

SYSTÈMES ET MATRICES

AVANT-PROPOS

La modélisation informatique fournit des prévisions en météorologie, remplace à peu de frais des expériences en soufflerie dans le domaine aéronautique, permet de simuler des explosions nucléaires ou thermonucléaires interdites, de construire des images en trois dimensions en imagerie médicale (tomographie). . .

Le principe consiste à discrétiser, c'est-à-dire à remplacer une partie finie de l'espace par un certain nombre de points et à approcher les expressions mathématiques par des combinaisons linéaires de valeurs en ces points, pour obtenir un grand nombre de systèmes avec plus ou moins d'inconnues (voir l'exemple 2.3.2). La précision du résultat et le temps de calcul dépendent de la finesse de la discrétisation et de la qualité de l'approximation. La modélisation nucléaire nécessite des milliards d'inconnues, mais le temps de calcul peut être de plusieurs semaines, ce qui n'est pas le cas de la météorologie et encore moins de la tomographie.

On comprend qu'il est nécessaire de trouver des méthodes de résolution des systèmes demandant le moins possible de calculs et adaptées à chaque type de problème et à la forme des systèmes obtenus.

La résolution des systèmes mène à la riche étude des matrices.

Table des matières

1	Prérequis	4
2	Survol historique	6
3	Généralités	7
3.1	Les premiers corps de caractéristique 2	7
3.2	Système linéaire	8
3.3	Matrice	10
3.4	Application linéaire, endomorphisme	14
3.5	Produit matriciel	20
3.6	Changement de base	24
3.7	Forme linéaire, dual	26
3.8	Application linéaire transposée	28
4	Déterminants	29
4.1	Permutations et transpositions	29
4.2	Formes n -linéaires alternées	30
4.3	Déterminant de n vecteurs	32
4.4	Déterminant d'une matrice	34
4.5	Développement d'un déterminant	35
4.6	Déterminant d'un endomorphisme	39
4.7	Déterminant de Vandermonde	39
4.8	Vecteur et valeur propres, polynôme caractéristique	40
4.9	Trigonalisation	41
5	Méthodes de résolution des systèmes	42
5.1	Système de Cramer	42
5.2	Système de Cramer triangulaire	43
5.3	Méthode de Gauss et de Gauss-Jordan	44
5.4	Décomposition triangulaire : méthode LU	51
5.5	Méthode de Cholesky	53
5.6	Diagonale dominance stricte	55
5.7	Méthode itérative de Jacobi	56
5.8	Méthode des moindres carrés	58
6	Espace euclidien, hermitien	59
6.1	Produit scalaire et norme	59
6.2	Base orthonormée	61
6.3	Orthogonalité	65
6.4	Isométrie	66
6.5	Angle de deux vecteurs	67
6.6	Isométries de $\mathbb{R}^2$	68
6.7	Le produit vectoriel	69
6.8	Complexifié, espace hermitien	70

7	Réduction des matrices	81
7.1	Polynôme caractéristique	81
7.2	Diagonalisation	84
7.3	Algorithme de Leverrier	85
7.4	Méthode des puissances	87
7.5	Endomorphismes et matrices nilpotents	90
7.6	Forme de Jordan	93
7.7	Diagonalisation simultanée de deux endomorphismes	99
7.8	Endomorphismes symétriques et antisymétriques	101
7.9	Endomorphismes hermitiens	105
7.10	Isométries de $\mathbb{R}^3$	106
7.11	Exemple de diagonalisation dans $\mathbb{F}_2$	110
8	Réduction des systèmes par des isométries	112
8.1	Isométries de $\mathbb{R}^n$	112
8.2	Rotations de Givens (1954)	114
8.3	Réflexions de Householder (1964)	116
9	Formes quadratiques	120
9.1	Formes quadratiques réelles	121
9.2	Formes quadratiques hermitiennes	126
9.3	En caractéristique 2	128
10	Systèmes différentiels linéaires	128
10.1	Réduction de M	128
10.2	Exponentielle de M	129
11	Interpolation de Lagrange	130
11.1	Polynômes de Lagrange	131
11.2	Matrices de Lagrange	131
12	Interpolation de Tchebychev	132
12.1	Polynômes et support de Tchebychev	132
12.2	Matrices de Tchebychev	133
13	Lissage par fonction spline cubique	134
13.1	Définition	134
13.2	Calcul	134
13.3	Résolution	135

1 Prérequis

Sont supposées connues :

- les notions élémentaires et les notations de la théorie des ensembles,
- les relations d'ordre, d'équivalence, les lois de composition,
- le raisonnement par l'absurde ($[A \implies \text{non}(A)] \implies [\text{non}(A)]$),
- le raisonnement par récurrence :
 $[P(0) \text{ vraie et } P(n) \implies P(n+1)] \implies [P(n) \text{ vraie pour tout } n \in \mathbb{N}]$.

Rappelons les définitions des structures fondamentales de l'Algèbre.

Un **groupe** est un ensemble G muni d'une loi de composition « $\star$ » associative ($g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3$, quels que soient les éléments écrits, ce qui permet d'oublier les parenthèses), possédant un **élément neutre** « e » ($x \star e = e \star x = x$ pour tout x) (donc G est non vide) et telle que tout élément x admette un **inverse** x^{-1} ($x \star x^{-1} = x^{-1} \star x = e$).

Il y a deux notations principales : l'additive, l'opération étant notée « $+$ », le neutre « 0 », et l'inverse, on dit dans ce cas l'opposé, « $-x$ », et la multiplicative, notée sans symbole (xy), de neutre « 1 ». Les deux peuvent être utilisées indifféremment, mais l'une des deux s'impose en général. Le groupe est commutatif si sa loi l'est. Un sous-groupe est une partie du groupe qui est un groupe pour la même loi, avec le même neutre. Il contient au moins l'élément neutre du groupe. Un groupe est **cyclique** s'il est engendré par l'un de ses éléments, comme $(\mathbb{Z}, +)$ est engendré par 1.

Un **anneau** est un groupe additif muni d'une seconde loi (multiplicative), possédant un élément neutre « 1 », associative et distributive par rapport à l'addition ($a(b+c) = ab+ac$). Si cette loi est commutative, l'anneau est commutatif. Il est **intègre** si un produit ne peut être nul que si l'un des facteurs est nul. L'ensemble des entiers $\mathbb{Z}$, l'ensemble $\mathbb{R}[X]$ des polynômes à une indéterminée et à coefficients réels sont des anneaux commutatifs. L'ensemble des fonctions continues réelles sur $\mathbb{R}$ est un anneau commutatif non intègre.

Un **corps** est un anneau intègre dont tous les éléments non nuls sont inversibles; $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps commutatifs, et $\mathbb{C}$ est **algébriquement clos**, en ce sens que tout polynôme de $\mathbb{C}[X]$ de degré n se factorise en un produit de n polynôme de degré 1 (les seuls irréductibles, avec les constants). Tout corps peut être plongé dans un corps algébriquement clos, sa **clôture algébrique**. Le **corps des fractions** d'un anneau intègre A est l'ensemble $A \times (A \setminus \{0\})$ quotienté par la relation d'équivalence $(a,b) \equiv (c,d) \iff ad = bc$, muni de l'addition $(a,b) + (c,d) = (ad+bc, bd)$ et de la multiplication $(a,b)(c,d) = (ac, bd)$. L'inverse de (a,b) si $a \neq 0$ est (b,a) . Ainsi $\mathbb{Q}$ est le corps des fractions de $\mathbb{Z}$. Un corps $\mathbb{K}$ est de **caractéristique** p (nombre premier), si la somme de p termes égaux à x est nulle pour tout $x \in \mathbb{K}$; $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ est un corps de caractéristique p .

Un **espace vectoriel** sur un corps $\mathbb{K}$, ou $\mathbb{K}$ -espace vectoriel est un groupe commutatif $(E, +)$ sur lequel opère $\mathbb{K}$:

$$\begin{aligned}\mathbb{K} \times E &\rightarrow E \\ (\lambda, x) &\mapsto \lambda x\end{aligned}$$

avec les règles suivantes :

$$\begin{aligned}\lambda(\mu x) &= (\lambda\mu) x \\ (\lambda + \mu) x &= \lambda x + \mu x \\ \lambda(x + y) &= \lambda x + \lambda y \\ 1 x &= x\end{aligned}$$

Une **combinaison linéaire** d'éléments de E (de **vecteurs**) à coefficients dans $\mathbb{K}$ est une somme finie $\lambda_1 x_1 + \dots + \lambda_r x_r$, dans laquelle les λ_i sont des éléments de $\mathbb{K}$ (des **scalaires**) et les x_i des éléments de E (des vecteurs).

Un **sous-espace vectoriel** d'un espace vectoriel V est une partie non vide de V stable par combinaisons linéaires. C'est donc un espace vectoriel.

L'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel, mais pas la réunion (sauf si l'un contient l'autre).

L'ensemble des combinaisons linéaires d'éléments $x_1, \dots, x_k$ de E à coefficients dans $\mathbb{K}$ est un sous-espace vectoriel de E noté $\text{Vect}(x_1, \dots, x_k)$ ce qui se lit **vectoriel** $x_1, \dots, x_k$.

Une famille $(x_i)_{1 \leq i \leq r}$ de vecteurs est **libre** si la relation $\sum_i \lambda_i x_i = 0$ implique que tous les λ_i sont nuls (la relation est dite **triviale**). Elle est **génératrice** si tout vecteur est une combinaison linéaire de ses membres. C'est une **base**, notée $\mathfrak{B}$, si elle est libre et génératrice. Si $x = \sum_i \lambda_i x_i$, les λ_i sont les **composantes**, ou les coordonnées, de x dans cette base; elles sont uniques.

De toute famille génératrice on peut extraire une base, toute famille libre peut être complétée en une base. Toutes les bases ont le même nombre d'éléments; ce nombre est la **dimension** de l'espace. Pour qu'une famille de n vecteurs soit une base d'un espace vectoriel de dimension n il suffit qu'elle soit libre ou génératrice.

Si $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ est une base de l'espace vectoriel E et si $x = \sum_{1 \leq i \leq n} x_i \epsilon_i$ est un vecteur de E , on notera :

$$x = x_{\mathfrak{B}} = (\lambda_i) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = {}^t(\lambda_1, \dots, \lambda_n),$$

la **transposition** $\ll {}^t \gg$ transformant une ligne en une colonne, et réciproquement. La notation en ligne tient moins de place que celle en colonne.

L'espace vectoriel $\mathbb{K}^n = \{{}^t(a_1, \dots, a_n) \mid a_i \in \mathbb{K}\}$, est le modèle d'espace vectoriel de dimension n . Sa **base canonique**, canonique car provenant de sa définition même, est la famille $(e_i)_{1 \leq i \leq n}$, e_i étant le vecteur dont les composantes sont nulles sauf la i -ème, égale à 1.

Si E_1 et E_2 sont deux sous-espaces vectoriels d'un espace vectoriel E , de dimensions respectives n_1 et n_2 , la **somme** est l'ensemble des éléments de E somme d'un élément de E_1 et d'un élément de E_2 . C'est un sous-espace vectoriel de E noté $E_1 + E_2$. On a :

$$\dim(E_1 + E_2) = \dim(E_1) + \dim(E_2) - \dim(E_1 \cap E_2).$$

Si $E_1 \cap E_2 = \{0\}$, la somme est **directe**, notée $E_1 \oplus E_2$.

Une **algèbre** sur un corps $\mathbb{K}$, ou $\mathbb{K}$ -algèbre, est un $\mathbb{K}$ -espace vectoriel muni d'une loi de composition interne qui en fait un anneau. Ainsi l'ensemble $\mathbb{K}[X]$ des polynômes à une indéterminée et à coefficients dans $\mathbb{K}$, muni de l'addition et du produit par les scalaires, est un $\mathbb{K}$ -espace vectoriel, et, muni de l'addition et du produit des polynômes, c'est un anneau, et donc une $\mathbb{K}$ -algèbre.

Deux suites numériques (u_n) et (v_n) sont équivalentes, $u_n \sim v_n$, si u_n/v_n tend vers 1 quand n tend vers l'infini. Ainsi, $n^2 + 4n \sim n^2$. Elles sont du même ordre si leur rapport tend vers un réel non nul, comme $(n^3 + 2n)$ et $2n^3 + n^2$, ces deux suites étant d'ordre 3 : $O(u_n) = O(v_n) = 3$. La suite (u_n) est négligeable devant (v_n) , c'est un $o(v_n)$, si u_n/v_n tend vers 0. Ainsi, n est un $o(n^2)$ (petit o de n^2).

2 Survol historique

La résolution de systèmes d'équations linéaires a intéressé les mathématiciens depuis l'Antiquité. Mais, sans la symbolique mathématique, c'est un problème très difficile, beaucoup plus qu'une équation du deuxième degré dont on savait, il y a quatre ou cinq millénaires, trouver les solutions positives.

Il y a plus de deux mille ans les mathématiciens Chinois savaient résoudre des systèmes linéaires, et la **méthode de Gauss** était connue du mathématicien Chinois Liu Hui au troisième siècle de notre ère.

Le mathématicien Arabe Abu Kamil (850-930) poursuit les travaux d'Al Khawarizmi (790-850) sur les équations quadratiques et rédige un ouvrage sur les systèmes linéaires.

La symbolique mathématique, créée par le mathématicien français François Viète (1540-1603), se construit progressivement jusqu'à nos jours. Elle réduit les problèmes à leur difficulté intrinsèque et permet un progrès ininterrompu.

Leibniz (1646-1716) et MacLaurin (1698-1746) résolvent les systèmes linéaires en dimension 2 et 3, puis le suisse Gabriel Cramer (1704-1752) donne ses fameuses formules.

Les déterminants ont été introduits par Bézout (1730-1783), qui en a compris l'importance, mais on doit à Augustin Cauchy (1789-1857) leur théorie et leur notation. Le terme de **matrice** est dû (1848) à James Sylvester (1814-1897), et Cayley (1821-1895) a créé le calcul matriciel.

3 Généralités

Dans tout le livre, $\mathbb{K}$ désignera un corps commutatif, E , F et G des $\mathbb{K}$ -espaces vectoriels de dimensions, respectivement, n , p et q , et δ_{ij} le **symbole de Kronecker**, égal à 1 si $j = i$, à 0 sinon.

Beaucoup de résultats en caractéristique nulle restent vrais en caractéristique non nulle p lorsque ce nombre (premier) n'intervient pas dans leur théorie.

3.1 Les premiers corps de caractéristique 2

Le plus petit corps de caractéristique 2 est le corps à deux éléments, les éléments neutres de l'addition, 0, et de la multiplication, 1. On a évidemment $1 + 1 = 0$, car l'autre possibilité, $1 + 1 = 1$, impliquerait que $1 = 0$. On le note $\mathbb{F}_2$, et il est égal à l'ensemble des classe d'entiers modulo 2 : $\mathbb{Z}/2\mathbb{Z}$.

Un élément est **algébrique** sur $\mathbb{K}$ s'il annule un polynôme de $\mathbb{K}[X]$ irréductible dans $\mathbb{K}$. Ainsi, le complexe i , annulant le polynôme irréductible $X^2 + 1$, est algébrique sur $\mathbb{R}$; $\mathbb{C} = \mathbb{R}[i]$ est une **extension algébrique** de $\mathbb{R}$. De même, le réel $\sqrt{2}$, racine de $X^2 + 2 \in \mathbb{Q}[X]$, est algébrique sur $\mathbb{Q}$, et le corps $\mathbb{Q}[\sqrt{2}]$ est une extension algébrique de $\mathbb{Q}$.

Les polynômes de degré 2 dans $\mathbb{F}_2[X]$ sont :

$$X^2, X^2 + 1 = (X + 1)^2, X^2 + X = X(X + 1)$$

et $X^2 + X + 1$, qui est irréductible. Soit i un élément algébrique sur $\mathbb{F}_2$ tel que $i^2 + i + 1 = 0$, ou $i^2 = i + 1$. On alors $i^3 + 1 = (i + 1)(i^2 + i + 1) = 0$, et donc $i^3 = 1$. L'inverse de i est donc i^2 , que l'on peut noter $\bar{i}$, et qui est la seconde racine de $X^2 + X + 1$ puisque le produit des racines est égal à 1, comme leur somme. Nous avons ainsi construit le corps à quatre éléments $\mathbb{F}_4$, qui est une **extension algébrique** de $\mathbb{F}_2$, comme $\mathbb{C}$ vis à vis de $\mathbb{R}$. Mais ce corps n'est pas algébriquement clos (un tel corps ne peut être fini) car, par exemple, $X^3 + X + 1$ est irréductible. Les éléments de $\mathbb{F}_4$ s'écrivent $a + bi$, $a, b \in \mathbb{F}_2$, ce qui montre que $\mathbb{F}_4$ est un $\mathbb{F}_2$ -espace vectoriel de dimension 2.

Les polynômes $X^3 + X + 1$ et $X^3 + X^2 + 1$ sont les seuls polynômes irréductibles de $\mathbb{F}_2[X]$, de degré 3, n'ayant pas de racines dans $\mathbb{F}_4$. Soit j un élément algébrique sur $\mathbb{F}_2$, racine du polynôme $X^3 + X + 1$, donc tel que $j^3 + j + 1 = 0$, ou $j^3 = j + 1$. On a $j^4 = j^2 + j$, $j^5 = j^3 + j^2 = j^2 + j + 1$, $j^6 = j^2 + 1$ et $j^7 = 1$. Les sept premières puissances de j forment le groupe cyclique multiplicatif G , et en ajoutant 0 on obtient le corps $\mathbb{F}_8 = \mathbb{F}_2[j]$, qui est un $\mathbb{F}_2$ -espace vectoriel de base $(1, j, j^2)$. Pour calculer dans $\mathbb{F}_8$, on joue sur les deux structures. On vérifie aisément que :

$$X^3 + X + 1 = (X + j)(X + j^2)(X + j^4)$$

et que :

$$X^3 + X^2 + 1 = (X + j^3)(X + j^5)(X + j^6).$$

On peut par ce procédé construire des corps finis de caractéristique 2 à 2^m éléments, m étant un entier positif quelconque, utilisés dans la théorie du codage.

3.2 Système linéaire

Un système linéaire (p, n) , p et n étant des entiers positifs, est la donnée de p équations à n inconnues dans lesquelles les inconnues, éléments de $\mathbb{K}$, sont à la puissance 1. Nous nommerons L_1 la première équation (ou ligne), L_2 la deuxième, ainsi de suite. Si par exemple nous ajoutons deux fois la k -ième ligne à la i -ième, nous noterons cette transformation $L_i + 2L_k$ au lieu de $L_i \leftarrow L_i + 2L_k$. De même, nous noterons $3L_i$ pour $L_i \leftarrow 3L_i$.

Exemple 3.2.1. Le système $(3, 2)$ dans $\mathbb{R}$:

$$\begin{cases} 3x - y &= -1 \\ x + y &= 3 \\ 2x + 3y &= \lambda \end{cases}$$

admet la solution $x = 1, y = 2$ si $\lambda = 8$. Il est impossible si $\lambda \neq 8$. ▽

Exemple 3.2.2. Le système réel $(2, 3)$:

$$\begin{cases} x + y + z &= 6 \\ x + 2y + 3z &= 14 \end{cases}$$

est vérifié pour $x = -2 + z$ et $y = 8 - 2z$. Il admet donc une infinité de solutions, selon la valeur donnée à z . ▽

Exemple 3.2.3. Le système réel $(2, 2)$:

$$\begin{cases} 2x - y &= 3 \\ 3x + 2y &= 15 \end{cases}$$

admet l'unique solution $x = y = 3$. ▽

Exemple 3.2.4. Considérons enfin le système réel $(3, 3)$:

$$\begin{cases} x + y + z &= 3 \\ x + 2y + 2z &= 5 \\ 3x + 4y + 4z &= \lambda \end{cases}$$

et calculons $2L_1 + L_2$:

$$3x + 4y + 4z = 11$$

pour conclure que si $\lambda \neq 11$ le système est impossible et si $\lambda = 11$ la troisième ligne est une combinaison linéaire des deux premières :

$$2L_1 + L_2 - L_3 = 0$$

et n'apporte aucune information. Le système se réduit à L_1 et L_2 , et $x = 1, y = 2 - z$. L_1 et L_2 sont les **équations principales**, x et y les **inconnues principales** et z est une **inconnue auxiliaire**, ou un paramètre. ▽

Ces exemples illustrent les différents cas possibles.

Voyons des exemples dans d'autres corps.

Exemple 3.2.5 Les vecteurs $v_1 = {}^t(0, 1, 1)$, $v_2 = {}^t(1, 0, 1)$, $v_3 = {}^t(1, 1, 0)$ de $\mathbb{F}_2^3$ forment-ils une base ?

Annulons la combinaison linéaire $a_1 v_1 + a_2 v_2 + a_3 v_3$:

$$a_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_2 + a_3 \\ a_1 + a_3 \\ a_1 + a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

qui admet la solution $a_1 = a_2 = a_3 = 1$ (outre la solution triviale nulle). Ces vecteurs sont donc liés, par la relation $v_1 + v_2 + v_3 = 0$, d'où $v_3 = v_1 + v_2$. Comme v_1 et v_2 sont libres, ces trois vecteurs engendrent un sous-espace vectoriel de dimension 2. ∇

Exemple 3.2.6 Même question avec $v_1 = {}^t(1, 1, 1)$, $v_2 = {}^t(0, 1, 1)$, $v_3 = {}^t(0, 0, 1)$.

La première ligne de la combinaison linéaire $a_1 v_1 + a_2 v_2 + a_3 v_3$ donne $a_1 = 0$, la deuxième, $a_1 + a_2 = 0$, donne $a_2 = 0$, et la troisième, $a_1 + a_2 + a_3 = 0$, donne $a_3 = 0$. La famille est donc libre, et c'est une base puisqu'elle a autant d'éléments que la base canonique. Un vecteur quelconque s'écrit dans cette base :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (x_2 - x_1) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + (x_3 - x_2) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

∇

Exemple 3.2.7 Calculer la dimension du sous-espace vectoriel de $\mathbb{F}_2^7$ engendré par :

$$v_1 = {}^t(1, 0, 1, 0, 1, 0, 1),$$

$$v_2 = {}^t(0, 1, 1, 0, 0, 1, 1),$$

$$v_3 = {}^t(0, 0, 0, 1, 1, 1, 1),$$

$$v_4 = {}^t(1, 1, 0, 1, 0, 0, 1).$$

Les quatre premières composantes de $a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$ sont $a_1 + a_4$, $a_2 + a_4$, $a_1 + a_2$ et $a_3 + a_4$. En les annulant, on trouve que les a_i sont égaux. Ceci donne, avec $a_i = 1$, $v_1 + v_2 + v_3 + v_4 = 0$, d'où $v_4 = v_1 + v_2 + v_3$, ce qui élimine v_4 . La famille (v_1, v_2, v_3) est-elle libre ? En annulant les deux premières composantes de $a_1 v_1 + a_2 v_2 + a_3 v_3$ on trouve $a_1 = a_2 = 0$. En annulant la quatrième, on trouve $a_3 = 0$. La famille est donc libre et engendre un sous-espace vectoriel de dimension 3. ∇

3.3 Matrice

Le tableau des coefficients des inconnues dans l'exemple 3.2.4 :

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

est la matrice du système. Les lignes de la matrice multiplient, élément par élément, le vecteur $v = {}^t(x, y, z)$ des variables :

$$Mv = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 3 & 4 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ x + 2y + 2z \\ 3x + 4y + 4z \end{pmatrix}.$$

Plus généralement, une **matrice** A de type (p, n) est un tableau d'éléments de $\mathbb{K}$, ou d'un anneau commutatif (de préférence intègre, comme $\mathbb{K}[X]$, ce qui permet de calculer dans son corps des fractions), à p lignes et n colonnes. L'élément à l'intersection de la i -ème ligne ($L_i(A)$) et de la j -ème colonne ($C_j(A)$) est noté a_{ij} , et la matrice est notée (a_{ij}) . Soit $B = (b_{ij})$ une autre matrice de même type et un élément quelconque λ de $\mathbb{K}$ (un **scalaire**). On définit la somme :

$$A + B = (c_{ij}), \quad c_{ij} = a_{ij} + b_{ij}$$

d'élément neutre $0_{p,n}$, ou simplement 0, dont toutes les composantes sont nulles, le produit :

$$\lambda A = (\lambda a_{ij}),$$

et les matrices E_{hk} , dont l'unique terme non nul, égal à 1, est obtenu pour $(i, j) = (h, k)$.

Une matrice à une seule ligne est une matrice-ligne, une matrice à une seule colonne (par exemple un vecteur) est une matrice-colonne.

L'ensemble des matrices de type (p, n) est noté $\mathfrak{M}_{p,n}(\mathbb{K})$.

Proposition 3.3.1. *L'ensemble $\mathfrak{M}_{p,n}(\mathbb{K})$ muni de l'addition et de la multiplication par les scalaires est un $\mathbb{K}$ -espace vectoriel de dimension pn .*

Démonstration. Pour toute matrice $A \in \mathfrak{M}_{p,n}(\mathbb{K})$ on a :

$$A = (a_{ij}) = \sum_{i,j} a_{ij} E_{ij}.$$

La famille $(E_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$ est donc génératrice. Elle est libre car :

$$\sum_{ij} a_{ij} E_{ij} = 0 \iff \forall (i, j), a_{ij} = 0.$$

Elle constitue sa base canonique et elle possède pn éléments. □

La **transposée** de la matrice $A = (a_{ij}) \in \mathfrak{M}_{p,n}(\mathbb{K})$ est la matrice ${}^tA = (b_{ij})$ avec $b_{ij} = a_{ji}$. Ses lignes sont les colonnes de A et elle est de type (n, p) .

Si $p = n$, les matrices sont dites **carrées**, et leur ensemble est noté $\mathfrak{M}_n(\mathbb{K})$.

Si les seuls termes éventuellement non nuls d'une matrice carrée D sont ses termes **diagonaux** d_{ii} , qui forment sa **diagonale principale**, elle est dite **diagonale**, et notée $\text{Diag}(d_{ii})$. Si de plus tous ses termes diagonaux valent 1, elle est notée I_n (pour **identité** en dimension n), et s'ils sont tous égaux à un scalaire λ , on a une **matrice scalaire**, λI_n .

Exemple 3.3.1. Une **suite récurrente** $u_n = \sum_{1 \leq i \leq k} a_i u_{n-i}$, les k premiers termes étant donnés, permet de définir une matrice carrée. Considérons par exemple la suite d'éléments de $\mathbb{F}_4$ (voir 2.1) définie par $x_4 = x_1 + x_3$ et de premiers termes 1, i , i^2 . La matrice faisant passer de (x_i, x_{i+1}, x_{i+2}) à $(x_{i+1}, x_{i+2}, x_{i+3})$ s'écrit :

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

On a $M^3 = M^2 + I_3$, $M^4 = M^3 + M = M^2 + M + I_3$, $M^5 = M + I_3$, $M_6 = M^2 + M$ et $M^7 = I_3$, ce qui montre que la suite est périodique :

$$1, i, i^2, i, 0, i^2, 1, 1, i, i^2, \dots$$

de période de longueur 7 : $(1, i, i^2, i, 0, i^2, 1)$. ▽

Exemple 3.3.2. Considérons l'équation différentielle $y' = y$ sur l'intervalle $[0, 1]$, avec la condition initiale $y(0) = 1$. Partageons l'intervalle en 10 parties égales (10 pas), et posons $y_i = y'_i = y(\frac{i}{10})$ et $h = \frac{1}{10}$, le **pas**. La formule de dérivation approchée d'ordre 3 (l'erreur est proportionnelle à h^3) :

$$y'_i = \frac{1}{2h} (y_{i-2} - 4y_{i-1} + 3y_i)$$

donne la suite récurrente :

$$2,8 y_i = 4 y_{i-1} - y_{i-2}.$$

La matrice :

$$M = \frac{1}{2,8} \begin{pmatrix} 0 & 1 \\ -1 & 4 \end{pmatrix}$$

fait passer de (y_i, y_{i+1}) à (y_{i+1}, y_{i+2}) , et y_{i+2} est la seconde composante de $M(y_i, y_{i+1})$, ou de $M^{i+1}(y_0, y_1)$.

Pour démarrer, il nous faut, calculée par une méthode quelconque, par exemple un développement limité puisque $y^{(k)}(0) = 1$, la valeur $y_1 = 1,1051709$. On trouve alors $y_2 = 1,2214\dots$ On a :

$$M^9 = \begin{pmatrix} -1,018303 & 3,152219 \\ -1,125792 & 3,484867 \end{pmatrix}$$

et la seconde composante de $M^9(y_0, y_1)$ donne $y_{10} = 2,7252$, d'où une erreur $\sim 7.10^{-3}$. On améliore la précision en augmentant le nombre de pas (avec 100 pas, on trouve une erreur $\sim 9.10^{-5}$. La convergence est lente) et l'ordre de la formule de dérivation approchée.

Cette méthode d'intégration est dite **multipas**. Elle peut ne pas être stable ni convergente : il y a des critères qui sont du domaine de l'Analyse Numérique. ∇

Deux matrices A et B de $\mathfrak{M}_n(\mathbb{K})$ **commutent** si $AB = BA$.

Si les termes t_{ij} d'une matrice carrée T sont nuls pour $i < j$, elle est dite **triangulaire inférieure**. S'ils sont nuls pour $i > j$, elle est **triangulaire supérieure**.

Exemple 3.3.3. Nous voulons résoudre l'équation $\sin x = 0,6$. Soit x_i une solution approchée. Posons $x_{i+1} = x_i + h_i$, $\sin(x_i) = s_i$, $\cos(x_i) = c_i$, d'où à l'ordre 3 :

$$\sin x = s_i + c_i h - \frac{1}{2} s_i h^2 - \frac{1}{6} c_i h^3 + O(h^4).$$

Négligeons le terme en h^4 , divisons par c_i et posons $d_i = \frac{0,6-s_i}{c_i}$. Il vient :

$$h - \frac{s_i}{2c_i} h^2 - \frac{1}{6} h^3 = d_i.$$

En élevant cette équation au carré et au cube, toujours en négligeant le terme en h^4 , nous obtenons $h^2 + \frac{s_i}{c_i} h^3 = d_i^2$ et $h^3 = d_i^3$, d'où le système triangulaire supérieur :

$$\begin{pmatrix} 1 & -\frac{s_i}{2c_i} & -\frac{1}{6} \\ 0 & 1 & -\frac{s_i}{c_i} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h \\ h^2 \\ h^3 \end{pmatrix} = \begin{pmatrix} d_i \\ d_i^2 \\ d_i^3 \end{pmatrix}.$$

Nous obtenons en remontant :

$$\begin{cases} h_i^3 &= d_i^3 \\ h_i^2 &= d_i^2 + \frac{s_i}{c_i} d_i^3 \\ h_i &= d_i - \frac{s_i}{2c_i} (+d_i^2 + \frac{s_i}{c_i} d_i^3) + \frac{d_i^3}{6}. \end{cases}$$

Comme $\sin(0,6) = 0,56$ et $\sin(0,7) = 0,64$, partons de solution approchée $x_0 = 0,65$. Nous obtenons $h_0 = -0,0065310\dots$ et $x_1 = x_0 + h_0 = 0,6434689\dots$. Evaluons l'erreur : $\arcsin(0,6) - x_1 = 3,2.10^{-5}$.

En reprenant la méthode à partir de x_1 , on trouve $h_1 = 0,00003215710\dots$ $x_2 = 0,643501108017\dots$ et $\arcsin(0,6) - x_2$ est voisin de 8.10^{-10} .

On obtient la précision que l'on veut en itérant la méthode, ou en calculant à l'ordre n pour un n assez grand, d'où un système $n \times n$. ∇

Une matrice carrée $M = (m_{ij})$ est **symétrique** si elle est égale à sa transposée : $m_{ij} = m_{ji}$; elle est **antisymétrique** si elle est l'opposée de sa transposée : $m_{ij} = -m_{ji}$.

Proposition 3.3.2. *L'ensemble $S_n(\mathbb{K})$ des matrices symétriques de $\mathfrak{M}_n(\mathbb{K})$ est un $\mathbb{K}$ -espace vectoriel de dimension $n(n+1)/2$. L'ensemble $A_n(\mathbb{K})$ des matrices antisymétriques de $\mathfrak{M}_n(\mathbb{K})$ est un $\mathbb{K}$ -espace vectoriel de dimension $n(n-1)/2$ si $\mathbb{K}$ n'est pas de caractéristique 2. Si $\mathbb{K}$ est de caractéristique 2, les notions de symétrie et d'antisymétrie sont confondues.*

Démonstration. Une matrice symétrique $S = (s_{ij})$ est déterminée par les scalaires s_{ij} pour $j \geq i$, au nombre de $n(n+1)/2$, et une matrice antisymétrique $A = (a_{ij})$, par les $n(n-1)/2$ scalaires a_{ij} pour $j > i$ (les a_{ii} sont nuls). Si $\mathbb{K}$ est de caractéristique 2, $m_{ij} = m_{ji}$ équivaut à $m_{ij} = -m_{ji}$. $\square$

Si $A \in \mathfrak{M}_n(\mathbb{K})$ commute avec sa transposée, elle est **normale**.

Une matrice $A \in \mathfrak{M}_n(\mathbb{R})$ est **orthogonale** si $A^t A = {}^t A A = I_n$. Les matrices symétriques, les matrices antisymétriques et les matrices orthogonales sont normales.

Si $A = (a_{ij}) \in \mathfrak{M}_n(\mathbb{C})$, on note $\bar{A}$ la matrice **conjuguée** $(\bar{a}_{ij})$. On définit alors la **matrice adjointe** A^* de A par :

$$A^* = {}^t \bar{A} = \overline{{}^t A} = (\bar{a}_{ji}).$$

On a évidemment $A^{**} = A$.

La matrice A est **hermitienne** si $A = A^*$. Elle est **normale** si elle commute avec son adjointe : $AA^* = A^*A$. Elle est **unitaire** si $AA^* = A^*A = I_n$. Les matrices hermitiennes sont normales, de même que les matrices unitaires.

Proposition 3.3.3. *L'ensemble $H_n(\mathbb{C})$ des matrices hermitiennes est un $\mathbb{R}$ -espace vectoriel de dimension n^2 .*

Démonstration. Soit $H = (h_{ij})$ une matrice hermitienne : $h_{ij} = \bar{h}_{ji}$. A quelle condition sur $\lambda \in \mathbb{C}$ la matrice λH est-elle hermitienne ? Comme $h_{ii} = \bar{h}_{ii}$ les h_{ii} sont réels, ce qui impose que λ soit réel. La somme de deux matrices hermitiennes étant hermitienne, $H_n(\mathbb{C})$ est un $\mathbb{R}$ -espace vectoriel.

Pour $j \neq i$ posons $h_{ij} = a_{ij} + ib_{ij}$, a_{ij} et b_{ij} réels, $a_{ji} = a_{ij}$, $b_{ji} = -b_{ij}$ ($b_{ii} = 0$). La matrice réelle :

$$S = \frac{1}{2} (H + {}^t H) = (a_{ij})$$

est symétrique. La matrice réelle :

$$A = \frac{1}{2i} (H - {}^t H) = (b_{ij})$$

est antisymétrique. La matrice $H = S + iA$ est donc définie par les $n(n+1)/2$ composantes de S et les $n(n-1)/2$ composantes de A . La dimension de $H_n(\mathbb{C})$ est donc égale à n^2 . $\square$

La somme des termes diagonaux d'une matrice carrée $M = (m_{ij})$ est sa **trace**, somme des termes diagonaux : $\text{tr}(M) = \sum_i m_{ii}$.

3.4 Application linéaire, endomorphisme

Dans cette partie sur les applications linéaires, λ et μ désigneront des scalaires quelconques.

Une **application linéaire**, ou **homomorphisme**, u de E dans F est une application qui conserve les combinaisons linéaires :

$$u(\lambda x + \mu y) = \lambda u(x) + \mu u(y),$$

ou, ce qui est équivalent (par récurrence), $u(\sum_i \lambda_i x_i) = \sum_i \lambda_i u(x_i)$, toujours pour des éléments de E et des scalaires quelconques.

Une application linéaire de E dans lui-même est un **endomorphisme**. L'endomorphisme i_E qui laisse invariants les éléments de E ($\forall x \in E, i_E(x) = x$), ou, ce qui est équivalent, les éléments d'une base de E , est l'**identité** de E .

Proposition 3.4.1. *Une application de E dans F est linéaire si et seulement si elle est linéaire sur les éléments d'une base quelconque de E .*

Démonstration. Soient (ϵ_i) une base de E et une application u de E dans F telle que $u(\sum_i \lambda_i \epsilon_i) = \sum_i \lambda_i u(\epsilon_i)$, (x_j) des éléments quelconques de E et (λ_j) des scalaires quelconques. Si $x_j = \sum_i x_{ij} \epsilon_i$, on a $u(x_j) = \sum_i x_{ij} u(\epsilon_i)$ et :

$$u(\sum_j \lambda_j x_j) = u(\sum_{i,j} \lambda_j x_{ij} \epsilon_i) = \sum_{i,j} \lambda_j x_{ij} u(\epsilon_i) = \sum_j \lambda_j u(x_j),$$

d'où la linéarité de u . La réciproque est triviale. $\square$

L'**image** de l'application linéaire u , $\text{Im}(u)$, est l'ensemble des vecteurs du but qui ont au moins un antécédent, et son **noyau**, $\text{Ker}(u)$, est l'ensemble des vecteurs de la source d'image nulle :

$$\begin{aligned} \text{Im}(u) &= \{y \in F \mid \exists x \in E : u(x) = y\}, \\ \text{Ker}(u) &= \{x \in E \mid u(x) = 0\}. \end{aligned}$$

Proposition 3.4.2. *Une application linéaire u de E dans F est déterminée par les images des vecteurs d'une base quelconque $\mathfrak{B}_E = (\epsilon_i)_{1 \leq i \leq n}$ de E dans une base quelconque $\mathfrak{B}_F = (\phi_j)_{1 \leq j \leq p}$ de F . Son noyau et son image sont des sous-espaces vectoriels de, respectivement, E et F .*

Démonstration. Si $\mathfrak{B}_E = (\epsilon_j)_{1 \leq j \leq n}$ est une base de E et $\mathfrak{B}_F = (\phi_i)_{1 \leq i \leq p}$ une base de F , un élément $x = \sum_j x_j \epsilon_j$ de E a pour image $u(x) = \sum_j x_j u(\epsilon_j)$. Si $u(\epsilon_j) = \sum_i u_{ij} \phi_i$, on a :

$$u(x) = \sum_j x_j \left(\sum_i u_{ij} \phi_i \right) = \sum_{i,j} x_j u_{ij} \phi_i$$

et les (u_{ij}) déterminent u .

Si $u(x) = u(y) = 0$, on aura $u(\lambda x + \mu y) = 0$; $\text{Ker}(u)$ est donc un sous-espace vectoriel de E .

Si y_1 et y_2 appartiennent à $\text{Im}(u)$, ces éléments ont des antécédents respectifs x_1 et x_2 ; on a alors :

$$u(\lambda x_1 + \mu x_2) = \lambda u(x_1) + \mu u(x_2) = \lambda y_1 + \mu y_2$$

et $\lambda y_1 + \mu y_2$ appartient à $\text{Im}(u)$, qui est donc un sous-espace vectoriel de F . $\square$

La dimension de l'image est le **rang** de l'application linéaire, noté $r_g(u)$.

Une application linéaire u est **injective** (c'est une injection) si $\text{Ker}(u) = \{0\}$. Elle est **surjective** (c'est une injection) si $\text{Im}(u) = F$, c'est-à-dire si $r_g(u) = p$. Si elle est injective et surjective, donc bijective (c'est une bijection), c'est un **isomorphisme**, ou un **automorphisme** si $F = E$; E et F sont alors **isomorphes**.

Un automorphisme u de E tel que $u \circ u = i_E$ est une **involution**. La conjugaison dans $\mathbb{C}$, espace vectoriel sur lui-même ou sur $\mathbb{R}$, une symétrie dans E , sont des involutions.

Proposition 3.4.3. *Deux $\mathbb{K}$ -espaces vectoriels sont isomorphes si et seulement si ils ont même dimension ($n = p$).*

Démonstration. Supposons que $n = p$, et soient $(\epsilon_i)_{1 \leq i \leq n}$ une base de E et $(\phi_i)_{1 \leq i \leq n}$ une base de F . L'application u de E dans F définie par :

$$u\left(\sum_i \lambda_i \epsilon_i\right) = \sum_i \lambda_i \phi_i$$

est linéaire (proposition 2.4.1), injective car si $u(\sum_i \lambda_i \epsilon_i) = 0$ on a $\sum_i \lambda_i \phi_i = 0$ et les λ_i sont tous nuls, les ϕ_i étant libres; elle est surjective car tout élément $y = \sum_i y_i \phi_i$ de F est l'image de $x = \sum_i y_i \epsilon_i$. $\square$

On a $r_g(u) \leq n$ car si (y_i) est une base de $\text{Im}(u)$ et x_i un antécédent de y_i pour $1 \leq i \leq r_g(u)$, les x_i sont libres et engendrent un sous-espace vectoriel E' de E , isomorphe à $\text{Im}(u)$, de dimension $r_g(u)$ au plus égale à $\dim(E) = n$.

Si elle est injective, on a forcément $p \geq n$ car les images par u des vecteurs d'une base quelconque de E sont libres et $\dim(\text{Im}(u)) = n$, or $\text{Im}(u)$ est un sous-espace de F , donc de dimension au plus égale à celle de F .

Si elle est surjective, comme $\dim(\text{Im}(u)) \leq \dim(E)$, on a $p = r_g(u) \leq n$.

Proposition 3.4.4. *L'application réciproque d'un isomorphisme est un isomorphisme.*

Démonstration. Soit $u : E \rightarrow F$ un isomorphisme; étant bijectif, u admet une application réciproque u^{-1} . Il reste à voir que u^{-1} est linéaire. Soient x_1 et

x_2 des éléments quelconques de E , $u(x_1) = y_1$, $u(x_2) = y_2$, $y_1 = u^{-1}(x_1)$ et $y_2 = u^{-1}(x_2)$. On a alors :

$$\begin{aligned} u^{-1}(\lambda y_1 + \mu y_2) = x & \iff u(x) = \lambda y_1 + \mu y_2 \\ & = \lambda u(x_1) + \mu u(x_2) \\ & = u(\lambda x_1 + \mu x_2) \end{aligned}$$

d'où $x = \lambda x_1 + \mu x_2$ et :

$$\begin{aligned} u^{-1}(\lambda y_1 + \mu y_2) & = \lambda x_1 + \mu x_2 \\ & = \lambda u^{-1}(y_1) + \mu u^{-1}(y_2). \end{aligned}$$

Linéaire et bijective, l'application u^{-1} est un isomorphisme. $\square$

Lorsque u est un isomorphisme, $M(u)$, dans des bases quelconques, est une matrice **régulière**, et $M(u^{-1})$ est son inverse pour une opération que nous définirons ultérieurement. Le système $u(x) = y_0$ est alors équivalent au système $x = u^{-1}(y_0)$ et admet une et une seule solution.

Proposition 3.4.5. *Toute base de E définit un isomorphisme entre E et $\mathbb{K}^n$, et réciproquement.*

Démonstration. Soit $\mathfrak{B} = (\epsilon_i)$ une base de E . L'application linéaire ϵ de $\mathbb{K}^n$ dans E définie par $\epsilon(e_i) = \epsilon_i$ est un isomorphisme. Réciproquement, si $\phi \in \mathcal{L}(\mathbb{K}^n, E)$ est un isomorphisme, la famille $(\phi(e_i))$ est libre et génératrice. $\square$

Un isomorphisme est **canonique** s'il est construit sans utilisation de bases. Nous en verrons un exemple un peu plus loin.

Proposition 3.4.6. *Soit u une application linéaire entre deux espaces vectoriels E et F de même dimension n . Si elle est injective ou surjective, c'est un isomorphisme. Si $F = E$, c'est un automorphisme.*

Démonstration. Supposons E , de base (ϵ_i) et F , de base (ϕ_i) de même dimension ($p = n$) et $u \in \mathcal{L}(E, F)$. Si u est injective, les $(u(\epsilon_i))$ sont libres et engendrent un sous-espace vectoriel de F de même dimension que F , donc égal à F , et u est surjective.

Si u est surjective, chaque ϕ_i a un antécédent ψ_i . Une relation $\sum_i \lambda_i \psi_i = 0$ donne, en appliquant u , $\sum_i \lambda_i \phi_i = 0$. Les (ϕ_i) étant libres les λ_i sont nuls et (ψ_i) , famille libre à n éléments, est une base de E . Définissons l'application linéaire v par $v(\phi_i) = \psi_i$ pour tout i . L'application $v \circ u$ est linéaire (proposition 2.3.7) et $(v \circ u)(\psi_i) = v(\phi_i) = \psi_i$ pour tout i ; c'est donc l'identité de E , $v = u^{-1}$, et u est un isomorphisme. $\square$

Proposition 3.4.7. *La composée de deux applications linéaires $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$ est une application linéaire. Si u et v sont injectives (respectivement surjectives, bijectives) $v \circ u$ est aussi injective (respectivement surjective, bijective). Si $v \circ u$ est injective, u l'est aussi. Si $v \circ u$ est surjective, v l'est aussi. Si u et v sont des isomorphismes, $v \circ u$ en est un. La réciproque est vraie si E , F et G ont même dimension.*

Démonstration. L'application $v \circ u : x \mapsto v(u(x))$ va de E dans G . Elle est linéaire car :

$$(v \circ u)(\lambda x + \mu y) = v(\lambda u(x) + \mu u(y)) = \lambda (v \circ u)(x) + \mu (v \circ u)(y).$$

Si u et v sont injectives, $v \circ u$ l'est aussi car $(v \circ u)(x) = 0$ implique $u(x) = 0$ et donc $x = 0$. Si $v \circ u$ est injective u est injective car $u(x) = 0$ implique $(v \circ u)(x) = 0$ donc $x = 0$, mais v peut ne pas l'être si $\text{Im}(u) \cap \text{Ker}(v) = \{0\}$. Si u et v sont surjectives, $v \circ u$ l'est aussi car tout $z \in G$ a un antécédant $y \in F$ qui a lui-même un antécédant $x \in E$. Si $v \circ u$ est surjective, v est surjective, mais u peut ne pas l'être car il n'est pas nécessaire que $\text{Im}(u)$ contienne $\text{Ker}(v)$.

D'après ce qui précède, si u et v sont des isomorphismes, $v \circ u$ en est un, et si $v \circ u$ est un isomorphisme, v est surjective, c'est donc un isomorphisme puisque F et G ont même dimension (proposition 3.4.6), et u est injective, et c'est un isomorphisme puisque $\dim(E) = \dim(F)$. $\square$

L'ensemble $\text{Aut}(E)$ des automorphismes de E muni de la composition est un groupe (non commutatif) dont l'élément neutre est i_E .

Proposition 3.4.8. Soient $u \in \text{Aut}(E)$ et $v \in \mathfrak{L}(E)$, tels que $v \circ u = i_E$. Alors v est l'automorphisme inverse de u ($v = u^{-1}$), et $u \circ v = v \circ u = i_E$.

Démonstration. Si (ϵ_i) est une base de E , $(\epsilon'_i = u(\epsilon_i))$ est une autre base. On a $v(\epsilon'_i) = \epsilon_i$, $(v \circ u)(\epsilon_i) = v(\epsilon'_i) = \epsilon_i$. $\square$

L'ensemble $\mathfrak{L}(E, F)$ des applications linéaires de E dans F est naturellement muni de l'addition :

$$u_1 + u_2 : x \mapsto u_1(x) + u_2(x)$$

et de la multiplication par les scalaires :

$$\lambda u : x \mapsto \lambda u(x).$$

Définissons les applications linéaires u_{hk} :

$$u_{hk}(\epsilon_h) = \phi_k, \quad u_{hk}(\epsilon_i) = 0 \text{ si } i \neq h.$$

Proposition 3.4.9. L'ensemble $\mathfrak{L}(E, F)$ muni de l'addition et de la multiplication par les scalaires est un $\mathbb{K}$ -espace vectoriel de dimension pn .

Démonstration. On vérifie immédiatement que $\mathfrak{L}(E, F)$ est stable pour l'addition et la multiplication par les scalaires, et que les pn u_{hk} en constituent une base. $\square$

L'application linéaire nulle ($\forall x \in E, u(x) = 0$) est notée 0 .

Si $F = E$, on note $\mathfrak{L}(E)$ pour $\mathfrak{L}(E, E)$.

Si $x = a$ et $x = b$ sont des solutions de l'équation $u(x) = y_0$, alors $u(b - a) = 0$ et $b - a$ appartient au noyau de u . Si $c \in \text{Ker}(u)$, $a + c$ est solution. La solution

générale est notée $a + \text{Ker}(u)$.

Le système de l'exemple 3.2.1 peut s'écrire $u(v) = v_0$, le problème étant de trouver x dans $\mathbb{R}^2$. En effet, l'application :

$$\begin{aligned} u : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} 3x - y \\ x + y \\ 2x + 3y \end{pmatrix} \end{aligned}$$

est linéaire. Son image peut s'écrire :

$$\begin{pmatrix} 3x - y \\ x + y \\ 2x + 3y \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

le produit dans le membre de droite s'effectuant « ligne par colonne » comme on le voit clairement. Le tableau de nombres :

$$M(u) = \begin{pmatrix} 3 & -1 \\ 1 & 1 \\ 2 & 3 \end{pmatrix}$$

est la **matrice associée** à u dans les bases de la source et du but (ici, les bases canoniques). Sa première colonne est l'image du premier vecteur de la base source :

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

et la seconde celle du second. Ces images étant indépendantes forment une base de $\text{Im}(u)$, qui est donc de dimension 2. Les conditions $y = 3x$ et $y = -x$ impliquent que $\text{Ker}(u) = \{(0, 0)\}$.

Plus généralement, la matrice $M_{\mathfrak{B}', \mathfrak{B}}(u)$ associée à une application linéaire $u \in \mathfrak{L}(E, F)$ dans les bases $\mathfrak{B} = (\epsilon_i)$ de E et $\mathfrak{B}' = (\epsilon'_j)$ de F est la matrice dont la i -ème colonne C_i est l'image dans $\mathfrak{B}'$ du i -ème vecteur de $\mathfrak{B}$. Nous notons $M_{\mathfrak{B}', \mathfrak{B}}(u)$ et non $M_{\mathfrak{B}, \mathfrak{B}'}(u)$ pour respecter la contravariance des vecteurs, une convention¹ dont nous parlerons plus loin. Remarquons que, dans l'écriture $\mathfrak{M}_{p, n}(\mathbb{K})$, le premier indice, p , le nombre de lignes, est la dimension du but, et n , le nombre de colonnes, celle de la source. Il y a là aussi inversion.

Proposition 3.4.10. *L'application :*

$$\begin{aligned} M_{\mathfrak{B}', \mathfrak{B}} : \quad \mathfrak{L}(E, F) &\rightarrow \mathfrak{M}_{p, n}(\mathbb{K}) \\ u &\mapsto M_{\mathfrak{B}', \mathfrak{B}}(u) \end{aligned}$$

est un isomorphisme de $\mathbb{K}$ -espaces vectoriels.

1. Les conventions d'Einstein simplifient les notations et permettent le calcul tensoriel.

Démonstration. Notons M pour $M_{\mathfrak{B}', \mathfrak{B}}$. L'application M est linéaire car il découle de la définition de $M(u)$ que $M(u_1 + u_2) = M(u_1) + M(u_2)$ et que $M(\lambda u) = \lambda M(u)$. Elle est évidemment injective (si $M(u) = 0$, alors $u = 0$) et c'est donc un isomorphisme, $\mathfrak{L}(E, F)$ et $\mathfrak{M}_{p, n}(\mathbb{K})$ ayant même dimension pn (proposition 2.4.6). Cet isomorphisme dépend du choix des bases et n'est donc pas canonique.

Une matrice quelconque $M \in \mathfrak{M}_{p, n}(\mathbb{K})$ est associée à l'application linéaire $u \in \mathfrak{L}(E, F)$ définie par : $u(\epsilon_i) = \sum_j m_{ji} \epsilon'_j$ (l'image dans la base de F du i -ème vecteur de la base de E est la i -ème colonne de M). $\square$

Dans l'exemple 3.2.4 :

$$M(u) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 3 & 4 & 4 \end{pmatrix}$$

l'image de u est engendrée par les deux premiers vecteurs-colonnes, qui sont libres ; elle est donc de dimension 2. Un vecteur (x, y, z) appartient au noyau si et seulement si $x = -y - z$ et $x = -2y - 2z$, ces deux conditions impliquant la troisième. Ceci se traduit par $z = -y$ et $x = 0$. Le noyau est donc engendré par le vecteur $(0, 1, -1)$; il est de dimension 1.

Conclusion : le système $u(v) = w_0$ n'est possible que si $w_0 \in \text{Im}(u)$ et les solutions, lorsque cette condition est remplie, sont obtenues en ajoutant à l'une quelconque d'entre elles, v_0 , un élément du noyau. On peut noter l'ensemble des solutions $v_0 + \text{Ker}(u)$.

Revenons au cas général. Soient :

- r le rang de l'application linéaire $u \in \mathfrak{L}(E, F)$,
- $\mathfrak{B}_1 = (\epsilon_i)_{1 \leq i \leq r}$ une base d'un supplémentaire K' du noyau,
- $\mathfrak{B}_2 = (\epsilon_j)_{r+1 \leq j \leq n}$ une base du noyau,

de sorte que $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ est une base de E . La restriction de u à K' est injective car $u(x) = 0$ implique que $x \in \text{Ker}(u)$ et donc que $x \notin K'$. Les vecteurs $u(\epsilon_i)$ pour i allant de 1 à r sont libres dans F et constituent une base de $\text{Im}(u)$; complétons-les en une base $\mathfrak{B}'$ de F . Dans ces bases ($\mathfrak{B}$ et $\mathfrak{B}'$), la matrice de u s'écrit :

$$M(u) = \begin{pmatrix} I_r & 0_{r, n-r} \\ 0_{p-r, r} & 0_{p-r, n-r} \end{pmatrix}$$

ce qui est sa **forme canonique**.

Deux matrices $M_1, M_2 \in \mathfrak{M}_{p, n}(\mathbb{K})$ sont **équivalentes** si elles ont le même rang, donc la même forme canonique.

Théorème 3.4.1 (du rang). *Pour toute application linéaire $u \in \mathfrak{L}(E, F)$ on a :*

$$\dim(\text{Ker}(u)) + r_g(u) = \dim(E).$$

Démonstration. La restriction de u à un supplémentaire K' de $\text{Ker}(u)$ est un isomorphisme de K' sur $\text{Im}(u)$, et :

$$\dim(E) = \dim(\text{Ker}(u)) + \dim(K') = \dim(\text{Ker}(u)) + \dim(\text{Im}(u))$$

d'où le résultat. $\square$

Proposition 3.4.11. *L'ensemble $\mathfrak{L}(E)$ muni de la composition est une $\mathbb{K}$ -algèbre (non commutative).*

Démonstration. La composition est associative et distributive par rapport à l'addition car :

$$\begin{aligned} (w \circ (v \circ u))(x) &= w(v(u(x))) = ((w \circ v) \circ u)(x), \\ (w \circ (u + v))(x) &= w(u(x)) + w(v(x)) = (w \circ u + w \circ v)(x) \end{aligned}$$

quels que soient $x \in E$ et les endomorphismes u, v et w . L'élément unité est l'identité i_E . $\square$

En posant $u \circ u = u^2$ et $u \circ u^n = u^{n+1}$ pour tout entier naturel $n > 2$, on fait opérer la $\mathbb{K}$ -algèbre $\mathbb{K}[X]$ des polynômes à une indéterminée sur $\mathfrak{L}(E)$:

$$\begin{aligned} \mathbb{K}[X] \times \mathfrak{L}(E) &\rightarrow \mathfrak{L}(E) \\ (P = \sum_i a_i X^i, u) &\mapsto P(u) = \sum_i a_i u^i \quad (0 \leq i \leq \deg(P)). \end{aligned}$$

Un endomorphisme $p \in \mathfrak{L}(E)$ est un **projecteur** si $p^2 = p$; p est la projection sur $\text{Im}(p)$ (invariante par p) de direction $\text{Ker}(p)$. Remarquons que $q = i_E - p$ est aussi un projecteur, car :

$$q^2 = (i_E - p)^2 = i_E - 2p + p^2 = i_E - p = q,$$

que $pq = qp = 0$, et que q est la projection sur $\text{Ker}(p)$ de direction $\text{Im}(p)$.

La restriction de p à $\text{Im}(p)$ est l'identité de $\text{Im}(p)$, car, si $y \in \text{Im}(p)$, il existe $x \in E$ tel que $p(x) = y$ et alors $p(y) = p^2(x) = p(x) = y$.

On a :

$$E = \text{Im}(p) \oplus \text{Ker}(p) = \text{Im}(p) \oplus \text{Im}(q) = \text{Ker}(p) \oplus \text{Ker}(q).$$

3.5 Produit matriciel

Définissons d'abord le **produit formel** :

$$(a_1, \dots, a_n) * {}^t(b_1, \dots, b_n) = a_1 b_1 + \dots + a_n b_n.$$

Soient (ϵ_i) , (ϕ_i) et (γ_i) des bases, respectivement, des $\mathbb{K}$ -espaces vectoriels E , F et G . On a une application naturelle de $\mathfrak{L}(E, F) \times \mathfrak{L}(F, G)$ dans $\mathfrak{L}(E, G)$ définie par la composition $((u, v) \mapsto v \circ u)$.

Nous voulons définir une multiplication matricielle compatible avec cette application, c'est-à-dire telle que, si A , B et C sont les matrices associées respectivement à u , v et $v \circ u$ dans les bases ci-dessus, on ait $C = BA$.

Remarquons d'abord que le nombre de colonnes de B doit être égal au nombre de lignes de A (c'est la dimension p de F). Cette condition supposée remplie, identifions les deux écritures de $v \circ u(\epsilon_j)$:

$$\begin{aligned} v(u(\epsilon_j)) &= v(\sum_h a_{hj} \phi_h) = \sum_h a_{hj} v(\phi_h) = \sum_h a_{hj} \sum_i b_{ih} \gamma_i \\ v(u(\epsilon_j)) &= \sum_i c_{ij} \gamma_i \end{aligned}$$

pour trouver que c_{ij} est égal au produit $L_i(B) * C_j(A)$:

$$c_{ij} = \sum_h b_{ih} a_{hj}.$$

On dit qu'on fait le produit « ligne par colonne » .

Lorsque le produit MM' de deux matrices de $\mathfrak{M}_n(\mathbb{K})$ est égal à I_n on dit que M' est la **matrice inverse** de M (et réciproquement), ce qu'on note $M' = M^{-1}$.

Le calcul :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

montre que, si $ad - bc \neq 0$, on a :

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

Proposition 3.5.1. *La matrice $M(u)$ d'un automorphisme u de E dans une base $\mathfrak{B}$ est inversible et $M(u)^{-1} = M(u^{-1})$.*

Démonstration. Comme :

$$M(u^{-1} \circ u) = M(i_E) = I_n$$

et :

$$M(u^{-1} \circ u) = M(u^{-1}) M(u),$$

on a bien $M(u^{-1}) = M(u)^{-1}$. □

Le calcul du produit de deux matrices de $\mathfrak{M}_n(\mathbb{K})$ demande $n^2(2n - 1)$ opérations ($\sim 2n^3$).

Si $F = E$, $\mathfrak{L}(E)$, munie de la composition, et $\mathfrak{M}_n(\mathbb{K})$, munie du produit matriciel, sont des $\mathbb{K}$ -algèbres.

Proposition 3.5.2. *Si $M \in \mathfrak{M}_n(\mathbb{K})$ est inversible, son inverse est unique.*

Démonstration. Si M' et M'' sont inverses de M , on a :

$$MM' = MM'' = I_n,$$

d'où, en multipliant à gauche par M' : $M' = M''$. □

Proposition 3.5.3. *Si $MM' = I_n$, alors $M'M = I_n$.*

Démonstration. Si u est l'automorphisme de $\mathbb{K}^n$ défini par la matrice M , M' est la matrice de u^{-1} , MM' celle de $u \circ u^{-1}$ et $M'M$ celle de $u^{-1} \circ u$; or $u \circ u^{-1} = u^{-1} \circ u$ (proposition 2.4.8). $\square$

Proposition 3.5.4. *L'application M est un isomorphisme entre les $\mathbb{K}$ -algèbres $\mathfrak{L}(E)$ et $\mathfrak{M}_n(\mathbb{K})$.*

Démonstration. On sait déjà que M est un isomorphisme d'espaces vectoriels (proposition 2.4.10). Comme $M(v \circ u) = M(v)M(u)$, c'est aussi un isomorphisme d'anneaux, donc d'algèbres. $\square$

Les polynômes de $\mathbb{K}[X]$ opèrent donc sur $\mathfrak{M}_n(\mathbb{K})$, comme sur $\mathfrak{L}(E)$, avec la convention, pour toute matrice $M \in \mathfrak{M}_n(\mathbb{K})$: $M^0 = I_n$.

Proposition 3.5.5. *La matrice transposée du produit de matrices $A_1 \dots A_r$ ($r \geq 2$) est égale au produit ${}^tA_r \dots {}^tA_1$*

Démonstration. Montrons-le pour $r = 2$. Si $A_2A_1 = (c_{ij})$ avec :

$$c_{ij} = L_i(A_2) * C_j(A_1),$$

on a :

$${}^t(A_2A_1) = (c'_{ij}), \text{ avec } c'_{ij} = c_{ji} = L_j(A_2) * C_i(A_1)$$

et :

$$\begin{aligned} {}^tA_1 {}^tA_2 &= (c''_{ij}), \text{ avec } c''_{ij} = L_i({}^tA_1) * C_j({}^tA_2) \\ &= C_i(A_1) * L_j(A_2) \end{aligned}$$

d'où le résultat : ${}^t(A_2A_1) = {}^tA_1 {}^tA_2$. On passe au cas général par une récurrence immédiate. $\square$

Proposition 3.5.6. *Le produit $A_1 \dots A_r$ ($r \geq 2$) de r matrices inversibles de $\mathfrak{M}_n(\mathbb{K})$ est inversible. Son inverse est égal à $A_r^{-1} \dots A_1^{-1}$.*

Démonstration. Montrons-le lorsque $r = 2$: $A_2^{-1}A_1^{-1}A_1A_2 = A_2^{-1}A_2 = I_n$. On passe au cas général par une récurrence immédiate. $\square$

Proposition 3.5.7. *La matrice inverse $D^{-1} = (d'_{ij})$ d'une matrice diagonale régulière $D = (d_{ij})$ est diagonale (régulière) et $d'_{ii} = d_{ii}^{-1}$ pour tout i .*

Démonstration. La matrice D est associée à un endomorphisme u de $\mathbb{K}^n$ tel que $u(e_i) = d_{ii}e_i$, $d_{ii} \neq 0$ pour tout i . On a donc :

$$u^{-1}(e_i) = \sum_j d'_{ij}e_j = d_{ii}^{-1}e_i,$$

d'où $d'_{ii} = d_{ii}^{-1}$ et $d'_{ij} = 0$ si $j \neq i$. $\square$

Proposition 3.5.8. *La matrice inverse $T' = (t'_{ij}) \in \mathfrak{M}_n(\mathbb{K})$ d'une matrice triangulaire supérieure (respectivement inférieure) régulière $T = (t_{ij})$ est triangulaire supérieure (respectivement inférieure). On a $t'_{ii} = t_{ii}^{-1}$ pour tout i .*

Démonstration. Démontrons dans le cas où T est triangulaire supérieure, l'autre cas se prouvant de façon identique.

Le produit TT' étant égal à I_n , on a :

$$L_n(T) * C_n(T') = t_{nn} t'_{nn} = 1$$

d'où $t'_{nn} = t_{nn}^{-1}$, et, pour tout $j < n$:

$$L_n(T) * C_j(T') = t_{nn} t'_{nj} = 0$$

d'où $t'_{nj} = 0$. Faisons l'hypothèse de récurrence H_h :

$$\forall i \geq h : t'_{ii} = t_{ii}^{-1} \text{ et } t'_{ij} = 0 \text{ si } j < i,$$

et remarquons que nous venons de prouver que H_n est vérifiée; T' est donc supposée avoir la forme :

$$T' = \begin{pmatrix} t'_{11} & \dots & \dots & \dots & \dots & t'_{1n} \\ \vdots & \ddots & & & & \vdots \\ t'_{h-1\ 1} & \dots & t'_{h-1\ h-1} & \dots & \dots & t'_{h-1\ n} \\ 0 & \dots & 0 & t'_{hh} & \dots & t'_{hn} \\ \vdots & & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & t'_{nn} \end{pmatrix}$$

Nous avons alors :

$$\begin{aligned} L_{h-1}(T) &= \begin{pmatrix} 0 & \dots & 0 & t_{h-1\ h-1} & \dots & \dots & t_{h-1\ n} \end{pmatrix}, \\ C_{h-1}(T') &= {}^t \begin{pmatrix} t'_{1\ h-1} & \dots & \dots & t'_{h-1\ h-1} & 0 & \dots & 0 \end{pmatrix}, \\ C_j(T') &= {}^t \begin{pmatrix} t'_{ij} & \dots & \dots & t'_{h-1\ j} & 0 & \dots & 0 \end{pmatrix}, \end{aligned}$$

pour $j < h-1$, d'où :

$$\begin{aligned} 1 &= L_{h-1}(T) * C_{h-1}(T') = t_{h-1\ h-1} t'_{h-1\ h-1}, \\ 0 &= L_{h-1}(T) * C_j(T') = t_{h-1\ h-1} t'_{h-1\ j}, \end{aligned}$$

et la conclusion : $t'_{h-1\ h-1} = t_{h-1\ h-1}^{-1}$ et $t'_{h-1\ j} = 0$ pour $j < h-1$, ce qui prouve H_{h-1} et achève la récurrence. $\square$

Proposition 3.5.9. *Si $T = (t_{ij})$ et $T' = (t'_{ij})$ sont des matrices de $\mathfrak{M}_n(\mathbb{K})$ triangulaires supérieures (respectivement inférieures), $TT' = T'' = (t''_{ij})$ est triangulaire supérieure (respectivement inférieure), et $t''_{ii} = t_{ii} t'_{ii}$. Pour tout polynôme $P \in \mathbb{K}[X]$, $P(T) = M = (m_{ij})$ est triangulaire supérieure (respectivement inférieure) et $m_{ii} = P(t_{ii})$ pour tout i .*

Démonstration. Traitons le cas où T et T' sont triangulaires supérieures, l'autre cas se prouvant de façon identique. On a $t''_{ij} = L_i(T) * C_j(T')$. Si $i > j$:

$$\begin{aligned} L_i(T) &= \begin{pmatrix} 0 & \dots & \dots & \dots & 0 & t_{ii} & \dots & t_{in} \end{pmatrix} \\ C_j(T') &= {}^t(t'_{1j} \quad \dots \quad t'_{jj} \quad 0 \quad \dots \quad \dots \quad 0) \end{aligned}$$

nous voyons que $t''_{ij} = 0$, et que, si $j = i$, alors $t''_{ii} = t_{ii} t'_{ii}$.

Faisant $T' = T$, nous voyons, par récurrence sur $n \in \mathbb{N}$, que T^n est triangulaire supérieure et que ses termes diagonaux sont t_{ii}^n . La somme de matrices triangulaires supérieures étant triangulaire supérieure, $M = P(T)$ est triangulaire supérieure, et $m_{ii} = P(t_{ii})$. $\square$

Proposition 3.5.10. *Si $A \in \mathfrak{M}_n(\mathbb{K})$ est inversible, sa transposée est inversible et $({}^tA)^{-1} = {}^t(A^{-1})$.*

Démonstration. En transposant $AA^{-1} = I_n$ on obtient ${}^t(A^{-1}) {}^tA = I_n$ (proposition 2.5.2), ce qui montre que tA est inversible, et que son inverse, $({}^tA)^{-1}$, est égale à ${}^t(A^{-1})$. $\square$

3.6 Changement de base

Soient deux bases de E , $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ et $\mathfrak{B}' = (\epsilon'_i)_{1 \leq i \leq n}$, avec $\epsilon_i = \epsilon(e_i)$ et $\epsilon'_i = \epsilon'(e_i)$ (proposition 2.4.5), de sorte que $\epsilon' \circ \epsilon^{-1}(\epsilon_i) = \epsilon'_i$. La i -ème colonne de la matrice de l'application linéaire $\epsilon' \circ \epsilon^{-1}$ dans la base $\mathfrak{B}$ est donc le vecteur ϵ'_i écrit dans la base $\mathfrak{B}$.

On note cette matrice $P_{\mathfrak{B}, \mathfrak{B}'}$ et on l'appelle **matrice de passage** de $\mathfrak{B}$ à $\mathfrak{B}'$. Elle représente l'identité de E dans les deux bases :

$$P_{\mathfrak{B}, \mathfrak{B}'} = M_{\mathfrak{B}, \mathfrak{B}'}(i_E).$$

Quel est l'effet de $P_{\mathfrak{B}, \mathfrak{B}'}$ sur les composantes d'un vecteur x de E ? Si x s'écrit $x_{\mathfrak{B}} = (x_i)$ dans $\mathfrak{B}$ et $x_{\mathfrak{B}'} = (x'_i)$ dans $\mathfrak{B}'$, on a :

$$x = \sum x'_i C_i(P_{\mathfrak{B}, \mathfrak{B}'}) = \sum x_i \epsilon_i$$

or $\sum x'_i C_i(P_{\mathfrak{B}, \mathfrak{B}'}) = P_{\mathfrak{B}, \mathfrak{B}'} x'_{\mathfrak{B}}$. D'où l'on conclut :

$$x_{\mathfrak{B}} = P_{\mathfrak{B}, \mathfrak{B}'} x_{\mathfrak{B}'}.$$

La matrice de passage de $\mathfrak{B}$ à $\mathfrak{B}'$ fait donc passer des composantes dans $\mathfrak{B}'$ aux composantes dans $\mathfrak{B}$. Si l'on a une troisième base $\mathfrak{B}''$, on a :

$$P_{\mathfrak{B}, \mathfrak{B}''} = P_{\mathfrak{B}, \mathfrak{B}'} P_{\mathfrak{B}', \mathfrak{B}''}.$$

Ceci est une convention (d'Einstein) : les vecteurs sont « contravariants » et les formes linéaires, étudiées plus loin, sont « covariantes ». Ils doivent s'écrire en colonne, mais, dans le texte, on écrit leurs transposés (en ligne) par commodité.

La matrice inverse assure le passage de $\mathfrak{B}$ à $\mathfrak{B}'$:

$$P_{\mathfrak{B}, \mathfrak{B}'}^{-1} = P_{\mathfrak{B}', \mathfrak{B}}.$$

Exemple 3.6.1. Soient $\mathfrak{B} = (\epsilon_1, \epsilon_2)$ la base canonique de $\mathbb{R}^2$ et :

$$\mathfrak{B}' = (\epsilon'_1 = \epsilon_1 + 2\epsilon_2, \epsilon'_2 = 3\epsilon_1 + 5\epsilon_2)$$

une autre base. On a :

$$P_{\mathfrak{B}, \mathfrak{B}'} = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$$

et, comme $\epsilon_1 = -5\epsilon'_1 + 2\epsilon'_2$ et $\epsilon_2 = 3\epsilon'_1 - \epsilon'_2$:

$$P_{\mathfrak{B}', \mathfrak{B}} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}.$$

On vérifie que $P_{\mathfrak{B}, \mathfrak{B}'} P_{\mathfrak{B}', \mathfrak{B}} = I_2 = P_{\mathfrak{B}', \mathfrak{B}} P_{\mathfrak{B}, \mathfrak{B}'}$. Soit $x_{\mathfrak{B}} = {}^t(3, 4)$. Le calcul direct donne :

$$x = 3(-5\epsilon'_1 + 2\epsilon'_2) + 4(3\epsilon'_1 - \epsilon'_2) = -3\epsilon'_1 + 2\epsilon'_2$$

d'où : $x_{\mathfrak{B}'} = {}^t(-3, 2)$. On a bien :

$$x_{\mathfrak{B}'} = P_{\mathfrak{B}', \mathfrak{B}} x_{\mathfrak{B}} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}.$$

▽

Quel est l'effet d'un changement de base sur la matrice d'une application linéaire $u \in \mathfrak{L}(E, F)$? Soient $\mathfrak{B}(E)$ et $\mathfrak{B}'(E)$ deux bases de E , $\mathfrak{B}(F)$ et $\mathfrak{B}'(F)$ deux bases de F . Le diagramme commutatif :

$$\begin{array}{ccc} \mathfrak{B}(E) & \xrightarrow{M_{\mathfrak{B}(F), \mathfrak{B}(E)}(u)} & \mathfrak{B}(F) \\ P_{\mathfrak{B}(E), \mathfrak{B}'(E)} \uparrow & & \downarrow P_{\mathfrak{B}'(F), \mathfrak{B}(F)} \\ \mathfrak{B}'(E) & \xrightarrow{M_{\mathfrak{B}'(F), \mathfrak{B}'(E)}(u)} & \mathfrak{B}'(F) \end{array}$$

montre que :

$$M_{\mathfrak{B}'(F), \mathfrak{B}'(E)}(u) = P_{\mathfrak{B}'(F), \mathfrak{B}(F)} M_{\mathfrak{B}(F), \mathfrak{B}(E)}(u) P_{\mathfrak{B}(E), \mathfrak{B}'(E)}.$$

On remarquera (contravariance) que les indices se lisent de droite à gauche :

$$\mathfrak{B}'(F) \leftarrow \mathfrak{B}(F), \mathfrak{B}(F) \leftarrow \mathfrak{B}(E), \mathfrak{B}(E) \leftarrow \mathfrak{B}'(E).$$

Si $u \in \mathfrak{L}(E)$, en notant $P = P_{\mathfrak{B}, \mathfrak{B}'}$, $M = M_{\mathfrak{B}}(u)$ et $M' = M_{\mathfrak{B}'}(u)$ le diagramme commutatif :

$$\begin{array}{ccc} \mathfrak{B}(E) & \xrightarrow{M} & \mathfrak{B}(E) \\ P \uparrow & & \downarrow P^{-1} \\ \mathfrak{B}'(E) & \xrightarrow{M'} & \mathfrak{B}'(E) \end{array}$$

se traduit par $M' = P^{-1} M P$, ou $M = P M' P^{-1}$.

Deux matrices $M, M' \in \mathfrak{M}_n(\mathbb{K})$ sont **semblables** s'il existe une matrice régulière $P \in \mathfrak{M}_n(\mathbb{K})$ telle que $M = P M' P^{-1}$, c'est-à-dire si elles représentent le même endomorphisme d'un $\mathbb{K}$ -espace vectoriel de dimension n dans deux bases. Des matrices semblables sont donc équivalentes.

3.7 Forme linéaire, dual

Une application linéaire d'un $\mathbb{K}$ -espace vectoriel E , de dimension n , dans $\mathbb{K}$ est une **forme linéaire** sur E , ou un **covecteur**, c'est-à-dire un vecteur **co-variant**, comme nous le verrons. L'ensemble des formes linéaires sur E est son **dual** noté E^* . Le dual du dual est le **bidual** de E noté E^{**} .

Exemple 3.7.1. Soit E le $\mathbb{K}$ -espace vectoriel des polynômes de $\mathbb{K}[X]$ de degré au plus égal à 2. Les applications de E dans $\mathbb{K}$:

$$\begin{aligned} v_1(P) &= P(0), \\ v_2(P) &= P'(0), \\ v_3(P) &= \int_0^1 P(t) dt \end{aligned}$$

sont des formes linéaires sur E . Leurs matrices dans la base canonique de E $(1, X, X^2)$ sont :

$$\begin{aligned} M(v_1) &= (1, 0, 0), \\ M(v_2) &= (0, 1, 0), \\ M(v_3) &= (1, 0, 2/3), \end{aligned}$$

et elles sont linéairement indépendantes. ▽

Exemple 3.7.2. La matrice de la forme linéaire :

$$\begin{aligned} u : E = \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto ax + by = (a, b) \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

est $M(u) = (a, b) \in E^*$. C'est donc la transposée d'un vecteur de $\mathbb{R}^2$. On peut écrire en transposant :

$$ax + by = (x, y) \begin{pmatrix} a \\ b \end{pmatrix}.$$

Le vecteur ${}^t(x, y)$ est donc, à transposition près, un élément de E^{**} . ▽

Imaginons un client entrant dans un magasin pour la première fois. Le magasin vend n articles (a_i) dont les prix respectifs sont p_i . Le client désire acheter la quantité q_i de chaque a_i . La facture sera donc $f = \sum_i q_i p_i$. Pour le gérant du magasin, les p_i sont connus, et les q_i sont les inconnues de l'équation. Pour le client, la situation est inverse : il connaît les q_i mais pas les p_i , qui sont pour lui les variables. Cet exemple illustre la dualité des points de vue.

Théorème 3.7.1. *Le dual E^* et le bidual E^{**} d'un $\mathbb{K}$ -espace vectoriel E , de dimension n , sont des $\mathbb{K}$ -espaces vectoriels de dimension n , et E^{**} est canoniquement isomorphe à E , auquel on peut l'identifier.*

Démonstration. Que E^* et E^{**} sont des $\mathbb{K}$ -espaces vectoriels de dimension n découle de la proposition 3.4.9 pour $m = 1$. Tout élément x de E définit la forme linéaire sur E^* :

$$\begin{aligned} x^{**} &: E^* \rightarrow \mathbb{K} \\ u &\mapsto u(x). \end{aligned}$$

On a en effet quels que soient les scalaires λ et μ :

$$\begin{aligned} x^{**}(\lambda u + \mu v) &= (\lambda u + \mu v)(x) \\ &= \lambda u(x) + \mu v(x) \\ &= \lambda x^{**}(u) + \mu x^{**}(v) \end{aligned}$$

d'où la linéarité de x^{**} .

Montrons que l'application ϕ de E dans E^{**} définie par $\phi(x) = x^{**}$ est un isomorphisme.

Elle est linéaire car $\phi(\alpha x + \beta y) = (\alpha x + \beta y)^{**}$ quels que soient les éléments et les scalaires écrits, et :

$$\begin{aligned} (\alpha x + \beta y)^{**}(u) &= u(\alpha x + \beta y) \\ &= \alpha u(x) + \beta u(y) \\ &= (\alpha x^{**} + \beta y^{**})(u) \\ &= (\alpha \phi(x) + \beta \phi(y))(u) \end{aligned}$$

signifie que $\phi(\alpha x + \beta y) = \alpha \phi(x) + \beta \phi(y)$.

Elle est injective car $\phi(x) = 0$ signifie que pour tout $u \in E^*$ $x^{**}(u) = 0$, que $u(x) = 0$, donc que $x = 0$.

La bijectivité découle de la proposition 3.4.6 et de l'égalité des dimensions.

Cet isomorphisme, ne dépendant d'aucune base, est canonique. $\square$

Une forme linéaire $u \in E^*$ est représentée dans une base $B = (\epsilon_1, \dots, \epsilon_n)$ de E par une matrice à une ligne :

$$M_B(u) = (u(\epsilon_1), \dots, u(\epsilon_n)).$$

La **base duale** de B est la base $B^* = (\epsilon_1^*, \dots, \epsilon_n^*)$ de E^* définie par :

$$\epsilon_i^*(\epsilon_j) = \delta_{ij}.$$

Rappelons que δ_{ij} est le symbole de Kronecker, défini dans les Généralités.

Grâce à son produit scalaire canonique, $\mathbb{K}^n$ est son propre dual et sa base canonique est sa propre base duale.

Exemple 3.7.3. Calculons la base duale de la base $\mathfrak{B}$ de $\mathbb{R}^3$:

$$\mathfrak{B} = \{\epsilon_1 = {}^t(1, 1, 1), \epsilon_2 = {}^t(0, 1, 1), \epsilon_3 = {}^t(0, 0, 1)\}$$

exprimée dans la base canonique. Dans cette même base (canonique), qui est la base canonique du dual, on a :

$$M(\epsilon_1^*) = (a_1, a_2, a_3), \quad M(\epsilon_2^*) = (b_1, b_2, b_3), \quad M(\epsilon_3^*) = (c_1, c_2, c_3),$$

et, d'après la définition :

$$\begin{aligned} \epsilon_1^*(\epsilon_1) &= 1 &= a_1 + a_2 + a_3, \\ \epsilon_1^*(\epsilon_2) &= 0 &= a_2 + a_3, \\ \epsilon_1^*(\epsilon_3) &= 0 &= a_3, \end{aligned}$$

d'où $\epsilon_1^* = (1, 0, 0)$. Un calcul identique donne :

$$\epsilon_2^* = (-1, 1, 0), \quad \epsilon_3^* = (0, -1, 1).$$

▽

3.8 Application linéaire transposée

Si $u \in \mathcal{L}(E, F)$ et si $\phi \in F^*$, $\phi \circ u$ est une forme linéaire sur E . L'application :

$$\begin{aligned} {}^t u : F^* &\rightarrow E^* \\ \phi &\mapsto \phi \circ u \end{aligned}$$

est linéaire :

$$\begin{aligned} {}^t u(\lambda \phi_1 + \mu \phi_2) &= (\lambda \phi_1 + \mu \phi_2) \circ u \\ &= \lambda \phi_1 \circ u + \mu \phi_2 \circ u \\ &= \lambda {}^t u \circ \phi_1 + \mu {}^t u \circ \phi_2 \end{aligned}$$

et appartient donc à $\mathcal{L}(F^*, E^*)$. C'est la **transposée** de u . Nous voyons que la transposition inverse le sens des applications, et transforme donc les objets contravariants en objets covariants. Une forme linéaire (ou covecteur) est transposée d'un vecteur.

Proposition 3.8.1. *Les bases respectives de E et de F étant $\mathfrak{B} = (\epsilon_i)$ et $\mathfrak{B}' = (\phi_i)$, la matrice de la transposée de $u \in \mathcal{L}(E, F)$ dans les bases duales $\mathfrak{B}'^* = (\phi_i^*)$, de F^* , et $\mathfrak{B}^* = (\epsilon_i^*)$, de E^* , est la transposée de la matrice $M(u)$:*

$$M_{\mathfrak{B}'^*, \mathfrak{B}'^*}({}^t u) = {}^t M_{\mathfrak{B}', \mathfrak{B}}(u).$$

Démonstration. Posons $M_{\mathfrak{B}', \mathfrak{B}}(u) = M = (m_{ij})$ et $M_{\mathfrak{B}'^*, \mathfrak{B}'^*}({}^t u) = {}^t M$. Comme $u(\epsilon_j) = \sum_i m_{ij} \phi_i$, on a :

$${}^t u(\phi_i^*) = \phi_i^*(u) : \epsilon_j \mapsto m_{ij}.$$

Faisant aller j de 1 à $\dim(E) = m$, nous obtenons la i -ème colonne de $M({}^t u)$:

$$C_i(M({}^t u)) = (m_{i1}, \dots, m_{im}) = L_i(M)$$

ce qui est le résultat annoncé. □

Proposition 3.8.2. *Si $u_i \in \mathcal{L}(E_i, E_{i+1})$, $1 \leq i \leq r$, on a :*

$${}^t(u_r \circ \dots \circ u_1) = {}^t u_1 \circ \dots \circ {}^t u_r.$$

Démonstration. Montrons-le d'abord pour $r = 2$. Soient $\mathfrak{B}_1 = (\epsilon_i)$, $\mathfrak{B}_2 = (\phi_i)$, $\mathfrak{B}_3 = (\psi_i)$ les bases respectives de E_1 , E_2 , E_3 , et $\mathfrak{B}_1^*$, $\mathfrak{B}_2^*$, $\mathfrak{B}_3^*$ les bases duales. On a :

$${}^t(u_2 \circ u_1)(\psi_i^*) = \psi_i^* \circ u_2 \circ u_1 = {}^t u_2 \circ \psi_i^* \circ u_1 = {}^t u_2 \circ {}^t u_1 \circ \psi_i^*$$

quel que soit i , d'où le résultat dans ce cas. On passe au cas général par une récurrence immédiate. $\square$

Le diagramme commutatif précédent, où $u \in \mathcal{L}(E)$, $P = P_{\mathfrak{B}, \mathfrak{B}'}$, $M = M_{\mathfrak{B}}(u)$ et $M' = M_{\mathfrak{B}'}(u)$:

$$\begin{array}{ccc} \mathfrak{B}(E) & \xrightarrow{M} & \mathfrak{B}(E) \\ P \uparrow & & \downarrow P^{-1} \\ \mathfrak{B}'(E) & \xrightarrow{M'} & \mathfrak{B}'(E) \end{array}$$

donne par transposition :

$$\begin{array}{ccc} \mathfrak{B}(E)^* & \xleftarrow{{}^t M} & \mathfrak{B}(E)^* \\ \downarrow {}^t P & & \uparrow {}^t P^{-1} \\ \mathfrak{B}'^*(E) & \xleftarrow{{}^t M'} & \mathfrak{B}'^*(E) \end{array}$$

c'est-à-dire ${}^t M' = {}^t P {}^t M {}^t P^{-1}$, ou ${}^t M = {}^t P^{-1} {}^t M' {}^t P$.

La transposition transforme les objets contravariants en objets covariants, et inversement. Elle laisse les scalaires invariants, mais ils ne sont ni covariants ni contravariants.

Proposition 3.8.3. *La transposée de la transposée d'une application linéaire $u \in \mathcal{L}(E, F)$ est égale à u .*

Démonstration. Remarquons d'abord que ${}^t({}^t u)$ appartient à $\mathcal{L}(E^{**}, F^{**})$, donc à $\mathcal{L}(E, F)$ (théorème 3.7.1). On a, quels que soient $x \in E$ et $\phi \in F^*$:

$$\phi \circ u(x) = x^{**}(\phi \circ u) = x^{**}({}^t u \circ \phi) = {}^t({}^t u)(x^{**}) \circ \phi = \phi \circ {}^t({}^t u)(x),$$

d'où l'égalité annoncée. $\square$

4 Déterminants

4.1 Permutations et transpositions

Une **permutation** σ sur l'ensemble ordonné $[1, n] = \{1, 2, \dots, n\}$ est une bijection sur cet ensemble :

$$\sigma : (1, 2, \dots, n) \mapsto (\sigma(1), \sigma(2), \dots, \sigma(n)).$$

La composée (ou le produit) $\sigma_2 \circ \sigma_1$ de deux permutations est la permutation qui commute i et $\sigma_2(\sigma_1(i))$, pour $i = 1, \dots, n$. L'associativité est triviale. L'inverse σ^{-1} de σ est définie par $i \mapsto \sigma^{-1}(i)$. L'ensemble des permutations est donc un groupe, le **groupe symétrique** $\mathfrak{S}_n$, dont l'élément neutre est la permutation **identique** $\sigma(i) = i$.

La **transposition** $\tau_{h,k}$ est la permutation qui commute h et k et laisse invariants les autres entiers. Elle est son propre inverse (c'est une involution). Cette notion n'a qu'un rapport assez lointain avec son homonyme pour les applications linéaires et les matrices (${}^t({}^t u) = u$).

Proposition 4.1.1. *Les transpositions engendrent $\mathfrak{S}_n$.*

Démonstration. Il nous faut montrer qu'une permutation quelconque σ est la composée de transpositions. Le sous-ensemble $(i_1, i_2, \dots, i_k)$ est un **cycle** pour σ si $\sigma(i_1) = i_2$, $\sigma(i_2) = i_3$, ainsi de suite jusqu'à $\sigma(i_k) = i_1$. Deux cycles sont évidemment disjoints, et $[1, n]$ se décompose en une réunion de cycles.

La restriction σ_1 de σ au premier cycle $C_1 = (1, \sigma(1), \sigma^2(1), \dots, \sigma^k(1))$ se décompose en :

$$\sigma_1 = \tau_{k, \sigma(k)} \circ \dots \circ \tau_{1, \sigma(1)}$$

Le premier élément du deuxième cycle est le plus petit entier non inclus dans le premier cycle. On le décompose en produit de transpositions comme le premier. Ainsi de suite jusqu'au dernier, et σ , produit des σ_i , est mise sous la forme annoncée. $\square$

Par exemple, la permutation $\sigma = (3, 5, 4, 1, 2)$ se décompose en deux cycles : $C_1 = (1, 3, 4)$ et $C_2 = (2, 5)$. On a :

$$\sigma_1 = \tau_{3,4} \circ \tau_{1,3}.$$

En effet, $\tau_{1,3}$ fait passer de $(1, 2, 3, 4, 5)$ à $(3, 2, 1, 4, 5)$, puis $\tau_{3,4}$ fait passer de $(3, 2, 1, 4, 5)$ à $(3, 2, 4, 1, 5)$. Enfin, $\sigma_2 = \tau_{2,5}$ fait passer à $(3, 5, 4, 1, 2)$.

La décomposition n'est pas unique puisque $\tau_{i,j}^2$ est l'identité, mais nous verrons un peu plus loin que la parité du nombre de transpositions dans la décomposition d'une permutation est fixée. Il y a donc des permutations paires et des impaires.

4.2 Formes n -linéaires alternées

Une **forme n -linéaire** sur un $\mathbb{K}$ -espace vectoriel E de dimension n est une application $u : E^n \rightarrow \mathbb{K}$ linéaire sur chaque facteur, c'est-à-dire que si l'on fixe les vecteurs v_j pour $j \neq i$, l'application :

$$\begin{aligned} u_i & : E \rightarrow \mathbb{K} \\ v & \mapsto u(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n) \end{aligned}$$

est une forme linéaire, pour tous les indices i de 1 à n . Posons pour une permutation $\sigma \in \mathfrak{S}_n$ et tout n -uplet (v_i) de vecteurs de E :

$$\sigma(v_1, \dots, v_n) = (v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

La forme n -linéaire u est **alternée** si elle s'annule chaque fois que deux des v_i sont égaux. Notons $\mathfrak{A}_n(E)$ l'ensemble de telles formes.

Si, pour toute transposition τ , $u \circ \tau = u$, la forme u est **symétrique** et si $u \circ \tau = -u$ elle est **antisymétrique**.

Proposition 4.2.1. *Toute forme n -linéaire alternée sur E est antisymétrique.*

Démonstration. Posons $v(v_i, v_j) = u(v_1, \dots, v_n)$, les vecteurs v_k pour $k \neq i$ et $k \neq j$ étant fixés. On a par bilinéarité de v :

$$v(v_i + v_j, v_i + v_j) = v(v_i, v_i) + v(v_j, v_j) + v(v_i, v_j) + v(v_j, v_i).$$

Or par hypothèse les quantités $v(v_i + v_j, v_i + v_j)$, $v(v_i, v_i)$ et $v(v_j, v_j)$ sont nulles, et il reste $v(v_i, v_j) + v(v_j, v_i) = 0$, d'où la conclusion. $\square$

Notons que la réciproque est fautive en caractéristique 2, où $x = -x$, ce qui confond symétrie et antisymétrie. Ainsi, pour $n = 2$, la forme bilinéaire $u((x_1, x_2), (x_3, x_4)) = x_1x_3 + x_2x_4$ sur un corps de caractéristique 2 est antisymétrique (car symétrique) mais n'est pas alternée car $u((\alpha, \beta), (\alpha, \beta))$, égal à $\alpha^2 + \beta^2$, n'est pas nul si $\alpha \neq \beta$.

Proposition 4.2.2. *La parité du nombre de transpositions dans la décomposition d'une permutation est fixée.*

Démonstration. Soit $u \in \mathfrak{A}_n(E)$ et $\sigma \in \mathfrak{S}_n$. Supposons σ mise sous la forme, d'une part, du produit de r_1 transpositions et, d'autre part, du produit de r_2 transpositions. Comme chaque transposition change le signe du résultat, on aura :

$$u \circ \sigma = (-1)^{r_1} u = (-1)^{r_2} u$$

ce qui montre que les parités de r_1 et de r_2 sont égales. $\square$

Si la permutation $\sigma \in \mathfrak{S}_n$ peut se décomposer en un produit de r transpositions, la quantité $\epsilon_\sigma = (-1)^r$ (indépendante des valeurs possibles de r d'après la proposition précédente) s'appelle la **signature** de σ . La signature de l'identité est égale à 1.

Proposition 4.2.3. *La signature d'un produit de permutations est le produit de leurs signatures. Les signatures d'une permutation σ et de σ^{-1} sont égales.*

Démonstration. Il suffit de le prouver pour deux permutations σ_1 et σ_2 . Il est clair que si σ_1 est le produit des r_1 transpositions τ_{1i} et σ_2 celui des r_2 transpositions τ_{2i} on aura :

$$\sigma_2 \circ \sigma_1 = \prod_{i=1}^{r_2} \tau_{2i} \circ \prod_{i=1}^{r_1} \tau_{1i},$$

ce qui montre que $\sigma_2 \circ \sigma_1$ est le produit de $r_1 + r_2$ transpositions, d'où :

$$\epsilon_{\sigma_2 \circ \sigma_1} = (-1)^{r_1+r_2} = \epsilon_{\sigma_1} \epsilon_{\sigma_2}.$$

Si l'on a $\sigma = \tau_1 \dots \tau_r$ il est clair que $\sigma^{-1} = \tau_r \dots \tau_1$, et que $\epsilon_{\sigma^{-1}} = \epsilon_\sigma$. $\square$

Proposition 4.2.4. *L'ensemble $\mathfrak{A}_n(E)$ est un $\mathbb{K}$ -espace vectoriel de dimension 1.*

Démonstration. Il est immédiat que $\mathfrak{A}_n(E)$ est un $\mathbb{K}$ -espace vectoriel.

Soit $\mathfrak{B} = (b_1, \dots, b_n)$ une base de E et une forme $u \in \mathfrak{A}_n(E)$ telle que $u(b_1, \dots, b_n) = a \in \mathbb{K}$. On a alors pour toute permutation σ :

$$u \circ \sigma(b_1, \dots, b_n) = \epsilon_\sigma u(b_1, \dots, b_n).$$

Soit $v_j = \sum_{1 \leq i \leq n} a_{ij} b_i$ pour $1 \leq j \leq n$. Par n -linéarité, $\phi(v_1, \dots, v_n)$ est une combinaison linéaire des $u(b_{i_1}, \dots, b_{i_n})$, termes non nuls seulement si les i_k sont deux à deux distincts; ils définissent dans ce cas une permutation.

Il ne reste donc que $(u \circ \sigma)((b_i))$, dont le coefficient est égal au produit des $a_{i, \sigma(i)}$, et l'on a :

$$u((v_1, \dots, v_n)) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n a_{i, \sigma(i)} u \circ \sigma(b_1, \dots, b_n),$$

ce qui montre que la donnée de $u(b_1, \dots, b_n)$ détermine complètement u , et que $\mathfrak{A}_n(\mathbb{K})$ est de dimension 1.

Il admet pour base l'élément ϕ défini par $\phi(b_1, \dots, b_n) = 1$. □

4.3 Déterminant de n vecteurs

Le **déterminant de n vecteurs** $v_1, \dots, v_n$ de $\mathbb{K}^n$ est la forme n -linéaire alternée ϕ qui vaut 1 sur la base canonique :

$$\det(v_1, \dots, v_n) = \phi(v_1, \dots, v_n).$$

On a donc, si $v_i = (v_{i1}, \dots, v_{in})$ pour tout i :

$$\det(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n v_{i, \sigma(i)}.$$

Le produit étant commutatif, nous pouvons changer l'ordre des termes en faisant agir une permutation quelconque σ' . Choisissons $\sigma' = \sigma^{-1}$:

$$\det(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n v_{\sigma^{-1}(i), \sigma(\sigma^{-1}(i))},$$

d'où, puisque $\epsilon_{\sigma^{-1}} = \epsilon_\sigma$:

$$\det(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_{\sigma^{-1}} \prod_{i=1}^n v_{\sigma^{-1}(i), i}.$$

Il n'y a plus qu'à remarquer que, quand σ parcourt $\mathfrak{S}_n$, σ^{-1} en fait de même, et à remplacer la notation de l'élément courant σ^{-1} par σ :

$$\det(v_1, \dots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n v_{\sigma(i), i}.$$

On note :

$$\det((v_i)) = \begin{vmatrix} v_{11} & \dots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nn} \end{vmatrix}.$$

Le déterminant de n vecteurs d'un $\mathbb{K}$ -espace vectoriel E de dimension n ne peut être défini qu'à partir d'une base. Soit donc $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ une base de E . La forme $u \in \mathfrak{A}(E)$ telle que $u(\epsilon_1, \dots, \epsilon_n) = 1$ est alors le déterminant des n vecteurs dans la base $\mathfrak{B}$:

$$\det_{\mathfrak{B}}(v_1, \dots, v_n) = u(v_1, \dots, v_n).$$

Proposition 4.3.1. *Si $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ et $\mathfrak{B}'$ sont deux bases quelconques de E , on a pour tout n -uplet (v_i) de vecteurs de E :*

$$\det_{\mathfrak{B}'}((v_i)) = \det_{\mathfrak{B}'}((\epsilon_i)) \det_{\mathfrak{B}}((v_i)).$$

Démonstration. La dimension de $\mathfrak{A}_n(E)$ étant égale à 1, les déterminants sont proportionnels et on a pour tout n -uplet (v_i) de vecteurs de E :

$$\det_{\mathfrak{B}'}((v_i)) = \lambda \det_{\mathfrak{B}}((v_i)), \quad \lambda \in \mathbb{K},$$

d'où :

$$\det_{\mathfrak{B}'}((\epsilon_i)) = \lambda \det_{\mathfrak{B}}((\epsilon_i)).$$

Comme $\det_{\mathfrak{B}}((\epsilon_i)) = 1$, on a $\lambda = \det_{\mathfrak{B}'}((\epsilon_i))$. □

N'oublions pas qu'un déterminant est alterné, c'est-dire qu'il s'annule dès qu'il existe un couple (i, j) tel que $v_i = v_j$.

Proposition 4.3.2. *On ne change pas le déterminant de n vecteurs (v_i) en ajoutant à l'un des vecteurs une combinaison linéaire des $n - 1$ autres.*

Démonstration. Il suffit de montrer que $\det((v_i)) = \det((v'_i))$, où $v'_i = v_i$ pour $i \neq k$ et $v'_k = v_k + \lambda v_j$, $j \neq k$. On a par linéarité :

$$\det((v'_i)) = \det((v_i)) + \lambda \det(v_1, \dots, v_{k-1}, v_j, v_{k+1}, \dots, v_n),$$

et le deuxième terme du membre de droite est nul à cause de la répétition de v_j qui se trouve en j -ème et en k -ième position. □

Théorème 4.3.1. *Une famille de n vecteurs d'un $\mathbb{K}$ -espace vectoriel E de dimension n est une base si et seulement si le déterminant de ces n vecteurs dans une base quelconque de E n'est pas nul.*

Démonstration. Montrons d'abord que si ces vecteurs $v_1, \dots, v_n$ sont liés leur déterminant dans une base $\mathfrak{B} = (\epsilon_i)_{1 \leq i \leq n}$ de E est nul. Il existe dans ce cas des scalaires λ_i non tous nuls tels que :

$$\sum_i \lambda_i v_i = 0.$$

On peut supposer, quitte à changer l'ordre des ϵ_i , que $\lambda_n \neq 0$. On a alors :

$$v_n = \lambda_n^{-1} \sum_{1 \leq i \leq n-1} \lambda_i v_i$$

et :

$$\begin{aligned} \det_{\mathfrak{B}}(v_1, \dots, v_n) &= \det_{\mathfrak{B}}(v_1, \dots, \lambda_n^{-1} \sum_{1 \leq i \leq n-1} \lambda_i v_i) \\ &= \lambda_n^{-1} \sum_{1 \leq i \leq n-1} \lambda_i \det_{\mathfrak{B}}(v_1, \dots, v_{i-1}, v_i) \\ &= 0 \end{aligned}$$

puisque les $\det_{\mathfrak{B}}(v_1, \dots, v_{i-1}, v_i)$ sont nuls, le i -ème vecteur étant égal au dernier.

Réciproquement, montrons que si la famille (v_i) est une base $\mathfrak{B}$ de E son déterminant dans une base quelconque ne peut être nul. La forme $u \in \mathfrak{A}(E)$ définie par $u(v_1, \dots, v_n) = 1$ est un déterminant, $\det_{\mathfrak{B}}$. Soit $\mathfrak{B}' = (v'_1, \dots, v'_n)$ une quelconque autre base. On a d'après la proposition 3.3.1 :

$$\begin{aligned} \det_{\mathfrak{B}}(v_1, \dots, v_n) &= \det_{\mathfrak{B}}(v'_1, \dots, v'_n) \det_{\mathfrak{B}'}(v_1, \dots, v_n) \\ 1 &= \det_{\mathfrak{B}}(v'_1, \dots, v'_n) \det_{\mathfrak{B}'}(v_1, \dots, v_n) \end{aligned}$$

ce qui montre que $\det_{\mathfrak{B}'}(v_1, \dots, v_n)$ ne peut être nul. $\square$

4.4 Déterminant d'une matrice

Le **déterminant de la matrice** $A = (a_{ij}) \in \mathfrak{M}_n(\mathbb{K})$ est le déterminant des n vecteurs colonnes $C_j(A) = (a_{ij})_{1 \leq i \leq n}$. On le note :

$$\det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Le déterminant de la matrice identité I_n est égal à 1.

Le déterminant de n vecteurs (v_i) est linéaire sur chaque v_i mais n'est pas globalement linéaire : $\det((v_i) + (w_i)) \neq \det((v_i)) + \det((w_i))$, et, si A et B sont dans $\mathfrak{M}_n(\mathbb{K})$, $\det(A + B) \neq \det(A) + \det(B)$. Notons bien que, si $\lambda \in \mathbb{K}$, alors $\det(\lambda A) = \lambda^n \det(A)$.

Proposition 4.4.1. *Le déterminant d'un produit de matrices est égal au produit des déterminants de ces matrices.*

Démonstration. Il suffit, par récurrence, de le prouver pour deux matrices, A et B de $\mathfrak{M}_n(\mathbb{K})$. L'application :

$$(C_1(A), \dots, C_n(A)) \mapsto \det(BA) = \det(B(C_i(A)))$$

est linéaire sur chaque facteur $C_i(A)$. En effet l'application :

$$C_i(A) \mapsto C_i(BA) = \begin{pmatrix} (L_1(B) * C_i(A)) \\ \vdots \\ L_n(B) * C_i(A) \end{pmatrix}$$

est linéaire, donc aussi sa composée avec $\det$.

Elle est alternée puisque si $C_i(A) = C_j(A)$, alors $C_i(BA) = C_j(BA)$.

L'application $A \mapsto \det(BA)$ appartient donc à $\mathfrak{A}_n(\mathbb{K}^n)$, et elle est proportionnelle à $A \mapsto \det(A)$. Faisant $A = I_n$, nous voyons que le coefficient de proportionnalité est égal à $\det(B)$, d'où : $\det(BA) = \det(B) \det(A) = \det(AB)$. $\square$

Proposition 4.4.2. *Si $A \in \mathfrak{M}_n(\mathbb{K}^n)$ est inversible, alors $\det(A^{-1}) = \frac{1}{\det(A)}$.*

Démonstration. On a : $\det(A^{-1}) \det(A) = \det(A^{-1}A) = \det(I_n) = 1$. $\square$

Proposition 4.4.3. *Les déterminants de $M \in \mathfrak{M}_n$ et de tM sont égaux.*

Démonstration. Le déterminant de M est égal à :

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n m_{i \sigma(i)}.$$

ou à :

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n m_{\sigma(i) i}.$$

Or la seconde forme est obtenue à partir de la première en permutant m_{ij} et m_{ji} , ce qui correspond à la transposition. $\square$

4.5 Développement d'un déterminant

Le déterminant de la matrice $M = (m_{ij}) \in \mathfrak{M}_n$ (ou des n vecteurs colonnes $C_j(M)$) est égal à :

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{i=1}^n m_{i \sigma(i)}.$$

Le **mineur** Δ_{hk} est le déterminant de la matrice $M^{(hk)} = (m_{ij}^{(hk)})$ obtenue à partir de M en supprimant sa h -ième ligne et sa k -ième colonne.

Remarquons que $M^{(hk)}$ appartient à $\mathfrak{M}_{n-1}(\mathbb{K})$, et que :

$$\begin{aligned} m_{ij}^{(hk)} &= m_{ij} & \text{si } i < h \text{ et } j < k, \\ m_{ij}^{(hk)} &= m_{i j+1} & \text{si } i < h \text{ et } j \geq k, \\ m_{ij}^{(hk)} &= m_{i+1 j} & \text{si } i \geq h \text{ et } j < k, \\ m_{ij}^{(hk)} &= m_{i+1 j+1} & \text{si } i \geq h \text{ et } j \geq k. \end{aligned}$$

Exemple 4.5.1. Les $3! = 6$ permutations sur $(1, 2, 3)$ sont :

$$\begin{aligned} (1, 2, 3) &\mapsto (1, 2, 3), & \text{aucune transposition,} \\ (1, 2, 3) &\mapsto (1, 3, 2), & \text{une transposition : } (2, 3) \mapsto (3, 2), \\ (1, 2, 3) &\mapsto (2, 1, 3), & \text{une transposition : } (1, 2) \mapsto (2, 1), \\ (1, 2, 3) &\mapsto (2, 3, 1), & \text{deux transpositions : } (1, 2) \mapsto (2, 1), (1, 3) \mapsto (3, 1), \\ (1, 2, 3) &\mapsto (3, 2, 1), & \text{une transposition : } (1, 3) \mapsto (3, 1), \\ (1, 2, 3) &\mapsto (3, 1, 2), & \text{deux transpositions : } (1, 2) \mapsto (2, 1), (2, 3) \mapsto (3, 2). \end{aligned}$$

Il y a trois permutations paires et trois impaires.
Le déterminant de la matrice :

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

est, en respectant l'ordre précédent pour les permutations, leur signature étant -1 dans le cas d'une transposition et 1 dans le cas de deux :

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

Mettons en facteur les termes de la deuxième colonne :

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

Calculons les déterminants Δ_{h2} des matrices $M^{(h2)}$:

$$\Delta_{12} = \begin{vmatrix} m_{21} & m_{23} \\ m_{31} & m_{33} \end{vmatrix} = m_{21}m_{33} - m_{23}m_{31} ,$$

$$\Delta_{22} = \begin{vmatrix} m_{11} & m_{13} \\ m_{31} & m_{33} \end{vmatrix} = m_{11}m_{33} - m_{13}m_{31} ,$$

$$\Delta_{32} = \begin{vmatrix} m_{11} & m_{13} \\ m_{21} & m_{23} \end{vmatrix} = m_{11}m_{23} - m_{13}m_{21} .$$

Nous voyons que :

$$\det(M) = -m_{12} \Delta_{12} + m_{22} \Delta_{22} - m_{32} \Delta_{32} .$$

Des calculs analogues montrent que :

$$\begin{aligned} \det(M) &= m_{11} \Delta_{11} - m_{21} \Delta_{21} + m_{31} \Delta_{31} \\ &= m_{13} \Delta_{13} - m_{23} \Delta_{23} + m_{33} \Delta_{33} \end{aligned}$$

c'est-à-dire : $L_i(\widetilde{M}) * C_i(M) = \det(M)$ pour $i = 1, 2$ et 3 .

Posons :

$$\widetilde{M} = \begin{pmatrix} \Delta_{11} & -\Delta_{21} & \Delta_{31} \\ -\Delta_{12} & \Delta_{22} & -\Delta_{32} \\ \Delta_{13} & -\Delta_{23} & \Delta_{33} \end{pmatrix}$$

et remarquons que $m_{11} \Delta_{12} - m_{21} \Delta_{22} + m_{31} \Delta_{32} = 0$, c'est-à-dire :

$$L_2(\widetilde{M}) * C_1(M) = 0 .$$

En fait, dès que $i \neq j$ on a :

$$L_i(\widetilde{M}) * C_j(M) = 0 .$$

Les résultats précédents montrent finalement que :

$$\widetilde{M} M = (L_i(\widetilde{M}) * C_j(M))_{1 \leq i, j \leq n} = \det(M) I_3 ,$$

d'où l'on déduit lorsque $\det(M) \neq 0$:

$$M^{-1} = \frac{1}{\det(M)} \widetilde{M} .$$

▽

Nous allons généraliser tout ceci.

On appelle **développement** du déterminant de M suivant sa j -ème colonne l'expression :

$$\sum_{1 \leq i \leq n} (-1)^{i+j} m_{ij} \Delta_{ij} ,$$

et suivant sa i -ème ligne :

$$\sum_{1 \leq j \leq n} (-1)^{i+j} m_{ij} \Delta_{ij} .$$

Proposition 4.5.1. *Le développement du déterminant de M suivant sa j -ème colonne ou sa i -ème ligne est égal à son déterminant :*

$$\det(M) = \sum_{1 \leq i \leq n} (-1)^{i+j} m_{ij} \Delta_{ij} = \sum_{1 \leq j \leq n} (-1)^{i+j} m_{ij} \Delta_{ij} .$$

Démonstration. Il suffit de démontrer pour une colonne, le cas d'une ligne s'en déduisant par transposition. Le coefficient c_{11} de m_{11} dans $\det(M)$ est :

$$c_{11} = \sum_{\sigma \in \mathfrak{S}_n} \epsilon_\sigma \prod_{\substack{i > 1 \\ \sigma(i) > 1}} m_{i \sigma(i)} ,$$

σ parcourant l'ensemble des bijections de la suite des $n - 1$ premiers entiers naturels plus grands que 1 dans la même suite. En remplaçant les m_{ij} par les $m_{i'j'}^{(11)}$ correspondants et en utilisant les remarques précédentes, nous obtenons :

$$c_{11} = \sum_{\sigma \in \mathfrak{S}_{n-1}} \epsilon_\sigma \prod_{1 \leq i \leq n-1} m_{i \sigma(i)}^{(11)} ,$$

ce qui est justement le déterminant Δ_{11} de $M^{(11)}$. Quant au coefficient $(-1)^{i+j}$ il est égal à 1.

Cherchons maintenant le coefficient c_{hk} de m_{hk} pour $(h, k) \neq (1, 1)$. En permutant les lignes 1 et h et les colonnes 1 et k de M nous amenons m_{hk} en position (11) en conservant globalement les lignes et les colonnes de $M^{(hk)}$, mais pas leur ordre.

Etudions l'effet de la permutation des colonnes C_1 et C_k de M sur les colonnes C'_i de M^{hk} . Si $k = 2$, elle est sans effet. Si $k > 2$, C'_1 se place en position $k-1$ car C_1 vient en position k et la première colonne de la matrice principale est supprimée; C'_2 se place en position 1 ... et C'_{k-1} en position $k-2$; les suivantes conservent leur place. Les colonnes de $M^{(hk)}$ ont donc subi la permutation :

$$(1, 2, \dots, k-1) \mapsto (2, 3, \dots, k-1, 1),$$

produit de $k-2$ transpositions, chacune faisant avancer 1 d'un rang.

Si $(hk) = (12)$, $M^{(hk)}$ est inchangée mais la permutation des indices 1 et 2 change le signe de $\det(M)$, de sorte que le coefficient de m_{12} est égal à $-\Delta_{12}$; or $(-1)^{h+k} = -1$.

Si $h = 1$ et $k > 2$, Δ_{12} est multiplié par $(-1)^{k-2}$ et c_{hk} par $(-1)^{k-1}$ à cause de la permutation sur les colonnes de M , ou par $(-1)^{k+1}$, c'est-à-dire par $(-1)^{k+h}$.

Si $(h, k) = (2, 2)$, Δ_{12} est inchangé, comme le déterminant principal, or $(-1)^{k+h} = 1$.

Si $h > 2$ et $k > 2$, Δ_{hk} est multiplié par $(-1)^{h-2}(-1)^{k-2}$, c'est-à-dire par $(-1)^{h+k}$, et le déterminant principal est inchangé (deux permutations, une sur les lignes, une sur les colonnes).

Nous avons donc, quels que soient i et j , $c_{ij} = (-1)^{i+j} \Delta_{ij}$.

Le déterminant de M est la somme de $n!$ produits dans chacun desquels figure un et un seul m_{ij} pris dans chaque colonne (et dans chaque ligne). Choisissons la j -ème colonne. Si nous enlevons à $\det(M)$ les $(n-1)!$ produits ayant m_{1j} en facteur, il en reste $(n-1)((n-1)!)$, dont aucun ne contient m_{1j} . Enlevons ensuite les $(n-1)!$ produits ayant m_{2j} en facteur ... ainsi de suite jusqu'au dernier, m_{nj} , et nous aurons épuisé $\det(M)$. Ceci se traduit par :

$$\det(M) = \sum_i m_{ij} c_{ij} = \sum_i (-1)^{i+j} m_{ij} \Delta_{ij}.$$

□

Proposition 4.5.2. *Le déterminant d'une matrice triangulaire supérieure (respectivement inférieure) est égal au produit de ses termes diagonaux.*

Démonstration. La proposition est triviale en dimension 1. Supposons-la vraie jusqu'à la dimension $n-1$ et soit $T = (t_{ij}) \in \mathfrak{M}_n(\mathbb{K})$ une matrice diagonale supérieure. On a, d'après la proposition précédente, en développant suivant la première colonne :

$$\det(T) = t_{11} \Delta_{11}.$$

Or la matrice T_{11} est diagonale supérieure, appartient à $\mathfrak{M}_{n-1}(\mathbb{K})$, et, d'après l'hypothèse de récurrence, son déterminant est égal au produit des t_{ii} pour $i \geq 2$:

$$\Delta_{11} = t_{22} \dots t_{nn},$$

d'où la conclusion pour la dimension n (en multipliant par t_{11}), ce qui achève la récurrence.

Par transposition, on obtient le résultat dans le cas des matrices triangulaires inférieures. $\square$

4.6 Déterminant d'un endomorphisme

Soit E un $\mathbb{K}$ -espace vectoriel de dimension n et $u \in \mathcal{L}(E)$. On définit le **déterminant** de u comme étant le déterminant de $M_B(u)$ dans une base quelconque $\mathfrak{B}$ de E :

$$\det_{\mathfrak{B}}(u) = \det(M_{\mathfrak{B}}(u)).$$

En effet, dans une autre base $\mathfrak{B}'$, on aura :

$$\begin{aligned} \det_{\mathfrak{B}'}(u) &= \det(M_{\mathfrak{B}'}(u)) \\ &= \det(P M_{\mathfrak{B}} P^{-1}) \\ &= \det(P) \det(M_{\mathfrak{B}}) \det(P^{-1}) \\ &= \det_{\mathfrak{B}}(u), \end{aligned}$$

car $\det(P) \det(P^{-1}) = 1$ (proposition 3.4.2), ce qui montre que la définition est indépendante du choix de la base.

4.7 Déterminant de Vandermonde

Le déterminant de la matrice $A_n \in \mathfrak{M}_n(\mathbb{R})$:

$$A_n = \begin{pmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{pmatrix}$$

dans laquelle les (x_i) sont deux à deux distincts, et que l'on rencontre assez souvent, est dit **de Vandermonde**. Pour $n = 1$, il est égal à $x_1 - x_0$, et pour $n = 2$, à $(x_1 - x_0)(x_2 - x_0)(x_2 - x_1)$. Faisons l'hypothèse de récurrence :

$$\det(A_{p-1}) = \prod_{0 \leq i < j \leq p-1} (x_j - x_i)$$

et calculons $\det(A_p)$. Faisons $C_j - x_0 C_{j-1}$ pour $1 \leq j \leq p$:

$$\det(A_p) = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 1 & x_1 - x_0 & \dots & x_1^{p-1}(x_1 - x_0) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_p - x_0 & \dots & x_p^{p-1}(x_p - x_0) \end{vmatrix}$$

et développons ce déterminant suivant L_1 en remarquant le facteur $x_i - x_0$ sur chaque L_i :

$$\det(A_p) = \prod_{1 \leq i \leq p} (x_i - x_0) \begin{vmatrix} 1 & x_1 & \dots & x_1^{p-1} \\ 1 & x_2 & \dots & x_2^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_p & \dots & x_p^{p-1} \end{vmatrix}.$$

Appliquons l'hypothèse de récurrence en décalant les indices d'une unité :

$$\begin{aligned}\det(A_p) &= \prod_{1 \leq i \leq p} (x_i - x_0) \prod_{1 \leq i \leq j \leq p} (x_j - x_i) \\ &= \prod_{0 \leq i \leq j \leq p} (x_j - x_i)\end{aligned}$$

ce qui achève la récurrence, et montre qu'un déterminant de Vandermonde n'est pas nul et donc que A_n est inversible.

4.8 Vecteur et valeur propres, polynôme caractéristique

Soit $u \in \mathfrak{L}(E)$ ($\dim(E) = n$) de matrice $M = (m_{ij})$ dans une base quelconque $\mathfrak{B} = (\epsilon_i)$. Un **vecteur propre** de u est un élément x non nul de E tel que $u(x) = \lambda x$ pour un certain scalaire λ , qui est alors une **valeur propre** de u .

L'ensemble des valeurs propres de u est son **spectre**, noté $\text{Spec}(u)$.

Un vecteur propre doit être non nul car si $x = 0$ alors $u(x) = 0$ et donc $u(x) = \lambda x$ quel que soit le scalaire λ . Par contre, si $x \neq 0$ et $u(x) = 0$, x est vecteur propre associé à la valeur propre 0.

Deux vecteurs propres engendrant la même droite (proportionnels) sont équivalents, et nous ne considérerons que les classes d'équivalence : si x est un vecteur propre, ax , $a \neq 0$, n'est évidemment pas un autre vecteur propre.

L'ensemble des vecteurs propres associés à une valeur propre λ complété par le vecteur nul est un sous-espace vectoriel de E , l'**espace propre** E_λ .

Le polynôme $P_u(X) = \det(Xi_E - u)$ est le **polynôme caractéristique** de u , ou de sa matrice dans une base quelconque (voir 3.6).

Proposition 4.8.1. *Soit $u \in \mathfrak{L}(E)$. Les assertions suivantes sont équivalentes :*

- (a) $\lambda \in \mathbb{K}$ est valeur propre de u ,
- (b) il existe un vecteur non nul $x \in E$ tel que $u(x) = \lambda x$,
- (c) l'endomorphisme $\lambda i_E - u$ n'est pas injectif,
- (d) $\det(\lambda i_E - u) = 0$,
- (e) λ est racine de $P_u(X)$.

Démonstration. .

- (a) $\Rightarrow$ (b) Découle de la définition d'une valeur propre.
- (b) $\Rightarrow$ (c) Comme $x \in \text{Ker}(\lambda i_E - u)$, $x \neq 0$, $\lambda i_E - u$ n'est pas injectif.
- (c) $\Rightarrow$ (d) Comme $\lambda i_E - u$ n'est pas un isomorphisme $\det(\lambda i_E - u) = 0$.
- (d) $\Rightarrow$ (e) $P_u(\lambda) = \det(\lambda i_E - u) = 0$.
- (e) $\Rightarrow$ (a) Si $P_u(\lambda) = 0$, $\det(\lambda i_E - u) = 0$, $\lambda i_E - u$ n'est pas injectif, il existe un $x \neq 0$ dans E tel que $u(x) = \lambda x$, et $\lambda \in \mathbb{K}$ est valeur propre de u . $\square$

Exemple 4.8.1. Le calcul des éléments propres de la matrice :

$$A = \begin{pmatrix} -7 & 3 \\ -18 & 8 \end{pmatrix}$$

représentant l'endomorphisme u dans la base canonique de $\mathbb{R}^2$ va nous donner une base dans laquelle la matrice de u est diagonale (ce n'est pas toujours possible). Le polynôme caractéristique de A est :

$$P(X) = \begin{vmatrix} X+7 & -3 \\ 18 & X-8 \end{vmatrix} = X^2 - X - 2,$$

et ses valeurs propres sont $\lambda_1 = 2$ et $\lambda_2 = -1$. Un vecteur propre associé à λ_1 est solution du système $(u - \lambda_1)(x) = 0$:

$$\begin{cases} -9x + 3y = 0 \\ -18x + 6y = 0 \end{cases}$$

de rang 1 car de déterminant nul, et qui se réduit donc à l'équation $y = 3x$, d'où le vecteur propre $v_1 = {}^t(1, 3)$. Un calcul identique donne le second vecteur propre $v_2 = {}^t(1, 2)$. Comme $u(v_1) = \lambda_1 v_1$ et $u(v_2) = \lambda_2 v_2$, la matrice de u dans la base (v_1, v_2) est diagonale. Si P est la matrice de passage de la base canonique à la base (v_1, v_2) (voir 2.5), on obtient finalement :

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix} = PDP^{-1}.$$

Une matrice diagonale est évidemment de manipulation beaucoup plus simple qu'une matrice quelconque, d'autant plus que la dimension est grande. ∇

4.9 Trigonalisation

Proposition 4.9.1. *Soit $u \in \mathcal{L}(E)$. Il existe une base dans laquelle la matrice $M(u) = (m_{ij})$ est triangulaire supérieure. Les termes diagonaux de M sont les valeurs propres de u , répétées selon leur ordre.*

Démonstration. La proposition est triviale en dimension 1.

Supposons-la démontrée pour tout espace vectoriel de dimension inférieure ou égale à $n-1$, et soit E un espace vectoriel de dimension n . Soit ϵ_1 un vecteur propre associé à la valeur propre λ_1 , E' un supplémentaire de la droite D engendrée par ϵ_1 et p est la projection de E sur E' parallèlement à D . L'application $u' = p \circ u|_{E'}$ appartient à $\mathcal{L}(E')$ et comme $\dim(E') = n-1$ il existe une base $(\epsilon_i)_{2 \leq i \leq n}$ dans laquelle la matrice de u' est triangulaire supérieure (hypothèse de récurrence).

Or $u(\epsilon_i) = (p \circ u)(\epsilon_i) + a_i \epsilon_1$ pour $2 \leq i \leq n$. Dans la base $(\epsilon_i)_{1 \leq i \leq n}$ la matrice de u est triangulaire supérieure. Le déterminant de $X - M(u)$ étant égal au polynôme caractéristique de u , les termes diagonaux de $M(u)$ sont les valeurs propres de u , chacune apparaissant un nombre de fois égal à son ordre. $\square$

5 Méthodes de résolution des systèmes

5.1 Système de Cramer

Un système $Mx = y_0$, M dans $\mathfrak{M}_n(\mathbb{K})$ et x et y_0 dans $\mathbb{K}^n$, est **compatible** s'il admet au moins une solution $x = (x_i)$. C'est un **système de Cramer** si $p = n$ et si la matrice M est inversible; il admet alors l'unique solution $x = M^{-1}y_0$. Dans ce cas, soit M_i la matrice obtenue en remplaçant la i -ème colonne de M par la colonne des seconds membres. Les formules de Cramer :

$$x_i = \frac{\det(M_i)}{\det(M)}, \quad 1 \leq i \leq n$$

sont élégantes mais inapplicables car nécessitant trop de calculs (de l'ordre de $(n+2)!$). Si $n = 2$, elles sont cependant pratiques pour un calcul à la main.

Supposons le système $Mx = y_0$ compatible et $x = x_0$ une solution. Le **rang** du système est le rang de M . Le système $Mx = 0$ est le **système homogène** associé. Il admet toujours la solution $x = 0$, et elle est unique s'il est de Cramer. Toute solution du système est la somme de x_0 et d'une solution du système homogène associé. L'ensemble des solutions est la **solution générale**, de la forme $x_0 + S$, S étant la solution générale du système homogène associé.

Théorème 5.1.1 (de Rouché-Fontené). *Si le système compatible $Mx = y_0$, $M \in \mathfrak{M}_n(\mathbb{K})$, est de rang $r < n$, on isole r lignes indépendantes et r inconnues principales, on fait passer les autres inconnues (auxiliaires) au second membre, et on obtient un système de Cramer en les r inconnues principales.*

Démonstration. La matrice du système obtenu est régulière par construction. □

Proposition 5.1.1. *En ajoutant à une équation d'un système une combinaison linéaire des autres équations on ne change ni sa compatibilité ni sa solution s'il est compatible. De même si l'on permute les lignes ou les colonnes.*

Démonstration. Soit $(L_i)_{1 \leq i \leq n} = 0$ un système d'équations linéaires.

L'équivalence :

$$\begin{pmatrix} L_i = 0 \\ L_j = 0 \end{pmatrix} \iff \begin{pmatrix} L_i = 0 \\ L_j + \lambda L_i = 0 \end{pmatrix}$$

pour $j \neq i$ et tout scalaire λ est triviale. On passe à $L_j + \sum_{j \neq i} \lambda_i L_i = 0$ par une récurrence immédiate. Ceci prouve le premier point. Une permutation de lignes est évidemment sans effet sur la solution éventuelle (mais pas sur la matrice du système). Une permutation des colonnes C_i et C_j revient à poser $x'_i = x_j$ et $x'_j = x_i$. □

Exemple 5.1.1. En enlevant deux fois la première ligne à la deuxième et quatre fois la première ligne à la troisième on transforme le système :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 4 & 7 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 16 \\ \lambda \end{pmatrix}$$

en celui-ci :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ -4 \\ \lambda - 40 \end{pmatrix}.$$

En soustrayant la deuxième ligne à la troisième on obtient le système :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ -4 \\ \lambda - 36 \end{pmatrix}$$

de rang 2, compatible si $\lambda = 36$. Les inconnues principales sont x et y , d'où :

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 - 3z \\ -4 + 2z \end{pmatrix}$$

et la solution en fonction du paramètre z : $x = 2 + z$, $y = 4 - 2z$. La solution $x = 2$, $y = 4$, $z = 0$ est une solution particulière. Le système homogène associé a pour solution générale $x = z$, $y = -2z$. ∇

5.2 Système de Cramer triangulaire

Considérons le système $Tx = y_0$ dans lequel $T = (t_{ij})$ est une matrice triangulaire supérieure de type (n, n) , dont aucun des termes diagonaux n'est nul. La dernière ligne s'écrit $t_{nn}x_n = y_n$. Comme $t_{nn} \neq 0$, on a $x_n = t_{nn}^{-1}y_n$ en une opération. Reportant cette valeur dans l'avant-dernière ligne :

$$L_{n-1} : t_{n-1, n-1}x_{n-1} + t_{n-1, n}x_n = y_{n-1}$$

on obtient $x_{n-1} = t_{n-1, n-1}^{-1}(y_{n-1} - t_{n-1, n}x_n)$, avec une division, une multiplication et une soustraction. Ainsi de suite en remontant jusqu'à x_1 , en ajoutant à chaque pas une multiplication et une soustraction. La résolution demande donc n divisions, $n(n-1)/2$ multiplications et autant de soustractions, soit $\sim n^2$ opérations.

L'application linéaire t associée à T est un isomorphisme.

Si on suppose nul un certain t_{ii} , les images des i premiers vecteurs de la base source ont pour seules composantes non nulles les $i-1$ premières et engendrent donc un sous-espace vectoriel de dimension au plus égale à $i-1$. Les $n-i$ dernières engendrent un sous-espace vectoriel de dimension au plus égale à $n-i$. On en déduit que $\text{Im}(t)$ est au plus de dimension $n-1$, et donc que t , n'étant pas surjective, n'est pas un isomorphisme.

On a les mêmes résultats avec une matrice triangulaire inférieure, en commençant par la première ligne et en descendant au lieu de monter. Une application linéaire associée à une matrice triangulaire est donc un isomorphisme si et seulement si le produit des termes diagonaux de cette matrice est non nul.

5.3 Méthode de Gauss et de Gauss-Jordan

Cette méthode était déjà connue du mathématicien chinois du troisième siècle Liu Hui. Elle est très simple et très intuitive, mais aussi très économe en calculs ; c'est la plus utilisée. Elle est basée sur la proposition 5.1.1.

Elle utilise les transformations suivantes sur un système $Mx = y_0$ à p lignes et n colonnes :

- $L_i + \lambda L_j$ (ajout de λ fois la j -ème ligne à la i -ème), représentée par la multiplication à gauche de la matrice M par la matrice $I_p + \lambda E_{ji}$ (les E_{hk} sont définis dans la proposition 3.3.1), de déterminant égal à 1 ;
- $L_i \leftrightarrow L_j$ (permutation des lignes L_i et L_j), représentée par la multiplication à gauche de M par la matrice $E_{ij} + E_{ji} + \sum_{i \neq k \neq j} E_{kk}$, de déterminant égal à -1 ;
- $C_i + \lambda C_j$ (ajout de λ fois la j -ème colonne à la i -ème), représentée par la multiplication à droite de M par la matrice $I_n + \lambda E_{ij}$, de déterminant égal à 1 ;
- $C_i \leftrightarrow C_j$ (permutation des colonnes C_i et C_j), représentée par la multiplication à droite de M par la matrice $E_{ij} + E_{ji} + \sum_{i \neq k \neq j} E_{kk}$, de déterminant égal à -1 .

Les multiplications à gauche représentent un changement de base au but, les multiplications à droite représentent un changement de base à la source.

Partons d'un système $Mx = y_0$, $M = (m_{ij})$ et supposons que $p = n$. La transformation $L_i - m_{i1}^{-1} m_{i1} L_1$ pour $i > 1$ annule tous les m_{i1} ($\sim 2n^2$ opérations). On passe à la deuxième colonne : $L_i - m_{i2} m_{22}^{-1} L_2$ pour $i > 2$ annule tous les m_{i2} ($2(n-1)^2$ opérations). Ainsi de suite jusqu'à la dernière colonne : $L_n - m_{n-1}^{-1} m_{n-1} L_{n-1}$ (4 opérations). On obtient un système triangulaire supérieur après $\sim 2n^3/3$ opérations.

Si l'un des m_{ii} est nul, et si l'un des m_{ji} , $j > i$, n'est pas nul, on permute L_i et L_j ; s'ils sont tous nuls et si l'un des m_{ij} , $j > i$, n'est pas nul, on permute C_i et C_j ; s'ils sont tous nuls et si l'un des m_{jk} , $j > i$ et $k > i$, n'est pas nul, on permute L_i avec L_j et C_i avec C_k ; si tous les m_{jk} sont nuls, le processus s'achève, et le système est de rang $i - 1$.

Si le processus va jusqu'à la dernière ligne et si $m_{nn} \neq 0$, on obtient un système triangulaire supérieur régulier. On peut alors appliquer la méthode en remontant. Pour la n -ième colonne, on effectue une opération pour chaque ligne L_i , le calcul de $m_{ni}^{-1} m_{in}$, soit $n - 1$ opérations, jusqu'à une opération, le calcul de $m_{22}^{-1} m_{12}$, pour la deuxième colonne, car il est inutile de vérifier que les termes $m_{i,i+1}$ sont bien annulés. Le calcul des $m_{ii}^{-1} m_{ji}$, $j < i$, est nécessaire pour les seconds membres. Cela fait $\sim n^2/2$ opérations, ce qui est négligeable devant $2n^3/3$.

Si $p > n$, la méthode précédente permet d'isoler, après d'éventuelles permutations, les n premières lignes. On résout alors le système $n \times n$ obtenu, puis on vérifie que sa solution est compatible avec les $p - n$ équations restantes (exemples 3.1.1 et 3.1.4).

Si $p < n$, la méthode permet d'isoler les p premières inconnues. On fait alors passer au second membre les $n - p$ restantes, d'où un système $p \times p$ (exemple 3.2.2).

Si le système est de rang $r < n$, quand la transformation arrive à la colonne C_r les $n - r$ dernières lignes sont nulles et on les supprime, puis on fait passer les $n - r$ inconnues auxiliaires au second membre, avant de remonter, pour obtenir un système diagonal régulier $r \times r$.

Nous verrons les problèmes éventuels (un m_{ii} nul ou proche de 0) un peu plus loin.

Exemple 5.3.1. Considérons le système :

$$\begin{cases} x + 2y + 3z &= 10 \\ x + 3y + 4z &= 13 \\ 3x + 8y + 11z &= \lambda \\ x + y + 2z &= \mu \end{cases}$$

On peut l'écrire :

$$\begin{pmatrix} 1 & 2 & 3 & 10 \\ 1 & 3 & 4 & 13 \\ 3 & 8 & 11 & \lambda \\ 1 & 1 & 2 & \mu \end{pmatrix}$$

La transformation $[L_2 - L_1, L_3 - 3L_1, L_4 - L_1]$ donne :

$$\begin{cases} 1 & 2 & 3 & 10 \\ 0 & 1 & 1 & 3 \\ 0 & 2 & 2 & \lambda - 30 \\ 0 & -1 & -1 & \mu - 10 \end{cases}$$

Enfin $[L_3 - 2L_2, L_4 + L_2]$ mène à :

$$\begin{cases} 1 & 2 & 3 & 10 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & \lambda - 36 \\ 0 & 0 & 0 & \mu + 7 \end{cases}$$

d'où l'on déduit que le rang de M est égal à 2, que λ doit être égal à 36 et μ à -7 . On ne conserve que les équations et inconnues principales :

$$\begin{cases} x + 2y &= 10 - 3z \\ y &= 3 - z, \end{cases}$$

et on remonte, en faisant $L_1 - 2L_2$:

$$\begin{cases} x &= 4 - z \\ y &= 3 - z. \end{cases}$$

▽

En Analyse Numérique, on effectue les calculs en ne conservant qu'un nombre fini de décimales, selon les capacités de l'ordinateur ou la précision des mesures. Si ces erreurs se propagent en conservant le même ordre de grandeur, la méthode utilisée est dite stable. On peut s'attendre, d'après les formules de Cramer, à ce que la méthode de Gauss ne soit pas stable lorsque le déterminant du système est voisin de zéro. Voyons deux exemples.

Remarquons que le résultat dépend de l'ordre dans lequel on effectue les opérations. Ainsi, avec une décimale, on a $0,8/0,9 = 0,8$ et :

$$\begin{aligned} 0,9 \cdot 0,8/0,9 &= 0,8 \\ &= 0,9 \cdot 0,8 \\ &= 0,7. \end{aligned}$$

Exemple 5.3.2. Traitons le système suivant en ne conservant dans les calculs que les deux premières décimales :

$$\begin{cases} x + \frac{1}{2}y + \frac{1}{3}z = 32 \\ \frac{1}{2}x + \frac{1}{3}y + \frac{1}{4}z = 22 \\ \frac{1}{3}x + \frac{1}{4}y + \frac{1}{5}z = 17. \end{cases}$$

Son déterminant est égal à $463 \cdot 10^{-6}$. Il est donc nul si l'on ne conserve que les deux premières décimales.

La transformation $[L_2 - 0,50 L_1 \quad L_3 - 0,33 L_1]$ donne :

$$\begin{cases} x + 0,50 y + 0,33 z = 32 \\ 0,08 y + 0,09 z = 6 \\ 0,09 y + 0,09 z = 6,44 \end{cases}$$

puis $L_3 - \frac{0,09}{0,08} L_2$:

$$\begin{cases} x + 0,50 y + 0,33 z = 32 \\ 0,08 y + 0,09 z = 6 \\ -0,01 z = -0,31 \end{cases}$$

donne un résultat $z = 31$ au lieu de $z = 60$, puis $y = 40,12$ au lieu de 12 et $x = 1,71$ au lieu de 6 . L'ordre de grandeur de l'erreur finale est bien supérieur à celui des erreurs initiales. ∇

Cet exemple montre que la méthode de Gauss est **instable** lorsqu'elle mène à un m_{ii} , un **pivot**, voisin de 0, en ce sens que les erreurs d'arrondi (ou de mesure) se propagent en croissant.

On essaie de remédier à ce défaut, lorsque l'expression décimale conservée du déterminant n'est pas trop petite, par des permutations sur les lignes ou les colonnes pour obtenir un pivot raisonnable. Ceci ne marche pas pour l'exemple précédent (bien choisi pour être récalcitrant). On peut diviser la ligne du pivot

trop petit par ce pivot, mais ce n'est pas recommandé (division par un nombre trop petit).

Exemple 5.3.3. Le déterminant du système :

$$\begin{cases} 3x + y + z &= 8 \\ x + 3y + z &= 10 \\ x + y + 3z &= 12 \end{cases}$$

est égal à 20, et sa solution est $x = 1$, $y = 2$ et $z = 3$. Calculons avec deux décimales.

Les transformations successives $[L_2 - L_1/3, L_3 - L_1/3]$, $[L_3 - 0, 67L_2/2, 67]$, $[L_2 - 0, 67L_3/2, 51, L_1 - L_3/2, 51]$ et $[L_1 - L_2/2, 67]$ mènent à la solution :

$$x = 1,02, \quad y = 2,01, \quad z = 2,99,$$

et l'erreur est de l'ordre de 10^{-2} .

▽

La méthode de Gauss-Jordan est une variante consistant à choisir par permutation le pivot maximum en valeur absolue, à diviser la ligne par la valeur du pivot (qui devient ainsi égal à 1), à annuler les termes au-dessus du pivot et non seulement ceux qui se trouvent au-dessous. Elle transforme une matrice inversible en l'identité.

Voyons les différentes utilisations de la méthode de Gauss :

- résolution d'un système,
- calcul du rang d'une matrice,
- calcul d'un déterminant (voir la section 4),
- calcul de l'inverse d'une matrice carrée régulière.

Le **rang d'une matrice** est le nombre de lignes non nulles obtenues.

Les seules transformations modifiant le déterminant (cas $p = n$) sont les permutations (chacune changeant son signe) et la multiplication d'une ligne par une constante (ce qui multiplie le déterminant par cette constante). Le déterminant est donc le produit des termes diagonaux obtenus, en tenant compte des permutations éventuelles de lignes ou de colonnes et des multiplications de lignes.

Pour calculer l'inverse d'une matrice carrée régulière M , on la ramène à l'identité par des transformations décrites ci-dessus, et les mêmes transformations font passer de l'identité à M^{-1} (exemple 4.3.4).

Ces transformations sont linéaires et sont donc représentées chacune par une matrice. Leur composée est donc une application linéaire.

Exemple 5.3.4. Les transformations $[L_2 - L_1, L_3 - 3L_1, L_4 - L_1]$, et $[L_3 -$

$2L_2, L_4 + L_2]$ effectuée dans l'exemple 4.3.1 sont représentées par les matrices :

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

dont le produit (de droite à gauche) est :

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

Multipliant (à gauche) la matrice initiale par cette dernière on trouve :

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

Les transformations successives $L_1 - 2L_1$ puis $C_3 - (C_2 + C_1)$, de matrices respectives, la première agissant à gauche et la seconde à droite :

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

donnent la forme canonique :

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

et le rang, égal à 2,

▽

Exemple 5.3.5. Calcul de l'inverse de la matrice M par la méthode de Gauss-

Jordan.

$$\begin{array}{ccc}
M & & I_3 \\
\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
& L_2 - L_1, L_3 - L_1 & \\
\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & -2 & 2 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\
& L_3 \leftrightarrow L_2, -L_2/2 & \\
\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & -1/2 \\ -1 & 1 & 0 \end{pmatrix} \\
& L_1 - L_2, L_3/2 & \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} & & \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & -1/2 \\ -1/2 & 1/2 & 0 \end{pmatrix} \\
& L_2 + L_3 & \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & & \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & -1/2 \\ -1/2 & 1/2 & 0 \end{pmatrix}
\end{array}$$

La composée des transformations successives est une application linéaire dont la matrice T est telle que $TM = I_3$, et donc $T = M^{-1}$. ∇

Exemple 5.3.6. Résolution , dans $\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$, du système :

$$\begin{cases} 2x + 3y + z &= 0 \\ x + y + 2z &= 4 \\ 3x + 2y + 3z &= 2 \end{cases}$$

Partant de :

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

effectuons la transformation $L_2 + 2L_1, L_3 + L_1$:

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

puis $L_2 + 4L_3, L_1 + L_3$:

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

puis $L_1 + L_2$:

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

et enfin $3L_1, 3L_2, 4L_3$:

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

d'où $x = 2, y = 1$ et $z = 3$.

▽

Exemple 5.3.7. Calcul, dans $\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$, de l'inverse de la matrice :

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

Partons de :

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 6 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

appliquons $L_2 + 6L_1, L_3 + 5L_1$:

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 5 & 3 & 6 & 1 & 0 \\ 0 & 0 & 5 & 5 & 0 & 1 \end{pmatrix}$$

puis $3L_3$:

$$\begin{pmatrix} 3 & 4 & 5 & 1 & 0 & 0 \\ 0 & 5 & 3 & 6 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 3 \end{pmatrix}$$

puis $L_2 + 4L_3, L_1 + 2L_3$:

$$\begin{pmatrix} 3 & 4 & 0 & 3 & 0 & 6 \\ 0 & 5 & 0 & 3 & 1 & 5 \\ 0 & 0 & 1 & 1 & 0 & 3 \end{pmatrix}$$

puis $3L_2$:

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

puis $L_1 + 3L_2$:

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

enfin $5L_1$:

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

d'où :

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

▽

Exemple 5.3.8. Résolution dans $\mathbb{Z}_4$ du système :

$$\begin{cases} 2x + y + z &= 0 \\ x + 2y + 3z &= 1 \\ x + 3y + z &= 1 \end{cases}$$

Une première transformation $L_1 \leftrightarrow L_2$ donne :

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

puis $L_2 + 2L_1$, $L_3 + 3L_1$:

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

enfin $L_3 + 3L_2$ donne la forme triangulaire :

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

d'où $z = 2$, puis en remontant, $y = 0$ et $x = 3$.

▽

Lorsque les scalaires appartiennent à un anneau non intègre, comme $\mathbb{Z}_4$, si les pivots sont inversibles on a une solution unique. Sinon, on peut en avoir plusieurs (l'équation $2x = 2$ dans $\mathbb{Z}_4$ admet les solutions $x = 1$ et $x = 3$), ou aucune (comme $2x = 1$ dans $\mathbb{Z}_4$).

5.4 Décomposition triangulaire : méthode LU

Soit le système de Cramer $Mx = y_0$. Nous allons voir que l'on peut mettre $M = (m_{ij})$ sous la forme d'un produit LU , dans lequel $L = (l_{ij})$ est triangulaire inférieure (**low** en anglais), avec $l_{ii} = 1$ (donc régulière), associée à une application linéaire l , et $U = (u_{ij})$ est triangulaire supérieure (**up**), associée à

une application linéaire u . Si U n'était pas régulière, u ne serait pas un isomorphisme, et $l \circ u$ non plus : M ne serait pas régulière contrairement à l'hypothèse. Aucun des u_{ii} ne peut donc être nul.

Soient M_i , T_i et U_i les matrices obtenues en prenant les i premiers éléments des i premières lignes de, respectivement, M , L et U pour $i < n$.

Remarquons que $M_i = L_i U_i$, que $\det(L_i) = 1$, et que $\det(U_i) = \det(M_i)$. De $\det(U_i) = u_{ii} \det(U_{i-1})$ nous déduisons que $u_{ii} = \det(M_i) / \det(M_{i-1})$. Les $\det(M_i)$ doivent donc être non nuls.

Pour $i = 2$:

$$L_2 U_2 = \begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} \end{pmatrix} = M_2$$

on a $u_{12} = m_{12}$, puis $l_{21} = m_{21}/u_{11}$ et $u_{22} = m_{22} - l_{21} u_{12}$.

Pour $i = 3$, on pose $u_3 = {}^t(u_{13}, u_{23})$ et $m_3 = {}^t(m_{13}, m_{23})$, d'où :

$$u_3 = L_2^{-1} m_3$$

ce qui donne $u_{13} = m_{13}$, $u_{23} = m_{23} - l_{21} u_{13}$. On en déduit $l_{31} = m_{31}/u_{11}$, $l_{32} = (m_{32} - l_{31} u_{12})/u_{22}$ et enfin $u_{33} = m_{33} - l_{31} u_{13} - l_{32} u_{23}$.

Supposons connues les matrices T_{i-1} et U_{i-1} . Posons $u_i = {}^t(u_{1i}, \dots, u_{i-1i})$ et $m_i = {}^t(m_{1i}, \dots, m_{i-1i})$, pour obtenir :

$$u_i = L_{i-1}^{-1} m_i.$$

On déduit ensuite l_{ij} pour $1 \leq j \leq i-1$ de l'égalité entre m_{jk} et le produit formel de la dernière ligne de L_i avec la j -ème colonne de U_i .

Le système, écrit sous la forme $LUx = y_0$, se résout en deux temps : $Lz = y_0$, puis $Ux = z$.

La méthode est pratique lorsque M est **tridiagonale**, comme celle de l'exemple suivant, car dans ce cas, L et U n'ont que deux diagonales non nulles, ce qui allège considérablement les calculs.

Notons que si $M = LU$, on a $M^{-1} = U^{-1}L^{-1}$.

Exemple 5.4.1. La matrice S_n se rencontre souvent en Analyse Numérique :

$$S_n = \begin{pmatrix} 4 & 1 & 0 & \dots & 0 \\ 1 & 4 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 1 & 4 \end{pmatrix}.$$

Définissons les suites (p_k) , (q_k) , (l_k) et (u_k) :

$$\begin{cases} p_1 = 4, & p_2 = 15, & p_{k+2} = 4p_{k-1} - p_{k-2}, \\ q_1 = 1, & q_{k+1} = p_k, \\ l_k = p_k / q_k, \\ u_{k+1} = -1 / l_k. \end{cases}$$

Il faut montrer qu'aucun p_k n'est nul. Faisons l'hypothèse de récurrence (vérifiée pour p_3) :

$$4^k - 4^{k-2} < p_k < 4^k .$$

On a alors :

$$\begin{cases} 4^{k+1} - 4^{k-1} < 4p_k < 4^{k+1} \\ -4^{k-1} < -p_{k-1} < -4^{k-1} + 4^{k-3} \end{cases}$$

d'où l'on déduit :

$$4^{k+1} - 4^k < 4^{k+1} - 2 \cdot 4^{k-1} < p_{k+1} < 4^{k+1} - 3 \cdot 4^{k-2} < 4^{k+1}$$

ce qui achève la récurrence et montre que $p_k > 15 \cdot 4^{k-1}$.

On a alors :

$$S_n = L_n U_n = \begin{pmatrix} l_1 & 0 & 0 & \dots & 0 \\ -1 & l_2 & 0 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & l_{n-1} & 0 \\ 0 & \dots & 0 & -1 & l_n \end{pmatrix} \begin{pmatrix} 1 & u_2 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & u_n \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} .$$

Remarquons que $L_n^{-1} = (x_{ij})$ avec $x_{ii} = l_i^{-1}$ et $x_{ij} = (l_j \dots l_i)^{-1}$ pour $j < i$, les autres composantes étant nulles, et que $U_n^{-1} = (y_{ij})$ avec $y_{ii} = 1$ et $y_{ij} = (-1)^{i+j} u_i \dots u_j$ pour $j > i$, les autres composantes étant nulles. ∇

Exemple 5.4.2. La décomposition :

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 1 & 2 & 0 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ 0 & b & 1 & 0 \\ 0 & 0 & c & 1 \end{pmatrix} \begin{pmatrix} u_1 & v_1 & 0 & 0 \\ 0 & u_2 & v_2 & 0 \\ 0 & 0 & u_3 & v_3 \\ 0 & 0 & 0 & u_4 \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & v_1 & 0 & 0 \\ au_1 & av_1 + u_2 & v_2 & 0 \\ 0 & bu_2 & bv_2 + u_3 & v_3 \\ 0 & 0 & cu_3 & cv_3 + u_4 \end{pmatrix}$$

donne successivement $u_1 = 1$, $v_1 = 2$, $a = 3$, $u_2 = -5$, $v_2 = 2$, $b = -3/5$, $u_3 = 11/5$, $c = 15/11$, $v_3 = 2$ et $u_4 = -19/11$. ∇

5.5 Méthode de Cholesky

Cholesky était un officier polytechnicien, brillant mathématicien, spécialiste de géodésie et de topographie. C'est dans ce cadre qu'il a créé la méthode attachée à son nom.

Un système $Sx = y_0$ est symétrique si la matrice $S \in \mathfrak{M}_n(\mathbb{R})$ est symétrique. La matrice S est **définie positive** si :

$$\forall x \in \mathbb{R}^n \setminus \{0\}, {}^t x S x > 0.$$

Une telle matrice est régulière car on ne peut avoir $Sx = 0$. Ses valeurs propres, et donc son déterminant, sont positifs, de même que ses termes diagonaux, car $s_{ii} = {}^t e_i S e_i$.

Nous allons ramener la résolution d'un tel système (de Cramer) à celle de deux systèmes triangulaires (de Cramer).

Remarquons qu'un système de Cramer quelconque $Mx = y_0$ est équivalent au système de Cramer symétrique ${}^t M M x = {}^t M y_0$ (${}^t M M$ étant symétrique).

Proposition 5.5.1. *Toute matrice symétrique S définie positive peut se mettre sous la forme d'un produit $T {}^t T$ dans lequel $T = (t_{ij})$ est une matrice triangulaire inférieure réelle régulière (aucun des t_{ii} n'est nul).*

Démonstration. Définissons, comme précédemment, les matrices S_i et T_i en prenant les i premières lignes et les i premières colonnes, respectivement, de S et de T . Les S_i sont définies positives. Soit en effet :

$$\begin{cases} x' = {}^t(x_1, \dots, x_i) \in \mathbb{R}^i \\ x = {}^t(x_1, \dots, x_i, 0, \dots, 0). \end{cases}$$

On a alors :

$$0 < {}^t x S x = {}^t x' S_i x'$$

quel que soit $x' \neq 0$. Les T_i telles que $T_i {}^t T_i = S_i$, une fois construites, sont donc régulières.

Pour $i = 1$, nous obtenons $t_{11}^2 = s_{11}$, d'où t_{11} par extraction d'une racine carrée d'un réel positif. On prend $t_{11} > 0$. Le calcul de $S_2 = T_2 {}^t T_2$:

$$\begin{pmatrix} t_{11} & 0 \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} t_{11} & t_{21} \\ 0 & t_{22} \end{pmatrix} = \begin{pmatrix} t_{11}^2 & t_{11} t_{21} \\ t_{11} t_{21} & t_{21}^2 + t_{22}^2 \end{pmatrix}$$

nécessite 4 opérations supplémentaires. C'est le nombre d'opérations donnant $t_{21} = t_{11}^{-1} s_{12}$ et $t_{22} = (s_{22} - t_{21}^2)^{1/2}$, une division et une racine carrée remplaçant deux multiplications (et une soustraction une addition). On a :

$$0 < \det(S_2) = (\det(T_2))^2 = t_{11}^2 t_{22}^2$$

et donc $t_{22}^2 > 0$. On prend $t_{22} > 0$.

Supposons connue T_{i-1} pour $2 < i - 1 < n$, et calculons la i -ème ligne de S_i :

$$L_i(S_i) = (t_{i1} t_{11}, t_{i1} t_{21} + t_{i2} t_{22}, \dots, t_{i1}^2 + t_{i2}^2 + \dots + t_{ii}^2).$$

Les $i - 1$ premières composantes donnent par identification le système de Cramer :

$$T_{i-1} \begin{pmatrix} t_{i1} \\ \vdots \\ t_{ii-1} \end{pmatrix} = \begin{pmatrix} s_{i1} \\ \vdots \\ s_{ii-1} \end{pmatrix}$$

d'où les t_{ij} (réels) pour $j \leq i - 1$ grâce à i^2 opérations. On calcule ensuite :

$$t_{ii} = \sqrt{s_{ii} - t_{i1}^2 - \cdots - t_{ii-1}^2}.$$

Il reste à voir que t_{ii} est réel. On le prendra alors positif. On a :

$$0 < \det(S_i) = \det^2(T_i) = t_{ii}^2 \det^2(T_{i-1})$$

et t_{ii}^2 est bien un réel positif. $\square$

Le calcul complet demande donc :

$$\sum_{1 \leq i \leq n} i^2 = \frac{2n^3 + 3n^2 + n}{6}$$

soit $\sim n^3/3$ opérations. On termine comme précédemment, en résolvant deux systèmes triangulaires.

Exemple 5.5.1. Reprenons la matrice S_n de l'exemple 5.4.1, et déterminons T_n , triangulaire supérieure, telle que $S_n = T_n {}^t T_n$. On a :

$$T_2 = \begin{pmatrix} \sqrt{15}/2 & 1/2 \\ 0 & 2 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 2\sqrt{210}/15 & 2\sqrt{15}/15 & 0 \\ 0 & \sqrt{15}/2 & 1/2 \\ 0 & 0 & 2 \end{pmatrix},$$

puis :

$$T_k = \begin{pmatrix} a_k & v_k \\ 0_{k-1,1} & T_{k-1} \end{pmatrix}$$

avec $v_k = (b_k, 0, \dots, 0)$, $b_k a_{k-1} = 1$ et $a_k^2 + b_k^2 = 4$. ∇

5.6 Diagonale dominance stricte

Une matrice $M = (m_{ij}) \in \mathfrak{M}_n(\mathbb{K})$ est à **diagonale dominance stricte** si, pour tout i , $|m_{ii}| > \sum_{j \neq i} |m_{ij}|$. Ces matrices se rencontrent fréquemment en Analyse Numérique.

Proposition 5.6.1. *Toute matrice à diagonale dominance stricte est régulière.*

Démonstration. Supposons la matrice $M = (m_{ij}) \in \mathfrak{M}_n(\mathbb{K})$, à diagonale dominance stricte, non régulière. Il existe alors un vecteur non nul $x = (x_i) \in \mathbb{K}^n$ tel que $Mx = 0$, et soit x_k une composante telle que, pour tout $i \neq k$, on ait $|x_k| \geq |x_i|$. La k -ième composante de Mx , $\sum_j m_{kj} x_j$, étant supposée nulle, on doit avoir :

$$m_{kk} x_k = - \sum_{j \neq k} m_{kj} x_j$$

d'où :

$$|m_{kk}| |x_k| \leq \sum_{j \neq k} |m_{kj}| |x_j|,$$

or :

$$|m_{kk}| |x_k| > \sum_{j \neq k} |m_{kj}| |x_k| \geq \sum_{j \neq k} |m_{kj}| |x_j|,$$

ce qui est absurde. $\square$

5.7 Méthode itérative de Jacobi

Considérons le système $Mx = y_0$ dans lequel la matrice $M \in \mathfrak{M}_n(\mathbb{K})$ est à diagonale dominance stricte. Ecrivons $M = D + C$, D étant diagonale et régulière. La méthode de Jacobi consiste à considérer la suite $(x_n)_{n \in \mathbb{N}}$ de vecteurs de $\mathbb{K}^n$ définie, à partir d'un vecteur $x_0 \in \mathbb{K}^n$ donné, par la relation de récurrence :

$$Dx_n = -Cx_{n-1} + y_0$$

ou encore :

$$x_n = -D^{-1}Cx_{n-1} + D^{-1}y_0,$$

la matrice D^{-1} étant diagonale, ses termes diagonaux étant les inverses de ceux de D , de sorte que $D^{-1}D = DD^{-1} = I_N$. Posons $-D^{-1}C = A = (a_{ij})$ et, pour un vecteur quelconque $z = (z_1, \dots, z_n)$, $N(z) = \sup_i (|z_i|)$.

Cette application $N : \mathbb{K}^n \rightarrow \mathbb{K}^*$ est une **norme** car elle vérifie (pour tous z, z_1, z_2 dans $\mathbb{K}^n$ et λ dans $\mathbb{K}$) les axiomes :

$$\begin{aligned} N(z) &= 0 \iff z = 0, \\ N(\lambda z) &= |\lambda| N(z), \\ N(z_1 + z_2) &\leq N(z_1) + N(z_2) \end{aligned}$$

et elle est continue en 0 car on dit que $z \rightarrow 0$ si $N(z)$ tend vers 0.

Si $(x_n)_{n \in \mathbb{N}}$ tend vers une limite x , cette limite vérifie l'équation $Mx = y_0$. Or la diagonale dominance stricte assure la convergence de cette suite.

Nous avons successivement :

$$\begin{aligned} x_1 &= Ax_0 + y_0, \\ x_2 &= A(Ax_0 + y_0) + y_0 = A^2x_0 + Ay_0 + y_0, \\ &\vdots \\ x_k &= A^kx_0 + (\sum_{0 \leq i \leq k-1} A^i)y_0. \end{aligned}$$

Comme $a_{ii} = 0$ et $a_{ij} = m_{ii}^{-1} m_{ij}$ pour $j \neq i$, la i -ème composante de Az est égale à :

$$L_i(A) * z = m_{ii}^{-1} \sum_{j \neq i} m_{ij} z_j$$

et donc :

$$|L_i(A) * z| \leq |m_{ii}^{-1}| \sup_{j \neq i} |z_j| \sum_{j \neq i} |m_{ij}|.$$

Soit $\lambda = \sup_i (|m_{ii}^{-1}| \sum_{j \neq i} |m_{ij}|)$. Par hypothèse $0 < \lambda < 1$, ce qui donne :

$$|L_i(A) * z| \leq \lambda \sup_{j \neq i} |z_j|$$

et :

$$N(Az) = \sup_i |L_i(A) * z| < \lambda \sup_i \sup_{j \neq i} |z_j| \leq \lambda \sup_i |z_i|,$$

c'est-à-dire : $N(Az) < \lambda N(z)$. Nous en concluons que $N(A^k z)$, inférieur à $\lambda^k N(z)$, tend vers 0 quand k tend vers l'infini, que $A^k z$ tend vers 0 quel que soit z , et que A^k tend vers la matrice nulle.

Donc, lorsque k tend vers l'infini, x_k tend vers $(\sum_{0 \leq i} A^i) y_0$. De :

$$(I_n - A)(I_n + A + A^2 + \dots + A^k) = I_n - A^{k+1}$$

nous déduisons que $A_k = \sum_{0 \leq i \leq k} A^i$ tend vers l'inverse de $I_n - A$, d'où la solution approchée :

$$x = (I_n - A)^{-1} y_0.$$

L'erreur commise à l'ordre k est :

$$e(k) = A^{k+1} \left(\sum_{0 \leq i} A^i \right) y_0 = A^{k+1} x.$$

Elle est proportionnelle à $\lambda^{-(k+1)}$. On peut donc approcher la solution d'aussi près qu'on le désire.

La méthode demande $\sim 2kn^3$ opérations. Le k choisi ne dépend que de λ , et la méthode est intéressante lorsque n est grand et que M a beaucoup de termes nuls. On peut l'améliorer en remarquant que :

$$(I_n - A)^{-1} = (I_n + A + A^2 + \dots + A^{k-1})(I_n + A^k + A^{2k} + A^{3k} + \dots).$$

Exemple 5.7.1. La matrice **tridiagonale** symétrique dans $\mathfrak{M}_n(\mathbb{R})$, $n \geq 2$, $S(n) = (s(n)_{ij})$, dont la diagonale principale est constituée de 4, les diagonales qui l'encadrent, de -1, les autres termes étant nuls, est un cas particulier intéressant. Etudions $S(4)$:

$$S(4) = \begin{pmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{pmatrix} = 4(I_4 - A), \quad A = \frac{1}{4} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Calculons les puissances 2 et 3 de A :

$$A^2 = \frac{1}{16} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad A^3 = \frac{1}{64} \begin{pmatrix} 0 & 2 & 0 & 1 \\ 2 & 0 & 3 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 0 & 2 & 0 \end{pmatrix}.$$

Supposons que :

$$A^{2k} = \frac{1}{4^{2k}} \begin{pmatrix} a_{2k} & 0 & b_{2k} & 0 \\ 0 & a_{2k} + b_{2k} & 0 & b_{2k} \\ b_{2k} & 0 & a_{2k} + b_{2k} & 0 \\ 0 & b_{2k} & 0 & a_{2k} \end{pmatrix},$$

et que :

$$A^{2k+1} = \frac{1}{4^{2k+1}} \begin{pmatrix} 0 & a_{2k+1} & 0 & b_{2k+1} \\ a_{2k+1} & 0 & a_{2k+1} + b_{2k+1} & 0 \\ 0 & a_{2k+1} + b_{2k+1} & 0 & a_{2k+1} \\ b_{2k+1} & 0 & a_{2k+1} & 0 \end{pmatrix}.$$

Multipliant ces matrices par A , nous voyons par identification que :

$$\begin{aligned} a_{2k+1} &= a_{2k} + b_{2k}, & b_{2k+1} &= b_{2k}, \\ a_{2k+2} &= a_{2k} + b_{2k}, & b_{2k+2} &= a_{2k} + 2b_{2k}, \\ a_{2k+3} &= 2a_{2k} + 3b_{2k}, & b_{2k+3} &= a_{2k} + 2b_{2k}, \end{aligned}$$

ce qui permet de calculer les termes des deux suites à partir de $a_2 = b_2 = 1$. ∇

5.8 Méthode des moindres carrés

Si des paramètres sont déterminés par des mesures et si l'on a plus de mesures, entachées d'erreur, que de paramètres, on peut obtenir un système sans solution :

$$\sum a_{ij} x_j = y_i, \quad 1 \leq i \leq p, \quad 1 \leq j \leq n, \quad p > n.$$

On va alors minimiser l'expression $\sum_i (\sum_j a_{ij} x_j - y_i)^2$, à défaut de la rendre nulle. Il suffit pour cela d'annuler les dérivées partielles par rapport aux x_j , ce qui donne p équations à autant d'inconnues, admettant une solution puisqu'une fonction à n variables, positive, dérivable, à dérivée continue, tendant vers l'infini à l'infini, admet un minimum sur $\mathbb{R}^n$.

Exemple 5.8.1. Le système :

$$\begin{cases} 2x + 3y &= 4,5 \\ x + 2y &= 3,3 \\ x + y &= 1 \end{cases}$$

n'admet pas de solution. La quantité :

$$\begin{aligned} q(x, y) &= (2x + 3y - 4,5)^2 + (x + 2y - 3,3)^2 + (x + y - 1)^2 \\ &= 6x^2 + 14y^2 + 18xy - 26,6x - 42,2y + 32,14 \end{aligned}$$

a pour dérivées partielles :

$$\begin{aligned} q'_x &= 12x + 18y - 26,6, \\ q'_y &= 18x + 28y - 42,2, \end{aligned}$$

système ($q'_x = q'_y = 0$) dont la solution est $x = -1,233\dots$ et $y = 2,3$. Ces valeurs, reportées dans les premiers membres du système initial, donnent $4,433\dots$, $3,366\dots$ et $1,066\dots$, d'où une erreur, pour chaque équation, égale à $66\dots 10^{-3}$ en valeur absolue. ∇

6 Espace euclidien, hermitien

6.1 Produit scalaire et norme

Un **espace euclidien** E est un $\mathbb{R}$ -espace vectoriel de dimension finie muni d'un **produit scalaire**, c'est-à-dire d'une forme bilinéaire :

$$\begin{aligned} E \times E &\rightarrow \mathbb{R} \\ (x, y) &\mapsto x \cdot y \end{aligned}$$

symétrique ($x \cdot y = y \cdot x$), définie ($x \cdot x = 0 \iff x = 0$), positive ($x \cdot x \geq 0$).

Tout sous-espace d'un espace vectoriel euclidien est évidemment euclidien.

Dans la suite, E est un espace euclidien de dimension n et son produit scalaire est noté $x \cdot y$.

Proposition 6.1.1. *On désigne par x, x' et y des éléments de E . Si $x \cdot y = 0$ quel que soit y , alors $x = 0$. Si $x \cdot y = x' \cdot y$ quel que soit y , alors $x = x'$.*

Démonstration. Faisant $y = x$, on obtient $x \cdot x = 0$, d'où $x = 0$. De $x \cdot y = x' \cdot y$ quel que soit y , on déduit que $(x - x') \cdot y = 0$ et que $x = x'$ d'après le premier point. $\square$

Deux vecteurs de E sont **orthogonaux** si leur produit scalaire est nul.

L'application $x \mapsto (x \cdot x)^{1/2}$ de E dans $\mathbb{R}_+$ est une norme, notée $\|x\|$. En effet, les axiomes définissant une norme :

$$\begin{aligned} \|x\| = 0 &\iff x = 0, \\ \|x + y\| &\leq \|x\| + \|y\|, \\ \|\lambda x\| &= |\lambda| \|x\|, \end{aligned}$$

quels que soient les vecteurs x et y et le scalaire λ , sont vérifiés. Seul le deuxième, l'inégalité triangulaire, n'est pas évident ; prouvons-le :

$$\begin{aligned} \|z - x\|^2 &= (z - y + y - x) \cdot (z - y + y - x), \\ &= \|z - y\|^2 + \|y - x\|^2 + 2(z - y) \cdot (y - x), \\ &\leq \|z - y\|^2 + \|y - x\|^2 + 2\|z - y\| \|y - x\|, \\ &\leq (\|z - y\| + \|y - x\|)^2, \end{aligned}$$

d'où le résultat en prenant les racines carrées de ces quantités positives.

Théorème 6.1.1 (de Pythagore (sixième siècle avant J-C)). *Les vecteurs x et y de H sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Démonstration. Comme :

$$\|x + y\|^2 = (x + y) \cdot (x + y) = x \cdot x + y \cdot y + x \cdot y + y \cdot x = \|x\|^2 + \|y\|^2 + 2x \cdot y$$

les assertions $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ et $x \cdot y = 0$ sont équivalentes. $\square$

Cette propriété était connue des Babyloniens un millénaire avant Pythagore. La première démonstration écrite qui nous soit parvenue figure dans le Livre I des Elements d'Euclide (troisième siècle avant J-C), propositions 47 (Dans un triangle rectangle ...) et 48 (réciproque). Mais il pourrait avoir transcrit une démonstration due à Pythagore.

Si les vecteurs $x_1, \dots, x_k$ sont deux à deux orthogonaux, on a :

$$\|x_1 + \dots + x_k\|^2 = \|x_1\|^2 + \dots + \|x_k\|^2$$

mais la réciproque est fautive. Si en effet v_1, v_2 et v_3 sont deux à deux orthogonaux et de norme 1, et si $x = v_1, y = v_2$ et $z = v_1 - v_2 + v_3$, on a :

$$\|z\|^2 = \|v_1\|^2 + \|v_2\|^2 + \|v_3\|^2$$

d'où :

$$\|x\|^2 + \|y\|^2 + \|z\|^2 = 2\|v_1\|^2 + 2\|v_2\|^2 + \|v_3\|^2 = 5$$

et :

$$\|x + y + z\|^2 = \|2v_1 + v_3\|^2 = 4\|v_1\|^2 + \|v_3\|^2 = 5$$

alors que que z n'est orthogonal ni à x ni à y ($z \cdot x = z \cdot y = 1$).

Le produit scalaire est la **forme polaire** de la norme euclidienne.

Proposition 6.1.2. *Le produit scalaire s'exprime en fonction de la norme euclidienne grâce à l'identité de polarisation :*

$$x \cdot y = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

Démonstration. On a en effet :

$$\begin{cases} \|x + y\|^2 = (x + y) \cdot (x + y) = \|x\|^2 + \|y\|^2 + 2x \cdot y \\ \|x - y\|^2 = (x - y) \cdot (x - y) = \|x\|^2 + \|y\|^2 - 2x \cdot y \end{cases}$$

d'où le résultat. □

On déduit des relations précédentes l'égalité du parallélogramme :

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

Normer un vecteur non nul est le diviser par sa norme ; le résultat est alors un vecteur **normé** ou **unitaire**, c'est-à-dire de norme 1.

Théorème 6.1.2 (Inégalité de Cauchy-Schwarz²). *La valeur absolue du produit scalaire de deux vecteurs x et y de E est au plus égale au produit de leurs normes. Il y a égalité si et seulement s'ils sont proportionnels.*

2. Karl Hermann Amandus Schwarz, mathématicien allemand (1843-1921)

Démonstration. Quel que soit le scalaire a , on a $\|ax + y\| \geq 0$, c'est-à-dire :

$$0 \leq \|ax + y\|^2 = (ax + y) \cdot (ax + y) = a^2\|x\|^2 + 2a x \cdot y + \|y\|^2.$$

Ce polynôme du deuxième degré en a est positif ou nul quel que soit a ; son discriminant réduit $\delta' = (x \cdot y)^2 - \|x\|^2\|y\|^2$ est donc négatif ou nul, d'où :

$$|x \cdot y| \leq \|x\| \|y\|$$

en prenant les racines carrées. L'égalité $|x \cdot y| = \|x\| \|y\|$ équivaut à l'existence d'une racine double, donc à $\|x + ay\| = 0$ et à $x = -ay$. $\square$

6.2 Base orthonormée

Une famille $(v_i)_{1 \leq i \leq m}$, $m \leq n$, de vecteurs de l'espace euclidien E est **orthonormée** (ou orthonormale) si ses éléments sont normés et deux à deux orthogonaux, c'est-à-dire si $v_i \cdot v_j = \delta_{ij}$ quels que soient les indices i et j ($1 \leq i, j \leq m$).

Proposition 6.2.1. *Une famille orthonormée de E est libre. Si elle a n vecteurs, c'est une base.*

Démonstration. Si une famille orthonormée $(v_i)_{1 \leq i \leq m}$ ($2 \leq m \leq n$) était liée, quitte à changer l'ordre des vecteurs, on aurait $v_m = \sum_{1 \leq i \leq m-1} \lambda_i v_i$; comme par hypothèse $v_m \cdot v_i = 0$ pour $1 \leq i \leq m-1$ et que $v_m \cdot v_i = \lambda_i$, les λ_i seraient tous nuls et v_m serait nul, ce qui est absurde puisqu'il est normé.

Si $m = n$, la famille est une base. $\square$

Une **base orthonormée** de l'espace euclidien E est une famille orthonormée de n vecteurs.

Proposition 6.2.2. *Dans une base orthonormée (v_i) le produit scalaire de deux vecteurs $x = (x_i)$ et $y = (y_i)$ est égal à $\sum_i x_i y_i$.*

Démonstration. Calculons le produit scalaire :

$$x \cdot y = \sum_i x_i v_i \cdot \sum_j y_j v_j = \sum_{ij} x_i y_j v_i \cdot v_j = \sum_{ij} x_i y_j \delta_{ij} = \sum_i x_i y_i.$$

$\square$

Corollaire Le $\mathbb{R}$ -espace vectoriel $\mathbb{R}^n$, $n \geq 1$ est muni de sa base canonique $\mathfrak{B} = (e_i)_{1 \leq i \leq n}$. La forme bilinéaire symétrique définie par $e_i \cdot e_j = \delta_{ij}$ est un produit scalaire. Muni de ce produit, $\mathbb{R}^n$ est euclidien, et la base $\mathfrak{B}$ est orthonormée. $\square$

Le produit scalaire précédent est **canonique** et la norme associée est la **norme euclidienne** de $\mathbb{R}^n$. Une base orthonormée (et ordonnée) de $\mathbb{R}^n$ est **directe** si le déterminant de ses éléments dans la base canonique est positif. Sinon elle est **indirecte**.

Théorème 6.2.1 (Procédé d'orthonormalisation de Gram-Schmidt). *L'espace euclidien E admet des bases orthonormées.*

Démonstration. Soit (ϵ'_i) une base quelconque de E . Le vecteur ϵ_1 est obtenu en normant ϵ'_1 . Le vecteur $\epsilon''_2 = \epsilon'_2 - (\epsilon'_2 \cdot \epsilon_1) \epsilon_1$ est orthogonal à ϵ_1 :

$$\epsilon''_2 \cdot \epsilon_1 = (\epsilon'_2 - (\epsilon'_2 \cdot \epsilon_1) \epsilon_1) \cdot \epsilon_1 = \epsilon'_2 \cdot \epsilon_1 - \|\epsilon_1\|^2 \epsilon'_2 \cdot \epsilon_1 = 0$$

et non nul car ϵ'_2 n'est pas proportionnel à ϵ'_1 ; on obtient ϵ_2 en le normant.

Supposons ainsi construits les ϵ_i pour $i \leq h$, normés et deux à deux orthogonaux. Si $h < n$, posons :

$$\epsilon''_{h+1} = \epsilon'_{h+1} - \sum_{1 \leq i \leq h} (\epsilon'_{h+1} \cdot \epsilon_i) \epsilon_i ;$$

ϵ''_{h+1} est orthogonal aux ϵ_i pour $1 \leq i \leq h$, non nul car ϵ'_{h+1} n'appartient pas au sous-espace engendré par les ϵ_i pour $i \leq h$ qui est égal au sous-espace engendré par les ϵ'_i pour $i \leq h$ (il est inclus et a même dimension), et on obtient ϵ_{h+1} en le normant. Si $h = n$, les ϵ_i sont n vecteurs libres de E et forment donc une base orthonormée. $\square$

Corollaire. La matrice de passage de la base (ϵ'_i) à la base (ϵ_i) est triangulaire supérieure puisque ϵ_j appartient à $\text{Vect}(\epsilon'_1, \dots, \epsilon'_j)$. De même pour la matrice de passage réciproque (proposition 2.5.8). $\square$

Théorème 6.2.2. *L'espace euclidien E est son propre dual.*

Démonstration. Soit $x \in E$; l'application :

$$\begin{aligned} x^* : E &\rightarrow \mathbb{K} \\ y &\mapsto x \cdot y \end{aligned}$$

est une forme linéaire telle que $x^*(x) = \|x\|^2$. Elle est donc nulle si et seulement si $x = 0$, et l'application $\phi : E \rightarrow E^*$ définie par $\phi(x) = x^*$ est injective.

Elle est linéaire car, quels que soient les scalaires a et b et les vecteurs x_1 et x_2 , on a :

$$\begin{aligned} (ax_1 + bx_2)^*(y) &= (ax_1 + bx_2) \cdot y \\ &= ax_1 \cdot y + bx_2 \cdot y \\ &= ax_1^*(y) + bx_2^*(y), \end{aligned}$$

c'est-à-dire : $\phi(ax_1 + bx_2) = a\phi(x_1) + b\phi(x_2)$. Les dimensions de E et de son dual étant égales (théorème 2.7.1), ϕ est un isomorphisme (proposition 2.3.6) canonique, car il provient de la structure euclidienne de l'espace, qui identifie E et son dual. $\square$

Proposition 6.2.3. *Une base orthonormée de E est sa propre base duale.*

Démonstration. La base duale d'une base orthonormée (ϵ_i) de E est une base (ϵ'_i) de E puisque $E^* = E$ (théorème 5.2.2) telle que $\epsilon'_j \cdot \epsilon_i = \delta_{ij}$ quels que soient les indices i et j . On a donc $\epsilon'_j \cdot \epsilon_i = \epsilon_j \cdot \epsilon_i$, $(\epsilon'_j - \epsilon_j) \cdot \epsilon_i = 0$ et $\epsilon'_j = \epsilon_j$ (proposition 5.1.1). $\square$

Soient $u \in \mathfrak{L}(E)$ et $x \in E$. La forme linéaire :

$$\begin{aligned} E &\rightarrow \mathbb{K} \\ y &\mapsto x \cdot u(y) \end{aligned}$$

est représentée par un vecteur $x' \in E$ (théorème 5.2.2). L'application $x \mapsto x'$ est notée u^* de sorte que :

$$x \cdot u(y) = x' \cdot y = u^*(x) \cdot y.$$

Cette application u^* est linéaire car, quels que soient les éléments et les scalaires écrits, le calcul :

$$\begin{aligned} u^*(ax + a'x') \cdot y &= (ax + a'x') \cdot u(y) \\ &= ax \cdot u(y) + a'x' \cdot u(y) \\ &= au^*(x) \cdot y + a'u^*(x') \cdot y \\ &= (au^*(x) + a'u^*(x')) \cdot y \end{aligned}$$

montre que $u^*(ax + a'x') = au^*(x) + a'u^*(x')$. Cet endomorphisme u^* est l'**adjoint** de u .

Théorème 6.2.3. *L'adjoint d'un endomorphisme u d'un espace euclidien E est égal à son transposé : $u^* = {}^t u$.*

Démonstration. Rappelons la définition du transposé de $u \in \mathfrak{L}(E)$: si ϕ est une forme linéaire sur E , ${}^t u \circ \phi = \phi \circ u$. Maintenant, ϕ est remplacée par un quelconque $x \in E : \phi(y) = x \cdot y = x^*(y)$. On a donc ${}^t u \circ x^* = x^* \circ u$, et, quel que soit $y \in E$:

$$({}^t u \circ x^*)(y) = (x^* \circ u)(y) = x \cdot u(y) = u^*(x) \cdot y = (u^* \circ x^*)(y),$$

c'est-à-dire $u^* = {}^t u$. □

Corollaire. Si u et v sont des endomorphismes d'un espace euclidien, on a $(u^*)^* = u$ et $(v \circ u)^* = v^* \circ u^*$ (proposition 2.8.2). Ces résultats sont d'ailleurs faciles à montrer directement. □

Proposition 6.2.4. *Dans la situation précédente, on a, dans une base ortho-normée, $M(u^*) = {}^t M(u)$.*

Démonstration. Cette égalité découle des propositions 2.7.1 et 5.2.3 :

$$M(u^*) = M({}^t u) = {}^t M(u).$$

□

Exemple 6.2.1. Soit E le $\mathbb{R}$ -espace vectoriel des polynômes de degré inférieur ou égal à 1 à coefficients réels muni du produit scalaire $f \cdot g = \int_0^1 f(t)g(t) dt$, et $u : E \rightarrow E$ l'opérateur de dérivation $u(f) = f'$.

Calculons d'abord dans la base canonique $(1, X)$ de E , qui n'est pas orthonormée :

$$M(u) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Soit $g = cX + d$ un élément quelconque. Si $u^*(1) = pX + q$, on a :

$$\begin{aligned} 1 \cdot u(g) &= \int_0^1 c \, dt = c, \\ u^*(1) \cdot g &= \int_0^1 (pt + q)(ct + d) \, dt \\ &= c\left(\frac{p}{3} + \frac{q}{2}\right) + d\left(\frac{p}{2} + q\right), \end{aligned}$$

d'où, en égalant $1 \cdot u(g)$ et $u^*(1) \cdot g$ quels que soient c et d , $p = 12$, $q = -6$.

Si $u^*(X) = rX + s$ on a :

$$\begin{aligned} X \cdot u(g) &= \int_0^1 ct \, dt = \frac{c}{2}, \\ u^*(X) \cdot g &= \int_0^1 (rt + s)(ct + d) \, dt \\ &= c\left(\frac{r}{3} + \frac{s}{2}\right) + d\left(\frac{r}{2} + s\right), \end{aligned}$$

d'où, en égalant $X \cdot u(g)$ et $u^*(X) \cdot g$, $r = 6$, $s = -3$. On a donc :

$$M(u^*) = \begin{pmatrix} -6 & -3 \\ 12 & 6 \end{pmatrix}$$

Calculons maintenant dans une base orthonormée. Commençons par en construire une. Comme $1 \cdot 1 = 1$, prenons $e_1 = 1$, et posons $e_2 = aX + b$. On a alors :

$$\begin{aligned} e_1 \cdot e_2 &= \int_0^1 (at + b) \, dt = \frac{a}{2} + b = 0 \\ e_2 \cdot e_2 &= \int_0^1 (at + b)^2 \, dt = \frac{a^2}{3} + ab + b^2 = 1 \end{aligned}$$

et $e_2 = \sqrt{3}(2X - 1)$ convient. Dans cette base, la matrice de u s'écrit :

$$M(u) = \begin{pmatrix} 0 & 2\sqrt{3} \\ 0 & 0 \end{pmatrix}$$

Soit $g = ce_1 + de_2 = 2d\sqrt{3}X + c - d\sqrt{3}$ et $u(g) = 2d\sqrt{3}e_1$. Avec $u^*(e_1) = pe_1 + qe_2$ on a :

$$\begin{aligned} e_1 \cdot u(g) &= e_1 \cdot 2d\sqrt{3}e_1 = 2d\sqrt{3} \\ u^*(e_1) \cdot g &= (pe_1 + qe_2) \cdot (ce_1 + de_2) = pc + qd \end{aligned}$$

d'où en identifiant (quels que soient c et d) $p = 0$ et $d = 2\sqrt{3}$. La matrice de u^* est donc :

$$M(u^*) = \begin{pmatrix} 0 & 0 \\ 2\sqrt{3} & 0 \end{pmatrix}.$$

C'est bien la transposée de $M(u)$.

▽

6.3 Orthogonalité

Deux sous-espaces F_1 et F_2 d'un espace euclidien E sont **orthogonaux**, ce qui se note $F_1 \perp F_2$, si tout élément de l'un est orthogonal à tout élément de l'autre. Il suffit (par bilinéarité) que les éléments d'une base de l'un soient orthogonaux aux éléments d'une base de l'autre.

Proposition 6.3.1. *Soit F un sous-espace de E de dimension $m < n$. Il existe des bases orthonormées de E dont les m premiers éléments forment une base orthonormée de F .*

Démonstration. Le sous-espace F étant euclidien, soit $(\epsilon_1, \dots, \epsilon_m)$ une base orthonormée de F . Complétons cette base par $(\epsilon'_{m+1}, \dots, \epsilon'_n)$ en une base de E , et appliquons le procédé d'orthonormalisation de Gram-Schmidt à partir de ϵ'_{m+1} , pour obtenir une base orthonormée de E dont les m premiers éléments forment une base orthonormée de F . $\square$

L'**orthogonal**, ou supplémentaire orthogonal, de F , noté $F^\perp$ est l'ensemble de tous les vecteurs de E orthogonaux à tous les vecteurs de F . Dans la situation de la proposition précédente, les vecteurs $(\epsilon_{m+1}, \dots, \epsilon_n)$ en constituent une base orthonormée :

$$F^\perp = \text{Vect}(\epsilon_{m+1}, \dots, \epsilon_n).$$

On a évidemment $E = F \oplus F^\perp$ et $(F^\perp)^\perp = F$.

Un **projecteur orthogonal** est un projecteur dont l'image et le noyau sont orthogonaux. Le projecteur orthogonal p_F sur un sous-espace F , de base orthonormée $(\epsilon_i)_{1 \leq i \leq m}$, est défini par :

$$p_F(x) = \sum_{1 \leq i \leq m} (x \cdot \epsilon_i) \epsilon_i.$$

On a en effet, en posant $x \cdot \epsilon_i = x_i$:

$$\begin{aligned} p_F^2(x) &= \sum_i ((\sum_j x_j \epsilon_j) \cdot \epsilon_i) \epsilon_i, \\ &= \sum_i \|\epsilon_i\|^2 x_i \cdot \epsilon_i \text{ puisque } \epsilon_i \cdot \epsilon_j = \delta_{ij}, \\ &= \sum_i x_i \cdot \epsilon_i \text{ puisque } \|\epsilon_i\| = 1, \\ &= p_F(x), \end{aligned}$$

ce qui montre que l'on a bien défini un projecteur. Si $x \in F$, alors $p_F(x) = x$.

Montrons que son image et son noyau sont orthogonaux.

Si $x \in \text{Ker}(p_F)$ on a $x \cdot \epsilon_i = 0$ pour tout i . Si $y \in \text{Im}(p_F) = F$, $y = \sum_i y_i \epsilon_i$, et $x \cdot y = \sum_i y_i x \cdot \epsilon_i = 0$.

On a donc $\text{Ker}(p_F) = F^\perp$; p_F projette sur F parallèlement à $F^\perp$; $i_E - p_F$ est aussi un projecteur orthogonal, de noyau F et d'image $F^\perp$, donc égal à $p_{F^\perp}$. On a de plus $p_{F^\perp} = p_F^*$ car, si $x = x_1 + x_2$ et $y = y_1 + y_2$, avec x_1 et y_1 dans F et x_2 et y_2 dans $F^\perp$ on a :

$$x \cdot p_F(y) = x \cdot y_2 = x_2 \cdot y_2 = x_2 \cdot y = p_{F^\perp}(x) \cdot y.$$

6.4 Isométrie

Une **isométrie** d'un espace euclidien E est un endomorphisme qui conserve la norme ou, ce qui est équivalent (proposition 5.1.2) le produit scalaire.

Théorème 6.4.1. *Soit $u \in \mathcal{L}(\mathbb{R}^3)$. Les assertions suivantes sont équivalentes :*

- (a) *u est une isométrie,*
- (b) *$u \circ u^* = u^* \circ u = i_E$,*
- (c) *$u^{-1} = u^*$.*

Une isométrie est donc un automorphisme.

Démonstration. (a) $\Rightarrow$ (b). Si u est une isométrie, on a quels que soient les vecteurs x et y :

$$x \cdot y = u(x) \cdot u(y) = x \cdot (u^* \circ u)(y),$$

d'où $x \cdot (y - (u^* \circ u)(y)) = 0$. D'après la proposition 5.1.1, y est égal à $(u^* \circ u)(y)$, et $u^* \circ u$ est égal à l'identité, de même que $u \circ u^* = (u^* \circ u)^*$, et u est un automorphisme.

(b) $\Rightarrow$ (a). Si $u u^*$ est égal à l'identité, $x \cdot y = x \cdot (u u^*)(y) = u(x) \cdot u(y)$ quels que soient x et y , et u est une isométrie.

(b) $\Rightarrow$ (c) $\Rightarrow$ (b). Trivial. $\square$

Proposition 6.4.1. *Si u est une isométrie d'un espace euclidien E , u^{-1} est une isométrie. Le produit de deux isométries de E est une isométrie. L'ensemble des isométries de E est un groupe (non commutatif).*

Démonstration. Tout y admet un antécédent $x = u^{-1}(y)$. Comme $y = u(x)$, x et y ont même norme : $\|u^{-1}(y)\| = \|x\| = \|y\|$, et u^{-1} est une isométrie.

Si u et v sont des isométries, $\|(v \circ u)(x)\| = \|u(x)\| = \|x\|$, et $v \circ u$ est une isométrie.

L'ensemble des isométries de E (non vide car il contient au moins l'identité) est donc un groupe dont l'élément neutre est l'identité i_E . $\square$

Le groupe des isométries de E est le **groupe orthogonal** de E , $O(E)$. Si $E = \mathbb{R}^n$, on le note $O(n)$.

Proposition 6.4.2. *L'image d'une base orthonormée par une isométrie est une base orthonormée. Réciproquement, si $\mathfrak{B} = (\epsilon_i)$ et $\mathfrak{B}' = (\epsilon'_i)$ sont deux bases orthonormées, l'endomorphisme de E défini par $u(\epsilon_i) = \epsilon'_i$ pour i allant de 1 à n est une isométrie.*

Démonstration. Soient (ϵ_i) une base orthonormée et u une isométrie de E . Posons $u(\epsilon_i) = \epsilon'_i$. Comme u conserve le produit scalaire, on a $\epsilon'_i \cdot \epsilon'_j = \epsilon_i \cdot \epsilon_j = \delta_{ij}$, et la base (ϵ'_i) est orthonormée.

Démontrons la réciproque. Quels que soient i et j on a d'une part :

$$u(\epsilon_i) \cdot u(\epsilon_j) = ({}^t u \circ u)(\epsilon_i) \cdot \epsilon_j$$

et d'autre part :

$$\begin{aligned} u(\epsilon_i) \cdot u(\epsilon_j) &= \epsilon'_i \cdot \epsilon'_j \\ &= \delta_{ij} \\ &= \epsilon_i \cdot \epsilon_j \end{aligned}$$

et donc $(({}^t u \circ u)(\epsilon_i)) \cdot \epsilon_j = 0$ (quel que soit j), d'où l'on déduit (proposition 5.1.1) que $({}^t u \circ u)(\epsilon_i) - \epsilon_i = 0$ (quel que soit i) et que ${}^t u \circ u = i_E$. $\square$

Une isométrie de $\mathbb{R}^n$ de déterminant positif dans la base canonique est une **rotation** et elle transforme la base canonique en une base **directe**. Si son déterminant est négatif c'est une **réflexion** et elle transforme la base canonique en une base **indirecte**.

Proposition 6.4.3. *Les valeurs propres d'une isométrie de $\mathbb{R}^n$ sont de module égal à 1 et son déterminant est égal à ± 1 .*

Démonstration. Soit u une isométrie de E , λ une valeur propre réelle et x un vecteur propre associé. On a donc $u(x) = \lambda x$ et $\|u(x)\| = \|x\|$, d'où :

$$\begin{aligned} \|u(x)\| &= \|x\| \\ &= \|\lambda x\| \\ &= |\lambda| \|x\| \end{aligned}$$

et la conclusion : $\lambda = \pm 1$.

Si λ est une valeur propre complexe, il existe dans le **complexifié** (voir 6.8) un vecteur propre associé, et le module de λ est égal à 1.

Le déterminant, produit des valeurs propres, est donc de module 1. Etant réel, il est égal à ± 1 . $\square$

Proposition 6.4.4. *Le déterminant de n vecteurs de $\mathbb{R}^n$ est indépendant, au signe près, du choix de la base orthonormée dans laquelle on le calcule.*

Démonstration. Soient $\mathfrak{B} = (\epsilon_i)$ et $\mathfrak{B}' = (\epsilon'_i)$ deux bases orthonormées de $\mathbb{R}^n$. L'endomorphisme u défini par $u(\epsilon'_i) = \epsilon_i$ pour tout i est une isométrie (proposition 6.4.2). Pour une famille quelconque (v_i) de n vecteurs de $\mathbb{R}^n$ on a d'après la proposition 4.3.1 :

$$\begin{aligned} \det_{\mathfrak{B}'}((v_i)) &= \det_{\mathfrak{B}'}((\epsilon_i)) \det_{\mathfrak{B}}((v_i)) \\ &= \det((u(\epsilon'_i)) \det_{\mathfrak{B}}((v_i)) \\ &= \det(u) \det_{\mathfrak{B}}((v_i)) \end{aligned}$$

or $\det(u) = \pm 1$ (proposition 6.4.3). $\square$

6.5 Angle de deux vecteurs

Si deux vecteurs x et y de l'espace euclidien E sont proportionnels, nous dirons que leur **angle** $\theta(x, y)$ est nul. Sinon, ils engendrent un espace euclidien E' de dimension 2, de base orthonormée $\mathfrak{B}' = (\epsilon_1, \epsilon_2)$, que l'on peut compléter en une base orthonormée $\mathfrak{B} = (\epsilon_i)$ de E (proposition 6.3.1).

Un vecteur non nul $v \in E'$ peut être défini par sa norme et son angle avec ϵ_1 . En effet, si $v/\|v\| = {}^t(v_1, v_2)$, il existe $\gamma \in [0, 2\pi[$ tel que $\cos \gamma = v_1$ et $\sin \gamma = v_2$, puisque $v_1^2 + v_2^2 = 1$. On a alors $v = \|v\| {}^t(\cos \gamma, \sin \gamma)$. Ceci définit une application de $E' \setminus \{0\}$ dans $\mathbb{R}_+^* \times [0, 2\pi[$, bijective, la réciproque étant $(a, b) \mapsto a {}^t(\cos b, \sin b)$.

Si $x = {}^t(x_1, x_2)$ et $y = {}^t(y_1, y_2)$ sont deux vecteurs quelconques non nuls de E' écrits dans la base $\mathfrak{B}'$, posons :

$$\begin{cases} \cos(\theta(x, y)) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{x_1 y_1 + x_2 y_2}{(x_1^2 + x_2^2)^{1/2} (y_1^2 + y_2^2)^{1/2}}, \\ \sin(\theta(x, y)) = \frac{\det_{\mathfrak{B}}(x, y)}{\|x\| \|y\|} = \frac{x_1 y_2 - x_2 y_1}{\|x\| \|y\|}, \end{cases}$$

On vérifie sans peine que $\cos^2(\theta) + \sin^2(\theta) = 1$ et que $\theta(x, y)$ ne dépend pas, au signe près (proposition 6.4.4), de la base orthonormée choisie pour le calcul, un changement de base étant dans ce cas une isométrie.

L'angle $\theta(x, y) \in [0, 2\pi[$ est ainsi bien défini dans toute base orthonormée obtenue par rotation à partir de $\mathfrak{B}$, au signe près dans les autres.

Remarquons que $\gamma = \theta(\epsilon_1, v)$.

Les vecteurs x et y étant mis sous la forme :

$$x = \|x\| {}^t(\cos \alpha, \sin \alpha) \quad \text{et} \quad y = \|y\| {}^t(\cos \beta, \sin \beta),$$

appliquons les formules précédentes :

$$\begin{cases} \cos(\theta(x, y)) = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\beta - \alpha), \\ \sin(\theta(x, y)) = \cos \alpha \sin \beta - \sin \alpha \cos \beta = \sin(\beta - \alpha), \end{cases}$$

d'où $\theta(x, y) = \beta - \alpha$, sans surprise.

6.6 Isométries de $\mathbb{R}^2$

Soit $u \in O(2)$ de matrice dans la base canonique :

$$M(u) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Par hypothèse, on a $a^2 + b^2 = c^2 + d^2 = 1$ et $ac + bd = 0$.

Posons $a = \cos \theta$, $b = \sin \theta$, $c = \sin \phi$ et $d = \cos \phi$.

L'équation $ac + bd = 0$ donne alors $\sin(\theta - \phi) = 0$, d'où $\theta = \phi$ ou $\theta = \pi - \phi$, et les deux formes possibles :

$$M(u) = R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad M(u) = S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Etudions la première. L'image par u de $x = \|x\|^t(\cos \phi, \sin \phi)$, vecteur quelconque de E :

$$\begin{aligned} u(x) &= \|x\|^t(\cos \theta \cos \phi - \sin \theta \sin \phi, \sin \theta \cos \phi + \cos \theta \sin \phi) \\ &= \|x\|^t(\cos(\phi + \theta), \sin(\phi + \theta)) \end{aligned}$$

fait un angle orienté θ avec x ; u est la rotation d'angle θ .

La seconde matrice représente un endomorphisme dont les vecteurs propres (orthonormés) sont $v_1 = {}^t(\cos(\theta/2), \sin(\theta/2))$ et $v_2 = {}^t(-\sin(\theta/2), \cos(\theta/2))$. La matrice de u dans la base (v_1, v_2) s'écrit :

$$M(u) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

et u est la symétrie orthogonale (ou réflexion) d'axe v_1 (et de direction v_2).

Les symétries orthogonales engendrent $O(2)$; en effet, on a pour une rotation quelconque et quel que soit ϕ :

$$R(\theta) = S(\phi)S(\phi - \theta).$$

L'ensemble des rotations de $\mathbb{R}^2$ est un sous-groupe de $O(2)$, le **groupe spécial orthogonal** de $\mathbb{R}^2$ noté $SO(2)$.

6.7 Le produit vectoriel

La base $\mathfrak{B} = (e_1, e_2, e_3)$ est la base canonique de $\mathbb{R}^3$:

$$e_1 = {}^t(1, 0, 0), \quad e_2 = {}^t(0, 1, 0), \quad e_3 = {}^t(0, 0, 1).$$

Elle est canonique car elle provient de la définition même de $\mathbb{R}^3 : \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

Le $\mathbb{R}$ -espace vectoriel $\mathbb{R}^3$ muni de son produit scalaire canonique est euclidien. Rappelons que pour des éléments quelconques $x = (x_i)$ et $y = (y_i)$ on a :

$$x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3,$$

et que le résultat est le même dans dans tout autre base orthonormée.

Soient x et y des vecteurs quelconques de $E = \mathbb{R}^3$. L'application :

$$\begin{aligned} E &\rightarrow \mathbb{R} \\ z &\mapsto \det_{\mathfrak{B}}(x, y, z) \end{aligned}$$

est une forme linéaire, donc un élément de E^* , et donc de E (qui est son propre dual). Cet élément est le **produit vectoriel** de x et de y , dans cet ordre, noté $x \times y$ (le symbole $\wedge$ désigne un autre produit, le **produit extérieur**).

Ce produit est bilinéaire, antisymétrique : $y \times x = -x \times y$, alterné : $x \times x = 0$, et indépendant du choix de la base, pourvu qu'elle soit orthonormée et directe (proposition 6.4.4).

On a d'après la définition : $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$. Cela donne pour des éléments quelconques $x = (x_i)$ et $y = (y_i)$, par bilinéarité et antisymétrie :

$$x \times y = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Proposition 6.7.1. *Le produit vectoriel de deux vecteurs de $\mathbb{R}^3$ est orthogonal à chacun des vecteurs.*

Démonstration. Le déterminant étant alterné, on a :

$$x \cdot (x \times y) = \det(x, x, y) = 0$$

et, pour la même raison, $y \cdot (x \times y) = 0$. □

Proposition 6.7.2. *Soient x et y deux vecteurs non proportionnels de $\mathbb{R}^3$ et z le vecteur normé orthogonal à x et à y tel que $\det(x, y, z) > 0$. Le produit vectoriel est égal à $\|x\| \|y\| \sin \theta(x, y) z$. Sa norme est au plus égale à $\|x\| \|y\|$; il y a égalité si x et y sont orthonaux.*

Démonstration. Soient $\epsilon_1 = x/\|x\|$, $y' = y - (y \cdot \epsilon_1) \epsilon_1$, $\epsilon_2 = y'/\|y'\|$ (méthode de Gram-Schmidt) et $\epsilon_3 = \epsilon_1 \times \epsilon_2 = z$, de sorte que $(\epsilon_1, \epsilon_2, \epsilon_3)$ est une base orthonormée directe. Dans cette base :

$$x = \|x\| (1, 0, 0), \quad y = \|y\| (\cos \theta(x, y), \sin \theta(x, y), 0)$$

et :

$$x \times y = \|x\| \|y\| (0, 0, \sin \theta(x, y)) = \|x\| \|y\| \sin \theta(x, y) z,$$

d'où la conclusion sur la norme de $x \times y$. □

Le produit vectoriel n'est pas associatif car, pour tout triplet de vecteurs (x, y, z) , on a :

$$\begin{aligned} x \times (y \times z) &= -(x \cdot y) z + (x \cdot z) y, \\ (x \times y) \times z &= -(y \cdot z) x + (x \cdot z) y. \end{aligned}$$

Pour tout couple de vecteurs (y, z) orthogonaux non nuls, l'équation $y \times x = z$ admet les solutions :

$$x = \lambda y - \frac{1}{\|y\|^2} y \times z, \quad \lambda \in \mathbb{R}.$$

6.8 Complexifié, espace hermitien

Soit E le $\mathbb{R}$ -espace vectoriel habituel, euclidien, de dimension n , muni d'une base orthonormée $\mathfrak{B} = (\epsilon_i)$. L'ensemble $E_{\mathbb{C}} = E \times E$ muni du produit :

$$\begin{array}{ccccc} \mathbb{C} & \times & E_{\mathbb{C}} & \rightarrow & E_{\mathbb{C}} \\ \alpha + i\beta & & (x, y) & \mapsto & (\alpha x - \beta y, \beta x + \alpha y) \end{array}$$

est un $\mathbb{C}$ -espace vectoriel. On écrit $x + i y$ l'élément (x, y) et on identifie $x \in E$ avec $(x, 0) \in E_{\mathbb{C}}$.

Notons bien que E n'est pas un sous-espace vectoriel de $E_{\mathbb{C}}$, les corps des scalaires étant différents. Les applications $(x, y) \mapsto x$ et $(x, y) \mapsto y$ ne sont donc pas des projecteurs.

Montrons que $\mathfrak{B}$ est une base de $E_{\mathbb{C}}$. L'écriture :

$$x + i y = \sum_i x_i \epsilon_i + i \sum_i y_i \epsilon_i = \sum_i (x_i + i y_i) \epsilon_i$$

prouve que c'est une famille génératrice. Supposons que :

$$\sum_i (a_i + i b_i) \epsilon_i = 0.$$

La partie réelle et la partie imaginaire sont nulles et donc aussi tous les a_i et tous les b_i , la famille étant libre sur $\mathbb{R}$. Elle est donc libre sur $\mathbb{C}$, c'est une base, et $E_{\mathbb{C}}$ est de dimension n sur $\mathbb{C}$.

Cet espace $E_{\mathbb{C}}$ est le **complexifié** de E . Le complexifié de $\mathbb{R}^n$ est $\mathbb{C}^n$, et la base canonique (e_i) de $\mathbb{R}^n$ est la base canonique de $\mathbb{C}^n$.

On désire munir $E_{\mathbb{C}}$ d'un produit scalaire qui coïncide avec le produit scalaire de E sur les vecteurs à composantes réelles et qui donne une norme telle que $\|x + i y\|^2 = \|x\|^2 + \|y\|^2$. Ce produit peut être défini par :

$$(x + i y | x' + i y') = x \cdot x' + y \cdot y' + i (x \cdot y' - y \cdot x'),$$

le point désignant le produit scalaire de E .

Voyons les propriétés de ce produit. On a pour tout $\alpha \in \mathbb{R}$ les identités :

$$(\alpha(x + i y) | x' + i y') = (x + i y | \alpha(x' + i y')) = \alpha(x + i y | x' + i y')$$

et :

$$\begin{aligned} (i(x + i y) | x' + i y') &= (-y + i x | x' + i y') \\ &= -y \cdot x' + x \cdot y' + i(-y \cdot y' - x \cdot x') \\ &= -i(x + i y | x' + i y'), \\ (x + i y | i(x' + i y')) &= (x + i y | -y' + i x') \\ &= y \cdot x' - x \cdot y' + i(y \cdot y' + x \cdot x') \\ &= i(x + i y | x' + i y'), \\ (x' + i y' | x + i y) &= \overline{x' \cdot x + y' \cdot y + i(x' \cdot y - y' \cdot x)} \\ &= \overline{(x + i y | x' + i y')}, \end{aligned}$$

d'où, quels que soient $\lambda, \mu \in \mathbb{C}$, $z, z', z'' \in E_{\mathbb{C}}$, les quatre relations :

$$\begin{aligned} (z' | z) &= \overline{(z | z')}, \\ (z | \lambda z' + \mu z'') &= \lambda(z | z') + \mu(z | z''), \\ (z | z) &\geq 0, \\ (z | z) &= 0 \iff z = 0. \end{aligned}$$

On déduit de ces quatre relations que :

$$\begin{aligned} (\lambda z' + \mu z'' | z) &= \overline{(z | \lambda z' + \mu z'')} \\ &= \bar{\lambda} \overline{(z | z')} + \bar{\mu} \overline{(z | z'')} \\ &= \bar{\lambda} (z' | z) + \bar{\mu} (z'' | z). \end{aligned}$$

Ce produit est donc linéaire sur le second facteur et **semi-linéaire** sur le premier. Un produit vérifiant ces quatre relations est un **produit scalaire hermitien**. Il permet de définir les notions d'orthogonalité et de norme comme dans le cas réel.

Un espace vectoriel complexe de dimension finie muni d'un tel produit scalaire est un **espace hermitien**.

Dans la suite, H désigne un espace hermitien de dimension n .

Théorème 6.8.1 (Inégalité de Cauchy-Schwarz). *Quels que soient x et y dans H on a :*

$$|(x | y)| \leq \|x\| \|y\|.$$

L'égalité a lieu si et seulement si x et y sont proportionnels.

Démonstration. Si $(x | y) = 0$, l'inégalité est triviale. Sinon, $(x | y)$ se met sous la forme $\alpha e^{i\theta}$, $\alpha = |(x | y)| \in \mathbb{R}_+^*$, et, quel que soit $\lambda \in \mathbb{C}$, on a :

$$\begin{aligned} 0 \leq \|x + \lambda y\|^2 &= \|x\|^2 + |\lambda|^2 \|y\|^2 + \lambda (x | y) + \bar{\lambda} (y | x) \\ &= \|x\|^2 + |\lambda|^2 \|y\|^2 + \alpha (\lambda e^{i\theta} + \bar{\lambda} e^{-i\theta}). \end{aligned}$$

Choisissons $\lambda = \rho e^{-i\theta}$, $\rho = |\lambda|$:

$$0 \leq \|y\|^2 \rho^2 + 2\alpha \rho + \|x\|^2.$$

Nous obtenons un trinôme du deuxième degré en ρ toujours positif, et son discriminant réduit $\alpha^2 - \|x\|^2 \|y\|^2$ est négatif ou nul, ce qui est l'inégalité annoncée. Il est nul en cas de racine double, c'est-à-dire lorsque $x + \lambda y = 0$. $\square$

Théorème 6.8.2. *Il existe des bases orthonormées de H . Dans une telle base, le produit scalaire hermitien de deux éléments quelconques $z = (z_i)$ et $z' = (z'_i)$ s'écrit :*

$$(z | z') = \sum_i \bar{z}_i z'_i.$$

Démonstration. On peut construire, à partir d'une base quelconque, une base orthonormée de H grâce au procédé de Gram-Schmidt (théorème 5.2.1). Dans une telle base (ϵ_i) on a, δ_{ij} étant le symbole de Kronecker :

$$\begin{aligned} (z | z') &= \sum_{ij} (z_i \epsilon_i | z'_j \epsilon_j) \\ &= \sum_{ij} \bar{z}_i z'_j (\epsilon_i | \epsilon_j) \\ &= \sum_{ij} \bar{z}_i z'_j \delta_{ij} \\ &= \sum_i \bar{z}_i z'_i \end{aligned}$$

$\square$

En posant :

$$(x + i y \mid x' + i y') = x \cdot x' + y \cdot y' - i (x \cdot y' - x' \cdot y)$$

et en changeant en conséquence la définition d'un produit scalaire hermitien (la linéarité étant maintenant sur le premier facteur), on définit un produit scalaire hermitien, conjugué du précédent, qui donne la même orthogonalité, la même norme, les mêmes bases orthonormées, dans lesquelles $(z \mid z') = \sum_i z_i \bar{z}'_i$.

Une base orthonormée de E est une base orthonormée de $E_{\mathbb{C}}$. La base canonique de $\mathbb{R}^n$ est la base canonique de $\mathbb{C}^n$.

Proposition 6.8.1. *Le produit scalaire hermitien s'exprime en fonction de la norme grâce à l'identité de polarisation :*

$$(z \mid z') = \frac{1}{4} (\|z + z'\|^2 - \|z - z'\|^2 + i \|z - iz'\|^2 - i \|z + iz'\|^2).$$

Démonstration. Considérons les relations :

$$\begin{aligned} \|z + z'\|^2 &= \|z\|^2 + \|z'\|^2 + (z \mid z') + (z' \mid z), \\ \|z - z'\|^2 &= \|z\|^2 + \|z'\|^2 - (z \mid z') - (z' \mid z), \\ \|z - iz'\|^2 &= \|z\|^2 + \|z'\|^2 - i (z \mid z') + i (z' \mid z), \\ \|z + iz'\|^2 &= \|z\|^2 + \|z'\|^2 + i (z \mid z') - i (z' \mid z), \end{aligned}$$

desquelles on déduit :

$$\begin{aligned} \|z + z'\|^2 - \|z - z'\|^2 &= 2((z \mid z') + (z' \mid z)) \\ &= 2((z \mid z') + \overline{(z \mid z')}) \\ &= 4 \Re(z \mid z'), \\ \|z - iz'\|^2 - \|z + iz'\|^2 &= -2i((z \mid z') - (z' \mid z)) \\ &= -2i((z \mid z') - \overline{(z \mid z')}) \\ &= 4 \Im(z \mid z'), \end{aligned}$$

d'où l'expression annoncée. $\square$

L'espace H est désormais muni d'une base orthonormée $\mathfrak{B} = (\epsilon_i)$, et le produit scalaire hermitien s'écrit $((z_i) \mid (z'_i)) = \sum_i \bar{z}_i z'_i$.

Théorème 6.8.3. *Un espace hermitien H est son propre dual.*

Démonstration. L'application $a \mapsto a^*$ qui transforme le vecteur $a \in H$ en la forme linéaire $z \mapsto (\bar{a} \mid z)$ est linéaire car $(\lambda a + \mu b)^* = \lambda a^* + \mu b^*$ quels que soient les éléments et les scalaires écrits :

$$\begin{aligned} (\lambda a + \mu b)^*(z) &= \overline{(\lambda a + \mu b \mid z)} \\ &= \lambda \overline{(a \mid z)} + \mu \overline{(b \mid z)} \\ &= (\lambda \bar{a} + \mu \bar{b} \mid z) \end{aligned}$$

et injective car $(\bar{a} \mid z) = 0$ quel que soit z implique $(\bar{a} \mid \bar{a}) = 0$, donc $\|a\| = 0$ et $a = 0$. C'est donc un isomorphisme, H et son dual ayant même dimension. $\square$

Soient $u \in \mathfrak{L}(H)$ et $z' \in H$, quelconques. L'application $z \mapsto (z' | u(z))$ est une forme linéaire sur H , et, d'après le théorème 5.8.3, il existe un élément z'' de H tel que $(z' | u(z)) = (z'' | z)$. Posons $z'' = u^*(z')$, et montrons que u^* est un endomorphisme de H . Quels que soient les scalaires et les vecteurs écrits, on a d'une part :

$$(\alpha z_1 + \beta z_2 | u(z)) = (u^*(\alpha z_1 + \beta z_2) | z)$$

et d'autre part :

$$\begin{aligned} (\alpha z_1 + \beta z_2 | u(z)) &= \bar{\alpha}(z_1 | u(z)) + \bar{\beta}(z_2 | u(z)) \\ &= (\alpha u^*(z_1) | z) + (\beta u^*(z_2) | z) \\ &= (\alpha u^*(z_1) + \beta u^*(z_2) | z) \end{aligned}$$

d'où $u^*(\alpha z_1 + \beta z_2) = \alpha u^*(z_1) + \beta u^*(z_2)$.

L'endomorphisme u^* est l'**adjoint** de u . Il est défini par :

$$(u(z) | z') = (z | u^*(z')), \quad \forall z, z' \in H.$$

L'endomorphisme u est **auto-adjoint** ou **hermitien** si $u = u^*$.

Théorème 6.8.4. *Soit $u \in \mathfrak{L}(H)$. Dans une base orthonormée, la matrice de l'adjoint u^* de u est la matrice adjointe de $M(u)$: $M(u^*) = M(u)^*$. On a $u^{**} = u$.*

Démonstration. Soit $z = (z_i)$ et $z' = (z'_i)$ des éléments quelconques de H , $M(u) = (u_{ij})$, $u(z)_i = \sum_j u_{ij} z_j$ et $u^*(z')_i = \sum_j u_{ij}^* z'_j$. On a :

$$\begin{aligned} (u(z) | z') &= \sum_{ij} \bar{u}_{ij} \bar{z}_j z'_i \\ &= \sum_{ij} {}^t \bar{u}_{ji} \bar{z}_j z'_i \\ &= \sum_{ij} \bar{z}_j {}^t \bar{u}_{ji} z'_i \end{aligned}$$

et :

$$\begin{aligned} (z | u^*(z')) &= \sum_j \bar{z}_j \sum_i u_{ji}^* z'_i \\ &= \sum_{ij} \bar{z}_j u_{ji}^* z'_i \end{aligned}$$

d'où $u_{ji}^* = {}^t \bar{u}_{ji}$ et $M(u^*) = {}^t \overline{M(u)} = M(u)^*$.

On passe de $M(u)$ à $M(u^{**})$ en transposant deux fois et en conjuguant deux fois, or ces transformations commutent et sont des involutions. On a donc $M(u^{**}) = M(u)$ et $u^{**} = u$ (proposition 3.5.4). $\square$

Proposition 6.8.2. *La matrice dans une base orthonormée d'un endomorphisme hermitien u de H est hermitienne. Réciproquement, si la matrice d'un endomorphisme u de H dans une base orthonormée, $U = (u_{ij})$, est hermitienne, u est hermitien.*

Démonstration. Si $u = u^*$, on a $M(u) = M(u^*) = M(u)^*$ (théorème 5.8.4) et $M(u)$ est hermitienne.

Pour la réciproque, partant de l'hypothèse ${}^t\bar{u}_{ij} = u_{ij}$, on a :

$$\begin{aligned}
(u(z) | z') &= \sum_{ij} \bar{u}_{ij} \bar{z}_j z'_i \\
&= \sum_{ij} {}^t\bar{u}_{ji} \bar{z}_j z'_i \\
&= \sum_{ij} \bar{z}_j {}^t\bar{u}_{ji} z'_i \\
&= \sum_{ij} \bar{z}_j u_{ji} z'_i \\
&= (z | u(z'))
\end{aligned}$$

□

Proposition 6.8.3. *Les valeurs propres d'un endomorphisme hermitien u de H sont réelles.*

Démonstration. Soit $(\lambda, x) \in \mathbb{C} \times (H \setminus \{0\})$ un couple tel que $u(x) = \lambda x$. On déduit des égalités :

$$\begin{aligned}
(x | u(x)) &= \lambda \|x\|^2 \\
(u(x) | x) &= \bar{\lambda} \|x\|^2 \\
(x | u(x)) &= (u(x) | x)
\end{aligned}$$

que $\lambda = \bar{\lambda}$.

□

Proposition 6.8.4. *L'ensemble $\Sigma(H)$ des endomorphismes hermitiens de H est un $\mathbb{R}$ -espace vectoriel de dimension n^2 .*

Démonstration. Soit $u \in \Sigma(H)$, $\lambda \in \mathbb{C}$ et $\lambda u \in \Sigma(H)$. Les valeurs propres λ_i de u sont réelles (proposition 5.8.3), celles de λu , égales à $\lambda \lambda_i$ aussi, et λ doit être réel. Dans ce cas, λu est hermitien car :

$$(\lambda u(z) | z') = \lambda (u(z) | z') = \lambda (z | u(z')) = (z | \lambda u(z')).$$

La somme de deux endomorphismes hermitiens étant évidemment un endomorphisme hermitien, $\Sigma(H)$ est un $\mathbb{R}$ -espace vectoriel.

L'application $M : u \mapsto M(u)$ est une bijection entre les espaces vectoriels $\Sigma(H)$ et $H_n(\mathbb{C})$ (proposition 5.8.2), donc un isomorphisme (Proposition 2.5.4), d'où la dimension de $\Sigma(H)$ (proposition 2.3.3). □

Proposition 6.8.5. *Quel que soit l'endomorphisme u de H , il existe deux endomorphismes hermitiens h_1 et h_2 tels que $u = h_1 + ih_2$.*

Démonstration. Posons $h_1 = (u + u^*)/2$ et $h_2 = -i(u - u^*)/2$. Le calcul :

$$\begin{aligned}
h_1^* &= (u^* + u^{**})/2 = (u^* + u)/2 = h_1, \\
h_2^* &= i(u^* - u^{**})/2 = i(u^* - u)/2 = h_2,
\end{aligned}$$

montre que h_1 et h_2 sont hermitiens. □

L'endomorphisme $u \in \mathfrak{L}(H)$ est une **isométrie** s'il conserve le produit scalaire hermitien ou la norme, c'est-à-dire si on a pour tout $z \in E_{\mathbb{C}}$:

$$\begin{aligned}
\|u(z)\| = \|z\| &\iff (u(z) | u(z)) = (z | z) \\
&\iff ({}^t\bar{u}(u(z)) | z) = (z | z) \\
&\iff ({}^t\bar{u}(u(z)) - z | z) = 0 \\
&\iff {}^t\bar{u}(u(z)) = z \\
&\iff {}^t\bar{u} \circ u = i_{E_{\mathbb{C}}}
\end{aligned}$$

Enfin, l'endomorphisme $u \in \mathfrak{L}(H)$ est **normal** s'il commute avec son adjoint : $u \circ u^* = u^* \circ u$. Un endomorphisme hermitien ou unitaire est normal. L'importance de cette notion est justifiée par le théorème 6.8.5, que préparent les deux propositions qui suivent.

Proposition 6.8.6. *Si u est un endomorphisme normal, on a pour tout x : $\|u(x)\| = \|u^*(x)\|$.*

Démonstration. Utilisant la définition de u^* puis sa normalité, on écrit :

$$\begin{aligned}
\|u(x)\|^2 &= (u(x) | u(x)) \\
&= (x | u^*(u(x))) \\
&= (x | u(u^*(x))) \\
&= (u^*(x) | u^*(x)) \\
&= \|u^*(x)\|^2
\end{aligned}$$

d'où le résultat. □

Proposition 6.8.7. *Si u est un endomorphisme normal, ses valeurs propres et celles de u^* sont conjuguées, et les vecteurs propres sont les mêmes :*

$$u(x) = \lambda x \iff u^*(x) = \bar{\lambda} x.$$

Démonstration. Si $v = u - \lambda i_H$, on a $v^* = u^* - \bar{\lambda} i_H$ et $v \circ v^* - v^* \circ v = u \circ u^* - u^* \circ u$, de sorte que si u est normal v l'est aussi, quel que soit $\lambda \in \mathbb{C}$. Si λ est valeur propre de u et x vecteur propre associé, la norme de $(u - \lambda)(x)$ est nulle, celle de $(u^* - \bar{\lambda})(x)$ aussi (proposition 6.8.6). On peut échanger les rôles de u et de u^* , puisque $(u^*)^* = u$, d'où la conclusion. □

Théorème 6.8.5. *Un endomorphisme normal et son adjoint sont diagonalisables dans une même base orthonormée.*

Démonstration. Il existe une base dans laquelle la matrice de u est triangulaire supérieure (proposition 4.9.1). Le procédé de Gram-Schmidt permet de construire une base orthonormée à partir de cette base, les matrices de passage étant triangulaires supérieures (corollaire du théorème 6.2.1). La matrice de u dans cette base orthonormée, $M(u)$, est triangulaire supérieure (proposition

2.5.9). Comme $M(u^*) = \overline{M(u)}$ (proposition 6.8.2), la matrice de u^* dans cette base est triangulaire inférieure :

$$M(u) = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & t_{nn} \end{pmatrix}, \quad M(u^*) = \begin{pmatrix} \bar{t}_{11} & 0 & \dots & 0 \\ \bar{t}_{12} & \bar{t}_{22} & \ddots & 0 \\ \vdots & & \ddots & 0 \\ \bar{t}_{1n} & \bar{t}_{2n} & \dots & \bar{t}_{nn} \end{pmatrix}$$

La première colonne de $M(u)$ montre que le premier vecteur de la base est un vecteur propre de u , donc que $t_{11} = \lambda_1$, si nous notons λ_i les valeurs propres de u . Ce vecteur est aussi vecteur propre de u^* , associé à la valeur propre $\bar{\lambda}_1$ (proposition 5.8.7) : la première colonne de $M(u^*)$ est de la forme $(\bar{\lambda}_1, 0, \dots, 0)$, d'où $t_{12} = \dots = t_{1n} = 0$. La première ligne de $M(u)$ est donc $(\lambda_1, 0, \dots, 0)$, et sa deuxième colonne : $(0, t_{22}, 0, \dots, 0)$, d'où $t_{22} = \lambda_2$. Le même raisonnement montre alors que la deuxième ligne de $M(u^*)$ est $(0, \bar{\lambda}_2, 0, \dots, 0)$... ainsi de suite jusqu'à la n -ième ligne et la n -ième colonne des deux matrices. $\square$

Proposition 6.8.8. *Un endomorphisme u de H est normal si et seulement s'il existe un polynôme $P \in \mathbb{C}[Z]$, de degré au plus égal à $n - 1$, tel que $u^* = P(u)$.*

Démonstration. Quel que soit $P \in \mathbb{C}[Z]$, il est immédiat que $uP(u) = P(u)u$.

Si u est normal, il existe une base orthonormée dans laquelle sa matrice $M(u)$ et celle de u^* , $M(u^*)$, sont diagonales (théorème 6.8.5), respectivement $D = (d_{ij})$ et $D^* = \overline{D}$, les termes diagonaux étant les valeurs propres : $d_{ii} = \lambda_i$ et $d_{ii}^* = \bar{\lambda}_i$. Si un polynôme $P(Z) = \sum_{0 \leq i \leq n-1} p_i Z^i$ est tel que $P(\lambda_i) = \bar{\lambda}_i$ pour tout i , on aura $P(D) = D^*$, c'est-à-dire $P(M(u)) = M(P(u)) = M(u^*)$ dans la base de diagonalisation, soit $P(u) = u^*$. Une valeur propre multiple donne lieu à une seule condition. Si u admet n' valeurs propres deux à deux distinctes notées $\mu_1, \dots, \mu_{n'}$, les p_i sont solution du système :

$$\begin{cases} p_0 + p_1 \mu_1 + p_2 \mu_1^2 + \dots + p_{n'-1} \mu_1^{n'-1} &= \bar{\mu}_1 \\ \dots & \\ p_0 + p_1 \mu_{n'} + p_2 \mu_{n'}^2 + \dots + p_{n'-1} \mu_{n'}^{n'-1} &= \bar{\mu}_{n'} \end{cases}$$

dont le déterminant est de Vandermonde (voir 4.7), donc non nul. La solution existe et elle est unique. Le degré de P est au plus égal à $n' - 1 \leq n - 1$.

Nous avons donc $P(M(u)) = M(u^*)$. Or, M étant un isomorphisme d'algèbre, $P(M(u)) = M(P(u))$, d'où $P(u) = u^*$. $\square$

Remarquons que si u est normal u^* l'est aussi, et il existe un polynôme Q , de degré au plus égal à $n' - 1$, tel que $u = Q(u^*)$. On a $P \circ Q(u) = Q \circ P(u) = u$, ce qui montre que les polynômes $P \circ Q(X)$ et $Q \circ P(X)$ sont congrus à X modulo le polynôme caractéristique de u .

Théorème 6.8.6. *Un endomorphisme u de H est diagonalisable dans une base orthonormée si et seulement s'il est normal.*

Démonstration. Si u est normal, il est diagonalisable dans une base orthonormée (théorème 6.8.5). Réciproquement, si la matrice de u dans une certaine base orthonormée est diagonale, $M(u) = D$, on a $M(U^*) = \overline{D}$. Il existe un polynôme P tel que $P(D) = D^*$ (voir la démonstration de la proposition 6.8.8), et donc $u^* = P(u)$. La proposition 6.8.8 permet de conclure. $\square$

Exemple 6.8.1. L'espace vectoriel complexe $\mathbb{C}^2$ est muni de sa structure hermitienne canonique. Une matrice $M \in \mathfrak{M}_2(\mathbb{C})$ est de la forme :

$$M = \begin{pmatrix} a e^{i\alpha} & c e^{i\gamma} \\ b e^{i\beta} & d e^{i\delta} \end{pmatrix}$$

où toutes les constantes écrites sont réelles et positives. On a :

$$M^* = {}^t\overline{M} = \begin{pmatrix} a e^{-i\alpha} & b e^{-i\beta} \\ c e^{-i\gamma} & d e^{-i\delta} \end{pmatrix}$$

et M est hermitienne si et seulement si $\alpha = \delta = 0$, $b = c$ et $\gamma = -\beta$.

Elle est unitaire si et seulement si :

$$\begin{pmatrix} a e^{i\alpha} & c e^{i\gamma} \\ b e^{i\beta} & d e^{i\delta} \end{pmatrix} \begin{pmatrix} a e^{-i\alpha} & b e^{-i\beta} \\ c e^{-i\gamma} & d e^{-i\delta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

c'est-à-dire si et seulement si :

$$\begin{pmatrix} a^2 + c^2 & ab e^{i(\alpha-\beta)} + cd e^{i(\gamma-\delta)} \\ ab e^{i(\beta-\alpha)} + cd e^{i(\delta-\gamma)} & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Posons d'abord $a = \cos \theta$, $c = \sin \theta$, $b = \sin \phi$ et $d = \cos \phi$ ($0 \leq \theta, \phi \leq \pi/2$). On a alors :

$$\cos \theta \sin \phi e^{i(\alpha-\beta)} + \sin \theta \cos \phi e^{i(\gamma-\delta)} = 0$$

et une deuxième relation conjuguée de celle-ci. Egalons les modules et les arguments des deux quantités égales $\cos \theta \sin \phi e^{i(\alpha-\beta)}$ et $\sin \theta \cos \phi e^{i(\gamma-\delta+\pi)}$:

$$\begin{aligned} \cos \theta \sin \phi - \sin \theta \cos \phi &= 0 \\ \alpha - \beta &= \gamma - \delta + \pi \end{aligned}$$

pour obtenir $\theta = \phi$ et $\delta = \beta - \alpha + \gamma + \pi$ (modulo 2π), et la forme générale d'une matrice unitaire :

$$U = \begin{pmatrix} \cos \theta e^{i\alpha} & \sin \theta e^{i\gamma} \\ \sin \theta e^{i\beta} & \cos \theta e^{i(\beta-\alpha+\gamma+\pi)} \end{pmatrix}$$

dont le déterminant est égal à $e^{i(\beta+\gamma+\pi)}$. Ces matrices forment le **groupe unitaire** $U(2)$. Si de plus $\beta + \gamma = \pi$, on obtient les matrices du **groupe spécial unitaire** $SU(2)$, de déterminant égal à 1..

La matrice M est normale si et seulement s'il existe un polynôme de degré inférieur à 2, donc de la forme $P(Z) = \lambda Z + \mu$, tel que $P(u) = u^*$. Ecrivons cette condition :

$$\begin{pmatrix} \lambda a e^{i\alpha} + \mu & \lambda c e^{i\gamma} \\ \lambda b e^{i\beta} & \lambda d e^{i\delta} + \mu \end{pmatrix} = \begin{pmatrix} a e^{-i\alpha} & b e^{-i\beta} \\ c e^{-i\gamma} & d e^{-i\delta} \end{pmatrix}$$

▽

Considérons d'abord le système :

$$\begin{cases} \lambda c e^{i\gamma} = b e^{-i\beta} \\ \lambda b e^{i\beta} = c e^{-i\gamma} \end{cases}$$

duquel nous déduisons :

$$\lambda = b c^{-1} e^{-i(\beta+\gamma)} = c b^{-1} e^{-i(\beta+\gamma)}$$

d'où $b = c$ et $\lambda = e^{-i(\beta+\gamma)}$.

Calculons enfin μ à partir du système :

$$\begin{cases} \mu = a(e^{-i\alpha} - e^{i(\alpha-\beta-\gamma)}) \\ \mu = d(e^{-i\delta} - e^{i(\delta-\beta-\gamma)}) \end{cases}$$

dont la condition de compatibilité est ($e^{i\pi} = -1$) :

$$a(e^{-i\alpha} + e^{i(\alpha-(\beta+\gamma+\pi))}) = d(e^{-i\delta} + e^{i(\delta-(\beta+\gamma+\pi))}).$$

Comme $e^{i\phi} + e^{i\psi} = 2 \cos((\phi - \psi)/2) e^{i(\phi+\psi)/2}$, on a :

$$\begin{cases} e^{-i\alpha} + e^{i(\alpha-\beta-\gamma-\pi)} = 2 \cos((2\alpha - \beta - \gamma - \pi)/2) e^{-i(\beta+\gamma+\pi)/2} \\ e^{-i\delta} + e^{i(\delta-\beta-\gamma-\pi)} = 2 \cos((2\delta - \beta - \gamma - \pi)/2) e^{-i(\beta+\gamma+\pi)/2} \end{cases}$$

d'où la condition : $a \sin(\alpha - (\beta + \gamma)/2) = d \sin(\delta - (\beta + \gamma)/2)$.

Exemple 6.8.2. Considérons la matrice A , qui représente l'endomorphisme u de $\mathbb{C}^2$ dans la base canonique, et son adjointe A^* , qui représente u^* dans la même base :

$$A = \begin{pmatrix} 5-i & -\sqrt{6} \\ i\sqrt{6} & 3+i \end{pmatrix}, \quad A^* = \begin{pmatrix} 5+i & -i\sqrt{6} \\ -\sqrt{6} & 3-i \end{pmatrix}.$$

Comme :

$$AA^* = \begin{pmatrix} 32 & -4\sqrt{6}(1+i) \\ -4\sqrt{6}(1-i) & 16 \end{pmatrix} = A^*A,$$

A est normale, et u aussi. Le polynôme caractéristique de A :

$$P(X) = X^2 - \text{tr}(A)X + \det(A) = X^2 - 8X + 16 + 8i$$

a pour racines les valeurs propres $\lambda_1 = 2 + 2i$ et $\lambda_2 = 6 - 2i$ associées respectivement aux vecteurs propres orthonormés :

$$\begin{aligned} e'_1 &= {}^t(i/2, (1+i)\sqrt{6}/4), \\ e'_2 &= {}^t(-(1+i)\sqrt{6}/4, 1/2). \end{aligned}$$

Dans la base (e'_1, e'_2) les matrices de u et de u^* sont respectivement :

$$D = \begin{pmatrix} 3+2i & 0 \\ 0 & 6-2i \end{pmatrix} \quad \text{et} \quad D^* = \begin{pmatrix} 3-2i & 0 \\ 0 & 6+2i \end{pmatrix}$$

Cherchons un polynôme $P(Z)$ de degré inférieur à 2, donc de la forme $aZ+b$, tel que $D^* = P(D)$, c'est-à-dire tel que $P(3+2i) = 3-2i$ et $P(6-2i) = 6+2i$. La solution du système :

$$\begin{cases} (3+2i)a + b = 3-2i \\ (6-2i)a + b = 6+2i \end{cases}$$

est donnée par les formules de Cramer :

$$a = \begin{vmatrix} 3-2i & 1 \\ 6+2i & 1 \end{vmatrix} \begin{vmatrix} 3+2i & 1 \\ 6-2i & 1 \end{vmatrix}^{-1} = \frac{1}{25}(-7+24i)$$

et :

$$b = \begin{vmatrix} 3+2i & 3-2i \\ 6-2i & 6+2i \end{vmatrix} \begin{vmatrix} 3+2i & 1 \\ 6-2i & 1 \end{vmatrix}^{-1} = \frac{1}{25}(144-108i).$$

On a alors $u^* = P(u) = au + b$. On a également $u = Q(u^*) = \bar{a}u^* + \bar{b}$, et $P \circ Q(X) = Q \circ P(X) = X$. ∇

Soit u un endomorphisme de $\mathbb{R}^n$; posons $c(u)(x+iy) = u(x) + iu(y)$ et vérifions que $c(u)$ est un endomorphisme de $\mathbb{C}^n$:

$$\begin{aligned} c(u)((a+ib)(x+iy)) &= c(u)(ax-by+i(bx+ay)) \\ &= u(ax-by) + iu(bx+ay) \\ &= (a+ib)u(x) + (ia-b)u(y) \\ &= (a+ib)(u(x)+iu(y)) \\ &= (a+ib)c(u)(x+iy) \end{aligned}$$

La matrice de $c(u)$ est la matrice de u (dans la même base).

Proposition 6.8.9. *Les valeur propre et le déterminant d'une isométrie u de l'espace hermitien H sont de module égal à 1.*

Démonstration. Si $u(x) = \lambda x$, on a $\|u(x)\| = |\lambda|\|x\| = \|x\|$, et $|\lambda| = 1$. Le déterminant, produit de toutes les valeurs propres, est de module égal à 1. $\square$

7 Réduction des matrices

7.1 Polynôme caractéristique

Nous noterons X l'homothétie Xi_E et sa matrice XI_n . Le déterminant de $X - u$, égal à $\det(X - M)$ si $M = M(u)$, est un polynôme de degré n , le **polynôme caractéristique** de u ou de $M(u)$, noté $P_u(X)$ ou $P_M(X)$, déjà rencontré en 3.8, et dont nous savons qu'il ne dépend pas du choix de la base (voir 3.8 et 3.6).

Si E est un espace vectoriel sur $\mathbb{C}$, $P_u(X)$ se scinde en n facteurs du premier degré. Si c'est un espace vectoriel sur $\mathbb{R}$ et s'il n'a pas toutes ses racines réelles, on travaillera dans son complexifié, en remarquant que $P_{c(u)}(X) = P_u(X)$, et donc que $P_{c(u)}(X)$ est à coefficients réels. Dans ce cas, si λ est une racine complexe, $\bar{\lambda}$ est aussi racine car $P_u(\bar{\lambda}) = \overline{P_u(\lambda)} = 0$. Si z est un vecteur propre associé à λ , le vecteur $\bar{z}$, dont les composantes sont les conjuguées de celles de z , est vecteur propre associé à $\bar{\lambda}$. En effet, $c(u)(\bar{z}) = \overline{c(u)(z)} = \bar{\lambda} \bar{z}$.

Ainsi, le polynôme caractéristique de la matrice de la rotation d'angle θ :

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

égal à $X^2 - 2 \cos \theta X + 1$, est irréductible dans $\mathbb{R}$ si $\theta \neq k\pi$, mais se factorise dans $\mathbb{C}$:

$$X^2 - 2 \cos \theta X + 1 = (X - e^{i\theta})(X - e^{-i\theta}).$$

Remarquons que les vecteurs $v_1 = {}^t(i, 1)$ et $v_2 = {}^t(-i, 1)$ ont une image qui leur est proportionnelle :

$$\begin{aligned} R(\theta) v_1 &= e^{i\theta} v_1, \\ R(\theta) v_2 &= e^{-i\theta} v_2, \end{aligned}$$

le coefficient de proportionnalité étant pour chacun l'une des valeurs propres.

Autre exemple, le polynôme caractéristique de la matrice :

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

est égal à $X^2 - 1$, et elle admet deux valeurs propres réelles : 1 et -1 . On vérifie que le vecteur $v_1 = (\cos(\theta/2), \sin(\theta/2))$, invariant, correspond à la valeur propre 1 et que le vecteur $v_2 = (-\sin(\theta/2), \cos(\theta/2))$ correspond à la valeur propre -1 . La matrice $S(\theta)$ représente la symétrie orthogonale d'axe engendré par v_1 .

Enfin, le polynôme caractéristique de la matrice :

$$M = \begin{pmatrix} -3 & 16 \\ -1 & 5 \end{pmatrix}$$

est égal à $X^2 - 2X + 1 = (X - 1)^2$, et 1 est valeur propre double. Le système $Mv = v$, $v = {}^t(x, y)$, donne l'équation $-x + 4y = 0$, d'où l'unique (à homothétie non nulle près) vecteur propre $v_1 = (4, 1)$. On complète v_1 en une base de $\mathbb{R}^2$ par un vecteur quelconque non proportionnel à v_1 , par exemple $v_2 = {}^t(1, 0)$.

On a alors $Mv_2 = {}^t(-3, -1) = v_2 - v_1$, et dans la base (v_1, v_2) la matrice s'écrit :

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

Existe-t-il une base dans laquelle la matrice s'écrit :

$$M'' = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}?$$

L'équation $Mv = v + v_1$ donne $-x + 4y = 4$, d'où par exemple $v_3 = (-1, 0)$, et dans la base (v_1, v_3) la matrice a la forme souhaitée.

Revenons au cas général; nous pouvons écrire :

$$P_u(X) = (X - \lambda_1)^{r_1} \dots (X - \lambda_k)^{r_k}, \quad \sum_i r_i = n,$$

r_i étant l'**ordre** de la valeur propre λ_i . On voit immédiatement que le terme constant de P_u est égal à $(-1)^n \det(u)$, et que le coefficient de X^{n-1} , égal à $-\sum_i m_{ii}$ d'après la définition de P_u , est aussi égal à $-\sum_i \lambda_i$. La somme des termes diagonaux de $M(u)$, sa trace, ne dépend donc pas du choix de la base, ce qui permet de définir la **trace** de u :

$$\text{tr}(u) = \text{tr}(M(u)) = \sum_i m_{ii} = \sum_i \lambda_i.$$

La trace et le déterminant d'une matrice sont des invariants dans sa classe de similitude. En effet, des matrices semblables représentant le même endomorphisme u dans différentes bases, leur trace est la somme des valeurs propres de u et leur déterminant est le produit de ces valeurs propres.

Proposition 7.1.1. *Les racines du polynôme caractéristique sont les valeurs propres, et à chaque valeur propre correspond au moins un vecteur propre.*

Démonstration. De $P_u(\lambda_i) = 0$ nous déduisons que $u - \lambda_i$, λ_i étant mis pour $\lambda_i \text{ id}_E$, n'est pas inversible, que son noyau n'est pas nul et contient un $x \neq 0$; de $(u - \lambda_i)(x) = 0$, ou $u(x) = \lambda_i x$, nous concluons que x est un vecteur propre et λ_i une valeur propre.

Réciproquement, si $\lambda_i \in \text{Spec}(u)$, il existe un $x \neq 0$ tel que $u(x) = \lambda_i x$, ou $(\lambda_i - u)(x) = 0$; $(\lambda_i - u)$ n'est pas inversible car non injective, son déterminant est nul, or $\det(\lambda_i - u) = P_u(\lambda_i)$; λ_i est donc racine de P_u . $\square$

Le noyau de l'application linéaire $u - \lambda$, si λ est une valeur propre de u , est un sous-espace vectoriel de E , appelé **espace caractéristique** relatif à cette valeur propre, noté K_λ .

Rappelons que l'ensemble des vecteurs propres associées à une valeur propre λ complété par 0 est un sous-espace vectoriel de E , appelé **espace propre** relatif à cette valeur propre, noté E_λ .

Les valeurs propres d'un projecteur p sont 1 et 0, les sous-espaces propres correspondants sont respectivement $\text{Im}(p)$ et $\text{Ker}(p)$.

Proposition 7.1.2. *Si $\lambda_1, \dots, \lambda_k$ sont des valeurs propres distinctes, des vecteurs propres associés, respectivement $x_1, \dots, x_k$, sont libres.*

Démonstration. C'est trivialement vrai pour $k = 1$ puisqu'un vecteur propre n'est pas nul. Supposons la proposition prouvée jusqu'à $k = j$ et passons au cas $j + 1$. Supposons qu'il existe une relation $\sum_{1 \leq i \leq j+1} a_i x_i = 0$, avec, quitte à changer l'ordre des x_i , $a_{j+1} \neq 0$. On aurait alors :

$$x_{j+1} = a_{j+1}^{-1} (a_1 x_1 + \dots + a_j x_j)$$

et :

$$u(x_{j+1}) = a_{j+1}^{-1} (a_1 \lambda_1 x_1 + \dots + a_j \lambda_j x_j) = \lambda_{j+1} x_{j+1}.$$

Si $\lambda_{j+1} = 0$, les autres λ_i ne sont pas nuls et cette équation lie $x_1, \dots, x_j$ qui sont libres par hypothèse, ce qui est absurde.

Si $\lambda_{j+1} \neq 0$, on a :

$$x_{j+1} = \lambda_{j+1}^{-1} a_{j+1}^{-1} (a_1 \lambda_1 x_1 + \dots + a_j \lambda_j x_j)$$

ce qui fournit une seconde relation liant les x_i . La différence de ces deux relations donne, en les multipliant par a_{j+1} :

$$a_1 (\lambda_{j+1}^{-1} \lambda_1 - 1) x_1 + \dots + a_j (\lambda_{j+1}^{-1} \lambda_j - 1) x_j = 0.$$

Comme les λ_i , $1 \leq i \leq j$, sont tous différents de λ_{j+1} , les coefficients ne sont pas nuls, ce qui contredit encore l'hypothèse de récurrence. $\square$

Soit $K_\lambda = \text{Ker}((u - \lambda)^r)$, **espace caractéristique** relatif à une valeur propre λ d'ordre r .

Remarquons que, si P et Q appartiennent à $\mathbb{K}[X]$ et u à $\mathfrak{L}(E)$, on a :

$$P(u) \circ Q(u) = (PQ)(u) = (QP)(u) = Q(u) \circ P(u).$$

Nous écrirons $(PQ)(u)$, ou simplement $PQ(u)$, pour $P(u) \circ Q(u)$.

Proposition 7.1.3. *L'ensemble K_λ est un sous-espace vectoriel de E contenant E_λ et stable par $P(u)$, $\forall P \in \mathbb{K}[X]$.*

Démonstration. L'ensemble K_λ contient E_λ , car si $x \in E_\lambda$, $(u - \lambda)(x) = 0$ et $(u - \lambda)^r(x) = 0$, puisque $r \geq 1$.

C'est un sous-espace vectoriel de E . En effet, si x et y appartiennent à K_λ , on a, quels que soient les scalaires α et β :

$$(u - \lambda)^r (\alpha x + \beta y) = \alpha (u - \lambda)^r (x) + \beta (u - \lambda)^r (y) = 0.$$

Il est stable par $P(u)$ car, si $x \in K_\lambda$:

$$((u - \lambda)^r P(u))(x) = (P(u)(u - \lambda)^r)(x) = P(u)(0) = 0$$

et $u(x)$ appartient à K_λ . □

Proposition 7.1.4. *Si λ une valeur propre de $u \in \mathfrak{L}(E)$, x un vecteur propre associé à λ et $P \in \mathbb{K}[X]$, alors $P(\lambda)$ est une valeur propre de $P(u)$ et x est un vecteur propre associé à $P(\lambda)$.*

Démonstration. On a $u^2(x) = u(\lambda x) = \lambda u(x) = \lambda^2 x$, d'où par récurrence $u^k(x) = \lambda^k x$. Si $P(u) = \sum_i a_i u^i$, on a :

$$P(u)(x) = \sum_i a_i u^i(x) = \sum_i a_i \lambda^i x = P(\lambda) x.$$

□

7.2 Diagonalisation

Un endomorphisme u est **diagonalisable** s'il existe une base dans laquelle sa matrice $M(u)$ est diagonale. Comme u et $M(u)$ ont même polynôme caractéristique, la diagonale de $M(u)$ est formée des valeurs propres de u , chacune figurant un nombre de fois égal à son ordre.

Proposition 7.2.1. *Soit $u \in \mathfrak{L}(E)$. Les assertions suivantes sont équivalentes :*

- (a) *u est diagonalisable,*
- (b) *la dimension de chaque espace propre E_{λ_i} est égale à l'ordre r_i de la valeur propre λ_i ,*
- (c) *E est somme directe des E_{λ_i} .*

Démonstration. (a) $\Rightarrow$ (b) Il existe une base (ϵ_i) dans laquelle $M(u)$ est diagonale, chaque ϵ_i est vecteur propre (d'après la définition de la matrice associée à un endomorphisme), et la dimension de E_{λ_i} est égale au nombre des ϵ_j associés λ_i , donc à son ordre de multiplicité r_i .

(b) $\Rightarrow$ (c) L'intersection de deux E_{λ_i} distincts est réduite à $\{0\}$ (proposition 6.1.2). La somme des E_{λ_i} est donc directe et sa dimension est égale à la somme des r_i , c'est-à-dire à n , la dimension de E .

(c) $\Rightarrow$ (a) La restriction de u à E_{λ_i} est l'homothétie de rapport λ_i . Elle est donc diagonalisable (d'ailleurs dans une base quelconque de E_{λ_i}), d'où r_i vecteurs propres appartenant à E_{λ_i} . Ceci pour chaque λ_i , ce qui donne n vecteurs propres et une base dans laquelle la matrice de u est diagonale. □

La différence entre cette diagonalisation, pour laquelle la source et le but ont la même base ($D = P^{-1}MP$) et celle obtenue par la méthode de Gauss consiste en ce que dans cette dernière (Gauss) les bases à la source et au but sont différentes ($D' = PMQ$).

Exemple 7.2.1. Considérons la matrice :

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

Les transformations $L_2 - 3/2 L_1$ et $L_1 - 2 L_2$ donnent la matrice :

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

D'autre part, le polynôme caractéristique de M étant $X^2 - 4X + 1$, les valeurs propres sont $2 + \sqrt{3}$ et $2 - \sqrt{3}$, d'où la forme diagonale :

$$D = \begin{pmatrix} 2 + \sqrt{3} & 0 \\ 0 & 2 - \sqrt{3} \end{pmatrix}.$$

Les matrices M et D' sont équivalentes alors que M et D sont semblables. On peut cependant remarquer que le produit des termes diagonaux, le déterminant, est le même pour D et D' .

7.3 Algorithme de Leverrier

Ce très remarquable algorithme dû à l'illustre astronome et mathématicien (1840) a été redécouvert dans les années 1948-1949 par des mathématiciens auxquels on l'attribue trop souvent (Souriau, Frame, Faddeev...). Il donne le polynôme caractéristique d'une matrice de $\mathfrak{M}_n(\mathbb{K})$ ($\mathbb{K}$ de caractéristique nulle), donc son déterminant, et son inverse si elle est inversible, après $\sim 2n^4$ opérations. Il est facile à programmer, même sur une calculatrice programmable.

Décrivons d'abord la méthode. Soit $M = (m_{ij}) \in \mathfrak{M}_n(\mathbb{K})$. On pose $M_0 = I_n$.

$$\begin{array}{lll} M'_1 = M M_0 & a_1 = \text{tr}(M'_1) & M_1 = M'_1 - a_1 I_n \\ M'_2 = M M_1 & a_2 = \frac{1}{2} \text{tr}(M'_1) & M_2 = M'_2 - a_2 I_n \\ \vdots & \vdots & \vdots \\ M'_{n-1} = M M_{n-2} & a_{n-1} = \frac{1}{n-1} \text{tr}(M'_{n-1}) & M_{n-1} = M'_{n-1} - a_{n-1} I_n \\ M'_n = M M_{n-1} & a_n = \frac{1}{n} \text{tr}(M'_n) & M_n = M'_n - a_n I_n = 0. \end{array}$$

Le polynôme caractéristique est alors $P_M = X^n - \sum_i a_i X^{n-i}$, le déterminant est a_n et, si $a_n \neq 0$, $M^{-1} = \frac{1}{a_n} M_{n-1}$.

Appliquons-la à un exemple ($M = M'_1$) :

$$\begin{aligned} M'_1 &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \\ 4 & 1 & 1 \end{pmatrix}, \quad a_1 = 2, \quad M_1 = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 1 \\ 3 & 2 & -1 \\ -2 & 1 & 1 \end{pmatrix}; \\ M'_2 &= \begin{pmatrix} 1 & 4 & 1 \\ 4 & 1 & 4 \\ 3 & 0 & 0 \end{pmatrix}, \quad a_2 = 6, \quad M_2 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 4 & 1 & -2 \end{pmatrix}; \\ M'_3 &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad a_3 = 3, \quad M_3 = 0. \end{aligned}$$

Le polynôme caractéristique de M est donc $X^3 - 2X^2 - 6X - 3$ et la matrice inverse :

$$M^{-1} = \frac{1}{a_3} M_2 = \frac{1}{3} \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 4 & 1 & -2 \end{pmatrix}.$$

Justifions la méthode. Les a_i , fonctions des valeurs propres λ_i , ne dépendent pas de la base choisie. Calculons donc dans une base dans laquelle $M(u)$ est triangulaire supérieure, et dont la diagonale est constituée des λ_i comptés selon leur ordre (proposition 3.9.1). Remarquons que les M'_i et les M_i sont aussi triangulaires supérieures. Il reste à calculer leurs termes diagonaux; écrivons le j -ième :

$$\begin{aligned} (M'_1)_{jj} &= \lambda_j, \\ (M_1)_{jj} &= -\sum_{i \neq j} \lambda_i, \\ (M'_2)_{jj} &= -\lambda_j \sum_{i \neq j} \lambda_i, \\ (M_2)_{jj} &= \sum_{i, h \neq j} \lambda_i \lambda_h, \\ &\dots \\ (M'_k)_{jj} &= (-1)^{k-1} \sum_{(i_1, \dots, i_k), \exists h: i_h = j} \lambda_{i_1} \dots \lambda_{i_k}, \\ (M_k)_{jj} &= (-1)^k (\prod_{h=i_1, \dots, i_k, h \neq k} \lambda_h), \\ &\dots \\ (M'_n)_{jj} &= (-1)^{n-1} \prod_i \lambda_i. \end{aligned}$$

Démontrons ceci par récurrence. Supposons que le j -ième terme diagonal de M'_k soit la somme de tous les produits $\lambda_{i_1} \dots \lambda_{i_k}$ étendue à tous les k -uples $i_1 \dots i_k$ contenant j , multipliée par $(-1)^{k-1}$. Le j -ième terme diagonal de M_k est alors le produit $\lambda_{i_1} \dots \lambda_{i_k}$, les i_h étant différents de j , multiplié par $(-1)^k$. Le j -ième terme diagonal de M'_{k+1} est le précédent multiplié par λ_j . Il est donc égal à la somme de tous les produits $\lambda_{i_1} \dots \lambda_{i_{k+1}}$ étendue à tous les $k+1$ -uples contenant j . L'hypothèse étant vérifiée pour M'_1 , la récurrence est achevée.

La trace de M'_k est égale à la somme des produits $\lambda_{i_1} \dots \lambda_{i_k}$ étendue à l'ensemble des k -uples contenus dans $\{1, 2, \dots, n\}$ multipliée par k , chaque produit apparaissant k fois, et par $(-1)^{k-1}$.

Il reste à comparer avec le développement de P_u :

$$P_u(X) = (X - \lambda_1) \dots (X - \lambda_n) = X^n + \sum_{1 \leq k \leq n} p_k X^{n-k}.$$

Chaque k -uple $i_1, \dots, i_k$ donne le monôme $(-1)^k \prod_{j=i_1, \dots, i_k} \lambda_j X^{n-k}$, et p_k est la somme de tous ces monômes, et la trace de M'_k est bien égale à p_k .

Montrons que $M_n = P_u(M)$. L'hypothèse de récurrence :

$$H(k) : M_k = M^k - a_1 M^{k-1} - \dots - a_{k-1} M - a_k I_n$$

est vérifiée pour $k = 1$, avec $a_0 = 0$, car $M_1 = M - a_1 I_n$. On a ensuite :

$$M_{k+1} = M^{k+1} - a_1 M^k - \dots - a_k M - a_{k+1} I_n,$$

ce qui est justement $H(k+1)$; la récurrence est achevée. Nous verrons plus loin que $P_u(M) = 0$ (théorème 6.6.1, de Cayley-Hamilton), ce qui prouve que $M_n = 0$. Quant à la matrice inverse, si $a_n \neq 0$, $M_n = 0$ implique $M'_n = a_n I_n$, or $M'_n = M M_{n-1}$.

7.4 Méthode des puissances

Soit une matrice $M \in \mathfrak{M}_n(\mathbb{R})$ dont les valeurs propres sont réelles et telles que $\lambda_1 > \lambda_2 > \dots > \lambda_n$, et $\mathfrak{B} = (e_i)$ la base canonique de $\mathbb{R}^n$. Nous nous proposons de déterminer les λ_i et les vecteurs propres associés (ϵ_i) .

On peut écrire $\epsilon_1 = \sum_j a_j e_j$, et supposer, à permutation près sur les e_i , que $a_1 \neq 0$ et même que $a_1 = 1$. On a alors $e_1 = \epsilon_1 + \sum_{2 \leq i \leq n} b_i e_i$ pour tout entier naturel p :

$$\frac{1}{\lambda_1^p} M^p(e_1) = \epsilon_1 + \sum_{i \geq 2} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^p \epsilon_i.$$

Puisque λ_1 est plus grand que les autres λ_j :

$$\lim_{p \rightarrow \infty} \frac{1}{\lambda_1^p} M^p(e_1) = \epsilon_1$$

ce qui donne le premier vecteur propre avec une approximation dépendant de p et du rapport λ_1/λ_2 . Si $p = 2^q$, ce calcul nécessite $\sim qn^3$ opérations.

Si $a_1 = 0$, la première colonne qui ne tend pas vers 0 donne ϵ_1 , ce qui indique la permutation à effectuer.

Dans la base $\mathfrak{B}_1 = (\epsilon_1, e_2, \dots, e_n)$, la matrice devient $M^{(1)}$ ($o(n^3)$ opérations). Sa première colonne est $C_1 = (\lambda_1, 0, \dots, 0)$. Nous allons construire une base dans laquelle, de plus, la première ligne sera $L_1 = (\lambda_1, 0, \dots, 0)$. Le problème sera ainsi ramené à la dimension $n - 1$.

Soit P la matrice obtenue à partir de I_n en remplaçant la première ligne par $(1, c_2, \dots, c_n)$. On obtient P^{-1} en changeant les c_i en $-c_i$. La première ligne de la matrice $M^{(2)} = P^{-1} M^{(1)} P$ ($o(n^3)$ opérations à cause des 0) est :

$$(\lambda_1, m_2, \dots, m_n)$$

avec :

$$m_j = m_{1j} + \lambda_1 c_j - m_{2j}^{(1)} c_2 - m_{3j}^{(1)} c_3 - \dots - m_{nj}^{(1)} c_n.$$

Le système $m_j = 0$ pour $2 \leq j \leq n$ donne, après $\sim 2n^3/3$ opérations, des c_i rendant nuls les m_j . On a alors, après $\sim (q + n/3)n^3$ opérations :

$$M^{(2)} = \begin{pmatrix} \lambda_1 & 0_{1n} \\ 0_{n1} & M' \end{pmatrix}$$

et on recommence avec $M' \in \mathfrak{M}_{n-1}(\mathbb{R})$. La réduction de la matrice sera complète après un nombre d'opérations équivalent à $(q+1/3) \sum_{1 \leq k \leq n} k^3 = (q+1/3) n^4/4$. L'expression des ϵ_i dans la base initiale nécessite $o(n^3)$ opérations.

Exemple 7.4.1 Nous nous proposons de diagonaliser la matrice :

$$M = \begin{pmatrix} 11 & 8 & -4 \\ 2 & 26 & -4 \\ -4 & 2 & 17 \end{pmatrix}.$$

Etudions l'évolution du rapport de la première composante et de la troisième sur la deuxième pour le premier vecteur colonne de M^{2^q} . Pour $q = 3$ on trouve 0,521 et $-0,080$, pour $q = 4$, 0,500762 et $-0,003$, pour $q = 5$, 0,5000011 et $-0,0000046$, pour $q = 6$, 0,5000000001 et -1.10^{-11} ... ce qui nous amène à la conclusion que le premier vecteur propre est $\epsilon_1 = {}^t(1, 2, 0)$. Comme $M\epsilon_1 = 27\epsilon_1$, la première valeur propre est $\lambda_1 = 27$. Dans la base (ϵ_1, e_2, e_3) , la matrice s'écrit :

$$M^{(1)} = \begin{pmatrix} 27 & 8 & -4 \\ 0 & 10 & 4 \\ 0 & 2 & 17 \end{pmatrix}.$$

Posons $e'_2 = e_2 + c_2\epsilon_1$ et $e'_3 = e_3 + c_3\epsilon_1$. Dans la base (ϵ_1, e'_2, e'_3) , la matrice s'écrit :

$$M^{(2)} = \begin{pmatrix} 27 & m_2 & m_3 \\ 0 & 10 & 4 \\ 0 & 2 & 17 \end{pmatrix}$$

avec $m_2 = 17c_2 - 2c_3 + 8$ et $m_3 = -4c_2 + 10c_3 - 4$. Le système $m_2 = m_3 = 0$ a pour solution $c_2 = -4/9$ et $c_3 = 2/9$, d'où, dans la base initiale, $e'_2 = \frac{1}{9} {}^t(-4, 1, 0)$ et $e'_3 = \frac{1}{9} {}^t(2, 4, 9)$. En restriction à $\text{Vect}(e'_2, e'_3)$, nous obtenons la matrice :

$$M' = \begin{pmatrix} 10 & 4 \\ 2 & 17 \end{pmatrix}.$$

Etudions le rapport des deux composantes de la première colonne de M'^{2^q} . Pour M'^8 , on trouve 1,931818, pour M'^{16} , 1,999725, pour M'^{32} , 1,999999996 ... d'où le vecteur propre de M' $\epsilon'_2 = {}^t(1, 2)$ dans la base (e'_2, e'_3) , et le vecteur propre $\epsilon_2 = {}^t(0, 1, 2)$ de $M^{(2)}$ dans la base (ϵ_1, e'_2, e'_3) . Comme $M'\epsilon'_2 = 18\epsilon'_2$, la valeur propre associée est $\lambda_2 = 18$. Exprimons ϵ_2 dans la base initiale :

$$\epsilon_2 = e'_2 + 2e'_3 = {}^t(0, 1, 2).$$

Dans la base (ϵ_2, e'_3) la matrice devient :

$$M'_1 = \begin{pmatrix} 18 & 4 \\ 0 & 9 \end{pmatrix},$$

et dans la base $(\epsilon_2, e'_3 + c\epsilon_2)$:

$$M'_2 = \begin{pmatrix} 18 & 4 + 9c \\ 0 & 9 \end{pmatrix}.$$

On obtient donc $\epsilon_3 = e'_3 - \frac{4}{9}\epsilon_2$, et $\lambda_3 = 9$. Exprimons ϵ_3 dans la base initiale :

$$\epsilon_3 = \frac{1}{9} {}^t(2, 4, 9) - \frac{4}{9} {}^t(0, 1, 2) = \frac{1}{9} {}^t(2, 0, 1).$$

On peut prendre $\epsilon_3 = {}^t(2, 0, 1)$. ▽

Si la matrice à réduire est symétrique, ses vecteurs propres sont orthogonaux (théorème 6.8.1), et, pour le calcul des e'_i on peut remplacer la résolution du système $(m_i = 0)$ ($\sim 2n^3/3$ opérations) par :

$$e'_i = e_i - \frac{e_i \cdot \epsilon_1}{\|\epsilon_1\|^2} \epsilon_1$$

($o(n^3)$ opérations).

Exemple 7.4.2 Réduisons la matrice symétrique :

$$S = \begin{pmatrix} 1 & 1,41 & 1,73 \\ 1,41 & 2 & 2,23 \\ 1,73 & 2,23 & 3 \end{pmatrix}.$$

Le calcul de S^{32} donne le premier vecteur propre :

$$\epsilon_1 = \begin{pmatrix} 1 \\ 1,352215 \\ 1,683416 \end{pmatrix}$$

et $S\epsilon_1 = 5,818933\epsilon_1 = \lambda_1\epsilon_1$ donne la première valeur propre. Si :

$$e'_2 = e_2 - \frac{e_2 \cdot \epsilon_1}{\|\epsilon_1\|^2} \epsilon_1 = \begin{pmatrix} -0,238807 \\ 0,677082 \\ -0,402012 \end{pmatrix}$$

et :

$$e'_3 = e_3 - \frac{e_3 \cdot \epsilon_1}{\|\epsilon_1\|^2} \epsilon_1 = \begin{pmatrix} -0,297299 \\ -0,402012 \\ 0,499523 \end{pmatrix}$$

la matrice devient dans la base (ϵ_1, e'_2, e'_3) :

$$S^{(1)} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0,093377 & -0,109332 \\ 0 & -0,143617 & 0,087689 \end{pmatrix}.$$

Nous sommes ramenés à l'étude de la matrice :

$$S' = \begin{pmatrix} 0,093377 & -0,109332 \\ -0,143617 & 0,087689 \end{pmatrix}.$$

Elle n'est pas symétrique, mais ses vecteurs propres sont orthogonaux car ce sont des vecteurs propres de S . Le calcul de S'^{16} nous donne le vecteur propre :

$$\epsilon_2 = \begin{pmatrix} 1 \\ -1,120400 \end{pmatrix}$$

dans la base (e'_2, e'_3) , d'où le deuxième vecteur propre de S dans la base initiale :

$$\epsilon_2 = \begin{pmatrix} 0,020354 \\ 0,243396 \\ -0,207600 \end{pmatrix}.$$

La valeur propre associée est $\lambda_2 = 0,215873$. Dans la base (ϵ_2, e'_3) , la matrice devient :

$$S'^{(1)} = \begin{pmatrix} \lambda_2 & -0,109332 \\ 0 & -0,034806 \end{pmatrix}.$$

Les derniers éléments propres sont :

$$\epsilon_3 = e'_3 - \frac{e'_3 \cdot \epsilon_2}{\|\epsilon_2\|^2} \epsilon_2 = \begin{pmatrix} -0,256176 \\ 0,089738 \\ 0,080094 \end{pmatrix}$$

dans la base initiale et $\lambda_3 = -0,034806$. ▽

7.5 Endomorphismes et matrices nilpotents

Un endomorphisme $u \in \mathcal{L}(E)$ ($\dim(E) = n$) est **nilpotent** s'il existe un entier naturel $r \geq 2$, le **degré** de u , tel que $u^r = 0$ et $u^{r-1} \neq 0$. Nous excluons ainsi le cas $u = 0$, sans intérêt.

Une matrice $M \in \mathfrak{M}_n(\mathbb{K})$ est **nilpotente** si elle est la matrice d'un endomorphisme nilpotent, c'est-à-dire s'il existe un entier naturel $r \geq 2$, le **degré** de M , tel que $M^r = 0$ et $M^{r-1} \neq 0$.

Proposition 7.5.1. *Un endomorphisme nilpotent $u \in \mathcal{L}(E)$ ($\dim(E) = n$) admet 0 comme unique valeur propre. Il existe une base dans laquelle la matrice de u est triangulaire supérieure de termes diagonaux nuls.*

Démonstration. Il existe $x \in E$ tel que $u^{r-1}(x) = y \neq 0$; comme $u(y)$, égal à $u^r(x)$, est nul, 0 est valeur propre de u . Remarquons que si $u(x) = \lambda x$, alors $0 = u^r(x) = \lambda^r x$, ce qui prouve que 0 est l'unique valeur propre de u . La seconde partie de la proposition est alors une conséquence de la proposition 4.9.1. $\square$

Proposition 7.5.2. *Si $u \in \mathfrak{L}(E)$ est nilpotent de degré r et si $x \in E$ est tel que $u^{r-1}(x) \neq 0$, alors $x, u(x), \dots, u^{r-1}(x)$ sont linéairement indépendants.*

Démonstration. La relation $u^{r-1}(x) \neq 0$ assure l'existence d'un $x \in E$ tel que $u^{r-1}(x) \neq 0$. Si nous appliquons u^{r-1} à une relation de dépendance linéaire :

$$a_0 x + a_1 u(x) + \dots + a_{r-1} u^{r-1}(x) = 0$$

nous obtenons $a_0 u^{r-1}(x) = 0$, d'où $a_0 = 0$. En lui appliquant maintenant u^{r-2} , nous obtenons $a_1 u^{r-1}(x) = 0$, d'où $a_1 = 0$... ainsi de suite jusqu'à $a_{r-1} = 0$. La relation est donc triviale et les éléments libres. $\square$

Proposition 7.5.3. *Le degré d'un endomorphisme nilpotent $u \in \mathfrak{L}(E)$ est au plus égal à la dimension n de E .*

Démonstration. Si u est de degré r il existe un $x \in E$ tel que $u^{r-1}(x) \neq 0$ et $u^r(x) = 0$; $x, u(x), \dots, u^{r-1}(x)$ sont libres (proposition 6.4.2) et engendrent un sous-espace vectoriel de E de dimension r , d'où la conclusion $r \leq n$. $\square$

Proposition 7.5.4. *Il existe une base dans laquelle la matrice $T = (t_{ij})$ de u , endomorphisme de E ($\dim(E) = n$) nilpotent de degré $r \leq n$ s'écrit :*

$$T = \begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & a_{n-1} \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

chaque a_i étant égal à 1 ou à 0.

Démonstration. Il existe un élément ϵ_n de E tel que $u^{r-1}(\epsilon_n) \neq 0$. On pose $\epsilon_{n-1} = u(\epsilon_n)$, et $u^{r-2}(\epsilon_{n-1}) \neq 0$... ainsi de suite, jusqu'à :

$$\epsilon_{n-r+1} = u^{r-1}(\epsilon_n) \neq 0$$

où :

$$u(\epsilon_{n-r+1}) = u^r(\epsilon_n) = 0.$$

Si $r = n$, nous obtenons ainsi une base $(\epsilon_i)_{1 \leq i \leq n}$ dans laquelle la matrice de u a la forme annoncée (les a_i sont tous égaux à 1).

Si $r < n$, soit E' le sous-espace engendré par les $(\epsilon_i)_{n-r+1 \leq i \leq n}$, stable par u ($u(E') \subset E'$), donc par $P(u)$ pour tout polynôme $P \in \mathbb{K}[X]$, et E'' un

supplémentaire de E' . Soit $x \in E''$, $x \neq 0$, et $u(x) = x' + x''$, $x' \in E'$, $x'' \in E''$, c'est-à-dire :

$$u(x) = x'' + \sum_{n-r+1 \leq j \leq n} \alpha_j \epsilon_j,$$

nous voyons que α_n doit être nul, car, sinon, $u^r(x)$ serait égal à $\alpha_n u^{r-1}(\epsilon_n)$, et ne serait pas nul alors que $u^r = 0$. Posons :

$$\epsilon_{n-r} = x - \sum_{n-r+1 \leq j \leq n-1} \alpha_j \epsilon_{j+1},$$

de sorte que $u(\epsilon_{n-r}) = x''$, puis $\epsilon_{n-r-1} = u(\epsilon_{n-r}) (= x'')$, jusqu'à un ϵ_i tel que $u(\epsilon_i) = 0$. Le sous-espace engendré par $\epsilon_{n-r}, \dots, \epsilon_i$ est stable par u par construction et sa somme avec E' est directe. Nous itérons la méthode jusqu'à trouver ϵ_1 . Dans la base (ϵ_i) la matrice de u a la forme annoncée, puisque ou bien $u(\epsilon_i) = \epsilon_{i-1}$ ou bien $u(\epsilon_i) = 0$. $\square$

La matrice de $\mathfrak{M}_r(\mathbb{K})$:

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

est le **bloc de Jordan** de dimension r et de valeur propre λ .

Exemple 7.5.1. La matrice :

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

d'un endomorphisme de $\mathbb{K}^4$ dans sa base canonique a un carré non nul, et $M^3 = 0$. Son degré (de nilpotence) est donc égal à 3. Cherchons un vecteur $x \in \mathbb{K}^4$ tel que $M^2 x \neq 0$. On voit que :

$$\begin{aligned} M e_4 &= e_2 + e_4, \\ M^2 e_4 &= e_2 + e_3 + e_4. \end{aligned}$$

Posons donc $\epsilon_4 = e_4$ et $\epsilon_3 = M e_4 = e_2 + e_4$, puis $\epsilon_2 = e_1$. Comme $M e_2 = 0$, posons $\epsilon_1 = e_2$. Dans la base (ϵ_i) la matrice devient :

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

et on peut l'écrire :

$$J = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$$

où B_1 est le bloc diagonal de dimension 1 réduit à 0 et B_2 le bloc de Jordan :

$$J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Les matrices de passage sont :

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

On peut aussi commencer par les vecteurs propres, ici au nombre de deux : $e'_1 = e_1$ et $e'_2 = e_2 + e_3 + e_4$. On cherche ensuite un vecteur $e'_3 = (a, b, c, d)$ tel que $M e'_3 = e'_2$, d'où $(0, c - b, d - b, d - b) = (0, 1, 1, 1)$, et, avec $b = 0$, $e'_3 = (0, 0, 1, 1)$, puis un vecteur e'_4 tel que $M e'_4 = e'_3$, d'où par exemple $e'_4 = e_4$. Les matrices de passage sont :

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

la seconde méthode est souvent plus pratique car il peut être difficile, par exemple en grande dimension, de trouver un x tel que $u^{r-1}(x) \neq 0$. ∇

Proposition 7.5.5. *Une matrice triangulaire supérieure T (ou inférieure) dont les termes diagonaux sont nuls est nilpotente.*

Démonstration. Soit $u \in \mathcal{L}(E)$ ($\dim(E) = n$) l'endomorphisme représenté par T dans la base (ϵ_i) . Posons $E_0 = \{0\}$, puis $E_i = \text{Vect}(\epsilon_1, \dots, \epsilon_i)$, de sorte que :

$$E_0 \subset E_1 \subset \dots \subset E_n = E.$$

On a alors $u(E_j) \subset E_{j-1}$, d'où $u^n(E) \subset E_0$, et $u^n = 0$. $\square$

7.6 Forme de Jordan

Proposition 7.6.1. *L'endomorphisme $P_u(u)$ est nilpotent, de degré $r \leq n$.*

Démonstration. Plaçons-nous dans une base dans laquelle la matrice $T = (t_{ij})$ de u est triangulaire supérieure (proposition 3.9.1), les t_{ii} étant égaux aux valeurs propres, chacune figurant (à la suite) un nombre de fois égal à son ordre. La matrice $P_u(T)$ est triangulaire supérieure (proposition 2.5.9) et ses termes diagonaux, égaux aux $P_u(t_{ii})$ sont nuls. D'où la conclusion (propositions 6.4.5 et 6.4.3). $\square$

Nous verrons plus loin que $r = 1$ (théorème de Hamilton-Cayley, 7.6.1). Posons :

$$\begin{aligned} P &= P_u^r, \\ P_i &= \prod_{j \neq i} (X - \lambda_j)^{r_j}, \\ Q_i &= (X - \lambda_i)^{r_i}, \end{aligned}$$

de sorte que $P_i Q_i = P$, et revenons aux $K_{\lambda_i} = \text{Ker}(Q_i)$, dont nous savons qu'ils sont stables par u (proposition 7.1.3).

Proposition 7.6.2. *L'intersection de deux $K_i = K_{\lambda_i}$ est réduite à $\{0\}$.*

Démonstration. Si $x \in K_i \cap K_j$ ($i \neq j$), on a $P_i(u)(x) = P_j(u)(x) = 0$, et P_i et P_j sont premiers entre eux; il existe donc des polynômes A et B tels que $AP_i + BP_j = 1$. On a alors $((AP_i + BP_j)(u))(x) = x$. Or $A(u)((P_i(u)(x))$ et $B(u)((P_j(u)(x))$ sont nuls et $x = 0$. $\square$

Proposition 7.6.3. *L'espace vectoriel E est somme directe des K_i .*

Démonstration. Les polynômes P_i sont premiers entre eux (ils n'ont pas de facteur commun), et l'identité de Bézout (pour les polynômes) nous assure de l'existence de polynômes R_i tels que $\sum_i R_i P_i = 1$. On a donc pour tout $x \in E$:

$$\left(\sum_i R_i(u) P_i(u) \right)(x) = x.$$

Posons $((R_i P_i)(u))(x) = x_i$, d'où $x = \sum_i x_i$. Comme :

$$Q_i(u)(x_i) = (R_i Q_i P_i)(u)(x) = 0$$

x_i appartient à K_i pour tout i et E est la somme des K_i . Il reste à voir que cette somme est directe.

Choisissons des éléments non nuls $x_i \in K_i$, $1 \leq i \leq n$, et considérons une relation $\sum_{1 \leq i \leq n} \alpha_i x_i = 0$. appliquons au premier membre de cette équation l'endomorphisme $(Q_1 \dots Q_{n-1})(u)$. On a, pour $i < n$:

$$(Q_1 \dots Q_{n-1})(u)(x_i) = \left(\prod_{j \neq i} Q_j(u) \right)(Q_i(u)(x_i)) = 0$$

puisque x_i appartient au noyau de $Q_i(u)$. Et pour x_n :

$$(Q_1 \dots Q_{n-1})(u)(x_n) \neq 0$$

car $y = (Q_2 \dots Q_{n-1})(u)(x_n)$ appartient à K_n , qui est stable par $P(u)$ quel que soit le polynôme $P \in \mathbb{K}[X]$, et $Q_1(u)(y)$ n'est pas nul car $y \notin K_1$. Il s'ensuit que $\alpha_n = 0$. On montre ensuite, par la même méthode, que $\alpha_{n-1} = 0$, et on descend ainsi jusqu'à $\alpha_1 = 0$. La relation est donc triviale, et la somme des K_i est directe. $\square$

Théorème 7.6.1 (de Hamilton-Cayley). *Tout endomorphisme u de E est racine de son polynôme caractéristique :*

$$P_u(u) = 0,$$

de même que toute matrice associée à u .

Démonstration. Tout élément $x \in E$ est somme d'éléments $x_i \in K_i$ (proposition 6.5.3) et donc $P_u(u)(x)$ est somme des $P_u(u)(x_i)$, et $(u - \lambda_i)^{r_i}(x_i) = 0$ par définition de K_i . On a donc :

$$\begin{aligned} P_u(u)(x_i) &= \left(\prod_j (u - \lambda_j)^{r_j} \right) (x_i) \\ &= \left(\prod_{j \neq i} (u - \lambda_j)^{r_j} \right) (u - \lambda_i)^{r_i} (x_i) \\ &= \left(\prod_{j \neq i} (u - \lambda_j)^{r_j} \right) (0) \\ &= 0, \end{aligned}$$

d'où $P_u(u)(x) = 0$ et $P_u(u) = 0$. La matrice associée à l'endomorphisme nul dans une base quelconque $\mathfrak{B}$ est nulle : $P_u(M_{\mathfrak{B}}(u)) = 0$. $\square$

Le **polynôme minimal** de u , ou des matrices qui le représentent, est le polynôme P non nul de plus petit degré tel que $P(u) = 0$. Le polynôme minimal de l'identité est X , celui de la matrice M de l'exemple 6.5.1 est X^3 . Lorsque u est diagonalisable, son polynôme minimal est le produit $\prod_{\lambda_i \in \text{Spec}(u)} (X - \lambda_i)$, en oubliant les ordres de multiplicité.

Étudions maintenant la restriction u_i de $u \in \mathfrak{L}(E)$ à K_i ; u_i est un endomorphisme de K_i ($u(K_i) \subset K_i$), son unique valeur propre est λ_i , d'ordre r_i .

Proposition 7.6.4. *La dimension de K_i est égale à r_i .*

Démonstration. D'après sa définition, le degré du polynôme caractéristique d'un endomorphisme est égal à la dimension de l'espace. Le polynôme caractéristique de u_i , $Q_i = (X - \lambda_i)^{r_i}$, est de degré r_i et donc $\dim(K_i) = r_i$. $\square$

Proposition 7.6.5. *Il existe une base de chaque K_i dans laquelle la matrice de u_i est la somme d'une matrice scalaire $\lambda_i I_{r_i}$ et d'une matrice du type T défini dans la proposition 7.5.4.*

Démonstration. Par définition de K_i l'application $(u_i - \lambda_i) \in \mathfrak{L}(K_i)$ est nilpotente, d'où la forme T de sa matrice associée dans une certaine base :

$$M(u_i) = M(u_i - \lambda_i) + M(\lambda_i i_{K_i}) = T + \lambda_i I_{r_i}$$

$\square$

Théorème 7.6.2 (de Jordan). *Soit $u \in \mathfrak{L}(E)$. Il existe une base de E dans laquelle $M(u) = (m_{ij})$ est diagonale par blocs, chaque bloc étant diagonal ou de*

Jordan :

$$M(u) = \begin{pmatrix} B_1 & 0 & \dots & \dots & 0 \\ 0 & B_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & B_h \end{pmatrix}$$

si u possède h valeurs propres distinctes. Précisément, la diagonale (principale) contient les valeurs propres, chaque λ_i figurant r_i fois à la suite, et la diagonale au-dessus ne contient que des 0 ou des 1 ($M(u)_{i,i+1} = 0$ ou 1).

Démonstration. Nous savons que E est somme directe de K_i (proposition 6.6.3) et que ceux-ci sont stables par u ($u(K_i) \subset K_i$, proposition 6.1.3). Prenons pour chaque K_i une base $\mathfrak{B}_i$ dans laquelle $M(u_i)$ prend la forme $D + T$ de la proposition 6.6.5, et pour base $\mathfrak{B}$ de E la réunion des $\mathfrak{B}_i$. Dans la base $\mathfrak{B}$ la matrice de u a la forme annoncée. Soit en effet un élément ϵ_i de $\mathfrak{B}$; il appartient à $\mathfrak{B}_j$ pour un certain j , par définition de $\mathfrak{B}$, et $u(\epsilon_i) = u_j(\epsilon_i)$ est égal à $\lambda_j \epsilon_i$ ou à $\lambda_j \epsilon_i + \epsilon_{j-1}$. La i -ème colonne de $M(u)$ est donc ou bien :

$$C_i = u(\epsilon_i) = {}^t(0, \dots, 0, \lambda_j, 0, \dots, 0)$$

ou bien :

$$C_i = {}^t(0, \dots, 0, 1, \lambda_j, 0, \dots, 0),$$

λ_j étant en i -ème position. □

Evaluons le nombre d'opérations pour obtenir la forme de Jordan d'une matrice $n \times n$. Le calcul du polynôme caractéristique par l'algorithme de Leverrier demande $\sim 2n^4$ opérations. La factorisation (approchée, mais d'aussi près qu'on le désire) de ce polynôme, qui sort du cadre de ce livre, en demande beaucoup moins (si n est grand) et le calcul de chacun des n vecteurs propres par la méthode de Gauss demande $\sim 2n^3/3$ opérations, soit en tout $\sim 8n^4/3$ opérations.

Exemple 7.6.1. Mettons la matrice :

$$M = \begin{pmatrix} 4 & 5 & -3 & -2 & -2 \\ 1 & 2 & 1 & -2 & -1 \\ 2 & 5 & -1 & -2 & -2 \\ -3 & -4 & 1 & 1 & 3 \\ 7 & 9 & -2 & -5 & -5 \end{pmatrix}$$

sous la forme de Jordan. Calculons $P_M(X)$ par l'algorithme de Leverrier, méthode la plus courte et la plus pratique. On a successivement :

$$M_1 = M - \text{tr}(M) I_5 = \begin{pmatrix} 3 & 5 & -3 & -2 & -2 \\ 1 & 1 & 1 & -2 & -1 \\ 2 & 5 & -2 & -2 & -2 \\ -3 & -4 & 1 & 0 & 3 \\ 7 & 9 & -2 & -5 & -6 \end{pmatrix},$$

$$M'_2 = MM_1 = \begin{pmatrix} 3 & 0 & 1 & -2 & -1 \\ 6 & 11 & -3 & -3 & -6 \\ 1 & 0 & 3 & -2 & -1 \\ 7 & 9 & -2 & -3 & -7 \\ 6 & 9 & -3 & -3 & -4 \end{pmatrix}, \quad \text{tr}(M'_2)/2 = 5,$$

$$M'_3 = M(M'_2 - 5I_5) = \begin{pmatrix} -7 & -6 & 5 & 5 & 1 \\ -9 & -15 & 0 & 9 & 9 \\ -1 & -6 & -1 & 5 & 1 \\ 8 & 12 & -4 & -1 & -8 \\ -27 & -36 & 9 & 18 & 21 \end{pmatrix}, \quad \text{tr}(M'_3)/3 = -1,$$

$$M'_4 = M(M'_3 + I_5) = \begin{pmatrix} -28 & -28 & 10 & 14 & 18 \\ -14 & -28 & 4 & 10 & 14 \\ -18 & -28 & 0 & 14 & 18 \\ -20 & -28 & 8 & 8 & 20 \\ -26 & -36 & 10 & 16 & 16 \end{pmatrix}, \quad \text{tr}(M'_4)/4 = -8,$$

$$M'_5 = M(M'_4 + 8I_5) = -4I_5, \quad \text{tr}(M'_5)/5 = -4,$$

d'où le polynôme caractéristique :

$$P_M(X) = X^5 - X^4 - 5X^3 + X^2 + 8X + 4.$$

La somme de ses coefficients étant nulle, -1 est racine. En divisant $P_M(X)$ par $(X + 1)$ on obtient $X^4 - 2X^3 - 3X^2 + 4X + 4$, -1 est encore racine, et une nouvelle division donne $X^3 - 3X^2 + 4$, qui s'annule pour $X = -1$, et une dernière division donne $X^2 - 4X + 4$, égal à $(X - 2)^2$. Les valeurs propres sont donc 2, double, et -1, triple.

Les vecteurs propres relatifs à la valeur propre 2 vérifient l'équation :

$$\begin{pmatrix} 2 & 5 & -3 & -2 & -2 \\ 1 & 0 & 1 & -2 & -1 \\ 2 & 5 & -3 & -2 & -2 \\ -3 & -4 & 1 & -1 & 3 \\ 7 & 9 & -2 & -5 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0.$$

Supprimons la première ligne (égale à la troisième). Les transformations successives : $L_2 - 2L_1$, $L_3 + 3L_1$, $L_4 - 7L_1$, $L_3 + 4L_2/5$, $L_4 - 9L_2/5$ mènent à un système dont les deux dernières lignes sont proportionnelles à $(0, 0, 0, 1, 0)$ et donnent $x_4 = 0$; les autres lignes donnent $x_1 + x_3 = x_5$ et $x_2 = x_3$. Ce sont les équations d'un plan ayant pour base, par exemple, $\epsilon_1 = {}^t(1, 0, 0, 0, 1)$ et $\epsilon_2 = {}^t(0, 1, 1, 0, 1)$. D'où le bloc diagonal $-I_2$.

Les vecteurs propres relatifs à la valeur propre 2 vérifient l'équation :

$$\begin{pmatrix} 5 & 5 & -3 & -2 & -2 \\ 1 & 3 & 1 & -2 & -1 \\ 2 & 5 & 0 & -2 & -2 \\ -3 & -4 & 1 & 2 & 3 \\ 7 & 9 & -2 & -5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0.$$

Les transformations successives :

$$(L_1 \leftrightarrow L_2), (L_2 - 5L_1, L_3 - 2L_1, L_4 + 3L_1L_5 - 7L_1),$$

$$(L_2 \leftrightarrow L_3), (L_3 - 10L_2, L_4 + 5L_2, L_5 - 12L_2),$$

$$(L_3 \leftrightarrow L_4/6), (L_4 + 12L_3, L_5 + 15L_3),$$

donnent un système de rang 4 avec $x_5 = 0$. La solution est donc une droite, engendrée par le vecteur $\epsilon_3 = {}^t(1, 0, 1, 1, 0)$. C'est l'unique vecteur propre associé à la valeur propre -1 .

Cherchons un vecteur ϵ_4 solution de l'équation $Mx = -x + \epsilon_3$:

$$\begin{pmatrix} 5 & 5 & -3 & -2 & -2 \\ 1 & 3 & 1 & -2 & -1 \\ 2 & 5 & 0 & -2 & -2 \\ -3 & -4 & 1 & 2 & 3 \\ 7 & 9 & -2 & -5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Les transformations précédentes donnent $\epsilon_4 = {}^t(0, 1, 0, 1, 1)$. On procède de même pour ϵ_5 ($M\epsilon_5 = -\epsilon_5 + \epsilon_4$) pour trouver $\epsilon_4 = {}^t(1, 0, 1, 0, 1)$. D'où le bloc de Jordan $J_3(-1)$. La forme réduite de M est finalement :

$$J = \begin{pmatrix} 2I_2 & 0 \\ 0 & J_3(-1) \end{pmatrix}.$$

▽

Proposition 7.6.6. *Un endomorphisme u de E est diagonalisable si et seulement si chacun de ses espaces propres est égal à l'espace caractéristique qui le contient.*

Démonstration. Si u est diagonalisable, E est somme directe des E_{λ_i} , de dimension r_i (proposition 6.2.1). Comme $E_{\lambda_i} \subset K_{\lambda_i}$ la dimension s_i de K_{λ_i} est supérieure ou égale à r_i pour tout i . D'autre part, E est somme directe des K_{λ_i} (Proposition 7.6.3) et sa dimension, n , somme de r_i , est aussi somme des s_i . On a donc pour tout i $s_i = r_i$ et $E_{\lambda_i} = K_{\lambda_i}$.

La réciproque découle de la proposition 7.2.1. □

7.7 Diagonalisation simultanée de deux endomorphismes

Proposition 7.7.1. *Des endomorphismes u et v commutent quel que soit v si et seulement si u est une homothétie ($u = \lambda i_E$, $\lambda \in \mathbb{K}$).*

Démonstration. Un homothétie commute évidemment avec tout endomorphisme. Réciproquement, si u n'est pas une homothétie, il existe des vecteurs x et y de E tels que $u(x) \neq u(y)$, et un endomorphisme v tel que $v(u(x)) = u(y)$ et $v(x) = x$. On a alors $u(v(x)) = u(x) \neq v(u(x))$. $\square$

Dorénavant, u et v sont deux endomorphismes diagonalisables de E .

Proposition 7.7.2. *La restriction de u à un sous-espace quelconque F (de dimension $m < n$) de E stable par u ($u(F) \subset F$) est diagonalisable dans une base de F que l'on peut compléter en une base de E dans laquelle la matrice de u est diagonale.*

Démonstration. Les valeurs propres de $u|_F : F \rightarrow F$ sont des valeurs propres de u . En effet, si $\lambda \in \mathbb{K}$ et $x \in F$ sont tels que $u|_F(x) = \lambda x$, comme $x \in E$, on a $u(x) = \lambda x$, et λ est valeur propre de u . Si $\lambda \in \text{Spec}(u|_F)$, l'espace caractéristique relatif à λ , notons-le $K_\lambda(F)$, est contenu dans K_λ , qui est égal à E_λ (proposition 6.6.6). Si $x \in K_\lambda(F)$, x appartient à E_λ , donc à l'espace propre de $u|_F$, noté F_λ ; ce dernier est donc égal à $K_\lambda(F)$, et F est somme directe des F_λ pour $\lambda \in \text{Spec}(u|_F)$, d'où l'on conclut que $u|_F$ est diagonalisable (proposition 6.2.1).

Soit $\lambda_i \in \text{Spec}(u)$, et G_i le supplémentaire de F_{λ_i} dans E_{λ_i} (si $\lambda_i \notin \text{Spec}(u|_F)$, on a $G_i = E_{\lambda_i}$). Formons une base $\mathfrak{B}_i$ de E_{λ_i} en complétant une base de F_{λ_i} par une base de G_i et soit $\mathfrak{B}$ la réunion des $\mathfrak{B}_i$. Plaçant d'abord les éléments de $\mathfrak{B}$ appartenant à F , nous obtenons une base de E dans laquelle la matrice de u est diagonale et dont les m premiers éléments forment une base de diagonalisation de $u|_F$. $\square$

Proposition 7.7.3. *Si u et v commutent, les espaces propres de l'un sont stables par l'autre.*

Démonstration. Soit $\lambda \in \text{Spec}(u)$, $x \in E_\lambda$. On a successivement $u(x) = \lambda x$, $(v \circ u)(x) = v(\lambda x) = \lambda v(x)$, $(u \circ v)(x) = \lambda v(x)$, d'où l'on conclut que $v(x)$ appartient à E_λ , qui est donc stable par v . $\square$

Théorème 7.7.1. *Les assertions suivantes sont équivalentes (u et v sont supposées diagonalisables) :*

- (a) u et v commutent,
- (b) il existe une base dans laquelle leurs matrices sont diagonales,
- (c) il existe un polynôme de $\mathbb{K}[X]$ tel que $v = P(u)$.

Démonstration. (a) $\Rightarrow$ (b) Si u et v commutent, les espaces propres de u sont stables par v (proposition 7.7.3). Si E_λ est un tel espace, la restriction de u à cet espace est l'homothétie de rapport λ , qui est diagonale dans toute base de E_λ . La restriction de v à E_λ est un endomorphisme de E_λ diagonalisable (proposition 7.7.2). Il existe donc une base de E_λ diagonalisant les restrictions

de u et de v . On termine en remarquant que E est somme directe des E_λ lorsque λ parcourt le spectre de u (proposition 7.2.1).

(b) $\Rightarrow$ (c) Soient (λ_i) les valeurs propres de u et μ_i celles de v . Il existe un polynôme P tel que, pour tout i , $P(\lambda_i) = \mu_i$ (même démonstration que pour la proposition 6.8.8), et $P(u) = v$.

(c) $\Rightarrow$ (a) Comme deux matrices diagonales commutent, on a dans la base de diagonalisation commune $M(u)M(v) = M(v)M(u)$. Or $M(u)M(v) = M(u \circ v)$ et $M(v)M(u) = M(v \circ u)$, et, d'après la proposition 3.5.4, ceci implique que $u \circ v = v \circ u$. $\square$

Remarque Lorsque les endomorphismes u et v commutent, il existe deux polynômes P et Q tels que $P(u) = v$ et $Q(v) = u$. On a donc $Q \circ P(u) = u$ et $P \circ Q(v) = v$. Ceci s'explique par le fait que $Q \circ P$ est égal à X plus un multiple du polynôme caractéristique de u (qui s'annule sur u , théorème 7.6.1, de Hamilton-Cayley), et que $P \circ Q$ est égal à un multiple de P_v plus X .

Exemple 7.7.1. Les endomorphismes u et v de $\mathbb{R}^3$ représentés dans la base canonique par les matrices :

$$U = \frac{1}{9} \begin{pmatrix} 1 & -10 & 14 \\ -10 & 10 & 4 \\ 14 & 4 & -2 \end{pmatrix} \quad \text{et} \quad V = \frac{1}{9} \begin{pmatrix} 10 & -46 & 50 \\ -46 & 37 & 4 \\ 50 & 4 & -11 \end{pmatrix}$$

commutent ($UV = VU$). Leurs polynômes caractéristiques :

$$P_u(X) = X^3 - X^2 - 4X + 4 \quad \text{et} \quad P_v(X) = X^3 - 4X^2 - 59X + 126$$

donnent leurs spectres : $\text{Spec}(u) = \{1, 2, -2\}$ et $\text{Spec}(v) = \{2, -7, 9\}$. Le vecteur propre de u relatif à la valeur propre 1 est solution du système :

$$\begin{pmatrix} 1 & -10 & 14 \\ -10 & 10 & 4 \\ 14 & 4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 9 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

dont les deux premières lignes sont libres et donnent pour $z = 1$ le système :

$$\begin{cases} 8x + 10y = 14 \\ 10x - y = 4 \end{cases}$$

de déterminant -108 dont la solution, donnée par les formules de Cramer, est :

$$x = \frac{-1}{108} \begin{vmatrix} 14 & 10 \\ 4 & -1 \end{vmatrix} = \frac{1}{2}, \quad y = \frac{-1}{108} \begin{vmatrix} 8 & 14 \\ 10 & 4 \end{vmatrix} = 1$$

d'où le vecteur propre ${}^t(1, 2, 2)$, dont l'image par V , ${}^t(2, 4, 4)$, correspond à la valeur propre 2 de v .

Le même calcul avec la valeur propre 2 de u donne le vecteur propre ${}^t(2, -2, 1)$, associé à la valeur propre 9 de v . Enfin, avec la valeur propre -2 de u on obtient le vecteur propre ${}^t(2, 1, -2)$, associé à la valeur propre -7 de v . La forme diagonale de v est donc $\text{Diag}(2, 9, -7)$ si celle choisie pour u est $\text{Diag}(1, 2, -2)$.

Le polynôme P , de degré inférieur à 3, donc de la forme $aX^2 + bX + c$, et tel que $v = P(u)$, est déterminé par les conditions $P(1) = 2$, $P(2) = 9$ et $P(-2) = -7$, qui donnent $P = X^2 + 4X - 3$.

Le polynôme Q tel que $P(v) = u$ est égal à $(-X^2 + 23X + 42)/84$. Le polynôme :

$$\begin{aligned} P \circ Q &= \frac{1}{7056}(X^4 - 46X^3 + 109X^2 + 9660X - 5292) \\ &= \frac{1}{7056}(X - 42)P_v + X \end{aligned}$$

est tel que $P \circ Q(v) = v$. De même :

$$\begin{aligned} Q \circ P &= -\frac{1}{84}(X^4 + 8X^3 - 13X^2 - 116X + 36) \\ &= -\frac{1}{84}(X + 9)P_u + X \end{aligned}$$

et $Q \circ P(u) = u$.

▽

7.8 Endomorphismes symétriques et antisymétriques

Une application ϕ de l'espace euclidien E dans lui-même est **symétrique** si, quels que soient les vecteurs x et y , on a $\phi(x) \cdot y = x \cdot \phi(y)$; elle est **antisymétrique** si $\phi(x) \cdot y = -x \cdot \phi(y)$.

Toute application symétrique, ou antisymétrique, est linéaire; c'est donc un endomorphisme. Montrons-le pour une application ϕ symétrique. On a, quels que soient les vecteurs et les scalaires écrits :

$$\begin{aligned} z \cdot (\phi(\lambda x + \mu y) - \lambda \phi(x) + \mu \phi(y)) &= \\ z \cdot \phi(\lambda x + \mu y) - z \cdot \lambda \phi(x) - z \cdot \mu \phi(y) &= \\ (\lambda x + \mu y) \cdot \phi(z) - \lambda x \cdot \phi(z) - \mu y \cdot \phi(z) &= 0, \end{aligned}$$

d'où $\phi(\lambda x + \mu y) - \lambda \phi(x) - \mu \phi(y) = 0$.

Un endomorphisme u de E est **symétrique**, ou **auto-adjoint**, si $u = u^*$.

Proposition 7.8.1. *L'espace vectoriel $\mathfrak{L}(E)$ est la somme directe des sous-espaces vectoriels $S(E)$, ensemble des endomorphisme symétriques, de dimension $n(n+1)/2$, et $A(E)$, ensemble des endomorphisme antisymétriques, de dimension $n(n-1)/2$.*

Démonstration. Tout endomorphisme u de E est la somme de l'endomorphisme symétrique $\frac{1}{2}(u + {}^t u)$ et de l'endomorphisme antisymétrique $\frac{1}{2}(u - {}^t u)$. Un endomorphisme à la fois symétrique et antisymétrique est évidemment nul, puisque le corps des scalaires n'est pas de caractéristique 2. On a donc $\mathfrak{L}(E) = S(E) \oplus A(E)$.

Considérons la matrice $M = (m_{ij})$ associée dans une base orthonormée à un endomorphisme u . Si u est symétrique, on a $u = {}^t u$ (théorème 5.2.3), $M = {}^t M$,

et $m_{ij} = m_{ji}$. Les $n(n+1)/2$ composantes m_{ij} pour $j \geq i$ sont libres, d'où la dimension de $S(E)$. Si u est antisymétrique, $m_{ij} = -m_{ji}$, $m_{ii} = 0$, et les $n(n-1)/2$ composantes m_{ij} pour $j > i$ sont libres. La dimension de $\mathfrak{L}(E)$ est égale à n^2 . $\square$

Remarquons que si $u \in \mathfrak{L}(E)$ est symétrique on a $(z | c(u)(z')) = \overline{(c(u)(z) | z')}$ et que si u est antisymétrique, $(z | c(u)(z')) = -(c(u)(z) | z')$. Ceci est une conséquence immédiate de l'expression :

$$(u(x) + i u(y) | x' + i y') = u(x) \cdot x' + u(y) \cdot y' + i(u(x) \cdot y' - x' \cdot u(y)).$$

Théorème 7.8.1. *Les valeurs propres d'un endomorphisme symétrique u sont réelles, et u est diagonalisable dans une base orthonormée.*

Démonstration. La définition (voir 5.8) du produit scalaire hermitien dans $E_{\mathbb{C}}$:

$$(x + i y | x' + i y') = x \cdot x' + y \cdot y' + i(x \cdot y' - y \cdot x')$$

donne pour $u \in \mathfrak{L}(E)$:

$$\begin{aligned} (u(x) + i u(y) | x' + i y') &= u(x) \cdot x' + u(y) \cdot y' + i(u(x) \cdot y' - u(y) \cdot x'), \\ (x + i y | u(x') + i u(y')) &= x \cdot u(x') + y \cdot u(y') + i(x \cdot u(y') - y \cdot u(x')). \end{aligned}$$

Si u est symétrique, on a donc pour $c(u) \in \mathfrak{L}(E_{\mathbb{C}})$:

$$\forall z, z' \in E_{\mathbb{C}} : (c(u)(z) | z') = (z | c(u)(z'))$$

Si λ est une valeur propre de $c(u)$ et z un vecteur propre associé, on a :

$$\begin{aligned} (z | c(u)(z)) &= (z | \lambda z) = \lambda (z | z) = \lambda \|z\|^2 \\ (c(u)(z) | z) &= (\lambda z | z) = \bar{\lambda} (z | z) = \bar{\lambda} \|z\|^2 \end{aligned}$$

d'où l'égalité $\lambda = \bar{\lambda}$, u étant symétrique ; les valeurs propres sont donc réelles, et ce sont donc les valeurs propres de u .

Les vecteurs propres x_1 et x_2 associés respectivement à des valeurs propres λ_1 et λ_2 distinctes sont orthogonaux. On a en effet $u(x_1) \cdot x_2 = \lambda_1 x_1 \cdot x_2$ et $x_1 \cdot u(x_2) = \lambda_2 x_1 \cdot x_2$, d'où $(\lambda_1 - \lambda_2) x_1 \cdot x_2 = 0$, et $x_1 \cdot x_2 = 0$.

Supposons la matrice de u sous la forme de Jordan (théorème 7.6.2), et l'existence, si λ est une valeur propre multiple d'ordre $r > 1$, de deux vecteurs y_1 et y_2 tels que $u(y_1) = \lambda y_1 + y_2$ et $u(y_2) = \lambda y_2$. On a alors $u(y_1) \cdot y_2 = y_1 \cdot u(y_2)$, avec $u(y_1) \cdot y_2 = \lambda y_1 \cdot y_2 + \|y_2\|^2$ et $y_1 \cdot u(y_2) = \lambda y_1 \cdot y_2$. La norme de y_2 étant nulle, y_2 est nul, et la matrice est diagonale. Le sous-espace propre E_{λ} est de dimension r et il admet une base orthonormée. Ceci est vrai pour chaque $\lambda_i \in \text{Spec}(u)$. Il reste à voir que les espaces propres E_{λ_i} sont deux à deux orthogonaux. Soit $y_i \in E_{\lambda_i}$ et $y_j \in E_{\lambda_j}$ ($\lambda_i \neq \lambda_j$). On a :

$$\begin{aligned} u(y_i) \cdot y_j &= \lambda_i y_i \cdot y_j \\ y_i \cdot u(y_j) &= \lambda_j y_i \cdot y_j \end{aligned}$$

d'où $y_i \cdot y_j = 0$. Nous avons ainsi n vecteurs orthonormés qui forment la base annoncée. $\square$

Exemple 7.8.1 Diagonalisons la matrice symétrique $S \in \mathfrak{M}_n(\mathbb{R})$ dont les termes diagonaux sont égaux à a et tous autres à b , a et b étant des réels non nuls.

Remarquons que si $a = b$ les n lignes de S sont égales et la matrice est de rang 1. Nous en déduisons la valeur propre $a - b$. Les vecteurs propres (x_i) correspondants sont solution de l'équation $\sum_i x_i = 0$, qui est l'équation de l'**hyperplan** H orthogonal au vecteur $v_n = {}^t(1, \dots, 1)$ (un hyperplan de $\mathbb{R}^n$ est un sous-espace de dimension $n - 1$; c'est un plan lorsque $n = 3$). Nous obtenons les $n - 1$ vecteurs propres deux à deux orthogonaux :

$$\begin{aligned} v_1 &= {}^t(1, -1, 0, \dots, 0), \\ v_2 &= {}^t(1, 1, -2, 0, \dots, 0), \\ &\dots \\ v_k &= {}^t(1, \dots, 1, -k, 0, \dots, 0), \\ &\dots \\ v_{n-1} &= {}^t(1, \dots, 1, 1 - n). \end{aligned}$$

Une fois normés, les vecteurs $v_1, \dots, v_n$ forment une base orthonormée. L'image par S de v_n étant égale à $(a + (n - 1)b) v_n$, v_n est vecteur propre.

Pour $a = (n - 2)/n$ et $b = -2/n$, S représente dans la base canonique de $\mathbb{R}^n$ la symétrie orthogonale par rapport à H . ∇

Théorème 7.8.2. *Soit u un endomorphisme antisymétrique de $\mathbb{R}^n$. Les valeurs propres de l'endomorphisme $c(u)$ de $\mathbb{C}^n$ sont imaginaires pures ou nulles et $c(u)$ est diagonalisable dans une base orthonormée.*

Démonstration. Reprenons le début de la démonstration du théorème 6.7.1. Si u est antisymétrique, on a :

$$\forall z, z' \in E_{\mathbb{C}} : (c(u)(z) | z') = -(z | c(u)(z')).$$

Le polynôme caractéristique de a possède n racines (comptées avec leur ordre de multiplicité) dans $\mathbb{C}$, chacune ayant (au moins) un vecteur propre associé. Soit λ une valeur propre et x un vecteur propre associé. On a :

$$\begin{aligned} (c(u)(z) | z) &= \bar{\lambda} (z | z) = \bar{\lambda} \|z\|^2, \\ (z | c(u)(z)) &= \lambda (z | z) = \lambda \|z\|^2. \end{aligned}$$

De $(z | c(u)(z)) = -(c(u)(z) | z)$, nous déduisons que $\bar{\lambda} = -\lambda$; λ est donc imaginaire pure (ou nulle).

Supposons qu'il existe une valeur propre λ et deux vecteurs z et z' non proportionnels de $\mathbb{C}^n$ tels que $c(u)(z) = \lambda z + z'$ et $c(u)(z') = \lambda z'$ (bloc de Jordan). On a alors :

$$\begin{aligned} (c(u)(z) | z') &= (\lambda z + z' | z'), \\ &= \bar{\lambda} (z | z') + \|z'\|^2, \\ &= -(z | c(u)(z')) \quad (\text{antisymétrie}), \\ &= -\lambda (z | z'), \\ &= \bar{\lambda} (z | z') \quad (\bar{\lambda} = -\lambda), \end{aligned}$$

et l'on déduit de la deuxième équation et de la dernière que $\|z'\| = 0$ et que $z' = 0$. La forme de Jordan de $c(u)$ est donc diagonale.

Si $\lambda = 0$ est valeur propre d'ordre $r \geq 1$, toute base de l'espace propre E_λ , de dimension r , diagonalise la restriction de u (à E_λ), qui est l'homothétie de rapport λ . On la choisit orthonormée.

Le polynôme caractéristique étant à coefficients réels, les valeurs propres non nulles sont conjuguées, donc aussi les composantes des vecteurs propres associés. Ces vecteurs propres sont orthogonaux. On a en effet $c(u)(z) = \lambda z$ et $c(u)(\bar{z}) = \bar{\lambda} \bar{z}$, avec $\lambda \neq 0$, et :

$$(\lambda z | \bar{z}) = (c(u)(z) | \bar{z}) = -(z | c(u)(\bar{z})) = -(z | \bar{\lambda} \bar{z}) = -(\lambda z | \bar{z}),$$

d'où $(z | \bar{z}) = 0$. Si une valeur propre est d'ordre $r > 1$, le sous-espace propre associé est de dimension r , et on peut choisir une base orthonormée de r vecteurs propres. Montrons enfin que deux vecteurs propres associés à deux valeurs propres distinctes non conjuguées sont orthogonaux. Si $c(u)(z) = \lambda z$ et $c(u)(z') = \mu z'$, $\bar{\lambda} \neq \mu$, on a :

$$\bar{\lambda} (z | z') = (c(u)(z) | z') = -(z | c(u)(z')) = -(z | \mu z') = -\mu (z | z').$$

On a donc construit une famille de n vecteurs propres orthonormés. □

Exemple 7.8.2. Diagonalisons la matrice antisymétrique à composantes réelles :

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

qui représente un endomorphisme antisymétrique u de $\mathbb{C}^3$ dans la base canonique.

Son polynôme caractéristique $X^3 + 3X$ se factorise en $X(X - \sqrt{3}i)(X + \sqrt{3}i)$. Ses vecteurs propres sont :

$$\begin{array}{lll} v_1 & = & {}^t(1, 1, 1) \quad \text{associé à la valeur propre } 0, \\ v_2 & = & {}^t(1 - i\sqrt{3}, 1 + i\sqrt{3}, -2) \quad \text{associé à la valeur propre } i\sqrt{3}, \\ v_3 & = & {}^t(1 + i\sqrt{3}, 1 - i\sqrt{3}, -2) \quad \text{associé à la valeur propre } -i\sqrt{3}. \end{array}$$

Ils sont deux à deux orthogonaux. En les normant (leurs normes respectives sont $\sqrt{3}$, $2\sqrt{3}$ et $2\sqrt{3}$) on obtient une base orthonormée (de $\mathbb{C}^3$) dans laquelle u est représenté par la matrice diagonale :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i\sqrt{3} & 0 \\ 0 & 0 & -i\sqrt{3} \end{pmatrix}.$$

On remarque que $\text{Im}(u)$, engendré par v_2 et v_3 et $\text{Ker}(u)$, engendré par v_1 , sont supplémentaires orthogonaux. ▽

Proposition 7.8.2. *Si u est un endomorphisme antisymétrique, on a, pour tout $x \in E$, $x \cdot u(x) = 0$. Les sous-espaces $\text{Im}(u)$ et $\text{Ker}(u)$ sont supplémentaires orthogonaux.*

Démonstration. Par symétrie du produit scalaire on a $x \cdot u(x) = u(x) \cdot x$, et par antisymétrie de u : $x \cdot u(x) = -u(x) \cdot x$, d'où le premier point.

Si $x \in \text{Ker}(u)$ et $y \in \text{Im}(u)$, il existe $z \in E$ tel que $y = u(z)$, et :

$$x \cdot y = x \cdot u(z) = -u(x) \cdot z = 0.$$

Les sous-espaces $\text{Im}(u)$ et $\text{Ker}(u)$ sont donc orthogonaux, et le théorème du rang (2.3.1) permet de conclure. $\square$

Proposition 7.8.3. *Si $u \neq 0$ est un endomorphisme antisymétrique de $\mathbb{R}^3$, il existe un (et un seul) vecteur non nul $z \in \mathbb{R}^3$ tel que $u(x) = z \times x$.*

Démonstration. La matrice de u dans une base orthonormée est antisymétrique :

$$M(u) = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

et, $\forall x \in \mathbb{R}^3$, $M(u)x = z \times x$ avec $z = {}^t(a, b, c)$. $\square$

7.9 Endomorphismes hermitiens

Soit H un espace hermitien. $u \in \mathcal{L}(H)$ est **hermitien**, ou **auto-adjoint**, si, quels que soient x et y dans H :

$$(x | u(y)) = (u(x) | y)$$

c'est-à-dire si $u = u^*$.

Théorème 7.9.1. *Un endomorphisme hermitien u est diagonalisable dans une base orthonormée.*

Démonstration. Nous savons que les valeurs propres sont réelles. Montrons que u est diagonalisable. S'il ne l'était pas, il existerait (théorème 7.5.4) une valeur propre λ et un couple (e_1, e_2) de vecteurs non nuls tels que :

$$\begin{aligned} u(e_2) &= \lambda e_2 + e_1 \\ u(e_1) &= \lambda e_1. \end{aligned}$$

La suite d'équations :

$$\begin{aligned} (u(e_2) | e_1) &= (e_2 | u(e_1)) \\ (\lambda e_2 + e_1 | e_1) &= (e_2 | \lambda e_1) \\ \lambda (e_2 | e_1) + \|e_1\|^2 &= \lambda (e_2 | e_1) \quad (\lambda = \bar{\lambda}) \end{aligned}$$

mène à la conclusion : $e_1 = 0$, ce qui contredit l'hypothèse.

Montrons enfin que les vecteurs propres sont deux à deux orthogonaux, et il ne restera plus qu'à les normer. Si e_1 et e_2 sont des vecteurs propres associés à deux valeurs propres distinctes λ_1 et λ_2 , on a :

$$\begin{aligned}(u(e_1) | e_2) &= \lambda_1 (e_1 | e_2) \\ (e_1 | u(e_2)) &= \lambda_2 (e_1 | e_2)\end{aligned}$$

d'où $(e_1 | e_2) = 0$. Si une valeur propre λ est de multiplicité $r > 1$, son espace propre E_λ , de dimension r , admet des bases orthonormées. $\square$

7.10 Isométries de $\mathbb{R}^3$

Soit $u \in \mathfrak{L}(\mathbb{R}^3)$. Nous savons (théorème 6.4.1) que les assertions suivantes sont équivalentes :

- (a) u est une isométrie,
- (b) $u \circ {}^t u = {}^t u \circ u = i_E$,
- (c) $u^{-1} = {}^t u$.

Soit u une isométrie de $\mathbb{R}^3$, différente de l'identité et de la symétrie centrale ($-i_E$), de matrice M dans la base canonique. Son polynôme caractéristique, de degré 3, a ou bien une racine réelle ou bien trois. Une valeur propre réelle d'une isométrie ne peut être égale qu'à 1 ou -1 puisque $u(x) = \lambda x$ implique :

$$\|x\| = \|u(x)\| = |\lambda| \|x\|.$$

Si elle admet trois valeurs propres réelles, il existe une base orthonormée de vecteurs propres dans laquelle la matrice de u prend, à permutation près, l'une des formes suivantes :

$$S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad DT = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

car si 1 est valeur propre triple, u est l'identité $i_{\mathbb{R}^3}$ et si c'est -1 , u est la symétrie centrale, et ces deux isométries ont la même écriture dans toutes les bases.

La première forme représente la **symétrie orthogonale** par rapport au plan propre associé à la valeur propre 1, ou **réflexion**. La seconde représente le **demi-tour** (ou la rotation d'angle π) d'axe engendré par le vecteur invariant.

Si elle admet une seule valeur propre réelle, $\lambda = \pm 1$, de vecteur propre associé v_1 (normé), le plan orthogonal à v_1 , de base orthonormée (positive) (v_2, v_3) telle que $v_2 \times v_3 = v_1$, est globalement invariant, puisque $u(v_2)$ et $u(v_3)$ sont orthogonaux à $u(v_1)$, donc à v_1 . La restriction de u à ce plan est donc une isométrie de $\mathbb{R}^2$.

Si $\lambda = 1$, u est la rotation d'axe $\text{Vect}(v_1)$ et d'angle θ tel que :

$$\cos \theta = v_2 \cdot u(v_2), \quad \sin \theta = u(v_2) \cdot v_3.$$

La matrice de u dans la base (v_i) est :

$$R(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Si $\lambda = -1$, u est la rotation précédente composée avec la réflexion de plan $\text{Vect}(v_2, v_3)$.

Proposition 7.10.1. *Les réflexions engendrent le groupe orthogonal $O(3)$.*

Démonstration. Il suffit de remarquer que :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

les deux premières matrices représentant des réflexions et la troisième, une rotation quelconque, toutes trois dans la base de réduction de cette dernière. $\square$

Précisons comment on étudie une isométrie donnée sous la forme d'une matrice M . On vérifie d'abord que les vecteurs colonnes sont orthonormés (ou, ce qui est équivalent, que $M {}^t M = I_3$).

L'endomorphisme v de matrice $a = \frac{1}{2}(M - {}^t M)$, antisymétrique dans le cas d'une rotation d'angle $\theta \neq k\pi$, nul dans le cas d'une symétrie ou d'un demi-tour, est intéressant. D'après la proposition 6.8.3, il existe un vecteur $z \neq 0$ tel que $v(x) = z \times x$. Voyons quelle est son expression dans une base de réduction dans le cas d'une rotation :

$$\frac{1}{2}(R(\theta) - {}^t R(\theta)) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sin \theta \\ 0 & \sin \theta & 0 \end{pmatrix}.$$

On a donc $z = \sin \theta v_1$. Si on choisit $\sin \theta = \|z\|$, on a $v_1 = z/\|z\|$.

Si $u(v_1) = v_1$, u est la rotation d'axe engendré et orienté par z et d'angle θ tel que $\sin \theta = \|z\|$, $\cos \theta = (\text{tr}(u) - 1)/2$.

Si $u(v_1) = -v_1$, u est la rotation d'axe engendré et orienté par z et d'angle θ tel que $\sin \theta = \|z\|$ et $\cos \theta = (\text{tr}(u) + 1)/2$ composée avec la symétrie (orthogonale) de même axe.

Dans les deux cas, on complète la base par un vecteur v_2 orthogonal à v_1 et par $v_3 = v_1 \times v_2$.

Si $a = 0$ et $\text{tr}(M) = -1$ u est le demi-tour, d'axe $\text{Ker}(u - 1)$.

Si $a = 0$ et $\text{tr}(M) = 1$ u est la réflexion, de plan $\text{Ker}(u - 1)$.

Dans le cas d'un demi-tour, il est plus rapide de calculer le plan orthogonal à l'axe, car le rang du système donnant $\text{Ker}(u + 1)$ est égal à 1, le système se réduit à une équation $ax_1 + bx_2 + cx_3 = 0$, et ${}^t(a, b, c)$ est un vecteur engendrant l'axe. Il reste à le normer et à compléter la base comme précédemment.

Dans les exemples suivants, les matrices sont données dans la base canonique de $\mathbb{R}^3$.

Exemple 7.10.1. La matrice :

$$M = \frac{1}{7} \begin{pmatrix} -3 & -6 & -2 \\ -2 & 3 & -6 \\ 6 & -2 & -3 \end{pmatrix}$$

($M {}^tM = I_3$) n'est pas symétrique. Sa partie antisymétrique est :

$$a(u) = \frac{1}{7} \begin{pmatrix} 0 & -2 & -4 \\ 2 & 0 & -2 \\ 4 & 2 & 0 \end{pmatrix},$$

d'où le vecteur $z = \frac{1}{7} {}^t(2, -4, 2)$, de norme $\frac{2\sqrt{6}}{7}$. Posons $v_1 = \frac{\sqrt{6}}{6} {}^t(1, -2, 1)$. Ce vecteur (normé) est invariant par M engendre et oriente l'axe de la rotation d'angle θ tel que $\sin \theta = \frac{2\sqrt{6}}{7}$ et $\cos \theta = -\frac{5}{7}$. On complète la nouvelle base en prenant $v_2 = \frac{\sqrt{2}}{2} {}^t(1, 0, -1)$ et $v_3 = v_1 \times v_2 = \frac{\sqrt{6}}{6} {}^t(1, 1, 1)$. ∇

Exemple 7.10.2. La matrice :

$$M(u) = \frac{1}{7} \begin{pmatrix} 3 & 2 & -6 \\ 2 & 6 & 3 \\ -6 & 3 & -2 \end{pmatrix}$$

($M {}^tM = I_3$) est symétrique, et sa trace est égale à 1 ; c'est une réflexion, de plan d'équation $2x - y + 3z = 0$. La direction de la réflexion est engendrée par le vecteur ${}^t(2, -1, 3)$, d'où le vecteur normé $v_1 = \frac{\sqrt{14}}{14} {}^t(2, -1, 3)$, que l'on complète en une base orthonormée positive par, par exemple :

$$v_2 = \frac{\sqrt{5}}{5} {}^t(1, 2, 0) \text{ et } v_3 = v_1 \times v_2 = \frac{\sqrt{70}}{70} {}^t(-6, 3, 5).$$

∇

Exemple 7.10.3. La matrice :

$$M(u) = \frac{1}{7} \begin{pmatrix} -3 & -2 & 6 \\ -2 & -6 & -3 \\ 6 & -3 & 2 \end{pmatrix}$$

($M {}^tM = I_3$) est symétrique, et sa trace est égale à -1 ; c'est un demi-tour, d'axe engendré par le vecteur ${}^t(2, -1, 3)$ (c'est l'opposé de la réflexion précédente). ∇

Exemple 7.10.4. Composons la rotation et la réflexion précédentes :

$$\frac{1}{7} \begin{pmatrix} -3 & -6 & -2 \\ -2 & 3 & -6 \\ 6 & -2 & -3 \end{pmatrix} \frac{1}{7} \begin{pmatrix} 3 & 2 & -6 \\ 2 & 6 & 3 \\ -6 & 3 & -2 \end{pmatrix} = \frac{1}{49} \begin{pmatrix} -9 & -48 & 4 \\ 36 & -4 & 33 \\ 32 & -9 & -36 \end{pmatrix}.$$

La partie antisymétrique de cette isométrie donne le vecteur :

$$z = \frac{1}{7} {}^t(-3, -2, 6), \quad ||z|| = 1,$$

d'où $\sin \theta = 1$ et $\theta = \pi/2$, $\cos \theta = 0$. La trace étant égale à -1 , la valeur propre réelle est égale à -1 . Ceci caractérise la composée d'une rotation et

d'une réflexion, ce que nous savons déjà, mais nous constatons que ni l'angle ni l'axe ne sont conservés. On a $v_1 = z$, et, par exemple, $v_2 = \sqrt{5}/5 {}^t(2, 0, 1)$ et $v_3 = v_1 \times v_2 = \sqrt{5}/35 {}^t(-2, -15, 4)$.

La composée dans l'ordre inverse est différente, alors que dans la forme réduite il y a commutativité car l'axe de la rotation est la direction de la réflexion. On obtient dans ce cas la matrice :

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

qui représente la composée de la rotation d'axe e_1 et d'angle $\pi/2$ avec la réflexion de direction e_1 . ∇

Exemple 7.10.5. Cherchons la matrice M de la rotation d'axe engendré et orienté par le vecteur ${}^t(1, 1, 0)$ et d'angle θ tel que $\cos \theta = -\frac{1}{3}$ et $\sin \theta = \frac{2\sqrt{2}}{3}$.

Construisons une base de réduction dans laquelle la matrice prendra la forme $R(\theta)$. Prenons $v_1 = \frac{\sqrt{2}}{2} {}^t(1, 1, 0)$, $v_2 = e_3$ et $v_3 = v_1 \times v_2 = \frac{\sqrt{2}}{2} {}^t(1, -1, 0)$. Nous obtenons la matrice de passage :

$$P = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \end{pmatrix},$$

d'où :

$$M = P R(\theta) {}^tP = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix}.$$

∇

Exemple 7.10.6. Cherchons la matrice M de la réflexion d'axe engendré par le vecteur ${}^t(1, 1, 0)$.

Le plan de la réflexion, orthogonal à l'axe, a pour équation $x+y=0$. Une base de réduction orthonormée positive est constituée des vecteurs $v_1 = {}^t(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$ (sur l'axe), $v_2 = {}^t(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$ (dans le plan) et $v_3 = v_1 \times v_2 = {}^t(0, 0, 1)$. D'où la matrice de passage :

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

$$M = P S {}^tP = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On pouvait trouver ce résultat géométriquement, sans calcul. ∇

7.11 Exemple de diagonalisation dans $\mathbb{F}_2$

Les corps $\mathbb{F}_2$ et $\mathbb{F}_4 = \mathbb{F}_2[i]$ ont été construits dans les Généralités. Rappelons que i et $\bar{i} = i^2 = i + 1$ sont les racines du polynôme $X^2 + X + 1$, irréductible sur $\mathbb{F}_2$. Etudions les matrices de $\mathfrak{M}_2(\mathbb{F}_2)$, au nombre de $2^4 = 16$, en excluant les quatre matrices diagonales, à savoir la matrice nulle, l'identité et :

$$D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad D_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Les matrices M dont le polynôme caractéristique est un carré ne sont pas diagonalisables, car, si λ est leur unique valeur propre, le rang de $M - \lambda I_2$ est inférieur à 2 et non nul, donc égal à 1, $\text{Ker}(M - \lambda I_2)$ est de dimension 1 et elles ont un seul vecteur propre. Il s'agit des matrices de polynôme caractéristique X^2 , donc nilpotentes :

$$N_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad N_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

et celles dont le polynôme caractéristique est $X^2 + 1 = (X + 1)^2$:

$$J_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \text{et} \quad J_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Remarquons que N_3 et J_1 , bien que symétriques, ne sont pas diagonalisables.

Les matrices de polynôme caractéristique $X^2 + X = X(X + 1)$:

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

et $X^2 + X + 1 = (X + i)(X + \bar{i})$:

$$M_5 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{et} \quad M_6 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

ayant deux valeurs propres, sont diagonalisables. Ce sont les six dernières.

Etudions M_5 . Les vecteurs propres associés aux valeurs propres i et $\bar{i}$ sont respectivement ${}^t(1, i)$ et ${}^t(1, \bar{i})$. La matrice de passage P est donc :

$$P = \begin{pmatrix} 1 & 1 \\ i & \bar{i} \end{pmatrix}.$$

Le calcul de P^{-1} :

$$PP^{-1} = \begin{pmatrix} 1 & 1 \\ i & \bar{i} \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a+b & c+d \\ ia+\bar{i}b & ic+\bar{i}d \end{pmatrix} = I_2$$

mène aux systèmes :

$$\begin{cases} a+b & = 1 \\ a i + b \bar{i} & = 0 \end{cases} \quad \text{et} \quad \begin{cases} c+d & = 0 \\ c i + d \bar{i} & = 1 \end{cases}$$

dont la transformation $L_2 + iL_1$ donne $b = i$ (car $i + \bar{i} = 1$), et $d = 1$, d'où $a = \bar{i}$ et $d = 1$. Finalement :

$$M_5 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ i & \bar{i} \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & \bar{i} \end{pmatrix} \begin{pmatrix} \bar{i} & 1 \\ i & 1 \end{pmatrix}$$

De même :

$$M_6 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i & \bar{i} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & \bar{i} \end{pmatrix} \begin{pmatrix} 1 & \bar{i} \\ 1 & i \end{pmatrix}.$$

La réduction des dernières matrices donne :

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$M_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

$$M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

$$M_4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Les huit polynômes caractéristiques des 512 matrices de $\mathfrak{M}_3(\mathbb{F}_2)$, se factorisent dans $\mathbb{F}_2$ (X^3 , $X^3 + X$, $X^3 + X^2$ et $X^3 + X^2 + X + 1$), $\mathbb{F}_4$ ($X^3 + 1$, $X^3 + X^2 + X$) et $\mathbb{F}_8$ pour $X^3 + X + 1$ et $X^3 + X^2 + 1$.

Etudions la matrice :

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

de polynôme caractéristique $X^3 + X + 1$. Calculons un vecteur propre associé à la valeur propre $j \in \mathbb{F}_8$:

$$\begin{cases} x + z &= jx \\ x + y + z &= jy \\ x + y &= jz \end{cases} \iff \begin{cases} j^3 x + z &= 0 \\ x + j^3 y + z &= 0 \\ x + y + jz &= 0 \end{cases} \iff \begin{cases} z = j^3 x \\ y = (1 + j^4) x = j^5 x. \end{cases}$$

Nous obtenons le vecteur ${}^t(1, j^5, j^3)$. Pour les valeurs propres j^2 et j^4 , on obtient, respectivement, ${}^t(1, j^3, j^6)$ et ${}^t(1, j^6, j^5)$, d'où la matrice de passage :

$$P = \begin{pmatrix} 1 & 1 & 1 \\ j^5 & j^3 & j^6 \\ j^3 & j^6 & j^5 \end{pmatrix}$$

dont nous allons calculer l'inverse par la méthode des mineurs, exceptionnellement simple en caractéristique 2 et en dimension 3 : il suffit de remplacer chaque terme par son mineur puis de transposer. Chaque mineur est la somme

de deux puissances de j ; le premier, par exemple, est égal à $j^{12} + j^8$, soit $j^5 + j$, $(j^2 + j + 1) + j$, $j^2 + 1$ et enfin j^6 . On obtient ainsi :

$$P^{-1} = \begin{pmatrix} j^6 & j & j^4 \\ j^5 & j^2 & j \\ j^3 & j^4 & j^2 \end{pmatrix}$$

ce qui achève la diagonalisation : $M = P \operatorname{Diag}(j, j^2, j^4) P^{-1}$.

8 Réduction des systèmes par des isométries

Nous allons définir des isométries permettant de triangulariser une matrice réelle quelconque et de mettre une matrice réelle symétrique sous forme tridagonale.

8.1 Isométries de $\mathbb{R}^n$

Toute isométrie u de $\mathbb{R}^n$ est un endomorphisme normal puisque commutant avec sa transposée ($u \circ {}^t u = i_{\mathbb{R}^n} = {}^t u \circ u$). Elle est donc diagonalisable (dans $\mathbb{C}$) dans une base orthonormée (théorème 5.8.6). Outre les termes égaux à ± 1 , la diagonale peut contenir des couples $(e^{i\theta}, e^{-i\theta})$, $\theta \neq k\pi$, associés à des rotations planes. Notons $E_{(1)}$, $E_{(-1)}$ et E_{rot} les sous-espaces propres correspondants, de dimensions respectives n_1 , n_2 et $2n_3$, avec $n_1 + n_2 + 2n_3 = n$.

Nous dirons que u est une rotation si elle est composée de rotations planes ($\det(u) = 1$), c'est-à-dire si n_2 est pair, car les paires de -1 sont associés à des demi-tours plans, donc des rotations. Elle peut être le produit d'un nombre quelconque de rotations planes, mais se décompose en un produit d'au plus partie entière de $n/2$ rotations de plans deux à deux orthogonaux ($2n_3 \leq n$).

Nous dirons que u est une réflexion si $n_3 = 0$ et si $n_2 = 1$ (${}^t u = u$ et $\det(u) = -1$).

Dans tous les autres cas, u est composée d'une rotation et d'une réflexion.

Remarquons que la composée d'un nombre pair de réflexions est une rotation.

Exemple 8.1.1. Le produit des rotations de $\mathbb{R}^4$ de matrices :

$$A = \frac{1}{7} \begin{pmatrix} -3 & -6 & 0 & -2 \\ -2 & 3 & 0 & -6 \\ 0 & 0 & 7 & 0 \\ 6 & -2 & 0 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

et :

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

est la rotation de matrice :

$$R = ABC = \frac{1}{21} \begin{pmatrix} 12 & 6 & -6 & 15 \\ 1 & -10 & -18 & -4 \\ 14 & 7 & 0 & -14 \\ -10 & 16 & -9 & -2 \end{pmatrix}.$$

Le polynôme caractéristique de R , $P = X^4 + \frac{8}{21}X^2 + 1$, n'a pas de racine réelle ($P(\pm 1) \neq 0$). Les valeurs propres de R sont donc de la forme $e^{i\theta}$:

$$\begin{aligned} \lambda = e^{i\theta} &\iff 21e^{4i\theta} + 8e^{2i\theta} + 21 = 0, \\ &\iff 21(e^{2i\theta} + e^{-2i\theta}) = -8, \\ &\iff 21\cos(2\theta) = -4, \\ &\iff \cos(2\theta) = -\frac{4}{21} = 2\cos^2\theta - 1, \\ &\iff \cos\theta = \pm\sqrt{\frac{17}{42}}, \quad \sin\theta = \pm\sqrt{\frac{5}{42}}. \end{aligned}$$

On obtient deux rotations d'angles respectifs $\theta = \arccos(\sqrt{\frac{17}{42}})$ et $\pi - \theta$, à plans orthogonaux.

Déterminons les plans de ces rotations. Celui de la première est engendré par deux vecteurs, v_1 et v_2 , tels que :

$$\begin{aligned} Rv_1 &= \cos\theta v_1 + \sin\theta v_2, \\ Rv_2 &= -\sin\theta v_1 + \cos\theta v_2, \end{aligned}$$

ce qui s'écrit dans $\mathbb{R}^8$:

$$\begin{pmatrix} R - \cos\theta I_4 & -\sin\theta I_4 \\ \sin\theta I_4 & R - \cos\theta I_4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = O_8.$$

Une solution est :

$$v_1 = \begin{pmatrix} 5\sqrt{17} + 10\sqrt{42} \\ 12\sqrt{17} + 24\sqrt{42} \\ -23\sqrt{17} - 46\sqrt{42} \\ 80\sqrt{17} + 9\sqrt{42} \end{pmatrix}, \quad v_2 = \begin{pmatrix} 12\sqrt{17}\sqrt{42} + 253 \\ -5\sqrt{17}\sqrt{42} + 184 \\ 151 \\ 0 \end{pmatrix}.$$

De même, le plan de la seconde est engendré par les vecteurs :

$$v_3 = \begin{pmatrix} -5\sqrt{17} + 10\sqrt{42} \\ -12\sqrt{17} + 24\sqrt{42} \\ 23\sqrt{17} - 46\sqrt{42} \\ -80\sqrt{17} + 9\sqrt{42} \end{pmatrix}, \quad v_4 = \begin{pmatrix} -12\sqrt{17}\sqrt{42} + 253 \\ 5\sqrt{17}\sqrt{42} + 184 \\ 151 \\ 0 \end{pmatrix}.$$

On peut vérifier que ces vecteurs sont deux à deux orthogonaux, que $\|v_1\| = \|v_2\|$ et que $\|v_3\| = \|v_4\|$. ▽

8.2 Rotations de Givens (1954)

L'automorphisme $I_{(p,q,\theta)}$, $1 \leq p < q \leq n$, $\theta \in [-\pi/2, \pi/2]$, de $\mathbb{R}^n$ est ainsi défini dans la base canonique (e_i) :

$$\begin{cases} I_{(p,q,\theta)}(e_p) = c e_p - s e_q, \\ I_{(p,q,\theta)}(e_q) = s e_p + c e_q, \\ I_{(p,q,\theta)}(e_i) = e_i \text{ pour } i \notin \{p, q\} \\ c = \cos(\theta), \quad s = \sin(\theta). \end{cases}$$

La restriction de $I_{(p,q,\theta)}$ à $\text{Vect}(e_p, e_q)$ est la rotation d'angle θ (voir 6.6). Remarquons que ${}^t I_{(p,q,\theta)} = I_{(p,q,-\theta)}$ et que $I_{(p,q,\theta)} I_{(p,q,-\theta)} = I_n$.

Soit $A \in \mathfrak{M}(\mathbb{R}^n)$. Le calcul de $I_{(p,q,\theta)} A$ demande $6n$ opérations :

$$\begin{cases} L_i(I_{(p,q,\theta)} A) = L_i(A) \text{ pour } i \notin \{p, q\}, \\ L_p(I_{(p,q,\theta)} A) = (c a_{pi} - s a_{qi})_{1 \leq i \leq n}, \\ L_q(I_{(p,q,\theta)} A) = (s a_{pi} + c a_{qi})_{1 \leq i \leq n}. \end{cases}$$

Comme $AI_{(p,q,\theta)} = {}^t({}^t I_{(p,q,\theta)} {}^t A)$, on a :

$$\begin{cases} C_i(AI_{(p,q,\theta)}) = C_i(A) \text{ pour } i \notin \{p, q\}, \\ C_p(AI_{(p,q,\theta)}) = (c a_{ip} - s a_{iq})_{1 \leq i \leq n}, \\ C_q(AI_{(p,q,\theta)}) = (s a_{ip} + c a_{iq})_{1 \leq i \leq n}. \end{cases}$$

Enfin le calcul de $B = I_{(p,q,\theta)} AI_{(p,q,\theta)} = (b_{ij})$ demande $12n$ opérations. Seules sont modifiées les lignes et les colonnes d'indices p et q . On a, pour $i, j \notin \{p, q\}$, $b_{ij} = a_{ij}$, puis pour $i \notin \{p, q\}$:

$$\begin{cases} b_{ip} = c a_{ip} + s a_{iq}, \\ b_{iq} = c a_{iq} - s a_{ip}, \\ b_{pi} = c a_{pi} + s a_{qi}, \\ b_{qi} = c a_{qi} - s a_{pi}, \end{cases}$$

enfin, pour $\{i, j\} = \{p, q\}$:

$$\begin{cases} b_{pp} = cs(a_{pq} + a_{qp}) + c^2 a_{pp} + s^2 a_{qq}, \\ b_{qq} = cs(a_{qp} - a_{pq}) + c^2 a_{qq} + s^2 a_{pp}, \\ b_{pq} = cs(a_{qq} - a_{pp}) + c^2 a_{pq} - s^2 a_{qp}, \\ b_{qp} = cs(a_{qp} - a_{pq}) + c^2 a_{qp} - s^2 a_{pq}. \end{cases}$$

Ces rotations permettent d'annuler des composantes de la matrice A . On peut la triangulariser (méthode QR par Givens, $\sim 3n^3$ opérations), ou, si elle est symétrique, la tridiagonaliser ($\sim 6n^3$ opérations).

Exemple 8.2.1. Pour triangulariser la matrice :

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

annulons le terme en position $(2, 3)$ dans $I_{(2,3,\theta_1)}M$, c'est-à-dire $-2(c+s)$, en posant $\cos \theta_1 = c$ et $\sin \theta_1 = s$, donc pour $\theta = -\pi/4$. On obtient :

$$M_1 = I_{(2,3,\pi/4)}M = \begin{pmatrix} 1 & 2 & -1 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -\sqrt{2}/2 & -\sqrt{2}/2 & 2\sqrt{2} \end{pmatrix}$$

Annulons le terme en position $(1, 3)$ dans $I_{(1,3,\theta_2)}M_1$, c'est-à-dire $-c - 2\sqrt{2}s$, d'où $c = 2\sqrt{2}/3$ et $s = -1/3$, et :

$$M_2 = I_{(1,3,\theta)}M_1 = \begin{pmatrix} \sqrt{2}/2 & 7\sqrt{2}/6 & 0 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ -1 & -4/3 & 3 \end{pmatrix}$$

Annulons enfin le terme en position $(1, 2)$ dans $I_{(1,2,\theta_3)}M_2$, $\sqrt{2}(7c-3s)/6$, avec $c = 3\sqrt{58}/58$ et $s = 7\sqrt{58}/58$:

$$M_3 = I_{(1,2,\theta)}M_2 = \begin{pmatrix} 5\sqrt{116}/58 & 0 & 0 \\ -\sqrt{2}/2 & \sqrt{116}/6 & 0 \\ -1 & -4/3 & 3 \end{pmatrix},$$

d'où $M = QR$ avec $Q = I_{(1,2,\theta_3)}I_{(1,3,\theta_2)}I_{(2,3,\theta_1)}$ et $R = M_3$. ▽

Pour tridiagonaliser une matrice symétrique de $S = (s_{ij}) \in \mathfrak{M}_n(\mathbb{R})$, on annule le terme s_{13} en remplaçant S par $I_{(2,3,\theta)}SI_{(2,3,-\theta)}$ pour un θ convenable, puis le terme s_{14} en faisant agir de même $I_{(3,4,\theta)}$ pour un θ convenable ..., jusqu'à s_{1n} en faisant agir $I_{(n-1,n,\theta)}$ pour un θ convenable. On passe alors à la deuxième ligne : on annule $s_{2,4}$ grâce à un $I_{(3,4,\theta)}$ bien choisi ..., jusqu'à s_{2n} grâce à un $I_{(3,n,\theta)}$..., ainsi de suite jusqu'à la ligne $n-2$, $s_{n-2,n}$ étant annulé par l'action d'un certain $I_{(n-1,n,\theta)}$.

Exemple 8.2.2. Tridiagonalisons la matrice symétrique :

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

L'action de $I_{2,3,-\pi/4}$ donne :

$$S_{23} = \begin{pmatrix} 4 & 2\sqrt{2} & 0 & 1 \\ 2\sqrt{2} & 2 & -3 & -\sqrt{2}/2 \\ 0 & -3 & 0 & \sqrt{2}/2 \\ 1 & -\sqrt{2}/2 & \sqrt{2}/2 & 2 \end{pmatrix}.$$

L'action de $I_{2,4,\theta}$ avec $\cos \theta = 2\sqrt{2}/3$ et $\sin \theta = -1/3$ donne :

$$S_{24} = \begin{pmatrix} 4 & 3 & 0 & 0 \\ 3 & 14/9 & -11\sqrt{2}/6 & -7\sqrt{2}/18 \\ 0 & -11\sqrt{2}/6 & 0 & 5/3 \\ 0 & -7\sqrt{2}/18 & 5/3 & 22/9 \end{pmatrix}.$$

L'action de $I_{3,4,\theta}$ avec $\cos \theta = 33/\sqrt{1138}$ et $\sin \theta = -7/\sqrt{1138}$ donne enfin la matrice tridiagonale :

$$S_{34} = \begin{pmatrix} 4 & 3 & 0 & 0 \\ 3 & 14/9 & -\sqrt{2276}/18 & 0 \\ 0 & -\sqrt{2276}/18 & 4004/5121 & 1149/569 \\ 0 & 0 & 1169/569 & 946/569 \end{pmatrix}.$$

▽

8.3 Réflexions de Householder (1964)

Soit $u = (u_i)$ un vecteur de $\mathbb{R}^n$ muni de sa base canonique $(e_i)_{1 \leq i \leq n}$. Nous allons construire une réflexion de matrice Q qui annule les $n - k - 1$ dernières composantes de u ($0 \leq k \leq n - 2$). Soit u' le vecteur ${}^t(0, \dots, 0, u_k, \dots, u_n)$.

Si $v = u' - \|u'\|e_k$, $w = v/\|v\|$ et $P = w {}^t w = (p_{ij})$ avec $p_{ij} = w_i w_j$, on a :

$$\begin{aligned} {}^t P &= P, \\ C_i(P) &= w_i w, \\ Pw &= w {}^t w w = w, \text{ car } w {}^t w = \|w\| = 1, \\ P^2 &= (w {}^t w)(w {}^t w) = w({}^t w w) {}^t w = w {}^t w = P, \end{aligned}$$

et P représente un projecteur symétrique (donc orthogonal), de rang égal à 1 (les colonnes de $w {}^t w$ sont proportionnelles à w) et d'image $\text{Vect}(w)$, diagonalisable dans une base orthonormée $\mathfrak{B} = (\epsilon_i)$ (théorème 7.8.1) :

$$M_{\mathfrak{B}}(P) = \text{Diag}(1, 0, \dots, 0).$$

Dans cette base, on a :

$$M_{\mathfrak{B}}(i_{\mathbb{R}^n} - 2P) = \text{Diag}(-1, 1, \dots, 1),$$

de sorte que $Q = I_n - 2P$ est la matrice, dans la base initiale, de la réflexion orthogonale de direction engendrée par ϵ_k .

On a $Qw = -w$, $Q(u - \|u\|e_k) = -u + \|u\|e_k$, et, pour tout vecteur x orthogonal à w , $Qx = x$, et donc $Qe_i = e_i$ pour $1 \leq i \leq k - 1$. Comme :

$$(u' + \|u'\|e_k) \cdot (u' - \|u'\|e_k) = \|u'\|^2 - \|u'\|^2 = 0$$

$u' + \|u'\|e_k$ est orthogonal à w , on a $Q(u' + \|u'\|e_k) = u' + \|u'\|e_k$, d'où :

$$\begin{aligned} Qu' &= \frac{1}{2}(Q(u' + \|u'\|e_k) + Q(u' - \|u'\|e_k)) \\ &= \frac{1}{2}(u' + \|u'\|e_k - u' + \|u'\|e_k) \\ &= \|u'\|e_k, \end{aligned}$$

et $Qu = {}^t(u_1, \dots, u_k, 0, \dots, 0)$, avec $u_k = \|u'\|$. Le calcul de P , et donc de Q , demande $\sim (n - k)^2/2$ opérations (le calcul des $w_i w_j$ pour $i \leq j$).

Si $A \in \mathfrak{M}_n(\mathbb{R})$, on a $QA = w {}^t w A$. Le calcul de ${}^t w A$ demande $\sim 2n(n - k)$ opérations et le produit par w en demande $n(n - k)$. Le calcul de QA demande donc $3n(n - k)$ opérations. De même pour le calcul de AQ .

Nous allons déduire de cette construction une méthode de triangularisation pour une matrice quelconque (méthode QR par Householder), et une méthode de tridiagonalisation pour une matrice symétrique.

Exemple 8.3.1. Soit $u = {}^t(0, 3, 4)$. Nous voulons annuler les deux dernières composantes de ce vecteur. On a successivement $\|u\| = 5$, $v = u - 5e_1$, $\|v\| = \frac{\sqrt{2}}{10}$ et $w = \frac{\sqrt{2}}{10}v$, puis :

$$P = v {}^t v = \frac{1}{50} \begin{pmatrix} -5 \\ 3 \\ 4 \end{pmatrix} (-5, 3, 4) = \begin{pmatrix} 1/2 & -3/10 & -2/5 \\ -3/10 & 9/50 & 6/25 \\ -2/5 & 6/25 & 8/25 \end{pmatrix},$$

et :

$$Q = I_3 - 2P = \begin{pmatrix} 0 & 3/5 & 4/5 \\ 3/5 & 21/25 & -12/25 \\ 4/5 & -12/25 & 9/25 \end{pmatrix},$$

avec $Qu = 5e_1$.

Pour obtenir $Qu = 5e_2$, on pose $v = u - 5e_2$, $\|v\| = 2\sqrt{5}$, $w = \frac{\sqrt{5}}{10} {}^t(0, -2, 4)$, et :

$$P = \frac{1}{20} \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix} (0, -2, 4) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 4/5 \\ 0 & 4/5 & -3/5 \end{pmatrix}$$

▽

Pour triangulariser une matrice M , on applique d'abord la méthode à $C_1(M)$, on obtient une matrice $M_1 = RM$ dont la première colonne est multiple de e_1 , puis au vecteur dont la première composante est nulle et les $n - 1$ dernières composantes sont celles de $C_2(M_1) \dots$ ainsi de suite jusqu'à l'avant-dernière colonne. On met ainsi M sous la forme d'un produit QR , où Q est une matrice orthogonale, composée de $n - 1$ réflexions, et R une matrice triangulaire supérieure. Ce travail nécessite $\sum_{1 \leq k \leq n-1} 3n(n - k) \sim 3n^3/2$ opérations.

Pour tridiagonaliser une matrice symétrique S , on annule les $n - 2$ dernières composantes de sa première colonne par la transformation $S \mapsto PS$. les $n - 2$ dernières composantes de la première ligne de SP sont donc nulles. La transformation $SP \mapsto PSP$ conserve la première ligne et annule les $n - 2$ dernières composantes de la première colonne. On annule ensuite, de la même façon, les $n - 3$ dernières composantes de la deuxième colonne et de la deuxième ligne $\dots$ ainsi de suite, jusqu'aux antépénultièmes ligne et colonne. Ce travail nécessite $\sim 3n^3$ opérations.

Exemple 8.3.2. Pour triangulariser la matrice :

$$M = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 1 & 4 & 1 & 5 \\ 1 & 0 & 1 & 7 \\ -1 & 0 & 5 & 5 \end{pmatrix}$$

posons d'abord :

$$\begin{aligned} u_1 &= {}^t(1, 1, 1, -1), \quad \|u_1\| = 2, \\ v_1 &= u_1 - \|u_1\|e_1 = {}^t(-1, 1, 1, -1), \quad \|v_1\| = 2, \\ w_1 &= v_1/\|v_1\| = \frac{1}{2} {}^t(-1, 1, 1, -1), \end{aligned}$$

$$Q_1 = I_4 - 2w_1 {}^t w_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}$$

d'où :

$$M_1 = Q_1 M = \frac{1}{2} \begin{pmatrix} 4 & 6 & -1 & 7 \\ 0 & 6 & 7 & 3 \\ 0 & -2 & 7 & 7 \\ 0 & 2 & 5 & 17 \end{pmatrix},$$

puis, dans un deuxième pas :

$$\begin{aligned} u_2 &= {}^t(0, 3, -1, 1), \quad \|u_2\| = \sqrt{11}, \\ v_2 &= u_2 - \|u_2\|e_2 = {}^t(0, 3 - \sqrt{11}, -1, 1), \quad \|v_2\| = \sqrt{22 - 6\sqrt{11}}, \\ w_2 &= v_2/\|v_2\| = \frac{1}{\sqrt{22 - 6\sqrt{11}}} {}^t(0, 3 - \sqrt{11}, -1, 1), \end{aligned}$$

$$Q_2 = I_4 - 2w_2 {}^t w_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0,904534 & -0,301511 & 0,301511 \\ 0 & -0,301511 & 0,047732 & 0,952267 \\ 0 & 0,301511 & 0,952267 & 0,047732 \end{pmatrix},$$

d'où :

$$M_2 = Q_2 M_1 = \begin{pmatrix} 2 & 3 & -0,5000000 & 3,500000 \\ 0 & 3,316624 & 2,864357 & 2,864357 \\ 0 & 0 & 1,492443 & 7,809068 \\ 0 & 0 & 4,507556 & 4,190931 \end{pmatrix},$$

et enfin, dans un troisième pas :

$$\begin{aligned} u_3 &= {}^t(0, 0, 1,492443, 4,507556), \quad \|u_3\| = 4,748205, \\ v_3 &= u_3 - \|u_3\|e_3 = {}^t(0, 0, -1,627881, 2,253778), \quad \|v_3\| = 2,780200, \\ w_3 &= v_3/\|v_3\| = {}^t(0, 0, -0,585526, 0,810653), \end{aligned}$$

$$Q_3 = I_4 - 2w_3 {}^t w_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0,314317 & 0,949317 \\ 0 & 0 & 0,949317 & -0,314317 \end{pmatrix},$$

d'où :

$$M_3 = Q_3 M_2 = \begin{pmatrix} 2 & 3 & -0,500000 & 3,500000 \\ 0 & 3,316624 & 2,864357 & 2,864357 \\ 0 & 0 & 4,748205 & 6,433052 \\ 0 & 0 & 0 & 6,096006 \end{pmatrix}.$$

Finalement, $M = QR$, avec $R = M_3$ et :

$$Q = Q_3 Q_2 Q_1 = \begin{pmatrix} 0,500000 & 0,500000 & 0,500000 & -0,500000 \\ 0,150755 & 0,753778 & -0,452267 & 0,452267 \\ 0,382919 & -0,191459 & 0,536087 & 0,727547 \\ -0,762000 & 0,381000 & 0,508000 & 0,127000 \end{pmatrix}.$$

Nous n'avons écrit que les 6 premières décimales. ▽

Exemple 8.3.3. Tridiagonalisons la matrice symétrique :

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

Les trois dernières composantes de la première colonne donnent le vecteur :

$$u_1 = {}^t(0, 2, 2, 1), \quad \|u_1\| = 3,$$

d'où les vecteurs :

$$v_1 = u_1 - 3e_2 = {}^t(0, -1, 2, 1), \quad \|v_1\| = \sqrt{6}, \quad w_1 = v_1 / \|v_1\|,$$

et les matrices :

$$P_1 = \frac{1}{6} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & -2 & 4 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix}$$

et :

$$R_1 = I_4 - 2P_1 = \frac{1}{3} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 2 & 1 \\ 0 & 2 & -1 & -2 \\ 0 & 1 & -2 & 2 \end{pmatrix},$$

La première transformée de S est donc :

$$S_1 = R_1 S R_1 = \frac{1}{9} \begin{pmatrix} 36 & 27 & 0 & 0 \\ 27 & 14 & 20 & 13 \\ 0 & 20 & 26 & -11 \\ 0 & 13 & -11 & -4 \end{pmatrix}.$$

Pour la seconde transformation, considérons les vecteurs :

$$\begin{cases} u_2 = \frac{1}{9} {}^t(0, 0, 20, 13), \quad \|u_2\| = \sqrt{569}/9, \\ v_2 = \frac{1}{9} {}^t(20 - \sqrt{569}, 13), \quad \|v_2\|^2 = 1138 - 40\sqrt{569}, \\ w_2 = v_2/\|v_2\|, \end{cases}$$

d'où la matrice, en n'écrivant que les 6 premières décimales :

$$R_2 = I_4 - 2w_2 {}^t w_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0,838443 & 0,544988 \\ 0 & 0 & 0,544988 & -0,838443 \end{pmatrix}$$

et enfin la matrice tridiagonale :

$$\begin{aligned} S_2 &= R_2 S_1 R_2 = \begin{pmatrix} 4 & 3 & 0 & 0 \\ 3 & 1,555555 & 2,650413 & 0 \\ 0 & 2,650413 & 0,781878 & 2,019332 \\ 0 & 0 & 2,019332 & 1,662566 \end{pmatrix} \\ &= R_2 R_1 S R_1 R_2 = R S R \end{aligned}$$

avec :

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0,666666 & 0,666666 & 0,333333 \\ 0 & 0,740625 & -0,642806 & -0,195636 \\ 0 & 0,083844 & 0,377299 & -0,922288 \end{pmatrix}.$$

▽

9 Formes quadratiques

La relation quadratique :

$$4xy = (x + y)^2 - (x - y)^2$$

était connue des Babyloniens au 17^{ème} siècle avant J-C, et sans doute bien avant, et leur permettait de résoudre les équations du deuxième degré (ils ne considéraient que les racines positives).

Le théorème de Pythagore (-580,-490) était connu des Mésopotamiens 1000 ans auparavant.

Les Pythagoriciens avaient calculé les solutions entières de l'équation diophantienne $x^2 + y^2 - z^2 = 0$.

Al Khawarizmi (790-850) fonde une théorie des équations quadratiques.

En géométrie, Descartes, Pascal et Desargues ramènent l'étude des coniques à celle de formes quadratiques dans le plan projectif.

En mécanique Newtonienne, l'énergie cinétique et les moments d'inertie sont des formes quadratiques.

9.1 Formes quadratiques réelles

Une **forme quadratique** sur E , $\mathbb{R}$ -espace vectoriel de dimension n muni d'une base (ϵ_i) , est une fonction polynomiale homogène de degré 2 :

$$q((x_1, \dots, x_n)) = \sum_{i,j \geq 1} q_{ij} x_i x_j$$

que l'on peut écrire :

$$q((x_1, \dots, x_n)) = \sum_i a_{ii} x_i^2 + \sum_{i,j \geq 1} a_{ij} (x_i x_j + x_j x_i).$$

On voit que l'ensemble FQ_n des formes quadratiques sur E est un $\mathbb{R}$ -espace vectoriel de base $(x_i x_j)_{i \leq j}$, donc de dimension $n(n+1)/2$.

Une **forme bilinéaire** sur E est une application $B : E \times E \rightarrow \mathbb{R}$ linéaire sur chaque facteur. Elle est **symétrique** si $B(x, y) = B(y, x)$ quels que soient x et y dans E . L'ensemble FBS_n des formes bilinéaires symétriques est un $\mathbb{R}$ -espace vectoriel. On a par bilinéarité et symétrie :

$$B((x_i), (y_i)) = \sum_i x_i y_i B(\epsilon_i, \epsilon_i) + \sum_{i < j} (x_i y_j + x_j y_i) B(\epsilon_i, \epsilon_j),$$

de sorte que la forme B est déterminée par les $n(n+1)/2$ scalaires $B(\epsilon_i, \epsilon_j)$ pour $i \leq j$: FBS est de dimension $n(n+1)/2$. L'ensemble $S_n(\mathbb{R})$ des matrices symétriques de $\mathfrak{M}_n(\mathbb{R})$ est aussi de dimension $n(n+1)/2$ (proposition 2.3.2).

Les applications linéaires :

$$\begin{aligned} u : \quad S_n &\rightarrow FBS_n \\ S = (s_{ij}) &\mapsto B((x_i), (y_i)) = \sum_i s_{ii} x_i y_i + \sum_{i < j} s_{ij} (x_i y_j + x_j y_i) \end{aligned}$$

et :

$$\begin{aligned} v : \quad S_n &\rightarrow FQ_n \\ S = (s_{ij}) &\mapsto q((x_i)) = \sum_i s_{ii} x_i^2 + 2 \sum_{i < j} s_{ij} x_i x_j \end{aligned}$$

sont des isomorphismes, d'où l'isomorphisme :

$$\begin{aligned} v \circ u^{-1} : \quad FBS_n &\rightarrow FQ_n \\ B &\mapsto q : x \mapsto B(x, x). \end{aligned}$$

On peut retrouver B , la **forme polaire** de q , à partir de q grâce à la matrice $S = v^{-1}(q)$. On peut aussi remarquer que :

$$\begin{aligned} q(x+y) &= B(x+y, x+y) = B(x, x) + B(y, y) + 2B(x, y) \\ q(x-y) &= B(x-y, x-y) = B(x, x) + B(y, y) - 2B(x, y) \end{aligned}$$

d'où l'on déduit les **identités de polarisation** :

$$\begin{aligned} B(x, y) &= \frac{1}{4}(q(x+y) - q(x-y)) \\ &= \frac{1}{2}(q(x+y) - q(x) - q(y)). \end{aligned}$$

La forme bilinéaire symétrique B , ou la forme quadratique associée, est **positive** si, pour tout $x \in E$, $B(x, x) = q(x) \geq 0$, et **définie**, ou non dégénérée, si de plus $B(x, x) = 0$ implique $x = 0$. Dans ce cas B est un produit scalaire et q la norme associée.

Théorème 9.1.1 (Diagonalisation simultanée). *Soit E un espace euclidien, q une forme quadratique et B sa forme polaire. Il existe une base orthonormée dans laquelle q est une combinaison linéaire de carrés.*

Démonstration. Soit $\mathfrak{B}$ une base orthonormée de E . La matrice symétrique S de q dans $\mathfrak{B}$ se met sous la forme $S = PD {}^tP$, P étant une matrice orthogonale (théorème 6.7.1). La nouvelle base, $\mathfrak{B}'$, image de $\mathfrak{B}$ par cette matrice orthogonale (représentant une isométrie), est orthonormée (proposition 6.4.2). On a dans $\mathfrak{B}$, quels que soient x et y dans E :

$$\begin{aligned} q(x) &= {}^t x P D {}^t P x = {}^t ({}^t P x) D ({}^t P x), \\ B(x, y) &= {}^t x P D {}^t P y = {}^t ({}^t P x) D ({}^t P y), \end{aligned}$$

et dans $\mathfrak{B}'$, avec $D = \text{Diag}(\lambda_i)$ et en posant ${}^t P x = x' = (x'_i)$ et ${}^t P y = y' = (y'_i)$:

$$\begin{aligned} q((x')) &= \sum_i d_i x_i'^2 \\ B(x', y') &= \sum_i d_i x'_i y'_i. \end{aligned}$$

□

Quand une forme quadratique est mise sous la forme $\sum_i \lambda_i x_i^2$, le nombre r de λ_i non nuls est son **rang**. S'il y a s λ_i positifs et t λ_i négatifs ($s + t = r$), on dit que sa **signature** est (s, t) . La norme dans un espace euclidien de dimension n est une forme quadratique de signature $(n, 0)$. Ces notions sont dues à Sylvester. Gauss a également travaillé sur ce sujet.

Exemple 9.1.1 La forme quadratique sur $\mathbb{R}^3$ $q(x, y, z) = xy + yz$ s'écrit :

$$q(x, y, z) = (x, y, z) \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

On obtient $S = PD {}^tP$ avec :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}, \quad P = \frac{1}{2} \begin{pmatrix} \sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 1 & 1 \end{pmatrix}$$

et la forme quadratique q s'écrit $\sqrt{2}(y'^2 - z'^2)$ dans la nouvelle base (orthonormée). Son rang est 2, sa signature $(1, 1)$. ▽

En divisant les vecteurs de la nouvelle base par $\sqrt{|\lambda_i|}$ lorsque $\lambda_i \neq 0$ et en mettant d'abord les vecteurs associés aux $\lambda_i > 0$ puis ceux associés aux $\lambda_i < 0$, on obtient une base dans laquelle :

$$q((x_i)) = \sum_{1 \leq i \leq r} x_i^2 - \sum_{r+1 \leq i \leq r+s} x_j^2.$$

La **méthode de Gauss** est un moyen rapide pour réduire la forme quadratique $q((x_i)) = \sum_{ij} q_{ij} x_i x_j$ à une combinaison linéaire de carrés. Il s'agit d'absorber les doubles produits dans des carrés.

Si un q_{ii} est différent de 0, on peut supposer, à une permutation de coordonnées près, que c'est q_{11} . On met alors q sous la forme :

$$q((x_i)) = q_{11}(x_1 + \sum_{2 \leq j \leq n} \frac{q_{1j}}{2q_{11}} x_j)^2 + q'((x_2, \dots, x_n)),$$

de sorte que l'on a, en posant $x'_1 = x_1 + \sum_{2 \leq j \leq n} \frac{q_{1j}}{2q_{11}} x_j$:

$$q((x'_1, x_2, \dots, x_n)) = q_{11}x_1'^2 + q'((x_2, \dots, x_n)),$$

et on recommence avec q' , qui a $n - 1$ variables.

Si q ne contient que des doubles produits, supposons q_{12} non nul, encore à une permutation de coordonnées près. Le changement de variables $x'_1 = x_1 + x_2$, $x'_2 = x_1 - x_2$, $x'_i = x_i$ pour $i > 2$, remplace $4x_1x_2$ par $x_1'^2 - x_2'^2$, et on a :

$$q((x'_i)) = \frac{q_{12}}{4} (x_1'^2 - x_2'^2) + q_1((x'_i)),$$

où q_1 ne contient pas de carrés. Nous sommes ramenés à la situation précédente.

Voici un exemple avec un carré :

$$\begin{aligned} xy + y^2 + yz &= (y + \frac{x}{2} + \frac{z}{2})^2 - \frac{1}{4}(x + z)^2 \\ &= x'^2 - y'^2, \end{aligned}$$

de signature $(1, 1)$, pour lequel la diagonalisation donne :

$$xy + y^2 + yz = \frac{1}{2}((\sqrt{3} + 1)y_1^2 - (\sqrt{3} - 1)z_1^2).$$

Puis un exemple sans carré :

$$\begin{aligned} q((x, y, z, t)) &= xy + yz + zt \\ &= \frac{1}{4}[(x + y)^2 - (x - y)^2] + yz + zt. \end{aligned}$$

Posons $x + y = x'$, $x - y = y'$, $z = z'$, $t = t'$:

$$\begin{aligned} q((x, y, z, t)) &= \frac{1}{4}(x'^2 - y'^2) + \frac{1}{2}(x'z' - y'z') + z't' \\ &= \frac{1}{4}((x'^2 + 2x'z' + z'^2) - z'^2 - (y'^2 + 2y'z' + z'^2) + z'^2 + z't') \\ &= \frac{1}{4}(x''^2 - y''^2 + z''^2 - t''^2) \end{aligned}$$

avec $x' + z' = x''$, $y' + z' = y''$, $z' + t' = z''$, $z' - t' = t''$. La signature de q est $(2, 2)$.

La diagonalisation donne :

$$q(x, y, z, t) = \frac{1}{4}((\sqrt{5} + 1)x_1^2 + (\sqrt{5} - 1)x_2^2 - (\sqrt{5} + 1)x_3^2 + (\sqrt{5} - 1)x_4^2).$$

Les coefficients sont différents mais le rang et la signature sont les mêmes. Nous allons le démontrer dans le cas général.

Théorème 9.1.2 (Loi d'inertie de Sylvester). *La signature d'une forme quadratique q est indépendante de la base dans laquelle elle est mise sous forme d'une combinaison linéaire de carrés.*

Démonstration. Il existe une base dans laquelle :

$$q(x) = \sum_{1 \leq i \leq r} x_i^2 - \sum_{r+1 \leq i \leq r+s} x_i^2.$$

Supposons que dans une autre base :

$$q(x) = \sum_{1 \leq i \leq r'} l_i^2 - \sum_{r'+1 \leq i \leq r'+s'} l_i^2,$$

les l_i étant des formes linéaires en les (x_i) indépendantes. Remarquons que $r + s = r' + s'$, le rang de q ne dépendant pas de la base. Des deux expressions de $q(x)$ on déduit :

$$\sum_{1 \leq i \leq r} x_i^2 + \sum_{r'+1 \leq i \leq r'+s'} l_i^2 = \sum_{1 \leq i \leq r'} l_i^2 + \sum_{r+1 \leq i \leq r+s} x_i^2.$$

Les membres de cette égalité s'annulant simultanément, on a l'équivalence :

$$\begin{cases} x_1 = 0 \\ \vdots \\ x_r = 0 \\ l_{r'+1} = 0 \\ \vdots \\ l_{r'+s'} = 0 \end{cases} \iff \begin{cases} l_1 = 0 \\ \vdots \\ l_{r'} = 0 \\ x_{r+1} = 0 \\ \vdots \\ x_{r+s} = 0 \end{cases}$$

et les deux systèmes ont le même rang. Montrons que le premier est de rang $r + s'$. Supposons qu'il existe une combinaison linéaire :

$$\sum_{1 \leq i \leq r} a_i x_i + \sum_{r'+1 \leq j \leq r'+s} a_j l_j = 0,$$

d'où $\sum_{1 \leq i \leq r} a_i x_i = - \sum_{r'+1 \leq j \leq r'+s} a_j l_j = 0$. On a alors :

$$q\left(\sum_{1 \leq i \leq r} a_i x_i\right) = q\left(\sum_{r'+1 \leq j \leq r'+s} b_j l_j\right).$$

Le premier membre de cette équation est positif, le second est négatif : ils sont donc nuls, et les a_i sont tous nuls. Les l_j étant indépendantes, les b_j sont également tous nuls.

Le premier système est donc formé de $r + s'$ équations indépendantes. Le même raisonnement s'applique au second, dont le rang est donc $r' + s$. Du système :

$$\begin{cases} r + s' &= r' + s \\ r + s &= r' + s' \end{cases}$$

on déduit que $s' - s = s - s'$, d'où $s = s'$ et $r = r'$. $\square$

Exemple 9.1.2. Mécanique : détermination des axes principaux d'inertie d'un solide S . Soit Δ une droite vectorielle de vecteur unitaire directeur (α, β, γ) , orthogonale au plan vectoriel Π (donc d'équation $\alpha x + \beta y + \gamma z = 0$). Le moment d'inertie de S par rapport à Δ est égal à :

$$\mathbb{I}_{\Delta}(S) = \frac{1}{2} \iiint_S d^2(P, \Delta) dm$$

De la distance d'un point P de coordonnées (x, y, z) dans un repère orthonormé au plan Π :

$$d(P, \Pi) = \|\alpha x + \beta y + \gamma z\|,$$

de sa distance à l'origine $(x^2 + y^2 + z^2)^{1/2}$ et du théorème de Pythagore nous déduisons le carré de la distance de P à Δ :

$$\begin{aligned} d^2(P, \Delta) &= x^2 + y^2 + z^2 - \|\alpha x + \beta y + \gamma z\|^2 \\ &= (1 - \alpha^2)x^2 + (1 - \beta^2)y^2 + (1 - \gamma^2)z^2 \\ &\quad - 2(\alpha\beta xy + \beta\gamma yz + \gamma\alpha zx) \\ &= (\beta^2 + \gamma^2)x^2 + (\alpha^2 + \gamma^2)y^2 + (\alpha^2 + \beta^2)z^2 \\ &\quad - 2(\alpha\beta xy + \beta\gamma yz + \gamma\alpha zx), \end{aligned}$$

d'où :

$$\mathbb{I}_{\Delta}(S) = A\alpha^2 + B\beta^2 + C\gamma^2 - 2(\alpha\beta F + \beta\gamma D + \gamma\alpha E)$$

avec :

$$\begin{aligned} A &= \iiint_S (y^2 + z^2) dm, & B &= \iiint_S (z^2 + x^2) dm, & C &= \iiint_S (x^2 + y^2) dm, \\ D &= \iiint_S yz dm, & E &= \iiint_S zx dm, & F &= \iiint_S xy dm. \end{aligned}$$

Cette forme quadratique a pour matrice :

$$I(S) = \begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix}$$

dont les vecteurs propres orthonormés engendrent les axes principaux d'inertie de S . ∇

Exemple 9.1.3. Les coniques dans le plan projectif réel. Le plan projectif $\mathbb{P}^2(\mathbb{R})$ est l'ensemble $\mathbb{R}^3 \setminus \{0\}$ quotienté par la relation d'équivalence :

$$(x, y, z) \sim (x', y', z') \iff \exists \lambda \in \mathbb{R}^* : (x', y', z') = \lambda(x, y, z).$$

C'est l'ensemble des droites vectorielles privées de l'origine. L'équation projective d'une conique est $x^2 + y^2 - z^2 = 0$ dans un repère orthogonal. Dans le plan $z = 1$ on obtient l'ellipse, dans le plan $y = 1$, l'hyperbole, et dans le plan $y + z = 1$ la parabole. ∇

9.2 Formes quadratiques hermitiennes

Une **forme quadratique hermitienne** q sur l'espace hermitien H est définie à partir d'un endomorphisme hermitien u :

$$\forall z \in H, \quad q(z) = (u(z) | z).$$

Pour tout $\lambda \in \mathbb{C}$ on a :

$$q(\lambda z) = (u(\lambda z) | \lambda z) = (\lambda u(z) | \lambda z) = \bar{\lambda} \lambda q(z) = \|\lambda\|^2 q(z)$$

et en particulier $q(iz) = q(z)$.

La matrice de u dans une base orthonormée est la matrice de q , ou de sa forme polaire $\phi(z, z') = (u(z) | z')$. Notons que ϕ est **sesquilinéaire** :

$$\begin{aligned} \phi(\lambda z, z') &= (\lambda u(z) | z') = \bar{\lambda} \phi(z, z') \\ \phi(z, \lambda z') &= (u(z) | \lambda z') = \lambda \phi(z, z') \end{aligned}$$

et possède la **symétrie hermitienne** :

$$\phi(z', z) = (u(z') | z) = (z' | u(z)) = \overline{(u(z) | z')} = \overline{\phi(z, z')}$$

quels que soient $\lambda \in \mathbb{C}$ et z et z' dans H .

Proposition 9.2.1. *Quel que soit $z \in H$, $q(z)$ est réel et l'ensemble FQH_n des formes quadratiques hermitiennes sur H est un $\mathbb{R}$ -espace vectoriel de dimension n^2 .*

Démonstration. L'endomorphisme u étant hermitien, on a :

$$q(z) = (u(z) | z) = (z | u(z)) = \overline{(u(z) | z)} = \overline{q(z)}$$

et $q(z)$ est réel. L'application $u \mapsto q$ du $\mathbb{R}$ -espace vectoriel $\Sigma(H)$, de dimension n^2 (proposition 5.8.4), dans FQH_n est linéaire, surjective par construction et injective ($q = 0$ implique $u = 0$). C'est donc un isomorphisme. $\square$

L'identité de polarisation s'écrit dans ce cas :

$$\phi(z, z') = \frac{1}{4} (q(z + z') - q(z - z') + i q(z - iz') - i q(z + iz')).$$

Exemple 9.2.1. La forme polaire de la forme quadratique hermitienne sur $\mathbb{C}^2$:

$$q((z_1, z_2)) = i(\bar{z}_1 z_2 - z_1 \bar{z}_2)$$

a pour matrice :

$$\Phi = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Remarquons que $q((x + iy, x' + iy')) = 2(x'y - xy')$ est bien réel.

Le polynôme caractéristique de Φ est $X^2 - 1$, ses valeurs propres sont 1 et -1 , ses vecteurs propres, respectivement, ${}^t(1, -i)$ et ${}^t(-i, 1)$, d'où les matrices de passage :

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = {}^t\bar{P} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

et la diagonalisation $\Phi = P \text{Diag}(1, -1) P^{-1}$. Dans la nouvelle base on a :

$$q((z_1, z_2)) = z_1 \bar{z}_1 - z_2 \bar{z}_2 = \|z_1\|^2 - \|z_2\|^2.$$

A partir de :

$$\begin{aligned} (z_1 + iz_2) \overline{(z_1 + iz_2)} &= \|z_1\|^2 + \|z_2\|^2 + i(\bar{z}_1 z_2 - z_1 \bar{z}_2), \\ (iz_1 + z_2) \overline{(iz_1 + z_2)} &= \|z_1\|^2 + \|z_2\|^2 - i(\bar{z}_1 z_2 - z_1 \bar{z}_2) \end{aligned}$$

et en posant $z'_1 = z_1 + iz_2$ et $z'_2 = iz_1 + z_2$ on obtient :

$$z'_1 \bar{z}'_1 - z'_2 \bar{z}'_2 = 2i(\bar{z}_1 z_2 - z_1 \bar{z}_2)$$

d'où $q(z'_1, z'_2) = \frac{1}{2}(\|z'_1\|^2 - \|z'_2\|^2)$. ▽

La loi d'inertie de Sylvester s'applique encore, ainsi que le concept de **signature**. On peut en effet reprendre la démonstration du théorème 9.1.2 en remplaçant les carrés (x_i^2, l_i^2) par les carrés des modules $(x_i \bar{x}_i = |x_i|^2, l_i \bar{l}_i = |l_i|^2)$, dire que ces carrés s'annulent simultanément, d'où la nullité des x_i et des l_i . La suite de la démonstration est inchangée.

Exemple 9.2.2. Appliquons la méthode de Gauss à la forme quadratique hermitienne sur $\mathbb{C}^3$:

$$q((z_1, z_2, z_3)) = |z_1|^2 + \bar{z}_1 z_2 + \bar{z}_2 z_1 - 2i(\bar{z}_1 z_3 - \bar{z}_3 z_1).$$

Si $l_1 = z_1 + z_2 - 2iz_3$, on a :

$$|l_1|^2 = |z_1|^2 + \bar{z}_1 z_2 + \bar{z}_2 z_1 - 2i(\bar{z}_1 z_3 - \bar{z}_3 z_1) + |z_2|^2 - 2i(\bar{z}_2 z_3 - \bar{z}_3 z_2) + 4|z_3|^2$$

et :

$$\begin{aligned} q((z_1, z_2, z_3)) - |l_1|^2 &= -|z_2|^2 + 2i(\bar{z}_2 z_3 - \bar{z}_3 z_2) - 4|z_3|^2 \\ &= -(z_2 - 2iz_3)(\bar{z}_2 - 2i\bar{z}_3) \end{aligned}$$

enfin :

$$q = |l_1|^2 - |l_2|^2$$

avec $l_2 = z_2 - iz_3$. La signature de q est $(1, 1)$. La forme diagonale de sa matrice associée :

$$Q = \begin{pmatrix} 1 & 1 & -2i \\ 1 & 0 & 0 \\ 2i & 0 & 0 \end{pmatrix}$$

est $\text{Diag}(1 + \frac{\sqrt{21}}{2}, 1 - \frac{\sqrt{21}}{2}, 0)$, ce qui confirme la signature.

9.3 En caractéristique 2

Les choses se passent plutôt mal. Prenons l'espace vectoriel $\mathbb{F}_2^2$. La forme quadratique $q((x, y)) = xy$ a pour matrice la matrice J_1 de 7.10, non diagonalisable. On ne peut la mettre sous forme d'une somme de carrés et elle n'a pas de forme polaire.

Le produit $(x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2$, que l'on admet en théorie du codage comme produit scalaire, est tel que $(1, 1) \cdot (1, 1) = 0$: le vecteur $(1, 1)$ est **isotrope**, c'est-à-dire orthogonal à lui-même.

10 Systèmes différentiels linéaires

Un système différentiel linéaire est une équation avec des conditions initiales :

$$x'(t) = M x(t), \quad x(0) = x_0,$$

où $x(t) = (x_i(t))$ est un vecteur de $\mathbb{R}^n$ fonction de t , $x'(t) = (x'_i(t))$ le vecteur dérivé en t , M une matrice de $\mathfrak{M}_n(\mathbb{R})$, et x_0 un vecteur de $\mathbb{R}^n$, appelé **conditions initiales**.

10.1 Réduction de M

Si $M = P J P^{-1}$ (J étant la forme de Jordan de M), on peut écrire ce système :

$$x'(t) = P J P^{-1} x(t), \quad x(0) = x_0$$

et en posant $y(t) = P^{-1} x(t)$, $y'(t) = P^{-1} x'(t)$:

$$y'(t) = J y(t), \quad y(0) = (c_i) = P^{-1} x_0.$$

Si les valeurs propres sont $\lambda_1, \dots, \lambda_s$, on commence par résoudre la dernière équation du système $y'_n = \lambda_s y_n$, ce qui donne $y_n(t) = c_n e^{\lambda_s t}$.

On a ensuite à résoudre des équations du type $y'_k(t) = \lambda_h y_k(t)$, $y_k(0) = c_k$, d'où $y_k(t) = c_k e^{\lambda_h t}$, ou du type $y'_k(t) = \lambda_h y_k(t) + y_{k-1}(t)$, $y_k(0) = c_k$, $y_{k-1}(t)$ étant déjà calculé. On cherche dans ce cas une solution particulière $z(t)$ de $z'(t) = y_{k-1}(t)$, $z(0) = z_0$ (z_0 quelconque, mais non nul) que l'on ajoute à la solution de l'équation $y'_k = \lambda_h y_k(t)$, $y_k(0) = c_k - z_0$.

Une fois $y(t)$ calculé, on revient à $x(t)$, égal à $P y(t)$. Remarquons que l'on n'a pas à calculer P^{-1} .

Le calcul du polynôme caractéristique nécessite $\sim 2n^4$ opérations par l'algorithme de Leverrier. Les autres calculs sont d'un ordre inférieur.

Nous allons présenter une autre méthode.

10.2 Exponentielle de M

On vérifie sans mal que l'application :

$$\begin{aligned} \mathfrak{M}_n(\mathbb{R}) &\rightarrow \mathbb{R}_+ \\ M &\mapsto \|M\| = \sup_{\|x\|=1} \|M x\| \end{aligned}$$

est une norme. On a évidemment $\|M^2\| \leq \|M\|^2$, car le x assurant le *sup* de $M^2 x$ n'est pas en général celui qui assure le *sup* de $M x$. D'où par récurrence $\|M^p\| \leq \|M\|^p$ pour tout $p > 1$. Posons, pour un réel t :

$$M_p(t) = I_n + t M + \frac{t^2}{2!} M^2 + \cdots + \frac{t^p}{p!} M^p \quad (I_n = M^0).$$

Nous avons :

$$\begin{aligned} \|M_p(t)\| &\leq 1 + |t| \|M\| + \frac{|t|^2}{2!} \|M\|^2 + \cdots + \frac{|t|^p}{p!} \|M\|^p \\ &\leq e^{|t| \|M\|}, \end{aligned}$$

ce qui assure la convergence de la série matricielle $\sum_{p \geq 0} \frac{t^p}{p!} M^p$, dont la somme est, par définition, l'**exponentielle** de tM , notée e^{tM} . Si M est diagonalisable ($M = P D P^{-1}$), on a pour tout n :

$$M^n = P D P^{-1} D D P^{-1} P \cdots P^{-1} P D P^{-1} = P D^n P^{-1},$$

et $e^{tM} = P e^{tD} P^{-1}$; e^{tD} est la matrice diagonale dont les termes diagonaux sont les exponentielles des valeurs propres de tM . On peut également calculer e^{tM} lorsque M est réductible à la forme de Jordan non diagonale. Ces méthodes exigent beaucoup de calculs.

La convergence étant **normale** (en norme), on peut dériver terme à terme :

$$\frac{d}{dt} e^{tM} = M + t M^2 + \frac{t^2}{2!} M^3 + \cdots = M e^{tM}.$$

appliquons au vecteur x_0 :

$$\frac{d}{dt} (e^{tM} x_0) = M e^{tM} x_0.$$

Le vecteur $x(t) = e^{tM} x_0$ est donc solution du système $x'(t) = M x(t)$, $x(0) = x_0$ (pour $t = 0$, $e^{tM} = I_n$).

Nous avons $e^{tM} = M_p(t) + R_p(t)$ avec :

$$\begin{aligned} \|R_p(t)\| &= \left| \frac{t^{p+1}}{(p+1)!} \right| \left\| 1 + \frac{tM}{p+2} + \frac{t^2 M^2}{(p+2)(p+3)} + \dots \right\| \\ &\leq \left| \frac{t^{p+1}}{(p+1)!} \right| \|e^{tM}\|. \end{aligned}$$

Ceci montre que $M_p(t)$ approche **d'aussi près qu'on veut** e^{tM} , avec p suffisamment grand. Si n est très grand, beaucoup plus que le p assurant la précision souhaitée, le calcul exige $\sim 2pn^3$ opérations, donc beaucoup moins que la première méthode ($\sim 2n^4$). Or un calcul approché, si l'erreur est inférieure aux erreurs d'arrondis d'une méthode exacte, ou à la précision des mesures, est équivalent à une méthode exacte.

Proposition 10.2.1. *Si les matrices A et B de $\mathfrak{M}_n(\mathbb{R})$ commutent, on a :*

$$e^{A+B} = e^A e^B.$$

Démonstration. La différence $M_{A,p} M_{B,p} - M_{A+B,p}$ est inférieure en module à $\left\| \frac{t^{p+1}}{(p+1)!} (M_{A,p} M_{B,p}) \right\|$ et tend donc vers 0 quand $p \rightarrow \infty$. $\square$

Exemple 10.2.1. Résolvons le système différentiel linéaire :

$$\begin{cases} x'(t) &= 2x(t) - 3y(t), \quad x(0) = 1, \\ y'(t) &= x(t) - 2y(t), \quad y(0) = 1. \end{cases}$$

La réduction de la matrice M du système donne :

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

On a :

$$e^{tM} = P e^{tD} P^{-1} = \frac{1}{2} \begin{pmatrix} 3e^t - e^{-t} & -3e^t + 3e^{-t} \\ e^t - e^{-t} & -e^t + 3e^{-t} \end{pmatrix},$$

et e^{tM} appliquée au vecteur $(1, 1)$ donne la solution $x(t) = y(t) = e^{-t}$. ∇

11 Interpolation de Lagrange

Voici une méthode cryptographique, le schéma à seuil de Shamir³, qui utilise l'interpolation de Lagrange. Un nombre réel a_0 est le secret. On génère aléatoirement des entiers naturels $a_1, \dots, a_{t-1}$, inférieurs à un nombre premier p , et on construit le polynôme $P = \sum_{i=0}^{t-1} a_i X^i \in \mathbb{F}_p[X]$, $\mathbb{F}_p$ étant le corps des entiers modulo p . Si n personnes (n beaucoup plus grand que t) sont numérotées de 1 à n , on donne à chacune le nombre $P(i)$, i étant son numéro. Alors, si t (au moins) d'entre ces personnes se réunissent, elles peuvent retrouver P par interpolation et donc a_0 . On sait en effet qu'un polynôme de degré $t-1$ est entièrement déterminé par sa valeur en t points distincts.

3. Handbook of Applied cryptography, A. Menezes, P. van Oorschot and S. Vanstone, CRC Press, 1996, ch.12, p. 526.

11.1 Polynômes de Lagrange

Revenons au cas général. Etant donné un intervalle $I = [a, b] \subset \mathbb{R}$ et $n + 1$ points $x_0 < x_1 < \dots < x_n$ équirépartis de I constituant le **support** de l'interpolation, les polynômes :

$$L_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i}$$

sont appelés **polynômes de Lagrange**. En posant $x_i = x_0 + ih$ ($h = (b - a)/n$) et $x = x_0 + sh$, on obtient :

$$L_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{s - i}{k - i} = l_k(s),$$

les $l_k(s)$ pouvant être mis dans une base de données.

11.2 Matrices de Lagrange

Cherchons un polynôme $P = \sum_{0 \leq i \leq n} \lambda_i X^i \in \mathbb{R}[X]$ (le polynôme d'interpolation) de degré au plus égal à n et tel que $P(x_i) = y_i$ pour $0 \leq i \leq n$, les (y_i) étant des réels donnés. Ceci se traduit par le système :

$$\begin{cases} \lambda_0 + \lambda_1 x_0 + \dots + \lambda_n x_0^n &= y_0, \\ \vdots \\ \lambda_0 + \lambda_1 x_n + \dots + \lambda_n x_n^n &= y_n, \end{cases}$$

qui s'écrit $A \Lambda = y$. Le déterminant de A est un déterminant de Vandermonde, donc non nul, et le système admet la solution unique $\Lambda = A^{-1}y$.

Le polynôme cherché s'écrit donc $P(X) = (1, X, \dots, X^n) A^{-1}y$.

Il nous reste à calculer A^{-1} . Si $y = e_k$, $A^{-1}e_k$ est la k -ième colonne de A^{-1} , $C_k(A^{-1})$. On a $L_j(A) C_k(A^{-1}) = \delta_{jk}$ d'après la définition de la matrice inverse. Or l'expression :

$$L_j(A) = (1, x_j, \dots, x_j^n) = (1, X, \dots, X^n)|_{X=x_j}$$

montre que $(1, X, \dots, X^n)A^{-1}e_k$ est le k -ième polynôme de Lagrange, L_k . La k -ième colonne de A^{-1} est donc formée des coefficients du polynôme L_k , selon les puissances croissantes du haut vers le bas.

Par exemple, pour le support $(0, 1, 2)$, on a :

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

puis :

$$\begin{aligned} L_1(x) &= -1/2(x-1)(x-2) = -1(x^2-3x+2)/2, \\ L_2(x) &= -x(x-2) = -x^2+2x, \\ L_3(x) &= x(x-1)/2 = (x^2-x)/2, \end{aligned}$$

d'où :

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

Cette méthode d'interpolation ne peut être utilisée quand n est grand, car x_i^n est pris pour 0 par le calculateur si $|x_0| < 1$ et pour l'infini si $|x_0| > 1$.

12 Interpolation de Tchebychev

12.1 Polynômes et support de Tchebychev

L'idée de départ est de poser $y = \cos \theta$ et $T_k(y) = \cos(k\theta)$. Il faut donc d'abord ramener l'intervalle d'interpolation $[a, b]$ à $[-1, 1]$, ce que l'on fait en posant :

$$y = \frac{2x - a - b}{b - a}.$$

Pour calculer les T_k , la formule :

$$T_{k+1}(y) = \cos((k+1)\theta) = 2 \cos \theta \cos(k\theta) - \cos((k-1)\theta),$$

déduite de la formule d'addition :

$$\cos p + \cos q = 2 \cos \frac{p+q}{2} \cos \frac{p-q}{2}, \text{ avec } p = (n+1)\theta, \quad q = (n-1)\theta,$$

donne la relation de récurrence :

$$T_{k+1}(y) = 2yT_k(y) - T_{k-1}(y)$$

qui engendre les T_k à partir de $T_1(y) = y$. Pour une raison qui se manifestera ultérieurement, on pose $T_0(y) = \sqrt{2}/2$ au lieu de 1. La formule donne :

$$T_2(y) = 2y^2 - 1, \quad T_3(y) = 4y^3 - 3y, \quad T_4(y) = 8y^4 - 8y^2 + 1 \dots$$

Elle montre que les T_{2k} sont pairs et les T_{2k+1} impairs, que $T_k(-1) = -1$ et $T_k(1) = 1$, et que les T_k prennent leur valeur entre -1 et 1 pour $-1 \leq y \leq 1$, ce qui évite les problèmes posés par les polynômes de Lagrange.

Les k racines de T_k sont réelles et comprises entre -1 et 1 .

En effet, $\cos(k\theta) = 0$ équivaut à $\theta = \pm(\pi/2k + h2\pi/k)$, $0 \leq h \leq k-1$, ce qui donne, pour y , k valeurs distinctes dans $] -1, 1[$.

Les $n+1$ points du support de Tchebychev sont :

$$y_j = -\cos \frac{(2j+1)\pi}{2n+2}, \quad 0 \leq j \leq n-1.$$

12.2 Matrices de Tchebychev

Passons à l'interpolation. Les valeurs réelles $f_0, \dots, f_n$ étant données, on veut construire un polynôme $P = \sum_{0 \leq i \leq n} a_i T_i$ tel que $P(y_k) = f_k$, $0 \leq k \leq n-1$. Ces conditions mènent au système :

$$\begin{cases} a_0 T_0(y_0) + \dots + a_n T_n(y_0) = f_0 \\ \vdots \\ a_0 T_0(y_n) + \dots + a_n T_n(y_n) = f_n \end{cases}$$

qui s'écrit $TA = F$ avec :

$$T = \begin{pmatrix} T_0(y_0) & \dots & T_n(y_0) \\ \vdots & & \vdots \\ T_0(y_n) & \dots & T_n(y_n) \end{pmatrix}, \quad A = \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}, \quad F = \begin{pmatrix} f_0 \\ \vdots \\ f_n \end{pmatrix}.$$

Calculons ${}^t T T$, en posant $\sum_k T_p(y_k) T_q(y_k) = \sum T_p T_q$:

$$\begin{aligned} {}^t T T &= \begin{pmatrix} T_0(y_0) & \dots & T_0(y_n) \\ \vdots & & \vdots \\ T_n(y_0) & \dots & T_n(y_n) \end{pmatrix} \begin{pmatrix} T_0(y_0) & \dots & T_n(y_0) \\ \vdots & & \vdots \\ T_0(y_n) & \dots & T_n(y_n) \end{pmatrix} \\ &= \begin{pmatrix} \sum T_0 T_0 & \dots & \sum T_0 T_n \\ \vdots & & \vdots \\ \sum T_n T_0 & \dots & \sum T_n T_n \end{pmatrix}. \end{aligned}$$

Les termes diagonaux valent $(n+1)/2$ et tous les autres sont nuls. En effet, pour $p > 0$ et $q > 0$, les égalités :

$$\begin{aligned} \sum_k T_p(y_k) T_q(y_k) &= \sum_k \cos\left(\frac{2k+1}{2n+2} p\pi\right) \cos\left(\frac{2k+1}{2n+2} q\pi\right) \\ &= \frac{n+1}{2} \text{ si } p = q \\ &= 0 \text{ si } p \neq q \end{aligned}$$

ainsi que $\sum T_0 T_0 = (n+1)/2$ et $\sum T_0 T_p = 0$, découlent du choix de T_0 et des formules de trigonométrie :

$$\begin{aligned} 2 \cos^2 x &= \cos 2x + 1, \\ \cos(a+b) + \cos(a-b) &= 2 \cos a \cos b, \\ \sum_{0 \leq k \leq n} \cos\left(\frac{2k+1}{2n+2} r\pi\right) &= 0 \text{ si } r \neq 0. \end{aligned}$$

On a donc $T^{-1} = \frac{2}{n+1} {}^t T$.

Pour $n = 2$ on obtient le support $(-\sqrt{3}/2, 0, \sqrt{3}/2)$, et :

$$T = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{3}/2 & 1/2 \\ \sqrt{2}/2 & 0 & -1 \\ \sqrt{2}/2 & \sqrt{3}/2 & 1/2 \end{pmatrix}, \quad T^{-1} = \frac{2}{3} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{3}/2 & 0 & \sqrt{3}/2 \\ 1/2 & -1 & 1/2 \end{pmatrix}.$$

Pour les données $F = (6, 2, 3)$, on obtient $a = T^{-1}F = (4\sqrt{3}, -\sqrt{3}, 1)$ et :

$$P(y) = 4\sqrt{3}T_0 - \sqrt{3}T_1 + T_2 = 2y^2 - \sqrt{3}y + 3.$$

L'interpolation polynomiale permet d'approcher d'aussi près que l'on veut (avec n suffisamment grand) l'intégrale définie d'une fonction f dont on n'a pas de primitive par celle d'un polynôme P tel que $P = f$ sur les points du support.

13 Lissage par fonction spline cubique

13.1 Définition

Etant donné l'intervalle $I = [0, 1]$ de $\mathbb{R}$, le **support** $(x_i)_{0 \leq i \leq n}$, $x_i = i/n$, des valeurs réelles $(y_i)_{0 \leq i \leq n}$ et un entier $k \geq 1$, n étant bien supérieur à k , il s'agit de construire une fonction φ , polynomiale de degré inférieur ou égal à $2k - 1$ sur chaque $]x_i, x_{i+1}[$, de classe C^{2k-2} sur I , et valant y_i en x_i .

On prend généralement $k = 2$ pour avoir des polynômes de degré au plus égal à 3, ce qui donne le meilleur rendement (temps de calcul/précision) et minimise la courbure totale de la courbe représentant φ sur I . C'est ce que nous allons faire. On dit alors que la fonction φ est **spline cubique**.

Cette méthode évite l'inconvénient de l'interpolation de Lagrange pour laquelle il y a un seul polynôme dont le degré augmente avec le nombre de points du support, ce qui la rend impraticable si n est grand, comme nous l'avons remarqué.

13.2 Calcul

Si χ_m est la fonction caractéristique de l'intervalle $[x_m, x_{m+1}[$, c'est-à-dire la fonction valant 1 sur $[x_m, x_{m+1}[$ et 0 ailleurs, P_0 et P_{n-1} des polynômes de degré inférieur ou égal à 1, et P_i , $1 \leq i \leq n-2$, des polynômes de degré inférieur ou égal à 3, posons :

$$\varphi(x) = \sum_{m=0}^{n-1} \chi_m(x) P_m(x).$$

Sur l'intervalle $[x_m, x_{m+1}[$ la fonction φ est égale au polynôme P_m ; donc aussi en x_{m+1} par continuité; or en ce point elle est égale à $P_{m+1}(x_{m+1})$, d'où $P_m(x_{m+1}) = P_{m+1}(x_{m+1})$. Posons $\varphi''(x_m) = z_m$; comme $x_i = \frac{i}{n}$, nous aurons successivement :

$$\begin{aligned} P_m''(x) &= n\{-z_m(x - x_{m+1}) + z_{m+1}(x - x_m)\}, \\ P_m'(x) &= \frac{n}{2}\{-z_m(x - x_{m+1})^2 + z_{m+1}(x - x_m)^2\} + C_m, \\ P_m(x) &= \frac{n}{6}\{-z_m(x - x_{m+1})^3 + z_{m+1}(x - x_m)^3\} + C_m x + D_m. \end{aligned}$$

Comme $P_m(x_m) = y_m$ et $P_m(x_{m+1}) = y_{m+1}$, nous aurons aux bornes :

$$\begin{aligned} y_m &= \frac{n}{6} z_m (x_{m+1} - x_m)^3 + C_m x_m + D_m, \\ y_{m+1} &= \frac{n}{6} z_{m+1} (x_{m+1} - x_m)^3 + C_m x_{m+1} + D_m, \end{aligned}$$

c'est-à-dire :

$$\begin{aligned} y_m &= \frac{z_m}{6n^2} + \frac{mC_m}{n} + D_m, \\ y_{m+1} &= \frac{z_{m+1}}{6n^2} + \frac{(m+1)C_m}{n} + D_m, \end{aligned}$$

d'où nous tirons :

$$\begin{aligned} C_m &= n(y_{m+1} - y_m) - \frac{1}{6n}(z_{m+1} - z_m), \\ D_m &= (m+1)y_m - my_{m+1} + \frac{1}{6n^2}[mz_{m+1} - (m+1)z_m], \end{aligned}$$

Raccordons maintenant les dérivées premières en x_m :

$$\begin{aligned} P'_m(x_m) &= -\frac{1}{2n}z_m + C_m, \\ P'_{m-1}(x_m) &= \frac{1}{2n}z_m + C_{m-1}, \end{aligned}$$

soit en égalant, en remplaçant les C_i par leur valeur et en multipliant par n :

$$z_{m-1} + 4z_m + z_{m+1} = 6n^2(y_{m-1} - 2y_m + y_{m+1})$$

avec $z_1 = z_{n-1} = 0$, les deux polynômes extrêmes P_0 et P_{n-1} étant de degré au plus égal à 1.

Nous obtenons ainsi l'équation matricielle :

$$\begin{pmatrix} 4 & 1 & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} z_2 \\ z_3 \\ \vdots \\ z_{n-3} \\ z_{n-2} \end{pmatrix} = \begin{pmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \end{pmatrix}$$

soit $MZ = Y$.

La matrice M est à diagonale dominance stricte; elle est donc inversible, et le problème admet une et une seule solution.

13.3 Résolution

Les valeurs propres de la matrice sont $\lambda_p = 4 + 2 \cos \frac{p\pi}{n-2}$ pour $p = 1, \dots, n-3$.

Etant symétrique, elle se diagonalise dans une base orthonormée; la matrice de passage s'écrit pour $n \geq 4$:

$$P = P^{-1} = \sqrt{\frac{2}{n-2}} \begin{pmatrix} \sin \frac{\pi}{n-2} & \sin \frac{2\pi}{n-2} & \dots & \sin \frac{(n-3)\pi}{n-2} \\ \sin \frac{2\pi}{n-2} & \sin \frac{4\pi}{n-2} & \dots & \sin \frac{(n-3)2\pi}{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sin \frac{(n-3)\pi}{n-2} & \sin \frac{(n-3)2\pi}{n-2} & \dots & \sin \frac{(n-3)(n-3)\pi}{n-2} \end{pmatrix}.$$

L'équation $MZ = Y$ où $M = PDP$ se résoud finalement en $Z = PD^{-1}PY$, de sorte que, à partir de la donnée des y_i , on obtient successivement les z_i , les C_i et les D_i et enfin les P_i .

Si n est grand, certaines composantes de P seront très petites.

La décomposition triangulaire $M = TU$ est ici bien adaptée, car T et U sont bi-diagonales, et leur calcul nécessite $3n - 3$ opérations. Pour $n = 2$, on pose :

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

et on obtient $a = 1/4$ et $b = 15/4$ grâce à trois opérations. Pour passer à $n = 3$, on ajoute à T une troisième colonne $(0, 0, 1)$ et une troisième ligne $(0, a, 1)$, à U une troisième colonne $(0, 1, b)$ et une troisième ligne $(0, 0, b)$, d'où $a = 4/15$ et $b = 56/15$ grâce à trois opérations. Le passage de $n = k$ à $n = k + 1$ s'effectue de la même façon.

Index

- Adjoint (endomorphisme), 63
- Adjointe (matrice), 13
- Algébriquement clos (corps), 4
- Algèbre sur un corps, 6
- Angle de deux vecteurs, 67
- Anneau, 4
- Application linéaire, 14
- Auto-adjoint, 74
- Auto-adjoint (endomorphisme), 101, 105
- Automorphisme, 15

- Base, 5
- Base canonique, 5
- Base directe, 61
- Base duale, 27
- Base indirecte, 61
- Base orthonormée, 61
- Bidual, 26
- Bloc de Jordan, 92

- Clôture algébrique, 4
- Combinaison linéaire, 5
- Complexifié, 71
- Corps, 4
- Covecteur, 26
- Cycle, 30

- Déterminant d'un endomorphisme, 39
- Déterminant d'une matrice, 34
- Déterminant de n vecteurs, 32
- Déterminant de Vandermonde, 39
- Demi-tour, 106
- Diagonale dominance, 55
- Diagonalisable (endomorphisme), 84
- Dimension, 5
- Dual, 26

- Endomorphisme, 14
- Endomorphisme antisymétrique, 101
- Endomorphisme symétrique, 101
- Espace caractéristique, 82
- Espace euclidien, 59
- Espace hermitien, 72
- Espace propre, 40

- Espace vectoriel, 5
- Exponentielle d'une matrice, 129

- Famille génératrice, 5
- Famille libre, 5
- Forme n -lin. alternée, 31
- Forme n -lin. antisymétrique, 31
- Forme n -lin. symétrique, 31
- Forme n -linéaire, 30
- Forme bilinéaire, 121
- Forme linéaire, 26
- Forme polaire, 60
- Forme quadratique, 121
- Forme quadratique hermitienne, 126

- Groupe, 4
- Groupe orthogonal, 66
- Groupe symétrique, 30
- Groupe unitaire, 78

- Hermitien (endomorphisme), 74, 105
- Homomorphisme, 14
- Hyperplan, 103

- Image, 14
- Injection, 15
- Involution, 15
- Isométrie, 66
- Isométrie (hermitienne), 76
- Isomorphisme, 15
- Isomorphisme canonique, 16
- Isotrope (vecteur), 128

- Matrice, 10
- Matrice antisymétrique, 12
- Matrice associée, 18
- Matrice de passage, 24
- Matrice hermitienne, 13
- Matrice identité, 11
- Matrice inverse, 21
- Matrice normale, 13
- Matrice régulière, 16
- Matrice symétrique, 12
- Matrice triangulaire, 12

Matrice unitaire, 13
 Matrices équivalentes, 19
 Matrices semblables, 26
 Mineur, 35

 Nilpotent (endomorphisme), 90
 Nilpotente (matrice), 90
 Normé (vecteur), 60
 Normal (endomorphisme), 76
 Norme, 56
 Noyau, 14

 Ordre d'une valeur propre, 82
 Orthogonal, 65
 Orthogonalité, 59
 Orthogonaux (s-esp), 65

 Permutation, 29
 Pivot, 46
 Polarisation (identité de), 60
 Polynôme caractéristique, 40
 Polynôme minimal, 95
 Polynômes de Lagrange, 131
 Produit formel, 20
 Produit scalaire, 59
 Produit scalaire hermitien, 72
 Produit vectoriel, 69
 Projecteur, 20
 Projecteur orthogonal, 65

 Réflexion, 67, 106
 Rang d'un système, 42
 Rang d'une appl. lin., 15
 Rang d'une matrice, 47
 Rotation, 67

 Scalaire, 5
 Signature d'une forme quadratique, 122
 Signature d'une permutation, 31
 Somme (de sous-espaces), 6
 Somme directe, 6
 Sous-espace vectoriel, 5
 Spectre (d'un endomorphisme), 40
 Spline cubique, 134
 Support d'interpolation, 131
 Surjection, 15
 Symbole de Kronecker δ_{ij} , 7

 Système compatible, 42
 Système de Cramer, 42
 Système homogène, 42

 Trace d'un endomorphisme, 82
 Trace d'une matrice, 13
 Transposée (appl. lin.), 28
 Transposée (matrice), 11
 Transposition, 30

 Unitaire (vecteur), 60

 Valeur propre, 40
 Vecteur, 5
 Vecteur propre, 40
 Vectoriel $x_1, \dots, x_k$, 5