A et B sont deux matrices avec le même nombre de lignes, on notera $(A|B)$ la matrice obtenue en juxtaposant A à côté de B. Par exemple, si

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 9 & 7 & -1 \\ 5 & -2 & 8 \end{pmatrix} \text{ alors } (A|B) = \begin{pmatrix} 1 & 0 & 9 & 7 & -1 \\ 2 & 1 & 5 & -2 & 8 \end{pmatrix}.$$

Plus généralement, on notera parfois les matrices avec des blocs. Par exemple, si

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

alors

$$A = \left(\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 4 & \multicolumn{2}{c}{\multirow{2}{*}{B}} \\ 7 & & \end{array}\right) \text{ avec } B = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}$$

ou encore

$$A = \left(\begin{array}{c|c} 1 & \\ 4 & C \\ 7 & \end{array}\right) \text{ avec } C = \begin{pmatrix} 2 & 3 \\ 5 & 6 \\ 8 & 9 \end{pmatrix}.$$

<u>La méthode</u>

La méthode du pivot fournit un algorithme pour passer d'une matrice quelconque à une matrice échelonnée (voire échelonnée réduite) en utilisant des opérations élémentaires. Voici les différentes étapes pour une matrice A.

<u>Etape 1</u> (On ignore les colonnes nulles)

Si la première colonne de A est nulle, c'est-à-dire

$$A = \left(\begin{array}{c|c} 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

on reprend l'étape 1 avec A'. Sinon on passe à l'étape suivante.

<u>Etape 2</u> (Installation du pivot)

On choisit une ligne dont le premier coefficient est non nul et on la place en première ligne.

<u>Etape 3</u> (Nettoyage sous le pivot)

On remplace chacune des lignes L_i $(i \neq 1)$ par $L_i - \lambda_i L_1$ avec λ_i choisi de sorte que le premier coefficient de la nouvelle ligne L_i s'annule.

<u>Etape 4</u> (Itération)

La matrice qu'on obtient est maintenant de la forme

$$\left(\begin{array}{c|c} p & \times \cdots \times \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right).$$

Si A' est non nulle, on reprend l'étape 1 avec elle. Sinon, on arrête l'algorithme.

Puisque après une boucle, l'algorithme reprend avec une matrice plus petite et que la matrice d'origine est de taille finie, l'algorithme aboutit forcément. De cette manière, en partant de n'importe quelle matrice, on arrive à une matrice échelonnée. De là, on peut faire deux étapes supplémentaires pour obtenir une matrice échelonnée réduite.

Etape 5

On multiplie les lignes avec pivots pour qu'ils valent 1.

Etage 6
On nettoie au dessus des pivots.

Remarque 1.34. Il est toujours possible de faire des étapes intermédiaires si cela semble simplifier les calculs.

Exemple 1.35. Considérons la matrice suivante

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

Puisque la première colonne est nulle, on considère la sous-matrice en bleu

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

On choisit une ligne de la matrice en question dont le premier coefficient est non nul. Pour simplifier les calculs, il peut être astucieux de prendre un pivot simple, par exemple 1 dans notre cas.

$$\begin{pmatrix} 0 & 4 & 2 & 6 \\ 0 & 2 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 4 & 2 \end{pmatrix} \rightarrow \begin{array}{c} L_3 \\ \\ L_1 \\ \\ \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 2 & 2 & 3 \\ 0 & 4 & 2 & 6 \\ 0 & 0 & 4 & 2 \end{pmatrix}.$$

Puis on nettoie sous le pivot

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 2 & 2 & 3 \\ 0 & 4 & 2 & 6 \\ 0 & 0 & 4 & 2 \end{pmatrix} \rightarrow \begin{array}{c} \\ L_2 - 2L_1 \\ L_3 - 4L_1 \\ \\ \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 4 & 2 \end{pmatrix}.$$

Et on recommence

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 4 & 2 \end{pmatrix} \rightarrow \begin{array}{c} \\ \\ L_3 - L_2 \\ L_4 - 2L_2 \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

pour obtenir une matrice échelonnée. Il reste à la réduire

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{array}{c} \\ L_2/2 \\ \\ \\ \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{array}{c} L_1 - L_3 \\ L_2 - L_3/2 \\ \\ \\ \end{array} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Remarque 1.36. La forme échelonnée d'une matrice n'est pas unique mais par contre la forme échelonnée réduite l'est.

1.4 Résolution des systèmes linéaires

La première application de la méthode du pivot est de résoudre les systèmes linéaires. En effet, il est évident que les solutions d'un système linéaire ne change pas si on interchange deux équations, si on multiplie une équation par une constante non nulle ou encore si on remplace une équation par sa somme avec une autre équation.

Définition 1.37. Soit $AX = B$ et $A'X = B'$ deux systèmes linéaires. On dit qu'ils sont *équivalents* si l'on peut passer de la matrice $(A|B)$ à la matrice $(A'|B')$ par des opérations élémentaires.

La discussion ci-dessus implique le résultat suivant.

Théorème 1.38. *Deux systèmes équivalents ont même ensemble de solutions.*

Une manière plus formelle de le voir est d'introduire les matrices suivantes. Soit $n \geq 1$ et $1 \leq k, l \leq n$ trois entiers. La matrice $E_{kl} = (e_{ij}) \in \mathscr{M}_n(\mathbb{R})$ est définie par $e_{kl} = e_{lk} = 1$, $e_{kk} = e_{ll} = 0$, $e_{ii} = 1$ pour tout $i \neq k, l$ et $e_{ij} = 0$ sinon. Multiplier à gauche une matrice A par E_{kl} revient à inverser la ligne k avec la ligne l.

Exemple 1.39. Si $n = 2$ et $A = \begin{pmatrix} 1 & 0 & 9 & 8 \\ -2 & -3 & 4 & 5 \end{pmatrix}$ alors

$$E_{12}A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 9 & 8 \\ -2 & -3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} -2 & -3 & 4 & 5 \\ 1 & 0 & 9 & 8 \end{pmatrix}.$$

De même, la matrice $F_{kl} = (f_{ij}) \in \mathscr{M}_n(\mathbb{R})$ est définie par $f_{kl} = 1$, $f_{ii} = 1$ pour tout $1 \leq i \leq n$ et $f_{ij} = 0$ sinon. Multiplier à gauche A par F_{kl} revient à échanger la ligne k par sa somme avec la ligne l.

Exemple 1.40.

$$F_{12}A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 9 & 8 \\ -2 & -3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} -1 & -3 & 13 & 13 \\ -2 & -3 & 4 & 5 \end{pmatrix}.$$

Enfin, si $\alpha \in \mathbb{R}$, $\alpha \neq 0$, alors la matrice $G_k(\alpha) = (g_{ij}) \in \mathscr{M}_n(\mathbb{R})$ est définie par $g_{kk} = \alpha$, $g_{ii} = 1$ si $i \neq k$ et $g_{ij} = 0$ sinon. La multiplication à gauche de A par $G_k(\alpha)$ revient à multiplier la ligne k par α.

Exemple 1.41.

$$G_1(3)A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 9 & 8 \\ -2 & -3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 27 & 24 \\ -2 & -3 & 4 & 5 \end{pmatrix}.$$

Ces trois types de matrices s'appellent des matrices élémentaires.

Remarque 1.42. Ces matrices sont inversibles, d'inverse $E_{kl}^{-1} = E_{kl}$, $G_k(\alpha)^{-1} = G_k(\alpha^{-1})$ et $F_{kl}^{-1} = G_k(-1)F_{kl}G_k(-1)$.

Pour résumer, la multiplication à gauche par une matrice élémentaire revient exactement à faire une des trois opérations élémentaires sur les lignes d'une matrice. Par conséquent, appliquer la méthode du pivot, revient à faire une succession de multiplications à gauches par des matrices élémentaires. Le théorème découle alors de la proposition élémentaire suivante.

Proposition 1.43. *Si A et B sont des matrices avec n lignes et que C a n colonnes alors*

$$C(A|B) = (CA|CB).$$

Démonstration du théorème. Si on passe de $(A|B)$ à $(A'|B')$ par une suite d'opération élémentaires, cela veut dire qu'il existe une suite de matrices élémentaires $E_1, \ldots, E_N$ telle que $(A'|B') = E_1E_2 \cdots E_N(A|B)$. La Proposition 1.43 implique que $A' = E_1 \cdots E_N A$ et que $B' = E_1 \cdots E_N B$. Donc si X vérifie $AX = B$ il vérifie aussi

$$E_1 \cdots E_N AX = E_1 \cdots E_N B \quad \text{c'est-à-dire} \quad A'X = B'.$$

Inversement, si $A'X = B'$ on a aussi $AX = B$ car les E_i sont inversibles. □

Par ailleurs, à tout système $AX = B$, on peut associer grâce à la méthode du pivot un système $A'X = B'$ avec $(A'|B')$ échelonnée réduite. Il ne reste plus qu'a résoudre ce type de système.

Théorème 1.44. *Soit $AX = B$ un système linéaire tel que la matrice $(A|B)$ soit échelonnée réduite. Le système a au moins une solution si et seulement si il n'y a pas de pivot dans la dernière colonne.*

Dans ce cas, les inconnues correspondant à une colonne avec pivot sont dites contraintes et les autres sont dites libres. L'ensemble des solutions s'obtient en fixant arbitrairement chacune des inconnues libres et en calculant à partir de là les variables contraintes.

Exemple 1.45. Si $A = \begin{pmatrix} 0 & 1 & 2 & 0 & 9 \\ 0 & 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ alors la matrice $(A|B)$ est échelonnée réduite et correspond au système linéaire

$$\begin{cases} x_2 + \ 2x_3 \qquad +9x_5 &= 0 \\ \qquad\qquad x_4 \ +8x_5 &= 0 \\ \qquad\qquad\qquad 0 &= 1 \end{cases},$$

qui n'a évidemment pas de solution car la dernière ligne est contradictoire. Cela vient du fait que $(A|B)$ a un pivot dans ça dernière colonne.

En revanche, si $C = \begin{pmatrix} 2 \\ -4 \\ 0 \end{pmatrix}$ alors la matrice $(A|C)$ est aussi échelonnée et le système correspondant est

$$\begin{cases} x_2 + \ 2x_3 \qquad +9x_5 &= 2 \\ \qquad\qquad x_4 \ +8x_5 &= -4 \\ \qquad\qquad\qquad 0 &= 0 \end{cases}.$$

Cette fois il n'y a pas de pivot dans la dernière colonne. Les variables x_2 et x_4 correspondent à une colonne avec un pivot, elles sont donc contraintes, alors que les colonnes de x_1, x_3 et x_5 sont sans pivot et ces variables sont libres. L'ensemble des solutions de $AX = C$ est donc

$$S = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \,|\, x_2 = 2 - 2x_3 - 9x_5, x_4 = -4 - 8x_5\}.$$

Si l'on veut une solution particulière, il suffit de choisir les variables libres de manière arbitraire, par exemple ici $x_1 = 12, x_3 = -1$ et $x_5 = 1$, et d'en déduire les autres variables, ici $x_2 = -5$ et $x_4 = -12$.

Une conséquence de ce théorème est que l'ensemble des solutions d'un système linéaire est soit vide, soit réduit à un élément, soit infini. Dans le cas d'un système homogène, le premier cas n'est pas possible car il y a toujours au moins une solution qui est 0.

Corollaire 1.46. *Si un système homogène a (strictement) plus d'inconnues que d'équations alors il a une infinité de solutions.*

1.5 Rang et noyau d'une matrice.

On définit ici, le rang et le noyau d'une matrice, ainsi que la méthode pour les calculer. Ces notions seront reprise dans les deux chapîtres suivant pour parler et calculer le rang d'un système de vecteurs ou d'une application linéaire et déterminer le noyau d'une application linéaire.

Définition 1.47. Le rang d'une matrice A est le nombre de pivot d'une de ses formes échelonnées.

Pour calculer le rang d'une matrice A il suffit donc de la mettre sous-forme échelonnée réduite.

Définition 1.48. Le noyau d'une matrice A est l'ensemble des solutions du système homogène $AX = 0$. On le note $\mathrm{Ker}(A)$.

Trouver le noyau d'une matrice A revient donc à résoudre le système $AX = 0$ et encore une fois on échelonne et on réduit la matrice A.

On verra dans le chapître suivant que $Ker(A)$ est un exemple d'espace vectoriel, et obtenir une forme échélonnée réduite pour A permettra de donner des solutions de $AX = 0$ préférée à partir desquelles retrouver toutes les autres, cet ensemble de solutions préférées formeront ce qu'on appelera une base de $Ker(A)$.

Exemple 1.49. On considère :

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

dont une forme échélonnée réduite est donnée par :

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

Le rang de A est donc de 3. Trouver le noyau revient donc à résoudre

$$\begin{cases} x_1 + 2x_2 + 2x_4 = 0 \\ x_3 + x_4 = 0 \\ x_5 = 0 \end{cases}$$

En mettant à gauche les variables correspondants aux colonnes contenant un pivot et à droite les autres :

$$\begin{cases} x_1 = -2x_2 - 2x_4 \\ x_3 = -x_4 \\ x_5 = 0 \end{cases}$$

On peut choisir librement les valeurs de x_2 et x_4 et n'importe quelle choix détermine de manière unique les valeurs pour de x_1, x_3 et x_5 (x_5 étant déjà égal à zéro.)

On notera que la somme du nombre de variables pouvant être choisis librement dans le calcul du noyau et du rang est évidement égal au nombre de colonnes. On reverra cette propriété dans le chapître 3, sous le nom de Théorème du rang.

1.6 Inversion de matrices

Il n'est pas évident à première vue de voir si une matrice est inversible ni de trouver cet inverse s'il existe. La méthode du pivot permet de répondre à ces deux questions.

Théorème 1.50. *Soit $A \in \mathscr{M}_n(\mathbb{R})$. La matrice A est inversible si et seulement si la matrice échelonnée réduite associée à $(A|I_n)$ est de la forme $(I_n|B)$. Dans ce cas, l'inverse de A est précisément la matrice B, c'est-à-dire $A^{-1} = B$.*

Démonstration. Si la forme échelonnée réduite de $(A|I_n)$ est $(I_n|B)$, cela signifie qu'il existe une suite de matrices élémentaires $E_1, \dots, E_N$ telle que $(I_n|B) = E_1 \cdots E_N(A|I_n)$. Cela implique, d'après la Proposition 1.43, que $E_1 \cdots E_N A = I_n$ et que $B = E_1 \cdots E_N$, c'est-à-dire que $BA = I_n$ et donc que A est inversible avec $A^{-1} = B$.

Inversement, puisque la seule matrice échelonnée réduite qui est inversible est I_n, si A est inversible alors la forme échelonnée réduite de $(A|I_n)$ est forcement de la forme $(I_n|B)$. □

Exemple 1.51. 1) Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Vérifions si elle est inversible.

$$(A|I_3) = \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{pmatrix} \to \begin{matrix} \\ L_2 - 4L_1 \\ L_3 - 7L_1 \end{matrix} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -12 & -7 & 0 & 1 \end{pmatrix}$$

$$\to \begin{matrix} \\ \\ L_3 - 2L2 \end{matrix} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

A cette étape, on voit que l'on n'arrivera jamais à avoir I_n à gauche et donc que A n'est pas inversible.

2) Par contre si $A' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 3 & 3 \end{pmatrix}$ alors

$$(A'|I_n) = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 3 & 3 & 0 & 0 & 1 \end{pmatrix} \to \begin{matrix} \\ \\ L_3 - 2L_1 \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & -2 & 0 & 1 \end{pmatrix}$$

$$\to \quad \begin{array}{c} \\ \\ L_3 - L_2 \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{pmatrix} \to \quad \begin{array}{c} L_1 - L_2 - L_3 \\ \\ \\ \end{array} \begin{pmatrix} 1 & 0 & 0 & 3 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{pmatrix}.$$

Donc A' est inversible et $A^{-1} = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix}$.

Chapitre 2

Espaces Vectoriels

La notion d'espace vectoriel est centrale aussi bien en mathématiques qu'en physique ou en informatique. Elle cherche à formaliser l'intuition d'"espace ambiant" à trois dimensions et ce faisant, donne un cadre bien plus général.

L'algèbre linéaire fournit un ensemble de techniques de calcul dans les espaces vectoriels qui s'avèrent particulièrement efficaces dans de nombreux problèmes, aussi bien théoriques qu'appliqués.

2.1 Définitions et exemples

La définition d'un espace vectoriel est abstraite et est rarement utilisée directement. Néanmoins, il est important de bien se familiariser avec.

Définition 2.1. Un *espace vectoriel* $(E, +, \cdot)$ est un ensemble E muni de deux lois

— une loi interne, notée $+ : E \times E \to E$, $(u, v) \mapsto u + v$,
— une loi externe, notée $\cdot : \mathbb{R} \times E \to E$, $(\lambda, u) \mapsto \lambda \cdot u$,

telles que les relations suivantes soient vérifiées :

$1)\forall (u, v) \in E^2,$	$u + v = v + u,$	$(+$ *est commutative*$)$
$2)\forall (u, v, w) \in E^3,$	$u + (v + w) = (u + v) + w,$	$(+$ *est associative*$)$
$3)\exists 0_E \in E, \forall u \in E,$	$u + 0_E = u,$	$(0_E$ *est l'élément neutre de* $+)$
$4)\forall u \in E, \exists -u \in E,$	$u + (-u) = 0_E,$	(*tout élément de* E *a un inverse*)
$5)\forall \lambda \in \mathbb{R}, (u, v) \in E^2,$	$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$	
$6)\forall (\lambda, \mu) \in \mathbb{R}^2, \forall u \in E,$	$\lambda \cdot u + \mu \cdot u = (\lambda + \mu) \cdot u$	
$7)\forall (\lambda, \mu) \in \mathbb{R}^2, \forall u \in E,$	$\lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u$	
$8)\forall u \in E,$	$1 \cdot u = u.$	

Notation 2.2. On abrégera “espace vectoriel” par “e.v”.

Définition 2.3. Les éléments d'un e.v sont appelés des *vecteurs* et ceux de $\mathbb{R}$ des *scalaires*.

L'exemple le plus élémentaire d'e.v est $\mathbb{R}^n$.

Exemple 2.4. Soit $n \geq 1$ un entier. On note $\mathbb{R}^n$ l'ensemble des n-uplets d'éléments de $\mathbb{R}$, c-à-d les objets de la forme $(x_1, x_2, \ldots, x_n)$ où pour tout $1 \leq i \leq n$, x_i est un élément de $\mathbb{R}$ et tel que $(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)$ si et seulement si $x_i = x'_i$ pour tout $1 \leq i \leq n$.

L'ensemble $\mathbb{R}^n$ muni des lois

— $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$,
— $\lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$,

est un e.v.

En fait, la Définition 2.1 généralise les règles de calcul sur $\mathbb{R}^n$ sans utiliser de système de coordonnées.

Proposition 2.5. *Soit E un e.v, $u \in E$ et $\lambda \in \mathbb{R}$. Alors on a*
1) $\lambda \cdot u = 0_E$ si et seulement si $\lambda = 0_{\mathbb{R}}$ ou $u = 0_E$,
2) $-u = (-1) \cdot u$.

Démonstration. 1) Si $\lambda = 0_{\mathbb{R}}$ alors

$$0_{\mathbb{R}} \cdot u = (0_{\mathbb{R}} + 0_{\mathbb{R}}) \cdot u = 0_{\mathbb{R}} \cdot u + 0_{\mathbb{R}} \cdot u,$$

d'où $0_{\mathbb{R}} \cdot u = 0_E$. De manière similaire, si $u = 0_E$ alors

$$\lambda \cdot (0_E) = \lambda \cdot (0_E + 0_E) = \lambda \cdot 0_E + \lambda \cdot 0_E,$$

et on obtient $\lambda \cdot 0_E = 0_E$.

Inversement, si $\lambda \cdot u = 0_E$ et $\lambda \neq 0_{\mathbb{R}}$ alors

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

2) $u + (-1) \cdot u = (1 - 1) \cdot u = 0_{\mathbb{R}} \cdot u = 0_E$. □

Notation 2.6. Pour simplicité, s'il n'y a pas de confusion possible le symbole $\cdot$ sera omis et on écrira 0 au lieu de 0_E ou $0_{\mathbb{R}}$.

Voici d'autres exemples d'e.v.

Exemple 2.7. 1) Si $m \geq 1$ et $n \geq 1$ sont deux entiers, alors $\mathscr{M}_{m,n}(\mathbb{R})$ munit des deux lois définies dans la Définition 1.10 et la Définition 1.12 est un e.v.
2) L'ensemble $\mathscr{F}(\mathbb{R}, \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ est un e.v. muni des lois suivantes. Si f et g sont deux fonctions de $\mathscr{F}(\mathbb{R}, \mathbb{R})$ alors la valeur de $f + g$ en x est par définition $(f + g)(x) := f(x) + g(x)$. De même, si $\lambda \in \mathbb{R}$ est un scalaire alors $(\lambda f)(x) := \lambda f(x)$ par définition.
3) Soit $a \in \mathbb{R}$. L'ensemble $\{(x, y) \in \mathbb{R}^2 \,|\, y = ax\}$ muni des lois induites par celles de $\mathbb{R}^2$ est un e.v.
4) Soit $(a, b) \in \mathbb{R}^2$. L'ensemble $\{(x, y) \in \mathbb{R}^2 \,|\, y = ax + b\}$ n'est pas un e.v si $b \neq 0$.

Définition 2.8. Soit E un e.v et $n \geq 1$ un entier. Soit $u_1, \dots, u_n$ des éléments de E. On appelle *combinaison linéaire* de $u_1, \dots, u_n$ tout vecteur v de la forme $v = \lambda_1 u_1 + \dots + \lambda_n u_n$ avec $\lambda_i \in \mathbb{R}$. On dira que la combinaison est *non triviale* si les λ_i ne sont pas tous nuls.

2.2 Sous-espaces vectoriels

Un moyen simple de vérifier qu'un ensemble est un e.v est de montrer qu'il est un sous-espace vectoriel d'un e.v qu'on connaît déjà.

Définition 2.9. Soit $(E, +, \cdot)$ un e.v. Un sous-ensemble F de E est une *sous-espace vectoriel* si
1) F est non vide,
2) F est stable pour $+$:
$\forall (u, v) \in F^2$, $u + v \in F$.
3) F est stable pour $\cdot$:
$\forall u \in F, \forall \lambda \in \mathbb{R}$, $\lambda u \in F$.

Si F est un sous-espace vectoriel de E, on le munit des lois induites par celles de E, ce qui en fait naturellement un e.v.

Notation 2.10. On abrégera "sous-espace vectoriel" par "s.e.v".

Exemple 2.11. 1) L'ensemble défini dans l'Exemple 2.7 3) est un s.e.v de $\mathbb{R}^2$.
2) Si $n \geq 1$ est un entier alors l'ensemble des matrices triangulaires supérieures de $\mathscr{M}_n(\mathbb{R})$ forme un s.e.v de $\mathscr{M}_n(\mathbb{R})$. Il en est de même avec l'ensemble de matrices triangulaires inférieures, les matrices symétriques ou encore les matrices anti-symétriques.
3) L'ensemble $\mathscr{C}(\mathbb{R})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un s.e.v de $\mathscr{F}(\mathbb{R}, \mathbb{R})$. En effet, $\mathscr{C}(\mathbb{R})$ n'est pas vide car les fonctions constantes sont dedans. De plus la somme de deux fonctions continues est continue et le produit d'une fonction continue par une constante est encore continue.
4) Si on considère l'équation différentielle $y'' + y = 0$ alors l'ensemble des solutions

$$S := \{f \in \mathscr{F}(\mathbb{R}, \mathbb{R}) \,|\, f \text{ est deux fois dérivable et } f'' + f = 0\}$$

est un s.e.v de $\mathscr{F}(\mathbb{R}, \mathbb{R})$.

Un moyen de détecter qu'un sous-ensemble n'est pas un s.e.v est donné par l'exercice suivant.

Exercice 2.12. Montrer que si $F \subset E$ est un s.e.v alors 0_E appartient à F.

Néanmoins, il n'est pas suffisant.

Exemple 2.13. Le sous-ensemble de $\mathbb{R}^2$ défini par

$$F := \{(x, y) \in \mathbb{R}^2 \,|\, y = x^2\}$$

n'est pas un s.e.v de $\mathbb{R}^2$ bien que $(0, 0) = 0_{\mathbb{R}^2} \in F$. En effet, $(1, 1)$ appartient à F mais $(-1) \cdot (1, 1) = (-1, -1)$ n'est pas dans F, donc F ne peut pas être un s.e.v.

La proposition suivante permet de vérifier plus rapidement qu'un sous-ensemble est un s.e.v.

Proposition 2.14. *Soit $(E, +, \cdot)$ un e.v. Un sous-ensemble $F \subset E$ est un s.e.v si et seulement si*
1) F est non vide.
2) $\forall (u, v) \in F^2, \forall \lambda \in \mathbb{R}, u + \lambda v \in F$.

Plus généralement, un s.e.v est un sous-ensemble stable par combinaisons linéaires.

Proposition 2.15. *Soit $u_1, \ldots, u_n$ des vecteurs d'un e.v E. Si F est un s.e.v de E contenant $u_1, \ldots, u_n$ alors F contient toutes les combinaisons linéaires de $u_1, \ldots, u_n$.*

Démonstration. On raisonne par récurrence sur n. Pour $n = 1$, une combinaison linéaire de u_1 est de la forme $\lambda_1 u_1$ avec $\lambda_1 \in \mathbb{R}$. Mais si u_1 appartient à F, alors, par définition d'un s.e.v, $\lambda_1 u_1$ aussi. Maintenant, supposons qu'il existe un $n \geq 1$ tel que la proposition soit vérifiée pour toute famille de n vecteurs. On en déduit que si $\sum_{i=1}^{n+1} \lambda_i u_i$ est une combinaison linéaire de $n+1$ vecteurs $u_1, \ldots, u_{n+1}$ de F alors par l'hypothèse de récurrence, $\sum_{i=1}^{n} \lambda_i u_i$ appartient à F. Mais, par la Proposition 2.14 cela implique que $\sum_{i=1}^{n+1} \lambda_i u_i = \sum_{i=1}^{n} \lambda_i u_i + \lambda_{n+1} u_{n+1}$ est aussi un élément de F. □

Il est facile de voir que l'intersection de s.e.v est encore un s.e.v.

Proposition 2.16. *L'intersection de deux s.e.v est un s.e.v. Plus généralement, l'intersection d'une famille de s.e.v est encore un s.e.v.*

Exemple 2.17. Dans $\mathscr{M}_n(\mathbb{R})$, l'intersection de l'ensemble de matrices triangulaires supérieures avec celui des matrices triangulaires inférieures est égale à l'ensemble de matrices diagonales qui est lui aussi un s.e.v.

Attention : En général, l'union de deux s.e.v n'en est pas un !

Exercice 2.18. Soit F et G deux s.e.v de E. Montrer que $F \cup G$ est un s.e.v si et seulement si $F \subset G$ ou $G \subset F$.

En revanche, il est possible de définir la somme de deux s.e.v F_1 et F_2 de E, qui est encore un s.e.v de E.

Définition 2.19. Soit E un e.v et $F_1, F_2 \subset E$ des s.e.v. La *somme* $F_1 + F_2$ de F_1 avec F_2 est définie par

$$F_1 + F_2 = \{u = u_1 + u_2 \,|\, u_1 \in F_1,\, u_2 \in F_2\}.$$

La somme est dite *directe* si $F_1 \cap F_2 = \{0\}$, et on écrira dans ce cas $F_1 \oplus F_2$. Dans ce cas, le couple $(u_1, u_2) \in F_1 \times F_2$ tel que $u = u_1 + u_2$ est unique.

Définition 2.20. Soit E un e.v et $F \subset E$ un s.e.v. Un supplémentaire $G \subset E$ de F est un s.e.v tel que $E = F \oplus G$. On dit aussi que F et G sont supplémentaires.

Attention : En général, le supplémentaire à F n'est pas unique !

Exemple 2.21. Soit $F = \{(x, 0) \in \mathbb{R}^2\}$. Les s.e.v $G_1 = \{(0, x) \in \mathbb{R}^2\}$ et $G_2 = \{(x, x) \in \mathbb{R}^2\}$ sont deux supplémentaires différents de F.

Exercice 2.22. Soit $\mathscr{P} \subset \mathscr{F}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions paires de $\mathbb{R}$ dans $\mathbb{R}$ et $\mathscr{I} \subset \mathscr{F}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions impaires.
1) Montrer que $\mathscr{P}$ et $\mathscr{I}$ sont des s.e.v de $\mathscr{F}(\mathbb{R}, \mathbb{R})$.
2) Montrer que $\mathscr{P}$ et $\mathscr{I}$ sont supplémentaires dans $\mathscr{F}(\mathbb{R}, \mathbb{R})$.

2.3 Familles génératrices, familles libres

Définition 2.23. Une *famille* d'un s.e.v E est simplement un sous-ensemble de E dont les éléments sont indexés. Dans ce cours, nous considérerons uniquement des familles finies, c-à-d de la forme $(u_1, \ldots, u_n)$ où chaque u_i est un vecteur de E pour $1 \leq i \leq n$.

Définition 2.24. Soit $(u_1, \ldots, u_n)$ une famille d'un e.v E. Le *sous-espace vectoriel engendré* par la famille $(u_1, \ldots, u_n)$ est l'ensemble de toutes les combinaisons linéaires possibles des $u_1, \ldots, u_n$. On le note $\mathrm{Vect}(u_1, \ldots, u_n)$.

Proposition 2.3.1. *L'ensemble* $\mathrm{Vect}(u_1, \ldots, u_n)$ *est, comme son nom l'indique, un s.e.v de* E.

Démonstration. En effet il est non vide car $u_1 \in \mathrm{Vect}(\{u_1, \ldots, u_n\})$ et si $v = \lambda_1 u_1 + \cdots + \lambda_n u_n$ et $v' = \lambda'_1 u_1 + \cdots + \lambda'_n u_n$ sont deux éléments de $\mathrm{Vect}(u_1, \ldots, u_n)$ et $\alpha \in \mathbb{R}$ alors $v + \alpha v' = (\lambda_1 + \alpha\lambda'_1)u_1 + \cdots + (\lambda_n + \alpha\lambda'_n)u_n$ est aussi une combinaison linéaire de $u_1, \ldots, u_n$ et donc un élément de $\mathrm{Vect}(u_1, \ldots, u_n)$. $\square$

Définition 2.25. Soit $(u_1, \ldots, u_n)$ une famille de vecteurs d'un e.v E. Soit F un sous-ensemble de E. On dit que la famille $(u_1, \ldots, u_n)$ *engendre* F si $F = \mathrm{Vect}(u_1, \ldots, u_n)$. Dans ce cas on dira aussi que la famille $(u_1, \ldots, u_n)$ est une *famille génératrice* de F. Lorsque $\mathrm{Vect}(u_1, \ldots, u_n) = E$ on dira simplement que $(u_1, \ldots, u_n)$ est une famille génératrice.

Exemple 2.26. 1) Soit $u = (1, 1) \in \mathbb{R}^2$. Une combinaison linéaire de u est simplement un vecteur de la forme λu où $\lambda \in \mathbb{R}$. On a donc que $\mathrm{Vect}(u) = \{(x, y) \mid x = y\}$ qui est la droite passant $(0, 0)$ dirigée par le vecteur u. Le vecteur u engendre donc cette droite.
2) Dans $\mathbb{R}^3$, les combinaisons linéaires des vecteurs $e_1 = (1, 0, 0)$ et $e_2 = (0, 1, 0)$ sont de la forme $v = \lambda_1 e_1 + \lambda_2 e_2$, c-à-d $v = (\lambda_1, \lambda_2, 0)$. Puisque λ_1 et λ_2 sont arbitraires, on voit que e_1 et e_2 engendre le plan d'équation $z = 0$:

$$\mathrm{Vect}(e_1, e_2) = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}.$$

3) Si nous considérons de nouveau l'ensemble S des solutions l'équation $y'' + y = 0$ défini dans l'Exemple 2.11 4) alors un résultat classique en équations différentielles dit que S est engendré par la famille $(\cos, \sin)$ de $\mathscr{F}(\mathbb{R}, \mathbb{R})$, c-à-d que toute solution f est de la forme $f(x) = \alpha \cos(x) + \beta \sin(x)$ où α et β sont deux constantes.

Définition 2.27. Soit E un e.v. Une famille de vecteur $(u_1, \ldots, u_n)$ est dite *libre* si toute combinaison linéaire vérifiant $\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$ vérifie aussi $\lambda_i = 0$ pour tout $1 \leq i \leq n$. Une famille qui n'est pas libre est dite *liée*.

Exemple 2.28. 1) Soit $u \in E$ avec $u \neq 0$. La famille à un élément (u) est alors une famille libre.
2) Toute famille contenant le vecteur nul 0 est liée.
3) Les vecteurs $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$ forment une famille libre dans $\mathbb{R}^3$.
4) La famille $(\cos, \sin)$ est une famille libre dans $\mathscr{F}(\mathbb{R}, \mathbb{R})$. En effet, supposons qu'il existe $(\lambda, \mu) \in \mathbb{R}^2$ tel que $\lambda \cos + \mu \sin = 0$, c'est-à-dire que pour tout x dans $\mathbb{R}$, $\lambda \cos(x) + \mu \sin(x) = 0$. En particulier, pour $x = 0$ on obtient que $0 = \lambda \cos(0) + \mu \sin(0) = \lambda$ et pour $x = \pi/2$, $0 = \lambda \cos(\pi/2) + \mu \sin(\pi/2) = \mu$. Donc $\lambda = \mu = 0$ dès que $\lambda \cos + \mu \sin = 0$, c'est-à-dire que notre famille est libre.

Définition 2.29. Deux vecteurs u et v d'un e.v E sont dit *colinéaires* s'il existe $\lambda \in \mathbb{R}$ tel que $u = \lambda v$ ou $v = \lambda u$.

Proposition 2.30. *Une famille de deux vecteurs* (u, v) *est liée si et seulement si* u *et* v *sont colinéaires.*

Attention : Il n'y a plus de tel critère pour une famille de plus de trois vecteurs !

Exemple 2.31. Les trois vecteurs $u_1 = (1, 0)$, $u_2 = (0, 1)$ et $u_3 = (1, 1)$ forment une famille liée bien qu'aucun d'eux n'est colinéaire à un autre. En effet $u_1 + u_2 - u_3 = 0$.

Proposition 2.32. *Soit* $(u_1, \ldots, u_n)$ *une famille libre d'un e.v* E *et* $v \in E$. *La famille* $(u_1, \ldots, u_n, v)$ *est liée si et seulement si* $v \in \mathrm{Vect}(u_1, \ldots, u_n)$.

Démonstration. Supposons que v est dans $\mathrm{Vect}(u_1, \ldots, u_n)$. Par définition, il existe des réels $\lambda_1, \ldots, \lambda_n$ tels que $v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$. Par conséquent, $\lambda_1 u_1 + \cdots + \lambda_n u_n + \lambda_{n+1} v = 0$ avec $\lambda_{n+1} = -1$ donne une combinaison linéaire non triviale de $u_1, \ldots, u_n, v$ qui s'annule et donc la famille est liée. Inversement, si la famille $(u_1, \ldots, u_n, v)$ est liée, il existe des réels $\lambda_1, \ldots, \lambda_{n+1}$, non tous nuls, tels que $\lambda_1 u_1 + \cdots + \lambda_n u_n + \lambda_{n+1} v = 0$. Le coefficient λ_{n+1} n'est pas nul car la famille $u_1, \ldots, u_n$ est libre. En divisant par λ_{n+1} on obtient

$$v = -\sum_{i=1}^{n} \frac{\lambda_i}{\lambda_{n+1}} u_i,$$

et donc $v \in \mathrm{Vect}(u_1, \ldots, u_n)$. □

Le théorème suivant est essentiel pour le prochain paragraphe.

Théorème 2.33. *Soit E un e.v engendré par une famille finie de n vecteurs $(u_1, \ldots, u_n)$. Alors, toute famille libre de E admet au plus n éléments.*

Démonstration. Supposons que E possède une famille de n vecteurs $(u_1, \ldots, u_n)$ qui l'engendre. Soit $m > n$ un entier et soit $(v_1, \ldots, v_m)$ une famille de m vecteurs. Notre but est de montrer que cette famille est liée, c-à-d qu'il existe des réels λ_i, $1 \leq i \leq m$, non tous nuls tels que

$$\sum_{i=1}^{m} \lambda_i v_i = 0.$$

Pour cela, nous allons considérer les λ_i comme des inconnues.

Par hypothèse, puisque la famille $(u_1, \ldots, u_n)$ engendre E, pour chaque $1 \leq i \leq n$ il existe n réels, $\alpha_{i,j}$ avec $1 \leq j \leq n$, tels que

$$v_i = \sum_{j=1}^{n} \alpha_{i,j} u_j.$$

En utilisant ces égalités, l'expression $\sum_{i=1}^{m} \lambda_i v_i = 0$ devient

$$\sum_{i=1}^{m} \lambda_i \left(\sum_{j=1}^{n} \alpha_{i,j} u_j \right) = 0.$$

En regroupant les termes contenant u_j on obtient

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} \lambda_i \alpha_{i,j} \right) u_j = 0.$$

Pour vérifier cette égalité, il suffit que les λ_i satisfassent

$$\sum_{i=1}^{m} \lambda_i \alpha_{i,j} = 0$$

pour chaque $1 \leq j \leq n$. On obtient ainsi un système homogène (le deuxième terme est nul) avec m inconnues (les $\lambda_1, \ldots, \lambda_m$) et n équations (une pour chaque j, $1 \leq j \leq n$). Puisque $m > n$, le Corollaire 1.46 dit qu'il y a une infinité de solution à ce système. En particulier, $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$ n'est pas la seule solution.

□

2.4 Bases et dimension

La dimension d'un e.v est l'invariant le plus important en algèbre linéaire. Dans une premier temps, nous voyons ce que signifie d'être de dimension finie.

Définition 2.34. Si un e.v E admet une famille génératrice finie, on dit qu'il est de *dimension finie*. Si ce n'est pas le cas, E est dit de *dimension infinie*.

Définition 2.35. Une famille libre et génératrice d'un e.v E est appelée une *base* de E.

Proposition 2.36. *Soit E un e.v et $(e_1, \dots, e_n)$ une base de E. Tout élément u de E s'écrit de manière unique comme combinaison linéaire de $(e_1, \dots, e_n)$. Les coefficients $\lambda_1, \dots, \lambda_n$ tels que $u = \lambda_1 e_1 + \cdots + \lambda_n e_n$ sont appelés les coordonnées de u dans la base $(e_1, \dots, e_n)$.*

Démonstration. Puisque la famille $(e_1, \dots, e_n)$ est une base, en particulier elle est génératrice. Donc, si u est un vecteur de E, il existe des réels $\lambda_1, \dots, \lambda_n$ tels que $u = \lambda_1 e_1 + \cdots + \lambda_n e_n$. Ceci montre l'existence des coordonnées.

Il reste à montrer l'unicité. Supposons que $\alpha_1, \dots, \alpha_n$ soient aussi des réels tels que $\alpha_1 e_1 + \cdots + \alpha_n e_n = u$. On en déduit que

$$\sum_{i=1}^{n} \alpha_i e_i = \sum_{i=1}^{n} \lambda_i e_i$$

et donc que

$$\sum_{i=1}^{n} (\alpha_i - \lambda_i) e_i = 0.$$

Or la famille $(e_1, \dots, e_n)$ est libre donc on doit forcément avoir $(\alpha_i - \lambda_i) = 0$ pour tout $1 \leq i \leq 0$, c'est-à-dire $\alpha_i = \lambda_i$. $\square$

Remarque 2.37. Il existe de nombreuses bases d'un même e.v E, qui donneront autant de systèmes de coordonnées différents.

Exemple 2.38. 1) Les vecteurs $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$ forment une base de $\mathbb{R}^3$.
2) Plus généralement, les vecteurs $e_i = (0, \dots, 0, 1, 0 \dots, 0)$, où le 1 se trouve à la i-ème place, donnent une base de $\mathbb{R}^n$, appelée base *canonique*.

Le Théorème 2.33 entraîne le résultat suivant.

Théorème 2.39. *Soit E une e.v de dimension finie. Toutes les bases de E ont le même cardinal (i.e. le même nombre d'éléments).*

Démonstration. Soit $(e_1, \dots, e_n)$ et $(f_1, \dots, f_m)$ deux bases de E. Par définition, elles sont toutes les deux libres et génératrices. Si $n < m$ alors par le Théorème 2.33, la famille $(f_1, \dots, f_m)$ est liée ce qui est une contradiction. Donc $n \leq m$ et par symétrie, $n = m$. $\square$

On peut maintenant poser la définition suivante.

Définition 2.40. Si E est un e.v de dimension finie alors la *dimension* de E est le cardinal d'une de ses bases. On la note $\dim(E)$ et par convention $\dim(E) = 0$ si $E = \{0\}$.

Exemple 2.41. 1) Puisque la famille des vecteurs e_i $1 \leq i \leq n$ est une base de $\mathbb{R}^n$, cet e.v est de dimension n, comme attendu.
2) Soit $a \in \mathbb{R}$. Le s.e.v de $\mathbb{R}^2$ défini par $\{(x, y) \in \mathbb{R}^2 \,|\, y = ax\}$ est de dimension 1 car le vecteur $(1, a)$ en est une base.

Les résultats suivants sont très utiles dans la pratiques.

Théorème 2.42. *Soit E un e.v de dimension finie, $\mathscr{G}$ une famille génératrice et $\mathscr{L}$ une famille libre telles que $\mathscr{L} \subset \mathscr{G}$. Alors, il existe une base $\mathscr{B}$ de E telle que $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$.*

Démonstration. Soit $\mathscr{B}$ un sous-ensemble de $\mathscr{G}$ tel que $\mathscr{L} \subset \mathscr{B}$, que $\mathscr{B}$ forme une famille libre et que $\mathscr{B}$ soit de cardinal maximal pour cette propriété. L'existence d'un tel $\mathscr{B}$ est garantie par le Théorème 2.33 et par le fait que E soit de dimension finie.

Montrons que $\mathscr{B}$ est une base de E. Puisque $\mathscr{B}$ est une famille libre, il reste à prouver qu'elle est aussi génératrice. Or, si $\mathrm{Vect}(\mathscr{B}) \neq E = \mathrm{Vect}(\mathscr{G})$, alors par la Proposition 2.32 il existe $u \in \mathscr{G} \setminus \mathscr{B}$ tel que $\mathscr{B}' := \mathscr{B} \cup \{u\} \subset \mathscr{G}$ est libre et vérifie $\mathrm{Card}(\mathscr{B}') = \mathrm{Card}(\mathscr{B}) + 1$ ce qui contredit la définition de $\mathscr{B}$. Donc, on obtient par l'absurde que la famille $\mathscr{B}$ est une base de E vérifiant $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$. □

Corollaire 2.43. *Soit E un e.v de dimension finie.*
1) Toute famille libre de E peut-être complétée en une base (Théorème de la base incomplète).
2) De toute famille génératrice, on peut extraire une base.
3) E possède une base.

Corollaire 2.44. *Soit E un e.v de dimension n.*
1) Toute famille libre a au plus n éléments, avec égalité si et seulement si c'est une base.
2) Toute famille génératrice a au moins n éléments, avec égalité si et seulement si c'est une base.

Ils permettent entre autre de calculer la dimension de la somme de s.e.v.

Proposition 2.45. *Soit F et G deux s.e.v d'un e.v E de dimension finie. Alors, la dimension de $F + G$ vérifie* $\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)$.

Corollaire 2.46. *Soit F et G deux s.e.v d'un e.v E de dimension finie. Les assertions suivantes sont équivalentes.*
1) F et G sont des supplémentaires.
2) $\dim(E) = \dim(F) + \dim(G)$ *et* $F \cap G = \{0\}$.
3) $\dim(E) = \dim(F) + \dim(G)$ *et* $E = F + G$.

Exemple 2.47. 1) L'ensemble des polynômes $\mathbb{R}[X]$ est de dimension infinie. En effet, pour tout $n \geq 1$, la famille $(1, X, \ldots, X^n)$ est libre ce qui implique par le Corollaire 2.44 que $\dim(\mathbb{R}[X]) > n$.
2) De même, l'e.v $\mathscr{F}(\mathbb{R}, \mathbb{R})$ est de dimension infinie, ainsi que son s.e.v des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$.

2.5 Rang d'un système de vecteurs et rang d'une matrice

Le rang d'une matrice est un invariant important qui a plusieurs interprétations différentes.

Définition 2.48. Soit $(v_1, \ldots, v_n)$ une famille de vecteurs. Le *rang* de $(v_1, \ldots, v_n)$ est le nombre maximal de v_i qui sont linéairement indépendant. Dit autrement, c'est la dimension de $\mathrm{Vect}(v_1, \ldots, v_n)$. On le note $\mathrm{rg}(v_1, \ldots, v_n)$.

On identifie les vecteurs v_i aux colonnes C_i d'une matrice $A \in \mathscr{M}_{m,n}(\mathbb{R})$, on a alors le résultat suivant :

Proposition 2.49. *Le rang d'une matrice A est égal au rang de son système de vecteurs colonnes, i.e.le nombre de pivot d'une de ses formes échelonnées.*

On obtient aussi immédiatement le résultat suivant :

Proposition 2.50. *Si $A \in \mathscr{M}_{m,n}(\mathbb{R})$ alors* $\mathrm{rg}(A) \leq \min(m,n)$.

La proposition suivante dit que l'espace engendré par les colonnes et celui engendré par les lignes d'une matrice sont de même dimension.

Proposition 2.51. *Soit $A \in \mathscr{M}_{m,n}(\mathbb{R})$. Le rang de A est égal au rang de ${}^tA$.*

Démonstration. Soit $A = (a_{ij}) \in \mathscr{M}_{m,n}(\mathbb{R})$. On note C_i, $1 \leq i \leq n$, les colonnes de A et L_j, $1 \leq j \leq m$ ses lignes. Rappelons qu'on identifie une colonne à un vecteur de $\mathbb{R}^m$ et une ligne à un vecteur de $\mathbb{R}^n$.

Le résultat équivaut à montrer que $\dim(\mathrm{Vect}(C_1,\ldots,C_n)) = \dim(\mathrm{Vect}(L_1,\ldots,L_m))$. Posons $r = \dim(\mathrm{Vect}(L_1,\ldots,L_m))$ et choisissons une base $(u_1,\ldots,u_r)$ de $\mathrm{Vect}(L_1,\ldots,L_m)$. Par définition, il existe des réels λ_{ij}, $1 \leq i \leq m$, $1 \leq i \leq r$, tels que

$$\begin{array}{lll} L_1 & = \lambda_{11}u_1 + \cdots + & \lambda_{1r}u_r \\ \vdots & & \vdots \\ L_m & = \lambda_{m1}u_1 + \cdots + & \lambda_{mr}u_r \end{array}.$$

Chacune de ces équations est une égalité entre vecteur de $\mathbb{R}^n$ et donc une égalité coordonnées par coordonnées. Cela implique que si $u_1 = (\alpha_{11},\ldots,\alpha_{1n}), \cdots, u_r = (\alpha_{r1},\ldots,\alpha_{rn})$ alors

$$C_1 = \alpha_{11}\begin{pmatrix}\lambda_{11}\\ \lambda_{21}\\ \vdots\\ \lambda_{m1}\end{pmatrix} + \alpha_{21}\begin{pmatrix}\lambda_{12}\\ \lambda_{22}\\ \vdots\\ \lambda_{m2}\end{pmatrix} + \cdots + \alpha_{r1}\begin{pmatrix}\lambda_{1r}\\ \lambda_{2r}\\ \vdots\\ \lambda_{mr}\end{pmatrix},$$

et de même pour chaque C_i, $1 \leq i \leq n$.

Les C_i sont donc engendrées par r vecteurs $\begin{pmatrix}\lambda_{11}\\ \lambda_{21}\\ \vdots\\ \lambda_{m1}\end{pmatrix}, \ldots, \begin{pmatrix}\lambda_{1r}\\ \lambda_{2r}\\ \vdots\\ \lambda_{mr}\end{pmatrix}$ ce qui implique que $\dim(\mathrm{Vect}(C_1,\ldots,C_n)) \leq r$. L'inégalité inverse se prouve de la même façon en raisonnant sur ${}^tA$ et on obtient bien une égalité entre ces deux dimensions. □

Voici deux applications du concept de rang.

Proposition 2.52. *Soit $u_1,\ldots,u_n$ des vecteurs de $\mathbb{R}^m$. La famille $(u_1,\ldots,u_n)$ est libre si et seulement si la matrice dont les colonnes sont les u_i est de rang n.*

Proposition 2.53. *Une matrice de $\mathscr{M}_n(\mathbb{R})$ est inversible si et seulement si sont rang est n.*

2.6 L'espace vectoriel des polynômes

Pour finir ce chapitre, nous introduisons un exemple particulièrement important d'espace vectoriel : l'ensemble des polynômes à coefficients réels. Les polynômes jouent un rôle central dans de nombreuses applications. Une des raisons est que ce sont des objets algébriques à la définition simple et facilement manipulable par un ordinateur.

Pour commencer, on choisit un symbole X qu'on appellera "indéterminé". Pour chaque entier $n \geq 1$ on notera formellement X^n la puissance n-ième de X. On prendra comme convention que $X^0 = 1$. Si $d \geq 0$ est un entier alors un *polynôme* P en X de *degré* d est une expression de la forme

$$P = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0 = \sum_{n=0}^{d} a_n X^n$$

où les a_n sont des nombres réels avec $a_d \neq 0$. Les a_n sont les *coefficients* de P et on pose par convention que $a_n = 0$ si $n > d$. On obtient alors qu'un polynôme P est une expression de la forme $P = \sum_{n \geq 0} a_n X^n$ où les a_n sont tous nuls sauf un nombre fini. On retrouve alors le degré de P par la formule

$$\deg(P) = \max\{n \in \mathbb{N} \mid a_n \neq 0\}.$$

Il est important de comprendre quand deux polynômes sont égaux. La définition, simple, est la suivante.

Définition 2.54. Deux polynômes $P = \sum_{n \geq 0} a_n X^n$ et $Q = \sum_{n \geq 0} b_n X^n$ sont égaux si $a_n = b_n$ pour tout $n \geq 0$.

Remarque 2.55. Le polynôme dont tous les coefficients sont nuls, le polynôme nul $P = 0$ joue un rôle particulier. Par convention on pose $\deg(0) = -\infty$.

On peut définir naturellement la somme de deux polynômes. Si $P = \sum_{n \geq 0} a_n X^n$ et $Q = \sum_{n \geq 0} b_n X^n$ alors on définie $P + Q$ par $P + Q = \sum_{n \geq 0} (a_n + b_n) X^n$. Il est facile de voir que $\deg(P + Q) \leq \max\{\deg(P), \deg(Q)\}$.

Exemple 2.56. Si $P = X^2 + 3X - 10$ et $Q = -X^{10} + 2$ alors $P + Q = -X^{10} + X^2 + 3X - 8$.

De même on peut multiplier un polynôme par un scalaire. Si $P = \sum_{n \geq 0} a_n X^n$ est un polynôme et $\lambda \in \mathbb{R}$ est un réel alors λP est défini par $\lambda P = \sum_{n \geq 0} (\lambda a_n) X^n$. Si $\lambda \neq 0$ alors $\deg(\lambda P) = \deg(P)$.

Exemple 2.57. Si $P = -X^6 + 3X^2 - 5X$ alors $-3P = 3X^6 - 9X^2 + 15$.

Proposition 2.6.1. *L'ensemble de tous les polynômes muni des deux lois ci-dessus est un espace vectoriel. On le note* $\mathbb{R}[X]$. *Son zéro est le polynôme nul* $P = 0$.

L'e.v $\mathbb{R}[X]$ possède les s.e.v naturels suivants.

Proposition 2.6.2. *Soit* $d \in \mathbb{N}$. *L'ensemble* $\mathbb{R}_d[X]$ *des polynômes réels de degré inférieur ou égal à* d *est un s.e.v de* $\mathbb{R}[X]$.

Démonstration. En effet, $\mathbb{R}_d[X]$ est non vide car le polynôme nul est dedans. De plus, si $(P, Q) \in (\mathbb{R}[X])^2$ et $\lambda \in \mathbb{R}$ alors $\deg(P + \lambda Q) \leq \max(\deg(P), \deg(Q))$. □

Remarque 2.58. Si $d \geq 1$, l'ensemble des polynômes de degré exactement d n'est pas un s.e.v de $\mathbb{R}[X]$ car il n'est pas stable par $+$. Par exemple $(X^d + 1) + (1 - X^d) = 2$ qui est de degré 0.

On peut illustrer plusieurs notions dans ces s.e.v

Exemple 2.59. Soit $F = \{aX + b \,|\, (a, b) \in \mathbb{R}^2\}$ et $G = \{cX^2 + d \,|\, (c, d) \in \mathbb{R}^2\}$. Ce sont des s.e.v de $\mathbb{R}[X]$, et on a que $F + G = \mathbb{R}_2[X]$. En effet, si $P \in \mathbb{R}_2[X]$ alors $P(X) = \alpha X^2 + \beta X + \gamma$ avec $\beta X + \gamma \in F$ et $\alpha X^2 \in G$, d'où $\mathbb{R}_2[X] \subset F + G$. L'autre inclusion est évidente. En revanche, la somme n'est pas directe car l'intersection de F et G est l'ensemble des polynômes constants.

Chacun des s.e.v $\mathbb{R}_d[X]$ possède une base naturelle.

Proposition 2.6.3. *La famille* $(1, X, X^2, \ldots, X^d)$ *est une base de* $\mathbb{R}_d[X]$. *Par conséquent* $\mathbb{R}_d[X]$ *est de dimension* $d + 1$.

Mais il existe de nombreuses autres bases.

Exemple 2.60. La famille (P_0, P_1, P_2) avec $P_0 = 2$, $P_1 = X + 2$, $P_2 = X^2 + X + 1$ est une base de $\mathbb{R}_2[X]$. En effet, la famille a 3 éléments et on a déjà vu que $\mathbb{R}_2[X]$ est de dimension 3. Donc par le Corollaire 2.44 il suffit de montrer que la famille est libre. Pour cela, choisissons une combinaison linéaire $\lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2$ qui s'annule. En utilisant les expressions des P_i on obtient

$$\lambda_2 X^2 + (\lambda_2 + \lambda_1)X + (\lambda_2 + 2\lambda_1 + 2\lambda_0) = 0.$$

Puisque un polynôme est nul si et seulement si tous ses coefficients sont nuls, on obtient les équations

$$\begin{cases} \lambda_2 = 0 \\ \lambda_2 + \lambda_1 = 0 \\ \lambda_2 + 2\lambda_1 + 2\lambda_0 = 0 \end{cases},$$

et on voit facilement que la seule solution est $\lambda_2 = \lambda_1 = \lambda_0 = 0$. Donc la famille est libre et c'est donc une base de $\mathbb{R}_2[X]$.

Chapitre 3

Applications linéaires

Les applications linéaires sont les applications entre espaces vectoriels qui préservent leur structure. De nombreuses transformations géométriques comme les rotations, les symétries ou les projections en font partie.

3.1 Définition et premières propriétés

Définition 3.1. Soit E et F deux e.v. Une application $f : E \to F$ est dite *linéaire* si
- $f(u+v) = f(u) + f(v)$ pour tout u et v dans E,
- $f(\lambda u) = \lambda f(u)$ pour tout $u \in E$ et $\lambda \in \mathbb{R}$.

Exemple 3.2. 1) L'application entre E et F qui envoie tout élément de E sur 0 est une application linéaire. On l'appelle l'application nulle.
2) L'application définie par

$$\begin{aligned} \mathrm{Id}_E : E &\to E \\ u &\mapsto u \end{aligned}$$

est une application linéaire appelée *identitée de E*. On la note Id s'il n'y a pas d'ambiguïté.
3) Soit $a \in \mathbb{R}$. L'application linéaire de $\mathbb{R}$ dans lui-même définie par $x \mapsto ax$ est linéaire. En fait, tout application linéaire de $\mathbb{R}$ dans lui-même est de cette forme.
4) L'application

$$\begin{aligned} \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (x,0) \end{aligned}$$

est une application linéaire. C'est la projection sur l'axe des x parallèlement à l'axe des y.
5) L'application de $\mathbb{R}[X]$ dans lui même qui à un polynôme P associe sont polynôme dérivé P' est linéaire. En effet $(P+Q)' = P' + Q'$ et $(\lambda P)' = \lambda(P)'$.

On obtient facilement ces premières propriétés.

Proposition 3.3. *Soit f une application entre deux e.v E et F.*
1) *Si f est linéaire alors $f(0) = 0$.*
2) *f est linéaire si et seulement si $f(u + \lambda v) = f(u) + \lambda f(v)$, pour tout $u, v \in E$ et $\lambda \in \mathbb{R}$.*

Exemple 3.4. L'application de $\mathbb{R}$ dans $\mathbb{R}$ défini par $x \mapsto x + 1$ ne peut pas être linéaire car elle vaut 1 en 0.

Proposition 3.5. *Si $f : E \to F$ est une application linéaire et que $\sum_{i=1}^{l} \lambda_i u_i$ est une combinaison linéaire dans E alors*

$$f\left(\sum_{i=1}^{l} \lambda_i u_i\right) = \sum_{i=1}^{l} \lambda_i f(u_i).$$

Proposition 3.6. *Si $\lambda \in \mathbb{R}$ et f, g sont deux applications linéaires entre E et F alors $f + g$ et λf aussi.*

De plus, l'ensemble des applications linéaires de E dans F est une e.v que l'on note $\mathscr{L}(E, F)$.

Proposition 3.7. *Si f et g sont deux applications linéaires telles que $f \circ g$ est bien défini alors c'est aussi une application linéaire.*

Un moyen d'obtenir beaucoup exemples est d'utiliser les matrices.

Proposition 3.8. *Si A est une matrice de $\mathscr{M}_{m,n}(\mathbb{R})$ alors on peut lui associer une application linéaire f_A de $\mathbb{R}^n$ dans $\mathbb{R}^m$ par*

$$f_A : \mathbb{R}^n \to \mathbb{R}^m$$
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Démonstration. Si u et v sont des vecteurs de $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$ alors $A(u + \lambda v) = Au + \lambda Av$. Donc f_A est linéaire. □

En fait, on obtient toutes les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ de cette manière.

Proposition 3.9. *Si f est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ alors il existe une unique matrice $A \in \mathscr{M}_{m,n}(\mathbb{R})$ telle que $f = f_A$.*

Démonstration. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. On définit la matrice A par

$$A = (f(e_1)| \cdots |f(e_n)),$$

où $f(e_i) \in \mathbb{R}^m$ est représenté comme un vecteur colonne. On obtient bien une matrice de taille $m \times n$. De plus, on observe facilement que $A \times e_i = f(e_i)$ pour tout $1 \leq i \leq n$. Par ailleurs, si $u \in \mathbb{R}^n$, il existe $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $u = \sum_{i=1}^{n} \lambda_i e_i$. La Proposition 3.5 implique alors que

$$f(u) = f\left(\sum_{i=1}^{n} \lambda_i e_i\right) = \sum_{i=1}^{n} \lambda_i f(e_i) = \sum_{i=1}^{n} \lambda_i A e_i = A\left(\sum_{i=1}^{n} \lambda_i e_i\right) = f_A(u),$$

ce qui montre bien que $f = f_A$. L'unicité de A vient du fait que si B vérifie aussi $f_B = f$ alors $Be_i = f(e_i) = Ae_i$ ce qui implique que la i-ème colonne de B coïncide avec celle de A. Donc $A = B$. □

La preuve ci-dessus donne en plus une formule explicite de A.

Définition 3.10. Si $f : \mathbb{R}^n \to \mathbb{R}^m$ est une application linéaire alors la matrice

$$A = (f(e_1)| \cdots |f(e_n))$$

s'appelle *la matrice associée à f* (dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$). Elle vérifie $f_A = f$.

Exemple 3.11. 1) L'application de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+z \\ -x+2y+5z \end{pmatrix}.$$

est linéaire est sa matrice est $\begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 5 \end{pmatrix}$
2) Soit $\theta \in [0, 2\pi[$. L'application associée à la matrice

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

est la rotation de centre 0 et d'angle θ.

Proposition 3.12. *Si $A, B \in \mathscr{M}_{m,n}(\mathbb{R})$, $C \in \mathscr{M}_{n,p}(\mathbb{R})$ et $\lambda \in \mathbb{R}$ alors*

$$f_{A+B} = f_A + f_B, \quad f_{\lambda A} = \lambda f_A \quad \text{et} \quad f_{AC} = f_A \circ f_C.$$

Remarque 3.13. L'égalité $f_{AC} = f_A \circ f_C$ est en quelque sorte la raison d'être de la définition, peu intuitive, du produit matriciel.

3.2 Noyau, image et rang

Puisque qu'une application linéaire respecte la structure d'e.v, elle envoie s.e.v sur s.e.v. Plus exactement, on a la proposition suivante.

Proposition 3.14. *Soit $f : E \to F$ une application linéaire.*
1) Si G est un s.e.v de E alors l'image de G par f,

$$f(G) := \{y \in F \mid \exists x \in G, f(x) = y\},$$

est un s.e.v de F.
2) Si H est un s.e.v de F alors l'image réciproque de H par f,

$$f^{-1}(H) := \{x \in E \mid f(x) \in H\},$$

est un s.e.v de E.

Démonstration. 1) Soit G un s.e.v de E. On a déjà vu que $0 \in G$ et que $f(0) = 0$ donc $f(G)$ contient 0 et est donc non-vide. Soit y_1, y_2 deux éléments de $f(G)$ et soit $\lambda \in \mathbb{R}$. Par définition, il existe deux éléments x_1 et x_2 de G tels que $f(x_1) = y_1$ et $f(x_2) = y_2$. Puisque G est un s.e.v on a que $x_1 + x_2 \in G$ et $\lambda x_1 \in G$. Par ailleurs, le fait que f soit linéaire implique que $f(x_1 + x_2) = f(x_1) + f(x_2)$ et $f(\lambda x_1) = \lambda f(x_1)$. Par conséquent, $y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2)$ avec $x_1 + x_2 \in G$ donc $y_1 + y_2 \in f(G)$. De même, $\lambda y_1 = \lambda f(x_1) = f(\lambda x_1)$ avec $\lambda x_1 \in G$ donc $\lambda y_1 \in f(G)$. On a bien vérifié les trois conditions pour avoir un s.e.v donc $f(G)$ est un s.e.v.

2) Soit H un s.e.v de F. Comme ci-dessus, on sait que 0 appartient à H et que $f(0) = 0$ donc $0 \in f^{-1}(H)$. On en déduit que $f^{-1}(H)$ est non-vide. Pour aller un peu plus vite, nous allons utiliser la Proposition 2.14 pour montrer qu'un coup la stabilité par addition et celle par multiplication. Soit x et x' deux éléments de $f^{-1}(H)$. Par définition, ils vérifient $f(x) \in H$ et $f(x') \in H$. De plus, si $\lambda \in \mathbb{R}$ alors $f(x + \lambda x') = f(x) + \lambda f(x')$ car f est linéaire. Donc, H étant un s.e.v, on a bien que $f(x + \lambda x') \in H$, c-à-d $x + \lambda x' \in f^{-1}(H)$. $\square$

Cela permet d'associer naturellement à f deux s.e.v.

Définition 3.15. Si $f : E \to F$ est une application linéaire alors on appelle *image de f* le s.e.v de F défini par

$$\operatorname{Im}(f) = f(E) = \{y \in F \mid \exists x \in E, f(x) = y\}.$$

La dimension de $\operatorname{Im}(f)$ s'appelle le rang de f, $\operatorname{rg}(f) = \dim(\operatorname{Im}(f))$.

Définition 3.16. Si $f : E \to F$ est une application linéaire alors on appelle *noyau de f* le s.e.v de E défini par

$$\operatorname{Ker}(f) = f^{-1}(0) = \{x \in E \mid f(x) = 0\}.$$

Certaines propriétés de f, comment l'injectivité ou la surjectivité, se lisent directement sur ces s.e.v. Mais rappelons d'abord ces notions.

Définition 3.17. Soit f une application entre deux ensembles E et F.
— f est dite *injective* si le fait que $f(x) = f(y)$ entraîne que $x = y$.
— f est dite *surjective* si $f(E) = F$.
— f est dite *bijective* si elle est à la fois injective et surjective.

Proposition 3.18. *Soit $f : E \to F$ une application linéaire.*
1) Elle est injective si et seulement si $\operatorname{Ker}(f) = \{0\}$.
2) Si la dimension de F est n alors f est surjective si et seulement si $\operatorname{rg}(f) = n$.

Démonstration. 1) Puisque f est linéaire, on a que $f(0) = 0$.

Supposons maintenant que f est injective. Si x est dans le noyau de f alors $f(x) = 0 = f(0)$ et donc par injectivité $x = 0$. Par conséquent, $\operatorname{Ker}(f) = \{0\}$.

Inversement, supposons que $\operatorname{Ker}(f) = \{0\}$. Si x et y sont deux vecteurs tels que $f(x) = f(y)$, alors par linéarité $f(x-y) = 0$. Cela implique que $x - y \in \operatorname{Ker}(f)$ et donc que $x = y$ car $\operatorname{Ker}(f) = \{0\}$.

2) On a toujours l'inclusion $f(E) \subset F$. Par conséquent, f est surjective si et seulement si $f(E) = F$ et donc si et seulement si $\operatorname{rg}(f) = \dim(\operatorname{Im}(f)) = \dim(f(E)) = n$. □

Le théorème suivant est important aussi bien théoriquement que dans la pratique.

Théorème 3.19 (Théorème du rang)**.** *Soit E et F deux e.v de dimension finie. Si $f : E \to F$ est une application linéaire alors*

$$\dim(E) = \operatorname{rg}(f) + \dim(\operatorname{Ker}(f)).$$

Démonstration. Soit $f : E \to F$ une application linéaire entre e.v de dimension finie. Soit $(u_1, \ldots, u_k)$ une base de $\operatorname{Ker}(f)$ où, par définition, $\dim(\operatorname{Ker}(f)) = k$. En particulier, la famille $(u_1, \ldots, u_k)$ est libre et on peut, par le théorème de la base incomplète, la compléter en une base $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ de E. On a alors $\dim(E) = k + r$. Pour obtenir le théorème il suffit donc de montrer que $r = \operatorname{rg}(f)$. Pour cela, on va montrer que la famille $(f(v_1), \ldots, f(v_r))$ est une base de $\operatorname{Im}(f)$, c'est-à-dire qu'elle est libre et génératrice de $\operatorname{Im}(f)$.

La première remarque est que $f(v_i) \in \operatorname{Im}(f)$ pour tout $1 \le i \le r$, donc

$$\operatorname{Vect}(f(v_1), \ldots, f(v_r)) \subset \operatorname{Im}(f).$$

Pour l'inclusion inverse, choisissons arbitrairement $w \in \mathrm{Im}(f)$. Par définition de l'image, il existe un vecteur $v \in E$ tel que $w = f(v)$. Or $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ est une base de E donc il existe des réels $\alpha_1, \ldots, \alpha_k$ et $\beta_1, \ldots, \beta_r$ tels que

$$v = \sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{r} \beta_j v_j.$$

Par linéarité, on a alors que

$$w = f(v) = f\left(\sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{r} \beta_j v_j\right) = \sum_{i=1}^{k} \alpha_i f(u_i) + \sum_{j=1}^{r} \beta_j f(v_j).$$

Mais les vecteurs u_i, $1 \leq i \leq k$, sont dans le noyau de f donc $f(u_i) = 0$. Cela implique que

$$w = \sum_{j=1}^{r} \beta_j f(v_j)$$

et donc par conséquent $w \in \mathrm{Vect}(f(v_1), \ldots, f(v_r))$. Puisque le choix de w était arbitraire, on obtient

$$\mathrm{Im}(f) \subset \mathrm{Vect}(f(v_1), \ldots, f(v_r)).$$

Les deux inclusions impliquent l'égalité et donc la famille $(f(v_1), \ldots, f(v_r))$ est bien génératrice de $\mathrm{Im}(f)$.

Il reste à montrer que cette famille est libre. Soit $\lambda_1, \ldots, \lambda_r$ des scalaires tels que $\lambda_1 f(v_1) + \cdots + \lambda_r f(v_r) = 0$. Par linéarité, on a

$$\sum_{i=1}^{r} \lambda_i f(v_i) = f\left(\sum_{i=1}^{r} \lambda_i v_i\right) = 0$$

donc $\sum_{i=1}^{r} \lambda_i v_i \in \mathrm{Ker}(f)$. Or $(u_1, \ldots, u_k)$ est une base de $\mathrm{Ker}(f)$ donc il existe des scalaires $\alpha_1, \ldots, \alpha_k$ tels que

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = \alpha_1 u_1 + \cdots + \alpha_k u_k,$$

ou écrit autrement

$$\lambda_1 v_1 + \cdots + \lambda_r v_r - \alpha_1 u_1 - \cdots - \alpha_k u_k = 0.$$

Mais la famille $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ est libre donc $\lambda_1 = \cdots = \lambda_r = -\alpha_1 = \cdots = -\alpha_k = 0$. On est parti d'une combinaison linéaire des $f(v_i)$ qui s'annulait et on en a conclu qu'elle était forcément triviale, ce qui signifie que la famille $(f(v_1), \ldots, f(v_r))$ est libre. $\square$

Corollaire 3.20. *Soit E et F deux e.v de dimension n et m respectivement. Soit f une application linéaire entre E et F.*
1) Si $n > m$ alors f n'est jamais injective.
2) Si $n < m$ alors f n'est jamais surjective.
3) Si $n = m$ alors f est injective si et seulement si elle est surjective et donc si et seulement si elle est bijective.

Démonstration. 1) Supposons que $n > m$. D'après le théorème du rang, $n = \mathrm{rg}(f) + \dim(\mathrm{Ker}(f))$. De plus $\mathrm{Im}(f) \subset F$ donc $\mathrm{rg}(f) \leq m$. Par conséquent, $\dim(\mathrm{Ker}(f)) = n - \mathrm{rg}(f) \geq n - m > 0$ et donc $\mathrm{Ker}(f) \neq \{0\}$. Par la Proposition 3.18, f n'est donc pas injective.
2) Supposons que $n < m$. Toujours d'après le théorème du rang, $\mathrm{rg}(f) = n - \dim(\mathrm{Ker}(f)) \leq n < m$ donc d'après la Proposition 3.18, f n'est pas surjective.
3) L'égalité $\mathrm{rg}(f) = n - \dim(\mathrm{Ker}(f))$ implique que si $n = m$ alors $\mathrm{rg}(f) = m$ si et seulement si $\dim(\mathrm{Ker}(f)) = 0$, et donc $\mathrm{Ker}(f) = \{0\}$. $\square$

3.2.1 Calculs pratiques

Dans cette section, nous allons voir comment déterminer, aussi bien en termes d'équations que de bases, l'image et le noyau d'une application linéaire.

Dans la suite, f sera une application linéaire entre $\mathbb{R}^n$ et $\mathbb{R}^m$ donnée par une matrice A.

Si l'on veut déterminer l'image de f, il faut trouver tous les $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m$ tels qu'il existe un $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ avec $AX = Y$. Pour cela, on échelonne et réduit la matrice $(A|Y)$ pour obtenir une matrice $(A'|Y')$.

-Le rang de f est le nombre de pivots dans la matrice A'.

-Les équations de l'image sont les équations en $y_1, \ldots, y_m$ qui garantissent qu'il n'y a pas de pivot dans la dernière colonne de $(A'|Y')$.

-Une base de $\mathrm{Im}(f)$ est donnée par la famille constituée des $f(e_i)$ pour lesquels il y a un pivot dans la i-ème colonne de A'.

Exemple 3.21. Si $f : \mathbb{R}^5 \to \mathbb{R}^6$ est définie par la matrice

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

alors la matrice échelonnée réduite correspondante est

$$(A'|Y') = \begin{pmatrix} 1 & 2 & 0 & 2 & 0 & y_1 \\ 0 & 0 & 1 & 1 & 0 & y_2 \\ 0 & 0 & 0 & 0 & 1 & y_3 - y_2 \\ 0 & 0 & 0 & 0 & 0 & y_4 - y_1 \\ 0 & 0 & 0 & 0 & 0 & y_5 - y_1 - y_2 \\ 0 & 0 & 0 & 0 & 0 & y_6 \end{pmatrix}.$$

Il y a 3 pivots dans A' donc le rang de f est 3 et $\mathrm{Im}(f)$ est donnée par les équations

$$\mathrm{Im}(f) = \left\{ \begin{pmatrix} y_1 \\ \vdots \\ y_6 \end{pmatrix} \in \mathbb{R}^6 \,\middle|\, \begin{array}{c} y_4 - y_1 = 0 \\ y_5 - y_1 - y_2 = 0 \\ y_6 = 0 \end{array} \right\}.$$

De plus, les pivots sont dans les colonnes 1, 3 et 5 donc les trois vecteurs

$$f(e_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \ f(e_3) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad f(e_5) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

forment une base de $\mathrm{Im}(f)$.

Pour le noyau, on prend $Y = 0$ et donc $Y' = 0$.
-Les équations de $\mathrm{Ker}(f)$ sont simplement données par le système $A'X = 0$.
-Si on note $k = \dim(\mathrm{Ker}(f))$, le système ci-dessus a exactement k inconnues libres (sans pivot dans leur colonne). Pour chacune de ces k inconnues, on considère le vecteur obtenue en prenant cette inconnue égale à 1 et toutes les autres inconnues libres égales à 0 (les inconnues contraintes étant déterminées par ces choix). On obtient k vecteurs qui forment une base de $\mathrm{Ker}(f)$.

Exemple 3.22. En reprenant l'exemple ci-dessus, on obtient grâce à A' que

$$\mathrm{Ker}(f) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix} \in \mathbb{R}^5 \;\middle|\; \begin{array}{c} x_1 = -2x_2 - 2x_4 \\ x_3 = -x_4 \\ x_5 = 0 \end{array} \right\}.$$

Les colonnes 2 et 4 n'ont pas de pivot donc une base de $\mathrm{Ker}(f)$ est donnée par les deux vecteurs de $\mathrm{Ker}(f)$ obtenus en prenant $x_2 = 1$ et $x_4 = 0$ puis $x_2 = 0$ et $x_4 = 1$, c'est-à-dire

$$\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Remarque 3.23. Dans le cas d'une application f entre e.v E et F de dimension finie, il est possible de faire la même chose dès que l'on a deux bases $\mathscr{A}$ et $\mathscr{B}$ de E et F respectivement. On regarde alors la matrice $M_{\mathscr{B},\mathscr{A}}(f)$ de f dans les bases $\mathscr{A}$ et $\mathscr{B}$ sans oublier alors que les différents vecteurs colonnes obtenus ne sont que des coordonnées dans $\mathscr{A}$ ou $\mathscr{B}$.

3.3 Matrices d'une application linéaire, matrices de passage

Pour une application entre $\mathbb{R}^n$ et $\mathbb{R}^m$, on a vu qu'il est possible de lui associer de manière "naturelle" une matrice. Ce n'est plus le cas pour une application entre deux e.v quelconques. Il faut faire des choix de bases et pour chaque choix, la matrice obtenue sera différente.

Définition 3.24. Soit E, F deux e.v de dimension finie et $\mathcal{A} = (u_1, \dots, u_n)$, $\mathcal{B} = (v_1, \dots, v_m)$ des bases de E et F respectivement.

Si f est entre application linéaire entre E et F alors on note $M_{\mathcal{B},\mathcal{A}}(f)$ la matrice dont la i-ème colonne est constituée des coordonnées du vecteur $f(u_i)$ dans la base $\mathcal{B}$. $M_{\mathcal{B},\mathcal{A}}$ s'appelle *la matrice de f dans les bases $\mathcal{A}$ et $\mathcal{B}$.*

Cette notion permet d'associer, dès qu'on a une base de E et une base de F, une matrice à n'importe quelle application linéaire entre E et F.

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$ tous deux équipés de leur base canonique, alors on retrouve les matrices de la Définition 3.10. Dans la pratique, prendre la base canonique n'est pas forcement un bon choix et trouver de "bonnes bases" pour une application donnée est au centre du programme de deuxième année, à travers la diagonalisation par exemple.

Un cas particulièrement important est le cas où $E = F$ et $f = \mathrm{Id}_E$.

Définition 3.25. Soit E un e.v de dimension finie et $\mathcal{A}$, $\mathcal{B}$ deux bases de E. *La matrice de passage de la base $\mathcal{A}$ à la base $\mathcal{B}$* est définie par

$$C_{\mathcal{A},\mathcal{B}} := M_{\mathcal{B},\mathcal{A}}(\mathrm{Id}_E).$$

Exemple 3.26. 1) Pour toute base $\mathcal{A}$, la matrice $C_{\mathcal{A},\mathcal{A}}$ est l'identitée.
2) Si $\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ et $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ sont deux bases de $\mathbb{R}^2$ on a

$$C_{\mathcal{A},\mathcal{B}} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}.$$

En effet, on peut vérifier que

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

Remarque 3.27. Il faut faire bien attention à l'ordre des bases dans la notation. La matrice $C_{\mathcal{A},\mathcal{B}}$ permet de passer de la base $\mathcal{A}$ à la base $\mathcal{B}$ et non l'inverse ! Cette convention est justifiée par la proposition suivante qui est d'une certaine manière une relation de Chasles.

Proposition 3.3.1. *Soit E, F et G trois e.v de dimension finie et $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ trois bases de E, F et G respectivement. Si $f: E \to F$ et $g: F \to G$ sont des application linéaires alors*

$$M_{\mathcal{C},\mathcal{B}}(g)M_{\mathcal{B},\mathcal{A}}(f) = M_{\mathcal{C},\mathcal{A}}(g \circ f).$$

En particulier, si $E = F$ alors

$$C_{\mathcal{A},\mathcal{B}} = (C_{\mathcal{B},\mathcal{A}})^{-1}$$

et

$$M_{\mathcal{A},\mathcal{A}}(f) = C_{\mathcal{A},\mathcal{B}}M_{\mathcal{B},\mathcal{B}}(f)C_{\mathcal{B},\mathcal{A}} = (C_{\mathcal{B},\mathcal{A}})^{-1}M_{\mathcal{B},\mathcal{B}}(f)C_{\mathcal{B},\mathcal{A}}.$$

Chapitre 4

Déterminant

Dans ce chapitre, nous allons introduire la notion de déterminant dans $\mathbb{R}^n$, qui à une famille $(u_1, \ldots, u_n)$ de n vecteurs de $\mathbb{R}^n$ associe un nombre réel. Il fournit deux informations sur cette famille. Sa valeur absolue s'interprète comme le "volume" du parallélépipède engendré par les vecteurs $u_1, \ldots, u_n$. Le terme "volume" doit être compris ici comme "longueur" si $n = 1$, "aire" si $n = 2$, "volume" si $n = 3$ et comme une notion similaire quand $n \geq 4$.

En particulier, cette interprétation géométrique implique que le déterminant est nul si et seulement si le parallélépipède en question est "plat", c'est-à-dire si et seulement si la famille $(u_1, \ldots, u_n)$ est liée. S'il est non nul alors $u_1, \ldots, u_n$ forment une base de $\mathbb{R}^n$ et le signe de leur déterminant indique si cette base est un repère direct ou indirect de $\mathbb{R}^n$.

La définition du déterminant en toute généralité est abstraite mais il intervient dans de très nombreuses situations pratiques.

4.1 Formes n-linéaires, formes alternées, déterminant

Définition 4.1. Soit E un e.v. Une *forme n-linéaire* sur E est une application

$$\begin{aligned} f : E^n = E \times \cdots \times E &\to \mathbb{R} \\ (u_1, \ldots, u_n) &\mapsto f(u_1, \ldots, u_n), \end{aligned}$$

telle que

$$\begin{aligned} &f(u_1, \ldots, u_{i-1}, (u_i + \lambda u_i'), u_{i+1}, \ldots, u_n) \\ &= f(u_1, \ldots, u_{i-1}, u_i, u_{i+1}, \ldots, u_n) + \lambda f(u_1, \ldots, u_{i-1}, u_i', u_{i+1}, \ldots, u_n), \end{aligned}$$

c'est-à-dire si f est linéaire en chacune des variables.

Si $n = 1$ on dira simplement que f est une *forme linéaire* et si $n = 2$ que f est une *forme bilinéaire.*

Définition 4.2. Soit E un e.v et f une forme n-linéaire. On dira que f est *symétrique* si elle est invariante par permutation des variables, c'est-à-dire

$$f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) = f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n),$$

et *alternée* (ou antisymétrique) si le fait de permuter deux variables change le signe :

$$f(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) = -f(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n).$$

Remarque 4.3. 1) Si f est une n-forme alternée, il est facile de voir que $f(u_1,\dots,u_n)=0$ dès que $u_i=u_j$ avec $i\neq j$.
2) Les formes bilinéaires symétriques et les formes linéaires vont jouer un rôle important par la suite. Dans ce chapitre, on se concentrera sur les formes n-linéaires alternées.

Exemple 4.4. 1) Le produit scalaire dans $\mathbb{R}^2$ défini par

$$\begin{aligned}\mathbb{R}^2\times\mathbb{R}^2 &\to \mathbb{R}\\ \left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}x_1'\\x_2'\end{pmatrix}\right) &\mapsto x_1x_1'+x_2x_2'\end{aligned}$$

est une forme bilinéaire symétrique sur $\mathbb{R}^2$.
2) Plus généralement, le produit scalaire de $\mathbb{R}^n$

$$\begin{aligned}\mathbb{R}^n\times\mathbb{R}^n &\to \mathbb{R}\\ \left(\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix},\begin{pmatrix}x_1'\\\vdots\\x_2'\end{pmatrix}\right) &\mapsto \sum_{i=1}^n x_ix_i'\end{aligned}$$

est une forme bilinéaire symétrique sur $\mathbb{R}^n$.
3) L'application

$$\begin{aligned}\mathbb{R}^2\times\mathbb{R}^2 &\to \mathbb{R}\\ \left(\begin{pmatrix}a\\b\end{pmatrix},\begin{pmatrix}c\\d\end{pmatrix}\right) &\mapsto ad-bc\end{aligned}$$

est une forme bilinéaire alternée.

La définition du déterminant repose sur le théorème suivant.

Théorème 4.5. *Soit E un e.v de dimension n.*
1) L'ensemble des formes n-linéaires alternées sur E forment un e.v de dimension 1.
2) De plus si $\mathscr{B}=(u_1,\dots,u_n)$ est une base de E et que f est une forme n-linéaire alternée non identiquement nulle alors

$$f(u_1,\dots,u_n)\neq 0.$$

Définition 4.6. Soit E un e.v de dimension n et $\mathscr{B}=(u_1,\dots,u_n)$. Le *déterminant* associé à la base $\mathscr{B}$ est l'unique forme n-linéaire alternée $\det_{\mathscr{B}}$ telle que

$$\det_{\mathscr{B}}(u_1,\dots,u_n)=1.$$

Dans le cas où $E=\mathbb{R}^n$ et $\mathscr{B}$ est la base canonique, on écrira simplement $\det$.

Remarque 4.7. Le point 1) du Théorème 4.5 implique l'unicité alors que les points 1) et 2) entraînent l'existence.

Exemple 4.8. 1) On va voir comment trouver une formule explicite du déterminant associé à la base canonique (e_1,e_2) dans $\mathbb{R}^2$.

Soit $u_1=\begin{pmatrix}a\\b\end{pmatrix}$ et $u_2=\begin{pmatrix}c\\d\end{pmatrix}$ deux vecteurs de $\mathbb{R}^2$. Ils peuvent s'écrire $u_1=ae_1+be_2$ et $u_2=ce_1+de_2$. On voit alors que

$$\begin{aligned}\det(u_1,u_2) &= \det(ae_1+be_2,u_2)\\ &= a\det(e_1,u_2)+b\det(e_2,u_2)\end{aligned}$$

en utilisant la linéarité en la première variable, puis que

$$\begin{aligned}a\det(e_1,u_2)+b\det(e_2,u_2)=&a(c\det(e_1,e_1)+d(\det(e_1,e_2))\\&+b(c\det(e_2,e_1)+d\det(e_2,e_2))\\=&ad-bc\end{aligned}$$

car $\det(e_1,e_2)=-\det(e_2,e_1)=1$ et $\det(e_1,e_1)=\det(e_2,e_2)=0$.
2) En effectuant le même genre de calcul dans $\mathbb{R}^3$, déjà plus laborieux, on obtiendrait la formule de Sarrus

$$\det\left(\begin{pmatrix}a\\b\\c\end{pmatrix},\begin{pmatrix}d\\e\\f\end{pmatrix},\begin{pmatrix}g\\h\\i\end{pmatrix}\right)=aei+dhc+gbf-gec-hfa-idb.$$

Des formules similaires peuvent être données en toute dimension mais elles sont très longues dès la dimension 4. Dans la prochaine section, on verra des méthodes pour calculer des déterminants en toute dimension.

Une des propriétés fondamentales du déterminant est donnée par le théorème suivant.

Théorème 4.9. *Soit E un e.v de dimension n et $\mathscr{B}$ une base de E. Une famille de n vecteurs $u_1,\dots,u_n$ est liée si et seulement si* $\det_{\mathscr{B}}(u_1,\dots,u_n)=0$.

Démonstration. Rappelons que dans un e.v de dimension n, une famille de n vecteurs est libre si et seulement si c'est une base. Donc si les vecteurs $u_1,\dots,u_n$ ne sont pas liés, ils forment une base et par le Théorème 4.5, $\det_{\mathscr{B}}(u_1,\dots,u_n)\neq 0$.

Inversement, si la famille est liée, il existe une combinaison linéaire non triviale entre les u_i, c-à-d

$$\sum_{i=1}^n \lambda_i u_i=0,$$

avec au moins un λ_i différent de 0. Sans perte de généralités, on peut supposer que $\lambda_1\neq 0$. On a alors

$$\begin{aligned}\det_{\mathscr{B}}(u_1,u_2,\dots,u_n)&=\det_{\mathscr{B}}\left(-\frac{1}{\lambda_1}\sum_{i=2}^n\lambda_iu_i,u_2,\dots,u_n\right)\\&=-\frac{1}{\lambda_1}\sum_{i=2}^n\lambda_i\det_{\mathscr{B}}(u_i,u_2,\dots,u_n)=0,\end{aligned}$$

où la dernière égalité vient du fait que $\det_{\mathscr{B}}(u_i,u_2,\dots,u_n)=0$ si $2\leq i\leq n$. □

En utilisant le déterminant dans $\mathbb{R}^n$ associé à la base canonique, il est aussi possible de définir le déterminant d'une matrice carrée.

Définition 4.10. Soit $A=(a_{ij})\in\mathscr{M}_n(\mathbb{R})$ une matrice carrée de taille n. Le déterminant de A est défini comme le déterminant des n vecteurs colonnes de A dans la base canonique de $\mathbb{R}^n$. On le note $\det(A)$ ou

$$\begin{vmatrix}a_{11}&\cdots&a_{1n}\\\vdots&\ddots&\vdots\\a_{n1}&\cdots&a_{nn}\end{vmatrix}$$

Exemple 4.11. 1) En utilisant la formule obtenue dans l'Exemple 4.8 1) on trouve que

$$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \times 4 - 2 \times 3 = -2.$$

2) La formule de Sarrus nous donne que

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 45 + 84 + 96 - 105 - 72 - 48 = 0$$

Le Théorème 4.9 se traduit de la manière suivante pour les matrices.

Théorème 4.12. *Un matrice $A \in \mathscr{M}_n(\mathbb{R})$ est inversible si et seulement si* $\det(A) \neq 0$.

Une propriété remarquable du déterminant est qu'il respecte le produit matriciel.

Proposition 4.1.1. *Le déterminant de matrices est multiplicatif, c'est-à-dire*

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

4.2 Calculs pratiques

Nous allons voir une méthode pour calculer les déterminants en toute dimension qui repose sur la formule de Laplace et les opérations élémentaires sur les lignes et les colonnes.

Le déterminant étant une forme n-linéaire alternée, on peut faire les opérations élémentaires suivantes sur les colonnes.

— Si on échange deux colonnes alors on change le signe du déterminant. Par exemple, dans le cas 2×2,
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} b & a \\ d & c \end{vmatrix}.$$

— Si $\alpha \in \mathbb{R}$ est un scalaire et qu'on change une colonne C_i par αC_i, alors on multiplie le déterminant par α.
$$\begin{vmatrix} a & \alpha b \\ c & \alpha d \end{vmatrix} = \alpha \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

— Si on remplace une colonne C_i par $C_i + \alpha C_j$ avec $i \neq j$, alors le déterminant reste inchangé.
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b + \alpha a \\ c & d + \alpha c \end{vmatrix}.$$

Le théorème suivant dit en particulier que ces trois règles restent valables si on remplace "colonne" par "ligne".

Théorème 4.13. *Soit $A \in \mathscr{M}_n(\mathbb{R})$. Le déterminant de A est égal au déterminant de ${}^tA$,*

$$\det(A) = \det({}^tA).$$

Avant d'introduire la formule de Laplace, on a besoin des notations suivantes.

Définition 4.14. Soit $A \in \mathscr{M}_n(\mathbb{R})$ et $1 \leq i, j \leq n$. On note $A_{ij} \in \mathscr{M}_{n-1}(\mathbb{R})$ la matrice obtenue à partir de A en supprimant la i-ième ligne et la j-ième colonne.

Exemple 4.15. Si
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$
alors
$$A_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} \quad , \quad A_{23} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} \quad \cdots$$

Théorème 4.16 (Formule de Laplace)**.** *Soit $A = (a_{ij}) \in M_n(\mathbb{R})$ et $1 \leq k \leq n$.*

— *Le déterminant peut se développer par rapport à la k-ième colonne :*
$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}).$$
— *Le déterminant peut se développer par rapport à la k-ième ligne :*
$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \det(A_{ki}).$$

Remarque 4.17. 1) Dans chacune de ces formules, il faut bien faire attention aux signes !
2) Dans la pratique, on appliquera la formule de Laplace après avoir fait des opérations élémentaires sur les lignes ou les colonnes afin de développer par rapport à une ligne ou une colonne simple (par exemple avec un seul coefficient non nul).

Exemple 4.18. 1) On considère la matrice
$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 3 \\ 4 & -1 & 0 \end{pmatrix}.$$
Si on développe par rapport à la deuxième ligne, on obtient
$$\begin{aligned} \det(A) &= (-1)^{2+1} a_{21} \det(A_{21}) + (-1)^{2+2} a_{22} \det(A_{22}) + (-1)^{2+3} a_{23} \det(A_{23}) \\ &= -1 \times \begin{vmatrix} 1 & 0 \\ -1 & 0 \end{vmatrix} + 2 \times \begin{vmatrix} 2 & 0 \\ 4 & 0 \end{vmatrix} - 3 \times \begin{vmatrix} 2 & 1 \\ 4 & -1 \end{vmatrix} \\ &= 0 + 0 + 18 = 18. \end{aligned}$$
Si on avait choisi de développer par rapport à la troisième colonne, alors on aurait obtenu
$$\det(A) = 0 \times \begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} - 3 \times \begin{vmatrix} 2 & 1 \\ 4 & -1 \end{vmatrix} + 0 \times \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 18,$$
ce qui est bien cohérent.
2) Dans les exemples de plus grandes tailles, il est souvent très préférable de faire des simplifications avant. Par exemple, si
$$B = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 2 & 4 \\ 2 & 1 & 2 & 5 \\ 1 & 1 & 1 & 6 \end{pmatrix}$$
alors
$$\det(B) = \begin{vmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 2 & 4 \\ 2 & 1 & 2 & 5 \\ 1 & 1 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & 2 \\ 0 & 3 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 4 \end{vmatrix} = 1 \times \begin{vmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 0 & 4 \end{vmatrix} = -1 \times \begin{vmatrix} 1 & 1 \\ 1 & 4 \end{vmatrix} = -3.$$

Une conséquence de la formule de Laplace est le théorème suivant.

Théorème 4.19. *Le déterminant d'une matrice* $A = (a_{ij})$ *triangulaire (supérieure ou inférieure) est le produit de ses éléments diagonaux.*

$$\det(A) = a_{11} \times a_{22} \times \cdots \times a_{nn}.$$

Chapitre 5

Valeurs propres, vecteurs propres

Un endomorphisme f d'un e.v E est une application linéaire $f\colon E \to E$. Dans ce dernier chapitre, nous allons introduire plusieurs objets associés à un endomorphisme d'un e.v qui va permettre de les étudier plus en profondeur. Dans tout ce chapitre, E sera un e.v de dimension finie et f un endomorphisme de E.

5.1 Valeurs propres, spectres et polynômes caractéristiques

Voici les définitions des objets de bases de ce chapitre.

Définition 5.1. *Une valeur propre* de f est un scalaire $\lambda \in \mathbb{R}$ tel qu'il existe un vecteur $u \in E$ non nul vérifiant
$$f(u) = \lambda u.$$
Un tel vecteur est dit *vecteur propre de f associé à λ*.

Définition 5.2. L'ensemble de toutes les valeurs propres de f s'appelle *le spectre de f* et se note $\mathrm{Spec}(f)$.

Définition 5.3. Si λ est une valeur propre de f, on appelle *espace propre associé à λ* l'ensemble des vecteurs propres associés à λ auquel on ajoute le vecteur nul. On le notera E_λ.

Remarque 5.4. En particulier, 0 est valeur propre si et seulement si $\mathrm{Ker}(f)$ n'est pas réduit à 0.

Remarque 5.5. On peut définir les même notions pour une matrice de $\mathscr{M}_n(\mathbb{R})$, en considérant l'endomorphisme de $\mathbb{R}^n$ qui lui est associé.

Proposition 5.1.1. *Soit $\lambda \in \mathrm{Spec}(f)$. L'espace propre E_λ vérifie*
$$E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id}).$$
En particulier, c'est un s.e.v.

Démonstration. Par définition, $u \in E_\lambda$ si et seulement si $f(u) = \lambda u$ ou $u = 0$. Mais puisque f est linéaire, $f(0) = 0 = \lambda \cdot 0$ donc $u \in E_\lambda$ si et seulement si $f(u) = \lambda u$ ou écrit autrement $(f - \lambda \mathrm{Id})(u) = 0$, c'est-à-dire si et seulement si $u \in \mathrm{Ker}(f - \lambda \mathrm{Id})$. $\square$

Proposition 5.1.2. *Si $\lambda, \mu \in \mathrm{Spec}(f)$ alors $\lambda \neq \mu$ si et seulement si $E_\lambda \cap E_\mu = \{0\}$.*

Démonstration. Supposons qu'il existe $u \in E_\lambda \cap E_\mu$ avec $u \neq 0$. On a alors que $f(u) = \lambda u$ mais aussi que $f(u) = \mu u$. D'où $\lambda u = \mu u$ et donc que $\lambda = \mu$ car $u \neq 0$.

La réciproque est évidente. □

D'après la Proposition 5.1.1, $\lambda \in \mathbb{R}$ est valeur propre si et seulement si $\mathrm{Ker}(f - \lambda \mathrm{Id}) \neq \{0\}$ ou dit autrement, puisqu'on est en dimension finie, si $f - \lambda \mathrm{Id}$ est non inversible. Cela amène, pour les matrices dans un premier temps, à la définition suivante.

Définition 5.6. Soit $A \in \mathscr{M}_n(\mathbb{R})$. *Le polynôme caractéristique de A* est le polynôme défini par

$$P_A(\lambda) = \det(A - \lambda I_n).$$

Ce polynôme vérifie plusieurs propriétés.

Proposition 5.1.3. *Soit $A \in \mathscr{M}_n(\mathbb{R})$. Comme son nom l'indique, le polynôme caractéristique de A est un polynôme, son degré est n.*

L'ensemble de ses racines coïncide avec le spectre de A. De plus, si on note

$$P_A(\lambda) = \sum_{i=1}^{n} a_i \lambda^i$$

alors

- $a_0 = \det(A)$,
- $a_n = (-1)^n$,
- $a_{n-1} = (-1)^{n-1}\mathrm{Tr}(A)$ *où* $\mathrm{Tr}(A)$ *désigne la trace de A, c'est à dire la somme des éléments diagonaux de A.*

Exemple 5.7. Considérons la matrice

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

On a alors

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

et

$$\begin{aligned} P_A(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 3 \\ 2 & 7-\lambda & 3 \\ 3 & 0 & 1-\lambda \end{vmatrix} &= (7-\lambda)\begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} \\ &= (7-\lambda)((1-\lambda)^2 - 9) = (7-\lambda)(4-\lambda)(-2-\lambda). \end{aligned}$$

Le spectre de A est donc $\mathrm{Spec}(A) = \{-2, 4, 7\}$.

Proposition 5.1.4. *Soit $A \in \mathrm{Gl}_n(\mathbb{R})$. Les polynômes caractéristiques de A et A^{-1} vérifient*

$$P_{A^{-1}}(\lambda) = (-1)^n \det(A^{-1})\lambda^n P_A(1/\lambda).$$

En particulier, si λ est valeur propre de A alors $\lambda \neq 0$ et λ^{-1} est valeur propre de A^{-1}.

Démonstration. Par définition,

$$\begin{aligned}
P_{A^{-1}}(\lambda) &= \det(A^{-1} - \lambda I_n) = \det(A^{-1} - \lambda(A^{-1}A)) \\
&= \det(\lambda A^{-1}(\lambda^{-1}I_n - A)) \\
&= \det(\lambda A^{-1}) \times \det(\lambda^{-1}I_n - 1) \\
&= \lambda^n \det(A^{-1})(-1)^n \det(A - \lambda^{-1}I_n) \\
&= (-1)^n \det(A^{-1})\lambda^n P_A(1/\lambda).
\end{aligned}$$

□

Proposition 5.1.5. *Soit $A \in \mathscr{M}_n(\mathbb{R})$ et $S \in \mathrm{Gl}_n(\mathbb{R})$. Les matrices A, ${}^tA$ et $S^{-1}AS$ ont même polynôme caractéristique.*

Démonstration. C'est une conséquence des égalités $\det(M) = \det({}^tM)$ et $\det(MN) = \det(M)\det(N)$. En effet, par définition

$$P_{{}^tA}(\lambda) = \det({}^tA - \lambda I_n) = \det({}^t(A - \lambda I_n)) = \det(A - \lambda I_n) = P_A(\lambda),$$

et

$$\begin{aligned}
P_{S^{-1}AS} &= \det(S^{-1}AS - \lambda I_n) = \det(S^{-1}AS - \lambda(S^{-1}I_nS)) \\
&= \det(S^{-1}(A - \lambda I_n)S) = \det(S^{-1})\det(A - \lambda I_n)\det(S) = P_A(\lambda).
\end{aligned}$$

□

On déduit de cette proposition que le polynôme caractéristique de la matrice $M_{\mathscr{B},\mathscr{B}}(f)$ de f dans la base $\mathscr{B}$ ne dépend pas de $\mathscr{B}$.

Définition 5.8. *Le polynôme caractéristique de f* est par définition le polynôme caractéristique de sa matrice dans une base quelconque.