Algorithmique mathématique 2

Arnaud Jehanne

28 septembre 2022

Table des matières

Chapitre 1

Introduction

Ce cours porte sur l'algorithmique des matrices à coefficients dans $K = \mathbb{R}$, $\mathbb{C}$ ou $\mathbb{Q}$. Nous travaillerons en particuliers sur la résolution d'équations, sur la recherche de valeurs propres et sur les aspects matriciels liés à l'algèbre bilinéaire.

Nous donnons ci-dessous quelques rappels et notations sur les deux premiers points.

1.1. Résolution d'équations

Nous nous intéresserons à la résolution de systèmes d'équations à coefficients dans K. Un tel système s'écrit

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

On peut aussi l'écrire sous forme matricielle :

$$AX = B$$

où $A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}$ appartient au K-espace vectoriel $M_{m,n}(K)$ des matrices à m lignes et n colonnes à coefficients dans K et où $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ appartient à $M_{m,1}(K)$.

Notations.

1. Dans la situation précédente, on écrira souvent $A = (a_{i,j})$ et $B = (b_i)$.

2. Le K-espace vectoriel $M_{n,n}(K)$ des matrices carrées à n lignes et n colonnes à coefficients dans K est aussi noté $M_n(K)$.

Dans un premier temps, nous étudierons le cas des matrices carrées inversibles. Nous verrons plusieurs méthodes pour cela. Par exemple, la méthode du pivot de Gauss et certaines méthodes itératives. Les méthodes itératives utilisent des suites de matrices. Pour étudier la convergence de ces suites, nous utiliserons des normes, c'est pourquoi une partie du cours portera sur cette notion.

1.2. Valeurs propres, vecteurs propres

Rappelons de quoi il s'agit.

Soit $A \in M_n(\mathbb{C})$ (l'espace vectoriel des matrices carrées à n lignes et n colonnes à coefficients dans $\mathbb{C}$).

Définition 1.2.1. *Un nombre complexe λ est appelée valeur propre de A s'il existe $X \in M_{n,1}(\mathbb{C})$ non nul tel que $AX = \lambda X$. Le vecteur X est alors appelé vecteur propre de A associé à la valeur propre λ.*

Notons $\chi_A(X) = \det(XI - A) \in \mathbb{C}[X]$ le polynôme caractéristique de A.

Théorème 1.2.2. *L'ensemble des valeurs propres de A est égal à l'ensemble des racines de χ_A. Cet ensemble est appelé le spectre de A, et il est noté $Spec(A)$.*

Les valeurs propres et vecteurs propres sont les notions qui interviennent lorsque l'on cherche à diagonaliser une matrice.

Si $A \in M_n(K)$, on dit que A est diagonalisable sur K s'il existe $P \in \mathrm{GL}_n(K)$ telle que $P^{-1}AP$ soit diagonale. Si l'on note

$$P^{-1}AP = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$$

(où les λ_i ne sont pas nécessairement deux à deux distinctes), ces λ_i sont les valeurs propres de A. Enfin, si $P_1, \dots, P_n$ sont les colonnes de P, ces vecteurs forment une base de $M_{n,1}(K)$ (c'est ce qui assure que P est inversible) et chaque P_i est un vecteur propre associé à la valeur propre λ_i.

1.3. Notations

Sauf avis contraire $e_i \in M_n(K)$ désignera le vecteur

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in M_{n,1}(K)$$

où le i-ème coefficient est égal à 1 et les autres coefficients sont égaux à 0. Par exemple, dans $M_{4,1}(K)$,

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Cette notation semble ambiguë puisque la taille n de e_i n'intervient pas dans la notation. Il doit être clair que si l'on travaille dans $M_{n,1}(K)$, e_i est un vecteur de taille n.

La famille des éléments $(e_i)_{i\in[[1,n]]}$ de $M_{n,1}(K)$ est une base du K-espace vectoriel $M_{n,1}(K)$ appelée base canonique. Cet espace vectoriel est donc de dimension n.

$E_{i,j}$ est la matrice de $M_n(K)$ dont tous les coefficients sont nuls sauf le coefficient (i, j), qui est égal à 1. Par exemple, dans $M_4(K)$,

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

Là encore, cette notation semble ambiguë puisque n n'intervient pas dans la notation $E_{i,j}$. Il doit là aussi être clair que si l'on travaille dans $M_n(K)$, $E_{i,j}$ est une matrice de taille $n \times n$.

La famille des éléments $(E_{i,j})_{(i,j)\in[[1,n]]^2}$ de $M_n(K)$ est une base du K-espace vectoriel $M_n(K)$ appelée base canonique. Cet espace vectoriel est donc de dimension n^2.

I_n est la matrice identité de $M_n(K)$. Par exemple,

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Si M est une matrice, le terme $M_{i,j}$ désigne le coefficient (i,j) de M. Si V est un vecteur, V_i est son i-ème coefficient. Dans $M_2(K)$, M peut donc s'écrire

$$M = \sum_{i=1}^{2} \sum_{j=1}^{2} M_{i,j} E_{i,j} = \begin{pmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{pmatrix}$$

Dans $M_{2,1}(K)$, V peut s'écrire

$$V = \sum_{i=1}^{2} V_i e_i = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

1.4. Multiplication des matrices

Je termine cette introduction par un rappel basique, mais important et qui ne semble pas inutile. Nous nous en servirons à maintes reprises.

Soient $A \in M_{m,n}(K)$ et $B \in M_{n,l}(K)$. On définit le produit $C = AB$. C'est la matrice de $M_{m,l}(K)$ dont le coefficient (i,j) est

$$C_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j}$$

où $(i,j) \in [[1,m]] \times [[1,l]]$.

Chapitre 2

Théorème de Gershgorin-Hadamard

Les résultats de ce chapitre servent à localiser les valeurs propres d'une matrice A. C'est-à-dire qu'ils permettent de déterminer un sous-ensemble borné de $\mathbb{C}$ qui contient $\mathrm{Spec}(A)$. Un corollaire donnera également un critère simple permettant de repérer certaines matrices inversibles.

On note $A = (a_{i,j}) \in M_n(\mathbb{C})$.

2.1. Le théorème

Théorème 2.1.1. *[Gershgorin-Hadamard] Pour tout $k \in [[1, n]]$, on pose*

$$R_k = \sum_{j \neq k} |a_{k,j}| \quad \textit{et} \quad D_k = \{z \in \mathbb{C} \ : \ |z - a_{k,k}| \leq R_k\}.$$

Alors

$$\mathit{Spec}(A) \subset \bigcup_{k=1}^{n} D_k.$$

Démonstration. Soit λ une valeur propre de A et soit X un vecteur propre associé. Donc

$$AX = \lambda X \quad \text{et} \quad X \neq 0.$$

On note $X = (x_i)$. Alors pour tout $i \in [[1, n]]$,

$$\sum_{j=1}^{n} a_{i,j} x_j = \lambda x_i$$

ce que l'on peut écrire

$$\sum_{j \neq i} a_{i,j} x_j = (\lambda - a_{i,i}) x_i. \tag{2.1}$$

Soit k tel que $|x_k| = \max\{|x_i| \ : \ i \in [[1,n]]\} = ||X||_\infty$. Alors $x_k \neq 0$ puisque $X \neq 0$. Comme l'égalité (2.1) est vraie pour tout i, elle est en particulier vraie si l'on remplace i par k :

$$\sum_{j \neq k} a_{k,j} x_j = (\lambda - a_{k,k}) x_k.$$

En divisant par le complexe non nul x_k, on obtient

$$\lambda - a_{k,k} = \sum_{j \neq k} a_{k,j} \frac{x_j}{x_k}.$$

En utilisant l'inégalité triangulaire,

$$|\lambda - a_{k,k}| \leq \sum_{j \neq k} |a_{k,j}| \frac{|x_j|}{|x_k|} \leq \sum_{j \neq k} |a_{k,j}| = R_k.$$

On en déduit que $\lambda \in D_k$. Ainsi, pour tout $\lambda \in \mathrm{Spec}(A)$, il existe $k \in [[1,n]]$ tel que $\lambda \in D_k$. □

Exemples. 1. Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Si l'on note $\mathcal{B}_F(c,r)$ la boule fermée de $\mathbb{C}$ de centre c et de rayon r,

$$\mathrm{Spec}(A) \subset \mathcal{B}_F(1,2) \cup \mathcal{B}_F(4,3).$$

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

$$\mathrm{Spec}(A) \subset \mathcal{B}_F(1,3) \cup \mathcal{B}_F(3,2).$$

Remarque. Comme A et ${}^tA$ ont même polynôme caractéristique, on a aussi

$$\mathrm{Spec}(A) \subset \bigcup_{k=1}^{n} D'_k$$

où

$$R'_k = \sum_{j \neq k} |a_{j,k}| \quad \text{et} \quad D'_k = \{z \in \mathbb{C} \ : \ |z - a_{k,k}| \leq R'_k\}.$$

2.2. Conséquences

Définition 2.2.1. *On appelle rayon spectral de A le réel positif*

$$\rho(A) = \max\{|\lambda| \ : \ \lambda \in Spec(A)\}.$$

Proposition 2.2.2.

$$\rho(A) \leq \max_{i \in [[1,n]]} \sum_{j=1}^{n} |a_{i,j}|.$$

Démonstration. Soit $\lambda \in \mathrm{Spec}(A)$. Il existe $k \in [[1,n]]$ tel que $\lambda \in D_k$, c'est-à-dire tel que

$$|\lambda - a_{k,k}| \leq \sum_{j \neq k} |a_{k,j}|.$$

Donc, en appliquant l'inégalité triangulaire à la somme $\lambda = (\lambda - a_{k,k}) + a_{k,k}$, on obtient

$$\begin{aligned} |\lambda| &\leq |\lambda - a_{k,k}| + |a_{k,k}| \\ &\leq \sum_{j=1}^{n} |a_{k,j}| \end{aligned}$$

□

Remarque. Comme $\mathrm{Spec}(A) = \mathrm{Spec}({}^tA)$, on a aussi

$$\rho(A) \leq \max_{j \in [[1,n]]} \sum_{i=1}^{n} |a_{i,j}|.$$

Proposition 2.2.3. *Si $0 \notin \bigcup_{k=1}^{n} D_k$, alors $A \in GL_n(\mathbb{C})$ (c'est-à-dire : A est inversible). De même, si $0 \notin \bigcup_{k=1}^{n} D'_k$, alors $A \in GL_n(\mathbb{C})$.*

Démonstration. Si $0 \notin \bigcup_{k=1}^{n} D_k$, alors 0 n'est pas valeur propre de A, donc A est inversible. □

Exercice. Montrer par un contre-exemple que la réciproque est fausse.

Nous allons pouvoir maintenant repérer facilement certaines matrices inversibles : les matrices à *diagonale strictement dominante* définie ci-dessous.

Définition 2.2.4.

1. *On dit que A est à diagonale dominante si pour tout $i \in [[1, n]]$,*
$$\sum_{j \neq i} |a_{i,j}| \leq |a_{i,i}|.$$
2. *On dit que A est à diagonale strictement dominante si pour tout $i \in [[1, n]]$,*
$$\sum_{j \neq i} |a_{i,j}| < |a_{i,i}|.$$

Théorème 2.2.5. *Si A est à diagonale strictement dominante, alors A est inversible.*

Démonstration. Si A est à diagonale strictement dominante, $\sum_{j \neq i} |a_{i,j}| < |a_{i,i}|$ pour tout $i \in [[1, n]]$, donc 0 n'appartient à aucun des D_i quand i parcourt $[[1, n]]$. D'après la proposition 2.2.3, cela montre que A est inversible. $\square$

Exercice. Donner un exemple de matrice à diagonale dominante non inversible.

Remarque. Si l'on définit les notions de diagonale dominante et strictement dominante par rapport aux colonnes, on obtient des résultats similaires.

Chapitre 3

Algorithme du pivot de Gauss

Nous travaillons toujours sur des matrices à coefficients dans $K = \mathbb{R}$, $\mathbb{C}$ ou $\mathbb{Q}$, mais ce chapitre resterait valable sur tout corps commutatif K.

3.1. Permutations sur les lignes et colonnes

Soit $\delta_{i,j}$ le symbole de Kronecker défini pour tout $(i,j) \in \mathbb{Z}^2$ par $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \neq j$.

Définition 3.1.1. *Soit $\sigma \in S_n$ une permutation de $\{1, \dots, n\}$. On appelle matrice de permutation de σ la matrice $P_\sigma \in M_n(K)$ définie par*

$$(P_\sigma)_{i,j} = \delta_{i,\sigma(j)}.$$

Exemple. Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$. On note aussi $\sigma = (1,2,3,4)$. Alors

$$P_\sigma = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Remarque. Une matrice de permutation est une matrice dont chaque coefficient appartient à $\{0,1\}$ et dans laquelle toute ligne et toute colonne contient un et un seul coefficient égal à 1.

Proposition 3.1.2. *Soient m et n deux entiers strictement positifs. Soit $M \in M_{m,n}(K)$.*

1. *Soit $\sigma \in S_m$ (donc $P_\sigma \in M_m(K)$).*

$$P_\sigma M = M' \quad \text{où} \quad M'_{i,j} = M_{\sigma^{-1}(i),j}$$

2. *Soit $\sigma \in S_n$ (donc $P_\sigma \in M_n(K)$).*

$$MP_\sigma = M'' \quad \text{où} \quad M''_{i,j} = M_{i,\sigma(j)}.$$

Démonstration.

$$\begin{aligned} M'_{i,j} &= \sum_{k=1}^{m} (P_\sigma)_{i,k} M_{k,j} \\ &= \sum_{k=1}^{m} \delta_{i,\sigma(k)} M_{k,j} \\ &= M_{\sigma^{-1}(i),j} \end{aligned}$$

La seconde assertion est laissée en exercice. □

Proposition 3.1.3. *Soient σ et σ' deux éléments de S_n.*

1. $P_\sigma e_i = e_{\sigma(i)}$.
2. $P_{\sigma\circ\sigma'} = P_\sigma P_{\sigma'}$.
3. $P_{Id} = I_n$, *où Id désigne l'application identique sur $\{1,\dots,n\}$.*
4. *P_σ est inversible d'inverse $P_{\sigma^{-1}} = {}^tP_\sigma$.*
5. $\det P_\sigma = \varepsilon(\sigma)$ *(où $\varepsilon(\sigma)$ désigne la signature de σ).*

Démonstration. Exercice. □

3.2. Opérations sur les lignes et colonnes

Proposition 3.2.1. *Soient i, j, i', j' des éléments de $[[1, n]]$.*

$$E_{i,j}E_{i',j'} = \begin{cases} E_{i,j'} & \textit{si } i' = j \\ 0 & \textit{sinon} \end{cases}$$

Démonstration.

$$\begin{aligned} (E_{i,j}E_{i',j'})_{k,l} &= \sum_{m=1}^{n} (E_{i,j})_{k,m}(E_{i',j'})_{m,l} \\ &= 0 \quad \text{si } k \neq i \text{ ou } l \neq j' \end{aligned}$$

$$\begin{aligned} (E_{i,j}E_{i',j'})_{i,j'} &= \sum_{m=1}^{n} (E_{i,j})_{i,m}(E_{i',j'})_{m,j'} \\ &= (E_{i',j'})_{j,j'} \\ &= \begin{cases} 1 & \text{si } i' = j \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

□

Proposition 3.2.2. *Soit $M \in M_n(K)$ et soient i, j des éléments de $[[1, n]]$.*

1. $E_{i,j}M = M'$ *où* $M'_{k,l} = \begin{cases} M_{j,l} & \text{si } k = i \\ 0 & \text{sinon} \end{cases}$

   Autrement dit, la multiplication à gauche par $E_{i,j}$ place la ligne j de M sur la ligne i de la nouvelle matrice et met des 0 ailleurs.

2. $ME_{i,j} = M''$ *où* $M''_{k,l} = \begin{cases} M_{k,i} & \text{si } l = j \\ 0 & \text{sinon} \end{cases}$

   Autrement dit, la multiplication à droite par $E_{i,j}$ place la colonne i de M sur la colonne j de la nouvelle matrice et met des 0 ailleurs.

Démonstration. Comme $M = \sum_{k,l} M_{k,l}E_{k,l}$, et comme les $M_{k,l}$ sont des éléments de K, $E_{i,j}M = \sum_{k,l} M_{k,l}E_{i,j}E_{k,l}$. En appliquant la règle de multiplication entre les $E_{i,j}$ de la proposition 3.2.1, on trouve que $E_{i,j}M = \sum_{l} M_{j,l}E_{i,l}$. C'est bien le résultat annoncé.

La seconde assertion se montre de la même façon. □

Remarque. Si $M \in M_{m,n}(K)$ est rectangulaire non carrée, on obtient des résultats similaires, mais dans la multiplication $E_{i,j}M$, il faut considérer $E_{i,j} \in M_m(K)$, alors que dans $ME_{i,j}$, il faut prendre $E_{i,j} \in M_n(K)$.

Corollaire 3.2.3. *Soient $M \in M_{m,n}(K)$, $\lambda \in K$. Pour tout entier k de $[[1, m]]$, on note l_k la k-ème ligne de M et pour tout $k \in [[1, m]]$, on note c_k la k-ème colonne de M.*

1. *Soient i, j des éléments distincts de $[[1, m]]$. On considère $E_{i,j} \in M_m(K)$. La transformation $M \longleftarrow (I_m + \lambda E_{i,j})M$ remplace la ligne l_i de M par $l_i + \lambda l_j$.*
2. *Soient i, j des éléments distincts de $[[1, n]]$. On considère $E_{i,j} \in M_n(K)$. La transformation $M \longleftarrow M(I_n + \lambda E_{i,j})$ remplace la colonne c_j de M par $c_j + \lambda c_i$.*

Remarque. Avec les mêmes notations, $\det(I_n + \lambda E_{i,j}) = 1$. Ainsi, quand M est une matrice carrée, une transformation décrite par ce corollaire laisse inchangé le déterminant de M.

Définition 3.2.4. *Soient $\lambda \in K$ et $i \in [[1, n]]$. On note*

$$D_{n,i}(\lambda) = Diag(1, \ldots, 1, \lambda, 1, \ldots, 1) \in M_n(K).$$

la matrice diagonale dont les coefficients diagonaux sont tous égaux à 1 sauf le coefficient (i, i) qui est égal à λ. Autrement dit, $D_{n,i}(\lambda) = I_n + (\lambda - 1)E_{i,i}$.

Proposition 3.2.5. *Soit $M \in M_{m,n}(K)$.*

1. *La transformation $M \longleftarrow D_{m,i}(\lambda)M$ remplace la ligne l_i de M par λl_i.*
2. *La transformation $M \longleftarrow MD_{n,i}(\lambda)$ remplace la colonne c_j de M par λc_j.*

Si $m = n$, dans chacune de ces transformations, le déterminant de M est multiplié par λ.

Démonstration. Exercice. □

3.3. Résolution d'équations et factorisation LU

Soient $A \in \mathrm{GL}_n(K)$ et $B \in M_{n,1}(K)$. On cherche à résoudre l'équation $AX = B$.

Pour résoudre une telle équation, on pourrait calculer

$$A^{-1} = \frac{1}{\det A}{}^t\mathrm{Com}(A),$$

ou bien utiliser les formules de Cramer

$$x_i = \frac{\det A_i}{\det A}$$

où A_i est la matrice A dans laquelle la i-ème colonne a été remplacée par B. Mais ces méthodes sont coûteuses : le calcul de $\det A$ par l'égalité

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) A_{\sigma(1),1} \dots A_{\sigma(n),n}.$$

demande une somme de $n!$ termes, chacun de ces termes étant obtenu comme un produit de n éléments de K.

La méthode du pivot de Gauss consiste à se ramener à un système triangulaire par des opérations élémentaires. Observons cette méthode à travers un exemple. Soient

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 3 \\ 6 & 5 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$$

Résoudre l'équation $AX = B$ équivaut à résoudre le système

$$(S) \quad \begin{cases} 2x + y + z = 1 \\ 4x + y + 3z = 4 \\ 6x + 5y + 2z = 2 \end{cases}$$

Appelons respectivement l_1, l_2, l_3 les lignes $1, 2$ et 3 de ce système.

On considère le pivot 2 de la première ligne et on fait les transformations

$$\begin{aligned} l_2 &\longleftarrow l_2 - 2l_1 \\ l_3 &\longleftarrow l_3 - 3l_1 \end{aligned}$$

on obtient le système équivalent

$$(S') \begin{cases} 2x + y + z = 1 \\ \quad -y + z = 2 \\ \quad 2y - z = -1 \end{cases}$$

On fait ensuite la transformation

$$l_3 \longleftarrow l_3 + 2l_2$$

pour obtenir

$$(S'') \begin{cases} 2x + y + z = 1 \\ \quad -y + z = 2 \\ \qquad z = 3 \end{cases}$$

Finalement

$$z = 3 \quad , \quad y = 3 - 2 = 1 \quad , \quad x = \frac{1}{2}(1 - 1 - 3) = \frac{-3}{2}$$

La solution de l'équation est donc $X = \begin{pmatrix} -3/2 \\ 1 \\ 3 \end{pmatrix}$.

Donnons une interprétation matricielle des opérations précédentes. Les transformations

$$\begin{aligned} l_2 &\longleftarrow l_2 - 2l_1 \\ l_3 &\longleftarrow l_3 - 3l_1 \end{aligned}$$

correspondent à la multiplication à gauche par $I_3 - 2E_{2,1}$, puis par $I_3 - 3E_{3,1}$. On a donc multiplié à gauche par

$$\begin{aligned} L_1' &= (I_3 - 3E_{3,1})(I_3 - 2E_{2,1}) \\ &= I_3 - 3E_{3,1} - 2E_{2,1} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \end{aligned}$$

Ce calcul est un cas très particulier du lemme suivant.

Lemme 3.3.1. *Soient i et j deux entiers de $[[1,n]]$ tels que $i \leq j$. Soient*

$$M_i = I_n + \sum_{k=i+1}^{n} m_{k,i} E_{k,i} \quad et \quad M'_j = I_n + \sum_{k=j+1}^{n} m'_{k,j} E_{k,j}$$

où les $m_{k,i}$ et les $m'_{k,j}$ sont des éléments de K. Alors

$$M_i M'_j = I_n + \sum_{k=i+1}^{n} m_{k,i} E_{k,i} + \sum_{k=j+1}^{n} m'_{k,j} E_{k,j}.$$

Démonstration. Exercice (utiliser la règle de multiplication des $E_{i,j}$). □

Nous utiliserons aussi le corollaire suivant.

Corollaire 3.3.2. *Soit $i \in [[1,n]]$. Si*

$$M_i = I_n + \sum_{k=i+1}^{n} m_{k,i} E_{k,i}$$

alors

$$M_i^{-1} = I_n - \sum_{k=i+1}^{n} m_{k,i} E_{k,i}.$$

Démonstration. Effectuer la multiplication des deux matrices en utilisant le lemme 3.3.1. □

Si l'on applique ce corollaire à L'_1, on voit que son inverse est

$$L_1'^{-1} = L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

Après cette transformation, l'équation est devenue

$$L'_1 AX = L'_1 B.$$

La transformation

$$l_3 \longleftarrow l_3 + 2l_2$$

multiplie à gauche par $L'_2 = I_3 + 2E_{3,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$ d'inverse $L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$.

L'équation est donc devenue

$$L'_2 L'_1 AX = L'_2 L'_1 B.$$

Soit
$$U = L_2' L_1' A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$
Alors
$$A = L_1 L_2 U = LU$$
où d'après la proposition 3.3.1,
$$\begin{aligned} L &= L_1 L_2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}. \end{aligned}$$
Les matrices L et U étant triangulaires, cette factorisation $A = LU$ pourra être utile pour différents calculs sur la matrice A.

Pour résoudre l'équation et calculer cette factorisation $A = LU$, on peut aussi disposer les calculs de la manière suivante.

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

$$\begin{pmatrix} 2 & 1 & 1 \\ (\mathbf{2}) & -1 & 1 \\ (\mathbf{3}) & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \qquad \begin{matrix} l_2 \longleftarrow l_2 - 2l_1 \\ l_3 \longleftarrow l_3 - 3l_1 \end{matrix}$$

Les coefficients entre parenthèses et en rouge remplacent des 0 dans la nouvelle matrice. Ce sont les opposés des multiplicateurs, c'est-à-dire les coefficients de la matrice L. Puis on continue

$$\begin{pmatrix} 2 & 1 & 1 \\ (\mathbf{2}) & -1 & 1 \\ (\mathbf{3}) & (\mathbf{-2}) & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad l_3 \longleftarrow l_3 + 2l_2$$

On retrouve la matrices U en haut à droite et la matrice L formée d'une diagonale de 1 et de la partie inférieure gauche (en rouge et entre parenthèses).

$$U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}$$

Par la suite, on utilisera les notations suivantes.

Notations.

1. $\mathcal{T}_{\mathrm{inf},n}(K)$ désigne l'ensemble des matrices triangulaires inférieures de $M_n(K)$.
2. $\mathcal{T}^1_{\mathrm{inf},n}(K)$ désigne l'ensemble des matrices de $\mathcal{T}_{\mathrm{inf},n}(K)$ dont les coefficients diagonaux sont égaux à 1.
3. $\mathcal{T}_{\mathrm{sup},n}(K)$ désigne l'ensemble des matrices triangulaires supérieures de $M_n(K)$.
4. $\mathcal{T}^1_{\mathrm{sup},n}(K)$ désigne l'ensemble des matrices de $\mathcal{T}_{\mathrm{sup},n}(K)$ dont les coefficients diagonaux sont égaux à 1.

L'algorithme du pivot de Gauss décrit sur cet exemple s'applique à condition qu'à chaque étape i, le pivot rencontré $A_{i,i}$ est non nul. On a alors le résultat suivant.

Théorème 3.3.3. *Soit $A \in GL_n(K)$. Si dans l'algorithme du pivot de Gauss, on ne rencontre pas de pivot nul, alors la matrice A peut s'écrire de façon unique $A = LU$ où $L \in \mathcal{T}^1_{inf,n}(K)$ et $U \in \mathcal{T}_{sup,n}(K)$.*

Démonstration. Exercice. □

Remarque. Comme $A \in \mathrm{GL}_n(K)$, les éléments diagonaux de U sont non nuls.

3.4. Factorisation PLU

Si l'on rencontre un pivot nul lors de l'exécution de l'algorithme du pivot de Gauss, on n'obtiendra pas une factorisation LU comme dans le paragraphe précédent, mais presque.

Voyons cela sur la matrice $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 4 & 1 \end{pmatrix} \in M_3(\mathbb{Q})$. Comme le coefficient $A_{1,1}$ est nul, la méthode ne s'applique pas telle quelle. On échange alors les lignes 1 et 2, ce qui correspond à la multiplication à gauche par la matrice de permutation $P_{(1,2)}$. Écrivons les transformations.

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 4 & 1 \end{pmatrix} \qquad l_1 \longleftrightarrow l_2$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 4 & 0 \end{pmatrix} \qquad l_3 \longleftarrow l_3 - l_1$$

$$\begin{pmatrix}1 & 0 & 1\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix} \qquad\qquad l_2 \longleftrightarrow l_3$$

On obtient

$$P_{(2,3)}L_1'P_{(1,2)}A = U = \begin{pmatrix}1 & 0 & 1\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix}$$

où

$$L_1' = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\-1 & 0 & 1\end{pmatrix}$$

Or, un calcul simple montre que $P_{(2,3)}L_1'P_{(2,3)} = L_1'' = \begin{pmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & 0 & 1\end{pmatrix}$. Ainsi, $P_{(2,3)}L_1' = L_1''P_{(2,3)}$ et $L_1''P_{(2,3)}P_{(1,2)}A = U$, donc

$$L_1''P_{(1,3,2)}A = U,$$

où dans L_1'', les coefficients $(2,1)$ et $(3,1)$ ont été inversés. Finalement, on obtient la décomposition

$$A = PLU,$$

où

$$P = P_{(1,2,3)} \quad , \quad L = L_1''^{-1} = \begin{pmatrix}1 & 0 & 0\\1 & 1 & 0\\0 & 0 & 1\end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix}1 & 0 & 1\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix}$$

On peut aussi disposer les calculs comme suit.

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

$$\begin{pmatrix}1 & 0 & 1\\0 & 0 & 1\\1 & 4 & 1\end{pmatrix} \qquad\qquad (1,2)$$

$$\begin{pmatrix}1 & 0 & 1\\(\mathbf{0}) & 0 & 1\\(\mathbf{1}) & 4 & 0\end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ (\mathbf{1}) & 4 & 0 \\ (\mathbf{0}) & 0 & 1 \end{pmatrix} \qquad\qquad (2,3)$$

Dans cette dernière opération, on a échangé l'intégralité des lignes 2 et 3, multiplicateurs compris.

$$\begin{pmatrix} 1 & 0 & 1 \\ (\mathbf{1}) & 4 & 0 \\ (\mathbf{0}) & (\mathbf{0}) & 1 \end{pmatrix}$$

Le point nouveau par rapport au paragraphe précédent réside dans la "commutation" de $P_{(2,3)}$ et L'_1, c'est-à-dire l'égalité $P_{(2,3)}L'_1P_{(2,3)} = L''_1$. C'est un cas particulier du lemme suivant.

Lemme 3.4.1. *Soient i, a, b trois éléments de $[[1, n]]$ tels que $b > a > i$. Si*

$$M_i = I_n + \sum_{k=i+1}^{n} m_{k,i}E_{k,i},$$

alors

$$P_{(a,b)}M_iP_{(a,b)} = I_n + \sum_{k\in\mathcal{I}} m_{k,i}E_{k,i} + m_{a,i}E_{b,i} + m_{b,i}E_{a,i}$$

où $\mathcal{I} = [[i + 1, n]] \smallsetminus \{a, b\}$. Autrement dit, cette opération échange les coefficients (a, i) et (b, i).

Démonstration.

$$P_{(a,b)}M_iP_{(a,b)} = I_n + \sum_{k=1}^{n} m_{k,i}P_{(a,b)}E_{k,i}P_{(a,b)}$$

Comme $i < a < b$, $E_{k,i}P_{(a,b)} = E_{k,i}$. De plus,

$$P_{(a,b)}E_{k,i} = \begin{cases} E_{k,i} & \text{si} \quad k \notin \{a, b\} \\ E_{b,i} & \text{si} \quad k = a \\ E_{a,i} & \text{si} \quad k = b \end{cases}$$

□

Grâce à ce lemme, et en s'inspirant de l'exemple ci-dessus, on peut démontrer le théorème principal de ce chapitre.

Théorème 3.4.2. *Soit $A \in GL_n(K)$. Il existe une matrice de permutation P, $L \in \mathcal{T}^1_{inf,n}(K)$ et $U \in \mathcal{T}_{sup,n}(K)$ telles que*

$$A = PLU.$$

De plus, si P est fixée, cette factorisation est unique.

Démonstration. Exercice. □

3.5. Applications

Si l'on connaît une factorisation $A = PLU$, alors il est facile de calculer le déterminant de A :

$$\det A = \det P \det U.$$

Résoudre une équation $AX = B$ revient à résoudre $PLUX = B$, c'est-à-dire $LUX = {}^tPB$. Pour cela, on peut commencer par résoudre $LY = {}^tPB$, qui est un système triangulaire. Une fois Y trouvé, il reste à résoudre $UX = Y$, qui est un autre système triangulaire.

Il y a plusieurs façons de calculer A^{-1}. L'une des méthodes consiste à résoudre $AX = e_i$ pour chacun des $i \in [[1, n]]$. Si l'on note S_i la solution de $AX = e_i$, alors A^{-1} est la matrice dont les colonnes sont les S_i :

$$A^{-1} = (S_1|\cdots|S_n).$$

3.6. Complexité

Nous allons compter le nombre d'opérations sur les éléments de K nécessaires dans l'algorithme du pivot de Gauss. Un échange de lignes revient à un changement d'indices. Nous ne compterons pas ces échanges. On peut donc supposer que les pivots rencontrés sont tous non nuls. On note $A = (a_{i,j})$. On note aussi $A_1 = A$.

Étape 1. La détermination de L'_1

$$\begin{aligned}
L'_1 &= I_n - \sum_{i=2}^{n} \frac{a_{i,1}}{a_{1,1}} E_{i,1} \\
&= \begin{pmatrix}
1 & 0 & 0 & \dots & 0 \\
-a_{2,1}/a_{1,1} & 1 & 0 & \dots & 0 \\
\vdots & & \ddots & \ddots & \vdots \\
\vdots & 0 & 0 & \ddots & 0 \\
-a_{n,1}/a_{1,1} & 0 & \dots & 0 & 1
\end{pmatrix}
\end{aligned}$$

demande $n-1$ divisions. Dans le calcul de

$$L'_1 A_1 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -a_{2,1}/a_{1,1} & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & 0 & 0 & \ddots & 0 \\ -a_{n,1}/a_{1,1} & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$
$$= \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a'_{2,2} & \dots & a'_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{n,2} & \dots & a'_{n,n} \end{pmatrix},$$

on sait que les coefficients sous le $a_{1,1}$ de la première colonne sont nuls. Il ne faut donc pas les calculer. D'abord dans un souci d'économie, mais aussi, dans le cas où les calculs sont approchés (dans $\mathbb{R}$ ou $\mathbb{C}$), un calcul mènerait à une approximation de 0, donc à une valeur non nulle.

Ce calcul demande $(n-1)^2$ multiplications et $(n-1)^2$ additions.

Étape k. Après la $(k-1)$-ème étape, on arrive à une matrice A_k de la forme

$$A_k = \begin{pmatrix} T_k & B_k \\ 0 & A'_k \end{pmatrix}$$

où $T_k \in \mathcal{T}\mathrm{sup}_{,k-1}(K)$, $B_k \in M_{k-1,n-k+1}(K)$ et $A'_k \in M_{n-k+1}(K)$. Le calcul de A_{k+1} se ramène à un calcul sur A'_k. Cela demande $n-k$ divisions pour le calcul de L'_k, puis de $(n-k)^2$ multiplications et $(n-k)^2$ additions pour le calcul de $A_{k+1} = L'_k A_k$.

En tout, le nombre d'opérations est égal à

$$\begin{aligned} C &= 2\sum_{k=1}^{n-1}(n-k)^2 + \sum_{k=1}^{n-1}(n-k) \\ &= 2\sum_{k=1}^{n-1} k^2 + \sum_{k=1}^{n-1} k \\ &= \frac{n(n-1)(2n-1)}{3} + \frac{n(n-1)}{2} \\ &= \frac{2}{3}n^3 + O(n^2) \end{aligned}$$

Théorème 3.6.1. *Le nombre d'opérations dans K nécessaires à l'exécution de l'algorithme du pivot de Gauss sur une matrice de $GL_n(K)$ est au plus équivalent à $\frac{2}{3}n^3$.*

Voyons maintenant le coût de la résolution d'un système triangulaire. Soit

$$(\mathcal{S})\ :\ \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ \qquad a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ \qquad\qquad a_{n,n}x_n = b_n \end{cases}$$

où les $a_{i,i}$ sont non nuls. Ce système est équivalent à

$$\begin{cases} x_n = a_{n,n}^{-1} b_n \\ x_{n-1} = a_{n-1,n-1}^{-1}\left(b_{n-1} - a_{n-1,n}x_n\right) \\ \quad \vdots \\ x_1 = a_{1,1}^{-1}\left(b_1 - \sum_{j=2}^{n} a_{1,j}x_j\right) \end{cases}$$

On calcule dans l'ordre $x_n, x_{n-1}, \ldots, x_1$. On peut donc se servir des valeurs de $x_n, \ldots, x_{k+1}$ pour le calcul de

$$x_k = a_{k,k}^{-1}\left(b_k - \sum_{j=k+1}^{n} a_{k,j}x_j\right).$$

Ce calcul demande au plus une division, $n-k$ additions et $n-k$ multiplications. En tout, le nombre maximal d'opérations est

$$\begin{aligned} C &= 2\sum_{k=1}^{n-1} k + n \\ &= n^2 \end{aligned}$$

Théorème 3.6.2. *La résolution d'un système triangulaire de taille n demande au plus n^2 opérations dans K.*

3.7. Problèmes de précision

Si l'on travaille dans $M_n(\mathbb{R})$ où $M_n(\mathbb{C})$, il est important de s'intéresser à la précision des calculs. Or si l'on a au départ une matrice dans $M_n(\mathbb{Z})$, un calcul exact nous mênera à des matrices de $M_n(\mathbb{Q})$ dont les coefficients auront un numérateur et dénominateur de plus en plus gros. Il est donc intéressant d'arrondir et de faire les calculs dans $M_n(\mathbb{R})$.

L'exemple simple suivant illustre le genre de problème que l'on pourra rencontrer.

Exemple. Soit l'équation

$$(\mathcal{E}) \;:\; \begin{pmatrix} 10^{-20} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

On cherche à résoudre $(\mathcal{E})$, en travaillant avec 10 chiffres de précision.

On prend 10^{-20} comme pivot. Cela donne

$$\begin{pmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2-10^{20} \end{pmatrix}$$

que l'on arrondit

$$\begin{pmatrix} 10^{-20} & 1 \\ 0 & -10^{20} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -10^{20} \end{pmatrix}.$$

On trouve donc $y = 1$ et $x = 0$, ce qui est complètement faux.

Commençons plutôt par échanger les lignes. $(\mathcal{E})$ s'écrit

$$\begin{pmatrix} 1 & 1 \\ 10^{-20} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Le pivot est maintenant 1.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1-10^{-20} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1-2.10^{-20} \end{pmatrix}$$

que l'on arrondit

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

donc $y = 1$, $x = 1$, ce qui est plus raisonnable.

Dans le premier cas, le pivot est petit : c'est 10^{-20}. Son inverse est donc très grand : 10^{20}. Quand on fait la différence de la ligne 2 moins 10^{20} fois la ligne 1, les nombres 1 (ligne 2 colonne 1) et 2 (membre de droite, ligne 2) deviennent négligeables.

De façon générale, on risque de perdre de l'information de cette manière lorsque le pivot est petit. C'est pourquoi il est judicieux d'utiliser un grand pivot. Pour cela, on peut utiliser l'une des deux méthodes suivantes.

Le pivot partiel. À chaque étape, on fait une permutation sur les lignes de telle sorte que le pivot soit le plus grand possible en module.

Le pivot total. À chaque étape, on fait une permutation sur les lignes et les colonnes de telle sorte que le pivot soit le plus grand possible en module. On obtient alors $AQ = PLU$ au lieu de $A = PLU$, où Q est une matrice de permutation. Pour résoudre l'équation $AX = B$, on peut poser $X = QY$, on résout $AQY = B$, donc $LUY = {}^tPB$, puis on trouve $X = QY$.

Chapitre 4

Espaces euclidiens

4.1. Formes bilinéaires symétriques

Ce paragraphe donne quelques définitions et propriétés de base d'algèbre bilinéaire sur $\mathbb{R}$.

Soit E un $\mathbb{R}$-espace vectoriel. On rappelle ci-dessous la définition des applications linéaires et formes linéaires.

Définition 4.1.1. *1. Soient F un autre $\mathbb{R}$-espace vectoriel et f une application de E dans F. On dit que f est une application linéaire si pour tout $(x, y) \in E \times E$ et tout $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}$,*

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

2. Une forme linéaire définie sur E est une application linéaire de E dans $\mathbb{R}$.

Exemple 1. L'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x_1, x_2) = 3x_1 - x_2$ est une forme linéaire sur $\mathbb{R}^2$.

Venons en maintenant aux formes bilinéaires. Ce sont des applications de $E \times E$ dans $\mathbb{R}$ qu'on définit plus bas, mais dont on donne tout de suite un exemple.

Exemple 2. L'application de $\mathbb{R}^2 \times \mathbb{R}^2$ dans $\mathbb{R}$ définie par $b((x_1, x_2), (y_1, y_2)) = 3x_1y_1 - x_1y_2 + x_2y_2$ est une forme bilinéaire sur $\mathbb{R}^2$.

Définition 4.1.2. *Soit b une application de $E \times E$ dans $\mathbb{R}$.*

1. *On dit que b est linéaire à droite si pour tout $x \in E$ fixé, l'application de E dans $\mathbb{R}$ qui à y associe $b(x, y)$ est linéaire.*
2. *On dit que b est linéaire à gauche si pour tout $y \in E$ fixé, l'application de E dans $\mathbb{R}$ qui à x associe $b(x, y)$ est linéaire.*
3. *On dit que b est une forme bilinéaire si elle est linéaire à gauche et à droite.*

Nous nous intéresserons plus particulièrement aux formes bilinéaires symétriques définies comme suit.

Définition 4.1.3. *Une forme bilinéaire b est dite symétrique si pour tout $(x, y) \in E \times E$,*

$$b(x, y) = b(y, x).$$

Remarque. La forme bilinéaire de l'exemple 2 n'est pas symétrique.

Exemple 3. L'application de $\mathbb{R}^2 \times \mathbb{R}^2$ dans $\mathbb{R}$ définie par $b((x_1, x_2), (y_1, y_2)) = 3x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$ est une forme bilinéaire symétrique sur $\mathbb{R}^2$.

Définition 4.1.4. *Une application q de E dans $\mathbb{R}$ est appelée une forme quadratique s'il existe une forme bilinéaire symétrique b sur E telle que $q(x) = b(x, x)$ pour tout $x \in E$. q est alors appelée forme quadratique associée à b.*

Proposition 4.1.5. *Une application q de E dans $\mathbb{R}$ est une forme quadratique si et seulement si elle vérifie les propriétés suivantes.*

1. *L'application b de $E \times E$ dans $\mathbb{R}$ définie par*

$$b(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

est bilinéaire.

2. *Pour tout $\lambda \in \mathbb{R}$ et tout $x \in E$,*

$$q(\lambda x) = \lambda^2 q(x).$$

La forme bilinéaire b est alors l'unique forme bilinéaire symétrique dont q est la forme quadratique associée. On dit aussi que b est la forme bilinéaire symétrique associée à q. On dit aussi que b est la forme polaire de q.

Démonstration. Supposons que q est une forme quadratique. Alors il existe une forme bilinéaire symétrique f telle que $q(x) = f(x, x)$ pour tout $x \in E$.

$$\begin{aligned} q(x+y) &= f(x+y, x+y) \\ &= f(x, x+y) + f(y, x+y) \quad \text{car } f \text{ est linéaire à gauche} \\ &= f(x,x) + f(x,y) + f(y,x) + f(y,y) \quad \text{car } f \text{ est linéaire à droite} \\ &= q(x) + 2f(x,y) + q(y) \quad \text{car } f \text{ est symétrique.} \end{aligned}$$

On obtient donc que

$$\begin{aligned} f(x,y) &= \frac{1}{2}\left(q(x+y) - q(x) - q(y)\right) \\ &= b(x,y) \end{aligned}$$

donc b est bilinéaire symétrique puisque f l'est. Enfin,

$$\begin{aligned} q(\lambda x) &= f(\lambda x, \lambda x) \\ &= \lambda f(x, \lambda x) \quad \text{car } f \text{ est linéaire à gauche} \\ &= \lambda^2 f(x,x) \quad \text{car } f \text{ est linéaire à droite.} \end{aligned}$$

Réciproquement, supposons les conditions 1 et 2 vérifiées. Il est clair que $b(x,y) = b(y,x)$ pour tout $(x,y) \in E$, donc b est une forme bilinéaire symétrique. Montrons que q est sa forme quadratique associée, c'est-à-dire $b(x,x) = q(x)$ pour tout $x \in E$.

$$\begin{aligned} b(x,x) &= \frac{1}{2}\left(q(2x) - 2q(x)\right) \quad \text{(par la propriété 1)} \\ &= \frac{1}{2}(4q(x) - 2q(x)) \quad \text{(par la propriété 2)} \\ &= q(x) \end{aligned}$$

□

Exemples.

1. $E = \mathbb{R}$, $q(x) = x^2$, $b(x,y) = xy$.
2. $E = \mathbb{R}^n$, $q(x) = \displaystyle\sum_{i=1}^n \lambda_i x_i^2$, $b(x,y) = \displaystyle\sum_{i=1}^n \lambda_i x_i y_i$.
3. $E = \mathbb{R}[x]$, $q(x) = \int_0^1 P(t)^2 dt$, $b(P,Q) = \int_0^1 PQ(t) dt$.

4.2. Produit scalaire

Soit E un $\mathbb{R}$-espace vectoriel.

Définition 4.2.1. *Soit q une forme quadratique définie sur E et b sa forme polaire.*

1. *On dit que q est positive si pour tout élément x de E, $q(x) \geq 0$.*
2. *On dit que q est définie positive si pour tout élément x non nul de E, $q(x) > 0$.*

Définition 4.2.2. *Un produit scalaire sur E est une forme bilinéaire symétrique dont la forme quadratique associée est définie positive.*

Définition 4.2.3. *Un espace euclidien est un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.*

Exemple. Soit $E = \mathbb{R}^n$. Pour $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$, on définit $b(x, y) = \sum_{i=1}^n x_i y_i$. Alors b est un produit scalaire. L'espace vectoriel E muni du produit scalaire b est donc un espace euclidien.

Théorème 4.2.4. *[Inégalité de Cauchy Schwarz] On suppose que E est muni d'un produit scalaire b. Soit q sa forme quadratique associée. Pour tout $(x, y) \in E^2$*

$$b(x, y)^2 \leq q(x)q(y). \tag{4.1}$$

De plus,

$$b(x, y)^2 = q(x)q(y) \iff x \text{ et } y \text{ sont colinéaires.} \tag{4.2}$$

Démonstration. Dans le cas où $y = 0$, cet énoncé est vérifié car tout élément x de E est colinéaire à 0 et vérifie $b(x, 0)^2 = 0 = q(x)q(0)$.

Supposons y non nul. Pour tout réel t, on pose $P(t) = q(x + ty)$. Alors $P(t) \geq 0$ pour tout $t \in \mathbb{R}$. De plus

$$P(t) = t^2 q(y) + 2tb(x, y) + q(x)$$

est un polynôme et t de degré 2 de signe constant. C'est donc que son discriminant est négatif ou nul. On en déduit que

$$b(x, y)^2 - q(x)q(y) \leq 0.$$

Cela montre bien l'inégalité de Cauchy-Schwarz (4.1).

Si x et y sont colinéaires, il existe $\lambda \in \mathbb{R}$ tel que $x = \lambda y$. Alors

$$b(x, y)^2 = \lambda^2 q(y)^2 = q(x)q(y).$$

Réciproquement, si $b(x,y)^2 = q(x)q(y)$, alors le discriminant du polynôme P est nul. Il existe donc $t_0 \in \mathbb{R}$ tel que $P(t_0) = 0$, c'est-à-dire tel que $q(x + t_0 y) = 0$. Comme q est définie positive, on en déduit que $x + t_0 y = 0$, donc que x et y sont colinéaires. $\square$

Remarque. Si la forme quadratique q est seulement supposée positive (et pas forcément définie positive), l'inégalité de Cauchy-Schwarz reste vraie. Par contre, l'équivalence (4.2) n'est plus valide.

Exercice. Trouver un contre-exemple qui justifie la remarque précédente.

À tout produit scalaire, on peut associer une norme. C'est l'objet du théorème 4.2.6 suivant. Mais rappelons tout d'abord ce qu'est une norme.

Définition 4.2.5. *Soient $K = \mathbb{R}$ ou $\mathbb{C}$ et E un K-espace vectoriel. Une application $x \mapsto ||x||$ de E dans $\mathbb{R}_+$ est une norme si les conditions suivantes sont vérifiées.*

1. $||x|| = 0$ *si et seulement si* $x = 0$.
2. *Pour tout* $\lambda \in K$ *et tout* $x \in E$,
$$||\lambda x|| = |\lambda| \cdot ||x||.$$
3. *Pour tout* $(x, y) \in E^2$,
$$||x + y|| \leq ||x|| + ||y||.$$
Cette inégalité est appelée "inégalité triangulaire".

Théorème 4.2.6. *Soit q une forme quadratique sur E définie positive. Pour tout élément x de E, on définit $||x|| = \sqrt{q(x)}$. Alors l'application de E dans $\mathbb{R}_+$ qui à x associe $||x||$ est une norme.*

Démonstration. Comme q est définie positive, $q(x) \geq 0$ pour tout $x \in E$, donc $||x|| = \sqrt{q(x)}$ est bien défini et appartient à $\mathbb{R}_+$. Vérifions les propriétés de la définition 4.2.5.

1. $||0|| = 0$ puisque $q(0) = 0$. Réciproquement, si $||x|| = 0$, alors $q(x) = 0$ donc $x = 0$ puisque q est définie positive.

2. $||\lambda x|| = \sqrt{q(\lambda x)} = \sqrt{\lambda^2 q(x)} = |\lambda| \sqrt{q(x)} = |\lambda| \cdot ||x||$.

3. Montrons que $||x+y|| \leq ||x|| + ||y||$. Pour cela, il suffit de montrer que
$$||x + y||^2 \leq ||x||^2 + ||y||^2 + 2||x|| \cdot ||y||. \tag{4.3}$$
En effet, si a et b sont des réels positifs, alors $a \leq b$ si et seulement si $a^2 \leq b^2$. Cette inégalité (4.3) s'écrit aussi
$$q(x + y) \leq q(x) + q(y) + \sqrt{q(x)q(y)}. \tag{4.4}$$

Or $q(x+y) = q(x) + q(y) + 2b(x,y)$. D'après l'inégalité de Cauchy-Schwarz, $|b(x,y)| \leq \sqrt{q(x)q(y)}$, donc $b(x,y) \leq \sqrt{q(x)q(y)}$. Cela prouve l'inégalité (4.4), qui elle même montre l'inégalité triangulaire. $\square$

4.3. Bases orthonormales

Soit E un espace euclidien. E est donc un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire que l'on notera $(*|*)$, c'est-à-dire que si $(x,y) \in E^2$, le produit scalaire de x par y est noté $(x|y)$. On note n la dimension de E.

Définition 4.3.1. *Soient $x_1, \dots, x_m$ des éléments de E. On dit que la famille $(x_i)_{i \in [[1,m]]}$ est orthogonale si $(x_i|x_j) = 0$ dès que $i \neq j$.*

Proposition 4.3.2. *Toute famille orthogonale d'éléments non nuls de E est libre. Par conséquent, le cardinal d'une telle famille est nécessairement inférieure ou égale à n.*

Démonstration. Laissée en exercice. $\square$

Définition 4.3.3.
1. *Une base de E est appelée une base orthogonale si c'est une famille orthogonale.*
2. *Une famille de E est dite orthonormale ou orthonormée si elle est orthogonale et si ses éléments sont de norme 1.*
3. *Une base de E est appelée une base orthonormale ou orthonormée si c'est une famille orthonormale.*

On sait qu'il existe des bases de E. Nous allons voir qu'il existe des bases orthogonales, et donc aussi des bases orthonormales. Nous donnons même un procédé pour construire une telle base : le procédé de Gram-Schmidt.

Définition 4.3.4. *Soit (b_i) une base de E. On dit que (b_i^*) est une base orthogonale de Gram-Schmidt associée à (b_i) si les conditions suivantes sont réalisées.*

1. *(b_i^*) est une base orthogonale pour $(\cdot|\cdot)$.*
2. *La matrice de passage de (b_i) à (b_i^*) est triangulaire supérieure à coefficients diagonaux strictement positifs.*

On dit aussi que (b_i^*) est une orthogonalisée de Gram-Schmidt de (b_i).

Notation. On note $\mathcal{T}_{sup,n}^{+}(\mathbb{R})$ (resp. $\mathcal{T}_{sup,n}^{++}(\mathbb{R})$) l'ensemble des éléments de $\mathcal{T}_{sup,n}(\mathbb{R})$ dont les coefficients diagonaux sont positifs (resp. strictement positifs).

Remarque. La seconde condition signifie que les b_i^* s'ecrivent de la manière suivante.

$$\begin{cases} b_1^* = a_{1,1}b_1 \\ b_2^* = a_{1,2}b_1 + a_{2,2}b_2 \\ \ \vdots \\ b_n^* = a_{1,n}b_1 + \cdots + a_{n,n}b_n \end{cases} \tag{4.5}$$

où les $a_{i,j}$ sont des réels et où $a_{i,i} > 0$ pour tout i.

Comme l'inverse d'un élément de $\mathcal{T}^{++}_{sup,n}(\mathbb{R})$ appartient à $\mathcal{T}^{++}_{sup,n}(\mathbb{R})$, cette condition signifie aussi que la matrice de passage de (b_i^*) à (b_i) appartient à $\mathcal{T}^{++}_{sup,n}(\mathbb{R})$, c'est-à-dire que les b_i s'écrivent comme suit.

$$\begin{cases} b_1 = y_{1,1}b_1^* \\ b_2 = y_{1,2}b_1^* + y_{2,2}b_2^* \\ \ \vdots \\ b_n = y_{1,n}b_1^* + \cdots + y_{n,n}b_n^* \end{cases} \tag{4.6}$$

où $y_{i,i} > 0$ pour tout i.

Cela implique en particulier que pour tout j, $\mathrm{Vect}(b_1, \ldots, b_j) = \mathrm{Vect}(b_1^*, \ldots, b_j^*)$.

Théorème 4.3.5 (Gram-Schmidt)**.** *Soit (b_i) une base de E.*

1. *Il existe une unique orthogonalisée de Gram-Schmidt (b_i^*) de (b_i) telle que $a_{i,i} = 1$ pour tout i (avec les notations de la remarque ci-dessus). Dans ce cours, on appellera cette base (b_i^*) l'orthogonalisée de Gram Schmidt de type 1 de b.*
2. *Il existe une unique orthogonalisée de Gram-Schmidt (b_i^*) de (b_i) qui est orthonormale. On appelle cette base l'orthonormalisée de Gram-Schmidt de (b_i).*

Démonstration. 1. On construit par récurrence l'orthogonalisée de Gram-Schmidt telle que $a_{i,i} = 1$ pour tout i. Nécessairement, $b_1^* = b_1$.

Soit $j \geq 2$. On suppose les éléments $b_1^*, \ldots, b_{j-1}^*$ construits et on cherche à construire b_j^*. Il s'écrit

$$b_j^* = b_j + \sum_{i=1}^{j-1} a_{i,j} b_i.$$

En vertu de la remarque précédente, il s'écrit aussi sous la forme

$$b_j^* = b_j - \sum_{i=1}^{j-1} y_{i,j} b_i^*.$$

Pour tout $i < j$, a condition $(b_j^*|b_i^*) = 0$ équivaut à

$$(b_j|b_i^*) = y_{i,j}(b_i^*|b_i^*)$$

c'est-à-dire

$$y_{i,j} = \frac{(b_j|b_i^*)}{(b_i^*|b_i^*)}.$$

Il existe donc un et un seul b_j^* convenable :

$$b_j^* = b_j - \sum_{i=1}^{j-1} \frac{(b_j|b_i^*)}{(b_i^*|b_i^*)} b_i^*.$$

2. Si l'on impose la condition $||b_i^*|| = 1$, il suffit de remplacer chaque b_i^* par $b_i^*/||b_i^*||$. Cela montre l'existence.

Réciproquement, si (b') est l'orthonormalisée de Gram-Schmidt de b. Soit $P = (p_{i,j})$ la matrice de passage de b à b'. On note $D = \text{diag}(p_{1,1}, \dots, p_{n,n})$. Si l'on pose $b_i^* = \dfrac{1}{p_{i,i}} b_i'$ pour tout i, $b^* = (b_i^*)$ est l'orthogonalisée de Gram-Schmidt de type 1 de b, puisque la matrice de passage de b à b^* est PD^{-1} qui appartient à $\mathcal{T}^1_{sup,n}(\mathbb{R})$. Cela montre l'unicité. □

Remarque. Si l'on pose $y_{i,i} = 1$ pour tout i et $y_{i,j} = 0$ pour $i > j$, la matrice $(y_{i,j})_{(i,j)\in[[1,n]]^2}$ est la matrice de passage de (b_i^*) à (b_i).

Algorithme 4.3.6. *1. Pour construire l'orthogonalisée de Gram-Schmidt b^* de type 1 de b, on pose $b_1^* = b_1$, puis pour tout $j \in [[2, n]]$*

$$b_j^* = b_j - \sum_{i=1}^{j-1} \frac{(b_j|b_i^*)}{(b_i^*|b_i^*)} b_i^*.$$

2. Une fois b^* construite, on peut construire l'orthonormalisée de Gram-Schmidt b' de b en posant pour $j \in [[1, n]]$

$$b_j' = \frac{1}{||b_j^*||} b_j^*$$

Corollaire 4.3.7. *Tout espace euclidien admet une base orthonormale.*

Proposition 4.3.8. *Soient E un espace euclidien et (ε_i) une base orthonormée de E. Pour tout $x \in E$,*

$$x = \sum_{i=1}^{n} (x|\varepsilon_i)\varepsilon_i.$$

Démonstration. On écrit x suivant la base (ε_i) : $x = \sum_{i=1}^{n} x_i \varepsilon_i$, où $x_i \in \mathbb{R}$ pour tout i. Soit $j \in [[1, n]]$.

$$\begin{aligned}(x|\varepsilon_j) &= \left(\sum_{i=1}^{m} x_i \varepsilon_i | \varepsilon_j \right) \\ &= \sum_{i=1}^{m} x_i (\varepsilon_i|\varepsilon_j) \\ &= x_j\end{aligned}$$

puisque $(\varepsilon_i|\varepsilon_j) = 0$ si $i \neq j$ et $(\varepsilon_j|\varepsilon_j) = 1$. □

Proposition 4.3.9. *Soit F un sous-espace vectoriel de E. Alors*

$$F^{\perp} = \{y \in E \ : \ (x|y) = 0 \ \forall x \in F\}$$

est un sous-espace vectoriel de E. De plus,

$$F \oplus F^{\perp} = E.$$

Soit p la projection sur F parallèlement à $F^{\perp}$ (appelée projection orthogonale sur F). Soit $(\varepsilon_1, \ldots, \varepsilon_m)$ une base orthonormale de F. Pour tout $x \in E$,

$$p(x) = \sum_{i=1}^{m} (x|\varepsilon_i)\varepsilon_i.$$

Démonstration. On laisse au lecteur le soin de montrer que $F^{\perp}$ est un sous-espace vectoriel de E. Soit $(\varepsilon_1, \ldots, \varepsilon_m)$ une base orthonormale de F.

Utilisons un raisonnement du type “analyse-synthèse”.

Analyse. Supposons que l'on puisse écrire $x = a + b$ où $a \in F$ et $b \in F^{\perp}$. Comme $b \in F^{\perp}$, et comme les ε_i appartiennent à F, $(b|\varepsilon_i) = 0$ pour tout $i \in [[1, m]]$. Donc si $x = a + b$, $(x|\varepsilon_i) = (a|\varepsilon_i)$ pour tout $i \in [[1, m]]$. Comme $a \in F$, en utilisant la proposition 4.3.8, $a = \sum_{i=1}^{m} (a|\varepsilon_i)\varepsilon_i = \sum_{i=1}^{m} (x|\varepsilon_i)\varepsilon_i$.

Synthèse. Montrons maintenant que si $x = a + b$ et si

$$a = \sum_{i=1}^{m} (x|\varepsilon_i)\varepsilon_i,$$

alors $a \in F$ et $b \in F^{\perp}$. Il est clair que $a \in F$ car les ε_i appertiennent à F. Pour montrer que $b \in F^{\perp}$, il suffit de montrer que pour tout $i \in [[1, m]]$, $(b|\varepsilon_i) = 0$ (voir le lemme 4.3.10 ci-dessous). Or

$$(b|\varepsilon_i) = (x|\varepsilon_i) - (a|\varepsilon_i).$$

Comme $a = \sum_{i=1}^{m}(x|\varepsilon_i)\varepsilon_i$ par hypothèse et $a = \sum_{i=1}^{m}(a|\varepsilon_i)\varepsilon_i$ par la proposition 4.3.8, $(x|\varepsilon_i) = (a|\varepsilon_i)$ pour tout i donc $(b|\varepsilon_i) = 0$.

Ce raisonnement permet de conclure que tout $x \in E$ s'écrit de façon unique $x = a + b$ ou $a \in F$ et $b \in F^\perp$. Cela montre bien que $E = F \oplus F^\perp$. On obtient par la même occasion que $p(x) = a = \sum_{i=1}^{m}(x|\varepsilon_i)\varepsilon_i$.

□

Lemme 4.3.10. *On reprend les notations de la proposition 4.3.9. Un élément y appartient à $F^\perp$ si et seulement si pour tout $i \in [[1, m]]$, $(y|\varepsilon_i) = 0$.*

Démonstration. Exercice. □

Exemple 1. On considère $E = \mathbb{R}^4$ muni du produit scalaire défini par $(x|y) = \sum_{i=1}^{4} x_i y_i$, où $x = (x_1, x_2, x_3, x_4)$ et $y = (y_1, y_2, y_3, y_4)$. Soient $b_1 = (1, 1, 1, 0)$ et $b_2 = (0, 1, 1, 1)$. On note $F = \text{Vect}(b_1, b_2)$. Alors $b = (b_1, b_2)$ est une base de F. On note p la projection orthogonale sur F. Soit $x = (1, 1, 1, 1)$.

Nous allons calculer l'orthogonalisée b^* de Gram-Schmidt de type 1 de b. Ensuite, nous en déduirons l'orthonormalisée de Gram-Schmidt de b. Enfin, nous calculerons $p(x)$.

Cherchons l'orthogonalisée de Gram-Schmidt de type 1 de b. Suivant l'algorithme 4.3.6, $b_1^* = b_1$, et

$$b_2^* = b_2 - \frac{(b_2|b_1^*)}{(b_1^*|b_1^*)} b_1^*$$

Comme $(b_2|b_1^*) = 2$ et $(b_1^*|b_1^*) = 3$, cela donne

$$b_2^* = b_2 - \frac{2}{3} b_1^*$$

donc $b_2^* = (0, 1, 1, 1) - \frac{2}{3}(1, 1, 1, 0) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}, 1\right)$. Pour vérifier le résultat, on calcule $(b_1^*|b_2^*)$ et on trouve bien 0.

On en déduit l'orthonormalisée $b' = (b_1', b_2')$ en divisant les b_i^* par leur norme.

$$b_1' = \frac{b_1^*}{||b_1^*||} = \frac{1}{\sqrt{3}}(1, 1, 1, 0)$$

$$b_2' = \frac{b_2^*}{||b_2^*||} = \frac{1}{\sqrt{15}}(-2, 1, 1, 3)$$

Ce sont bien des vecteurs de norme 1.

Pour calculer $p(x)$, on utilise la proposition 4.3.9, qui nous dit que

$$p(x) = (x|b'_1)b'_1 + (x|b'_2)b'_2$$

Or $(x|b'_1) = \frac{1}{\sqrt{3}}(1+1+1) = \sqrt{3}$ et $(x|b'_2) = \frac{1}{\sqrt{15}}(-2+1+1+3) = \frac{3}{\sqrt{15}}$ (on rappelle que $x = (1,1,1,1)$). Finalement,

$$\begin{aligned} p(x) &= (1,1,1,0) + \frac{3}{15}(-2,1,1,3) \\ &= (1,1,1,0) + \frac{1}{5}(-2,1,1,3) \\ &= \left(\frac{3}{5}, \frac{6}{5}, \frac{6}{5}, \frac{3}{5}\right) \end{aligned}$$

Rappelons la définition d'une distance sur E.

Définition 4.3.11. *Une application d de $E \times E$ dans $\mathbb{R}_+$ est une distance si pour tout $(x,y,z) \in \mathbb{R}^3$ elle vérifie les propriétés suivantes.*

1. *$d(x,y) = 0$ si et seulement si $x = y$.*
2. *$d(x,y) = d(y,x)$*
3. *$d(x,z) \leq d(x,y) + d(y,z)$.*

Proposition 4.3.12. *Soit $\mathcal{N}$ une norme sur E. L'application d de $E \times E$ dans $\mathbb{R}_+$ qui à (x,y) associe $\mathcal{N}(x-y)$ est une distance sur E. Elle est appelée distance associée à $\mathcal{N}$.*

Soit d la distance associée à $||\cdot||$ (donc pour tout $(x,y) \in E^2$, $d(x,y) = ||x-y||$).

Nous allons définir la distance d'un élément à un sous-ensemble non vide de E. Pour cela, rappelons la définition de la borne inférieure.

Définition 4.3.13. *Soit $\mathcal{D}$ un sous-ensemble non vide minoré de $\mathbb{R}$. La borne inférieure de $\mathcal{D}$, notée $\inf \mathcal{D}$, est le plus grand des minorants de $\mathcal{D}$. C'est-à-dire que $m = \inf \mathcal{D}$ si*

1. *pour tout $\alpha \in \mathcal{D}$, $m \leq \alpha$.*
2. *si $\beta \leq \alpha$ pour tout $\alpha \in \mathcal{D}$, alors $\beta \leq m$*

Venons en maintenant à la définition annoncée.

Définition 4.3.14. *Pour tout sous-ensemble non vide F de E et tout $x \in E$, on appelle distance de x à F le réel*

$$d(x, F) = \inf\{d(x, y) \ : \ y \in F\}.$$

Comme d est à valeurs dans $\mathbb{R}_+$, cet ensemble $\mathcal{D} = \{d(x, y) \ : \ y \in K\}$ est minoré par 0. Comme F est non vide, $\mathcal{D}$ aussi non vide donc $d(x, F)$ est bien défini dans $\mathbb{R}$.

Proposition 4.3.15. *Soient F un sous espace vectoriel de E et p la projection orthogonale sur F. Alors pour tout $x \in E$,*

$$d(x, p(x)) = d(x, F).$$

De plus, si $y \neq p(x)$,

$$d(x, y) > d(x, F).$$

Démonstration. Exercice. □

Dans la démonstration de ce résultat, on rencontre le célèbre théorème de Pythagore sous la forme suivante.

Théorème 4.3.16 (Pythagore)**.** *Soient x et y deux éléments de E tel que x est orthogonal à y. Alors*

$$||x + y||^2 = ||x||^2 + ||y||^2$$

Démonstration. Pour tout $(x, y) \in E^2$,

$$\begin{aligned} ||x + y||^2 &= (x + y|x + y) \\ &= (x|x) + (x|y) + (y|x) + (y|y) \\ &= ||x||^2 + 2(x|y) + ||y||^2 \end{aligned}$$

Si x et y sont orthogonaux, $(x|y) = 0$ et on trouve bien le résultat annoncé. □

Exemple 2. On reprend l'exemple 1 où $E = \mathbb{R}^4$ et $F = \mathrm{Vect}(b_1, b_2)$ où $b_1 = (1, 1, 1, 0)$ et $b_2 = (0, 1, 1, 1)$. On a vu que si $x = (1, 1, 1, 1)$, $p(x) = \left(\frac{3}{5}, \frac{6}{5}, \frac{6}{5}, \frac{3}{5}\right)$. Donc $x - p(x) = \left(\frac{2}{5}, \frac{-1}{5}, \frac{-1}{5}, \frac{2}{5}\right)$ et

$$||x - p(x)|| = \frac{1}{5}\sqrt{4 + 1 + 1 + 4} = \frac{\sqrt{10}}{5}$$

Finalement

$$d(x, F) = ||x - p(x)|| = \sqrt{\frac{2}{5}}$$

4.4. Matrices de Gram

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n. Soit b une forme bilinéaire symétrique définie sur E et q la forme quadratique associée.

Définition 4.4.1. *Soit $(f_1, \dots, f_r)$ une famille d'éléments de E. On appelle matrice de Gram de (f_i) pour b la matrice $M = (m_{i,j}) \in M_r(\mathbb{R})$ telle que $m_{i,j} = b(f_i, f_j)$ pour tout $(i,j) \in [[1,r]]^2$.*

Remarque. Comme b est symétrique, la matrice de Gram ainsi définie est symétrique, c'est-à-dire que $m_{i,j} = m_{j,i}$ pour tout (i,j).

Proposition 4.4.2. *Soient $(\varepsilon_1, \dots, \varepsilon_n)$ une base de E et M la matrice de Gram de (ε_i) pour b. Soient*

$$x = \sum_{i=1}^{n} x_i \varepsilon_i \quad et \quad y = \sum_{i=1}^{n} y_i \varepsilon_i$$

deux éléments de E écrits dans la base (ε_i). Soient $X = {}^t(x_1, \dots, x_n)$, $Y = {}^t(y_1, \dots, y_n)$,

$$b(x,y) = {}^tXMY = {}^tYMX.$$

Démonstration.

$$\begin{aligned}
b(x,y) &= b\left(\sum_{i=1}^{n} x_i \varepsilon_i, \sum_{j=1}^{n} y_j \varepsilon_j\right) \\
&= \sum_{i=1}^{n} x_i b\left(\varepsilon_i, \sum_{j=1}^{n} y_j \varepsilon_j\right) \quad \text{(linéarité à gauche)} \\
&= \sum_{i,j} x_i y_j b(\varepsilon_i, \varepsilon_j) \quad \text{(linéarité à droite).}
\end{aligned}$$

D'autre part, comme

$$MY = \begin{pmatrix} b(\varepsilon_1, \varepsilon_1) & \dots & b(\varepsilon_1, \varepsilon_n) \\ \vdots & & \vdots \\ b(\varepsilon_n, \varepsilon_1) & \dots & b(\varepsilon_n, \varepsilon_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} b(\varepsilon_1, \varepsilon_j) y_j \\ \vdots \\ \sum_{j=1}^{n} b(\varepsilon_n, \varepsilon_j) y_j \end{pmatrix},$$

$$\begin{aligned}
{}^tXMY &= (x_1, \dots, x_n) \begin{pmatrix} \sum_{j=1}^n b(\varepsilon_1, \varepsilon_j) y_j \\ \vdots \\ \sum_{j=1}^n b(\varepsilon_n, \varepsilon_j) y_j \end{pmatrix} \\
&= \sum_{i=1}^n x_i \sum_{j=1}^n b(\varepsilon_i, \varepsilon_j) y_j \\
&= \sum_{i,j} x_i y_j b(\varepsilon_i, \varepsilon_j).
\end{aligned}$$

□

Remarque. Si b est un produit scalaire la famille $(f_1, \dots, f_n)$ est une base orthonormée de E si et seulement si sa matrice de Gram pour b est égale à I_n.

Proposition 4.4.3. *Soient $(\varepsilon_1, \dots, \varepsilon_n)$ une base de E et Q la matrice de Gram de (ε_i) pour b. Soit $(f_1, \dots, f_r)$ une famille de E. On note P la matrice de $M_{n,r}(\mathbb{R})$ dont les colonnes sont les coordonnées de $f_1, \dots, f_r$ dans la base (ε_i). La matrice de Gram de (f_i) pour b est ${}^tPQP \in M_r(\mathbb{R})$.*

Démonstration.

$$\begin{aligned}
b(f_i, f_j) &= b\left(\sum_{k=1}^n P_{k,i}\varepsilon_k, \sum_{l=1}^n P_{l,j}\varepsilon_l\right) \\
&= \sum_{k,l} P_{k,i} P_{l,j} b(\varepsilon_k, \varepsilon_l) \\
&= \sum_{k,l} P_{k,i} P_{l,j} Q_{k,l} \\
&= \sum_{k} P_{k,i} (QP)_{k,j} \\
&= ({}^tPQP)_{i,j}
\end{aligned}$$

□

On note $\mathrm{Sym}_n(\mathbb{R})$ l'ensemble des matrices symétriques de $M_n(\mathbb{R})$.

Soit $M \in \mathrm{Sym}_n(\mathbb{R})$. Alors M définit une forme bilinéaire b sur $M_{n,1}(\mathbb{R})$:

$$b(X, Y) = {}^tXMY$$

et donc aussi une forme quadratique q. On dit que M est positive (resp. définie positive) si q l'est, comme l'indique la définition suivante.

Définition 4.4.4. *Soit $M \in Sym_n(\mathbb{R})$. On dit que M est positive si pour tout $X \in M_{n,1}(\mathbb{R})$, ${}^tXMX \geq 0$. On dit que M est définie positive si pour tout $X \in M_{n,1}(\mathbb{R}) \smallsetminus \{0\}$, ${}^tXMX > 0$.*

Remarque. Si $M \in \mathrm{Sym}_n(\mathbb{R})$, l'application

$$\begin{aligned} q \ : \ M_{n,1}(\mathbb{R}) &\to \mathbb{R} \\ X &\mapsto {}^tXMX \end{aligned}$$

est une forme quadratique sur $M_{n,1}(\mathbb{R})$ et M est la matrice de Gram de la base canonique de $M_{n,1}(\mathbb{R})$ pour q. Alors M est positive si q est positive. Elle est définie positive si q est définie positive.

Proposition 4.4.5. *Supposons que b soit un produit scalaire (donc (E, b) est un espace euclidien). Soit $(f_1, \dots, f_n)$ une famille d'éléments de E et soit M la matrice de Gram de (f_i) pour b.*

1. *Si la famille (f_i) est liée,* $\det M = 0$.
2. *Si elle est libre, il existe $R \in \mathcal{T}^{++}_{sup,n}(\mathbb{R})$ telle que $M = {}^tRR$. De plus,* $\det M > 0$ *et M est définie positive.*

Démonstration. 1. Si la famille est liée, il existe $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n \smallsetminus \{(0, \dots, 0)\}$ tel que $\sum_{i=1}^{n} \lambda_i f_i = 0$. Pour tout j, on note M_j la j-ème colonne de M.

$$\begin{aligned} \sum_{j=1}^{n} \lambda_j M_j &= \sum_{j=1}^{n} \lambda_j \begin{pmatrix} b(f_1, f_j) \\ \vdots \\ b(f_n, f_j) \end{pmatrix} \\ &= \begin{pmatrix} b\left(f_1, \sum_{j=1}^n \lambda_j f_j\right) \\ \vdots \\ b\left(f_n, \sum_{j=1}^n \lambda_j f_j\right) \end{pmatrix} \\ &= \begin{pmatrix} b\left(f_1, 0\right) \\ \vdots \\ b\left(f_n, 0\right) \end{pmatrix} \\ &= 0 \end{aligned}$$

2. Supposons que la famille (f_i) soit libre. C'est donc une base de E. Soit $(\varepsilon_1, \dots, \varepsilon_n)$ son orthonormalisée de Gram-Schmidt. La matrice de Gram

de (ε_i) pour b est I_n. Soit R la matrice de passage de (ε_i) à (f_i). Alors $R \in \mathcal{T}^{++}_{sup,n}(\mathbb{R})$. Grâce à la proposition 4.4.3, on sait que

$$M = {}^tRR.$$

Comme $\det R \neq 0$, $\det M = (\det R)^2 > 0$. De plus, si $X \in M_{n,1}(\mathbb{R}) \setminus \{0\}$, alors $RX \neq 0$ et

$${}^tXMX = {}^tX{}^tRRX = {}^t(RX)(RX) > 0.$$

□

Remarque. Cette factorisation $M = {}^tRR$ d'une matrice symétrique définie positive est appelée décomposition de Cholesky de M. Nous y reviendrons plus tard.

4.5. Matrices orthogonales

Définition 4.5.1. *Une matrice U de $M_n(\mathbb{R})$ est dite orthogonale si ${}^tUU = I_n$. On appelle $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales.*

Remarque. Si $U \in \mathcal{O}_n(\mathbb{R})$, $(\det U)^2 = 1$ donc $\det U \in \{-1, 1\}$.

Exercice. L'ensemble $\mathcal{O}_n(\mathbb{R})$ muni de la multiplication des matrices est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

Proposition 4.5.2. *Soit $U \in M_n(\mathbb{R})$. Alors $U \in \mathcal{O}_n$ si et seulement si les colonnes de U forment une base orthonormale de $M_{n,1}(\mathbb{R})$ pour le produit scalaire défini par $(X|Y) = {}^tXY$.*

Démonstration. Exercice. □

Exemples. Pour tout n, $I_n \in \mathcal{O}_n(\mathbb{R})$. Les matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ appartiennent à $\mathcal{O}_2(\mathbb{R})$ pour tout $t \in \mathbb{R}$.

Soit E un espace vectoriel euclidien. On note $(\cdot|\cdot)$ son produit scalaire.

Définition 4.5.3. *Soit u un endomorphisme de E. On dit que u est orthogonal si pour tout $(x, y) \in E^2$, $(u(x)|u(y)) = (x|y)$.*

Lemme 4.5.4. *Soient A et B dans $M_n(\mathbb{R})$. Pour tout $(i, j) \in [[1, n]]^2$, ${}^te_iAe_j = A_{i,j}$. En particulier, ${}^tXAY = {}^tXBY$ pour tout $(X, Y) \in M_{n,1}(\mathbb{R})^2$ si et seulement si $A = B$.*

Proposition 4.5.5. *Soit (ε_i) une base orthonormée de E. Soient u un endomorphisme de E et U sa matrice dans la base (ε_i). Alors u est orthogonal si et seulement si U est orthogonale.*

Démonstration. Soient x et y des éléments de E. Soient X et Y les vecteurs coordonnées de x et y. Alors

$$(x|y) = {}^tXY \quad \text{et} \quad (u(x)|u(y)) = {}^tX({}^tUU)Y.$$

On en déduit que u est orthogonal si et seulement si ${}^tX({}^tUU)Y = {}^tXY$ pour tout $(X,Y) \in M_{n,1}(\mathbb{R})^2$, c'est-à-dire si et seulement si ${}^tUU = I_n$. □

4.6. Endomorphismes auto-adjoints

Notation. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie. Soit a un endomorphisme de E. Soit (f_i) une base de E et soit A la matrice de a dans (f_i). On sait que $\mathrm{Spec}(A)$ ne dépend pas de la base choisie. On note

$$\mathrm{Spec}(a) = \mathrm{Spec}(A)$$

Les valeurs propres de a sont les éléments de $\mathbb{R}$ tels qu'il existe $x \in E \smallsetminus \{0\}$ vérifiant $a(x) = \lambda x$. Ces valeurs propres sont les éléments réels de $\mathrm{Spec}(a)$.

Soit $(E, (\cdot|\cdot))$ un espace euclidien de dimension n. Soit (ε_i) une base orthonormée de E.

Proposition 4.6.1. *Soit u un endomorphisme de E. Il existe un unique endomorphisme u^* de E tel que pour tout $(x,y) \in E^2$,*

$$(u(x)|y) = (x|u^*(y)).$$

Si A est la matrice de u dans la base (ε_i), la matrice de u^* dans cette base est égale à ${}^tA$.

Démonstration. Pour tout $(x,y) \in E^2$, on note X et $Y \in M_{n,1}(\mathbb{R})$ les vecteurs coordonnées respectifs de x et y dans la base (ε_i). Soient u et v deux endomorphismes de E, et soient A et B leur matrice respective dans la base (ε_i).

$$(u(x)|y) = {}^t(AX)Y = {}^tX{}^tAY \quad \text{et} \quad (x|v(y)) = {}^tXBY$$

donc $(u(x)|y) = (x|v(y)$ pour tout $(x,y) \in E^2$ si et seulement si ${}^tX{}^tAY = {}^tXBY$ pour tout $(X,Y) \in M_{n,1}(\mathbb{R})$, ce qui est équivalent à $B = {}^tA$ en vertu du lemme 4.5.4. □

Exemple. Sur $E = M_{n,1}(\mathbb{R})$, on considère le produit scalaire défini par $(X|Y) = {}^tXY$. Pour tout $A \in M_n(\mathbb{R})$, l'adjoint de $X \mapsto AX$ est l'endomorphisme $X \mapsto {}^tAX$.

Définition 4.6.2. *Soit $u \in End(E)$. On dit que u est auto-adjoint si $u^* = u$.*

Proposition 4.6.3. *Soit A une matrice symétrique réelle. Alors $Spec(A) \subset \mathbb{R}$.*

Démonstration. Soient $\lambda \in \mathbb{C}$ une valeur propre de A et $X \in M_{n,1}(\mathbb{C})$ un vecteur propre associé à cette valeur propre. Alors ${}^t\overline{X}AX = \lambda {}^t\overline{X}X$. D'autre part, ${}^t\overline{X}AX = {}^t(\overline{AX})X = \overline{\lambda}{}^t\overline{X}X$. Ainsi, $\lambda {}^t\overline{X}X = \overline{\lambda}{}^t\overline{X}X$. Comme ${}^t\overline{X}X > 0$, on en déduit que $\overline{\lambda} = \lambda$, et donc que $\lambda \in \mathbb{R}$. □

Corollaire 4.6.4. *Si $u \in End(E)$ est auto-adjoint, alors $Spec(u) \subset \mathbb{R}$.*

Proposition 4.6.5. *Soient u un endomorphisme de E et F un sous espace vectoriel de E stable par u (c'est-à-dire tel que $u(F) \subset F$). Alors $F^\perp$ est stable par u^*.*

Démonstration. Soit $y \in F^\perp$. Alors pour tout $x \in F$, $(x|y) = 0$ donc pour tout $x \in F$, $(u(x)|y) = 0$ (puisque $u(F) \subset F$), c'est-à-dire $(x|u^*(y)) = 0$ pour tout $x \in F$. On en déduit que $u(y) \in F^\perp$. □

Théorème 4.6.6. *[décomposition spectrale]*

1. *Soit u un endomorphisme auto-adjoint de E. Alors il existe une base orthonormée de E formée de vecteurs propres de u.*
2. *Toute matrice symétrique réelle A est diagonalisable. Plus précisément, il existe une matrice U orthogonale telle que ${}^tUAU$ est diagonale.*

Démonstration. On le démontre la première assertion par récurrence.

Si $n = 1$, soit $\eta \in E$ tel que $(\eta|\eta) = 1$. Alors (η) est une base de E, donc il existe λ tel que $u(\eta) = \lambda\eta$. η est donc bien un vecteur propre normé.

Supposons le résultat vérifié si $\dim E = n - 1$. Soit E de dimension n. L'endomorphisme u admet au moins une valeur propre réelle λ d'après le corollaire 4.6.4. Soit η_n un vecteur propre associé. Quitte à le diviser par $||\eta_n||$, on peut le supposer de norme 1. Comme $u(\mathbb{R}\eta_n) \subset \mathbb{R}\eta_n$, d'après le lemme précédent, le sous-espace vectoriel $F = \eta_n^\perp$ est stable par u. On a aussi $E = \mathbb{R}\eta_n \oplus F$. Soit v l'endomorphisme de F induit par u (c'est-à-dire tel que $v(x) = u(x)$ pour tout $x \in F$). D'après l'hypothèse de récurrence, F admet une base orthonormale $\eta_1, \ldots, \eta_{n-1}$ formée de vecteurs propres de v, donc de u. Finalement, $(\eta_1, \ldots, \eta_n)$ est une base orthonormée de E formée de vecteurs propres de u.

Pour démontrer la seconde assertion, on applique le 1 à $E = M_{n,1}(\mathbb{R})$ muni du produit scalaire $(X|Y) = {}^tXY$. Alors l'endomorphisme u de E qui à X associe AX est auto-adjoint. Il existe donc une base orthonormée $U_1, \dots, U_n$ de E formée de vecteurs propres de u, donc de la matrice A. Soit $U = (U_1|\dots|U_n)$ la matrice de $M_n(\mathbb{R})$ dont les colonnes sont les U_i. La matrice U est alors orthogonale et donc $U^{-1} = {}^tU$. La matrice ${}^tUAU = U^{-1}AU$ est diagonale. $\square$

Chapitre 5

Espaces hermitiens

Les espaces hermitiens sont des espaces vectoriels de dimension finie sur $\mathbb{C}$ munis d'un produit scalaire dit hermitien. Il existe une forte ressemblance entre espaces euclidiens et espaces hermitiens, du point de vue des résultats et aussi des démonstrations.

Nous nous contentons de donner ici les définitions et quelques résultats importants. Certaines preuves seront traitées en exercice.

5.1. Formes hermitiennes

Soit E un $\mathbb{C}$-espace vectoriel.

Définition 5.1.1. *Soient F un autre $\mathbb{C}$-espace vectoriel et f une application de E dans F. On dit que f est une application semi-linéaire si pour tout $(x, y) \in E \times E$ et tout $(\alpha, \beta) \in \mathbb{C} \times \mathbb{C}$,*

$$f(\alpha x + \beta y) = \overline{\alpha} f(x) + \overline{\beta} f(y).$$

Venons en maintenant aux formes hermitiennes.

Définition 5.1.2. *Soit h une application de $E \times E$ dans $\mathbb{C}$.*

1. *On dit que h est semi-linéaire à gauche si pour tout $y \in E$ fixé, l'application de E dans K qui à x associe $h(x, y)$ est semi-linéaire.*
2. *On dit que h est sesquilinéaire si elle est semi-linéaire à gauche et linéaire à droite.*
3. *on dit que h est une forme hermitienne si elle est sesquilinéaire et si pour tout $(x, y) \in E^2$,*
$$h(x, y) = \overline{h(y, x)} \tag{5.1}$$

Remarque. L'égalité (5.1) montre que si h est hermitienne, alors $h(x,x) \in \mathbb{R}$ pour tout élément x de E.

Définition 5.1.3. *Une forme quadratique hermitienne est une application de E dans $\mathbb{R}$ telle qu'il existe une forme hermitienne h vérifiant $h(x,x) = q(x)$ pour tout $x \in E$. On dit que q est la forme quadratique hermitienne associée à h.*

Remarque. L'égalité (5.1) montre que $\overline{q(x)} = q(x)$, donc que $q(x) \in \mathbb{R}$ pour tout x in E.

Proposition 5.1.4. *Une application q de E dans $\mathbb{R}$ est une forme quadratique hermitienne si et seulement si les conditions suivantes sont vérifiées.*

1. *L'application h de $E \times E$ dans $\mathbb{C}$ définie par*
$$h(x,y) = \frac{1}{4}\left(q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)\right)$$
est une forme sesquilinéaire.
2. *Pour tout $x \in E$ et tout $\lambda \in \mathbb{C}$,*
$$q(\lambda x) = |\lambda|^2 q(x).$$

L'application h est alors l'unique forme hermitienne dont q est la forme quadratique associée. On dit aussi que h la forme hermitienne associée à q.

Exemples.

1. $E = \mathbb{C}$, $q(x) = |x|^2$, $h(x,y) = \overline{x}y$.
2. $E = \mathbb{C}^n$, $q(x) = \sum_{i=1}^{n} \lambda_i |x_i|^2$ (où les λ_i sont des éléments de $\mathbb{R}$), $h(x,y) = \sum_{i=1}^{n} \lambda_i \overline{x_i} y_i$.
3. $E = \mathbb{C}[x]$, $q(P) = \int_0^1 |P(t)|^2 dt$, $h(P,Q) = \int_0^1 \overline{P}Q(t)dt$.

Définition 5.1.5. *Soit q une forme quadratique hermitienne, et soit h sa forme hermitienne associée.*

1. *On dit que q est positive si $q(x) \geq 0$ pour tout $x \in E$. On dit aussi que h est positive.*
2. *On dit que q est définінie positive si $q(x) > 0$ pour tout $x \in E$. On dit aussi que h est un produit scalaire hermitien (ou bien que h est hermitienne définie positive).*

Définition 5.1.6. *Un espace hermitien est un $\mathbb{C}$ espace vectoriel de dimension finie muni d'un produit scalaire hermitien.*

5.2. Orthogonalité

Soit E un espace hermitien et h son produit scalaire hermitien. Comme dans le cas réel, on dit que deux éléments x et y de E sont orthogonaux si $h(x,y)=0$.

On définit les familles et bases orthogonales et orthonormales comme dans le cas réel. Là encore, tout espace hermitien admet une base orthonormée. On peut obtenir une telle base par la méthode de Gram-Schmidt.

5.3. Matrices remarquables

Définition 5.3.1. *Une matrice U de $M_n(\mathbb{C})$ est dite unitaire si ${}^t\overline{U}U = I_n$. On appelle $\mathcal{U}_n(\mathbb{C})$ l'ensemble de ces matrices.*

Exercice. Soit $U \in M_n(\mathbb{C})$. Montrer que $U \in \mathcal{U}_n(\mathbb{C})$ si et seulement si les colonnes de U forment une base orthonormale de $M_{n,1}(\mathbb{C})$ pour le produit scalaire hermitien défini par $(X|Y) = {}^t\overline{X}Y$.

Exercice. 1. Montrer que l'ensemble $\mathcal{U}_n(\mathbb{C})$ muni de la multiplication des matrices est un groupe.

2. Montrer que $\mathcal{O}_n(\mathbb{R}) = \mathcal{U}_n(\mathbb{C}) \cap M_n(\mathbb{R})$.

Définition 5.3.2. *Une matrice A de $M_n(\mathbb{C})$ est dite hermitienne si ${}^t\overline{A} = A$. On note $Herm_n(\mathbb{C})$ l'ensemble de ces matrices.*

Remarque. $\mathrm{Herm}_n(\mathbb{C}) \cap M_n(\mathbb{R}) = \mathrm{Sym}_n(\mathbb{R})$.

Proposition 5.3.3. *Soit A une matrice hermitienne. Alors $Spec(A) \subset \mathbb{R}$.*

Théorème 5.3.4. *[décomposition spectrale] Toute matrice hermitienne A est diagonalisable. Plus précisément, il existe une matrice U unitaire telle que ${}^t\overline{U}AU$ est diagonale (et à coefficients réels d'après la proposition 5.3.3).*

Chapitre 6

Normes matricielles, suites de matrices

Dans ce chapitre, on étudie des suites de matrices de $M_n(\mathbb{C})$. et leur convergence. Pour cela, nous avons besoin d'une topologie. Quelques rappels sur les espaces vectoriels normés seront donc faits, avant de définir et d'étudier les normes matricielles. La troisième partie sera consacrée aux suites (A^k).

On note $K = \mathbb{R}$ ou $\mathbb{C}$.

6.1. Norme sur un espace vectoriel

Soit E un K-espace vectoriel. Nous avons déjà revu plus haut la définition d'une norme sur E.

Définition 6.1.1. *Une application $x \mapsto ||x||$ de E dans $\mathbb{R}_+$ est une norme si les conditions suivantes sont vérifiées.*

1. $||x|| = 0$ *si et seulement si* $x = 0$.
2. *Pour tout* $\lambda \in K$ *et tout* $x \in E$,
$$||\lambda x|| = |\lambda| \cdot ||x||.$$
3. *Pour tout* $(x, y) \in E^2$,
$$||x + y|| \leq ||x|| + ||y||.$$

La proposition qui suit donne des exemples de norme.

Proposition 6.1.2. *Sur K^n, les applications suivantes sont des normes.*

1. $||x||_\infty = \max_{i \in [[1,n]]} |x_i|$
2. $||x||_1 = \displaystyle\sum_{i=1}^{n} |x_i|$.

3. $||x||_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$.

Si $K = \mathbb{R}$, la norme $||\cdot||_2$ provient du produit scalaire $(x|y) = \sum_{i=1}^{n} x_i y_i$. Si $K = \mathbb{C}$, elle provient du produit scalaire hermitien $(x|y) = \sum_{i=1}^{n} \overline{x_i} y_i$.

Démonstration. Exercice. □

Définition 6.1.3. *Soient N et N' deux normes d'un K-espace vectoriel E. On dit que N et N' sont équivalentes s'il existe deux réels strictement positifs α et β tels que pour tout $x \in E$,*

$$\alpha N(x) \leq N'(x) \leq \beta N(x).$$

On note alors : $N \sim N'$.

Proposition 6.1.4. *La relation $\sim$ est une relation d'équivalence.*

Démonstration. Exercice. Rappelons qu'il s'agit de montrer les propriétés suivantes.

1. Réflexivité : $N \sim N$ pour toute norme N.
2. Symétrie : si N et N' sont deux normes sue E, et si $N \sim N'$, alors $N' \sim N$.
3. Transitivité : si N, N' et N'' sont des normes définies sur E qui vérifient $N \sim N'$ et $N' \sim N''$, alors $N \sim N''$.

□

Proposition 6.1.5. *Les normes $||\cdot||_1$, $||\cdot||_\infty$ et $||\cdot||_2$ sont équivalentes. Plus précisément, pour tout $x \in K^n$,*

$$||x||_\infty \leq ||x||_1 \leq n||x||_\infty \quad et \quad ||x||_\infty \leq ||x||_2 \leq \sqrt{n}||x||_\infty$$

Démonstration. Exercice. □

Plus généralement, on a le résultat suivant (admis).

Théorème 6.1.6. *Si E est un K-espace de dimension finie ($K = \mathbb{R}$ ou $\mathbb{C}$), toutes les normes définies sur E sont équivalentes.*

6.2. Normes matricielles

L'ensemble $M_n(K)$ est un K-espace vectoriel, et la notion de norme définie sur un espace vectoriel s'applique aussi à $M_n(K)$. Mais $M_n(K)$ est aussi munie d'un produit interne : la multiplication entre matrices. Les normes matricielles sont des normes définies sur un espace de matrices qui vérifient en plus une inégalité triangulaire pour le produit.

Définition 6.2.1. *On dit qu'une norme $||\cdot||$ définie sur $M_n(K)$ est une norme matricielle si pour tout $(A,B) \in M_n(K)^2$,*

$$||AB|| \leq ||A|| \cdot ||B||.$$

Remarque. Il existe des normes sur $M_n(K)$ qui ne sont pas des normes matricielles (voir les TD).

Théorème 6.2.2. *Soient $(E, ||\cdot||_E)$ et $(F, ||\cdot||_F)$ deux K-espaces vectoriels normés. Soit u une application linéaire de E dans F. Les propriétés suivantes sont équivalentes.*

1. *u est continue.*
2. *u est continue en 0.*
3. *$\exists C \geq 0 \ : \ \forall x \in E, \ ||u(x)||_F \leq C||x||_E$.*
4. *u est lipschitzienne, c'est-à-dire : $\exists C \geq 0 \ : \ \forall x, y \in E, \ ||u(x) - u(y)||_F \leq C||x - y||_E$.*

Démonstration. Exercice. □

Corollaire 6.2.3. *Soit A une matrice de $M_{m,n}(K)$. L'application de $M_{n,1}(K)$ dans $M_{m,1}(K)$ qui à X associe AX est continue pour toute norme sur $M_{n,1}(K)$ et $M_{m,1}(K)$.*

Remarque. Comme on est en dimension finie, pour que cet énoncé soit exact, il faut et il suffit qu'il soit exact pour une norme de $M_{n,1}(K)$ et de $M_{m,1}(K)$. En dimension finie, nous parlerons dorénavant de continuité sans spécifier de norme.

Démonstration. D'après la remarque précédente, on peut choisir les normes qu'on veut pour démontrer le théorème. Prenons $||\cdot||_\infty$ sur $M_{n,1}(K)$ et $M_{m,1}(K)$. Soit $X = {}^t(x_1, \ldots, x_n) \in M_{n,1}(K)$. Pour tout $k \in [[1, n]]$,

$$(AX)_k = \sum_{j=1}^{n} a_{k,j} x_j$$

donc

$$\begin{aligned}
|(AX)_k| &\leq \sum_{j=1}^{n} |a_{k,j}|.|x_j| \\
&\leq ||X||_\infty \sum_{j=1}^{n} |a_{k,j}| \\
&\leq ||X||_\infty \left(\max_{i\in[[1,n]]} \sum_{j=1}^{n} |a_{i,j}| \right) \\
&\leq C||X||_\infty
\end{aligned}$$

où

$$C = \max_{i\in[[1,n]]} \sum_{j=1}^{n} |a_{i,j}|.$$

Comme cette inégalité est vraie pour tout k, on en déduit que

$$||AX||_\infty \leq C||X||_\infty.$$

□

En dimension finie, les applications linéaires peuvent se décrire à l'aide de matrices. On peut donc déduire le résultat suivant.

Corollaire 6.2.4. *Si E et F sont des K-espaces vectoriels de dimension finie, toute application linéaire de E dans F est continue.*

Corollaire 6.2.5. *Soit $A \in M_{m,n}(K)$ (donc si $X \in M_{n,1}(K)$, alors $AX \in M_{m,1}(K)$. Pour toute norme $||\cdot||$ de $M_{n,1}(K)$, l'ensemble $\{AX \ : \ ||X|| = 1\}$ est borné, quelle que soit la norme de $M_{m,1}(K)$ considérée.*

Démonstration. On a vu dans la preuve du corollaire 6.2.3 que c'est vrai pour la norme $||\cdot||_\infty$. C'est donc vrai pour toute norme. □

Définition 6.2.6. *Soit $||\cdot||$ une norme sur $M_{n,1}(\mathbb{C})$. On appelle norme subordonnée à $||\cdot||$ l'application*

$$\begin{aligned}
M_n(\mathbb{C}) &\to \mathbb{R} \\
A &\mapsto |||A||| = \sup_{||X||=1} ||AX||
\end{aligned}$$

Proposition 6.2.7. *Avec les mêmes notations, pour tout $A \in M_n(\mathbb{C})$,*

$$|||A||| = \sup_{0<||X||\leq 1} ||AX|| = \sup_{X\neq 0} \frac{||AX||}{||X||}.$$

Démonstration. Pour montrer ces égalités, on montre les inégalités suivantes.

$$|||A||| \leq \sup_{0<||X||\leq 1} ||AX|| \leq \sup_{X\neq 0} \frac{||AX||}{||X||} \leq |||A|||.$$

Commençons par la première de ces inégalités, en partant de la gauche.

$$|||A||| = \sup_{||X||=1} ||AX|| \leq \sup_{0<||X||\leq 1} ||AX||$$

puisque $\{X \in M_{n,1}(\mathbb{C}) \ : \ ||X|| = 1\} \subset \{X \in M_{n,1}(\mathbb{C}) \ : \ 0 < ||X|| \leq 1\}$. La première inégalité est donc démontrée.

Montrons la seconde inégalité. Pour tout élément X de $\{X \in M_{n,1}(\mathbb{C}) \ : \ 0 < ||X|| \leq 1\}$, il est clair que $||AX|| \leq \dfrac{||AX||}{||X||}$ donc

$$||AX|| \leq \sup_{0<||Z||\leq 1} \frac{||AZ||}{||Z||}.$$

Comme c'est vrai pour tout X tel que $0 < ||X|| \leq 1$, on obtient

$$\sup_{0<||X||\leq 1} ||AX|| \leq \sup_{0<||Z||\leq 1} \frac{||AZ||}{||Z||},$$

ce qui bien sûr s'écrit aussi

$$\sup_{0<||X||\leq 1} ||AX|| \leq \sup_{0<||X||\leq 1} \frac{||AX||}{||X||},$$

Comme $\{X \in M_{n,1}(\mathbb{C}) \ : \ 0 < ||X|| \leq 1\} \subset M_{n,1}(\mathbb{C}) \setminus \{0\}$,

$$\sup_{0<||X||\leq 1} \frac{||AX||}{||X||} \leq \sup_{X\neq 0} \frac{||AX||}{||X||},$$

d'où la seconde inégalité.

Reste la troisième. Soit $X \in M_{n,1}(\mathbb{C}) \setminus \{0\}$, et soit $Y = \dfrac{X}{||X||}$ (donc $||Y|| = 1$).

$$\begin{aligned}
\frac{||AX||}{||X||} &= \left|\left|\frac{1}{||X||}AX\right|\right| \\
&= \left|\left|A\frac{X}{||X||}\right|\right| \\
&= ||AY|| \\
&\leq \sup_{||Z||=1} ||AZ|| \\
&\leq |||A|||.
\end{aligned}$$

On obtient bien l'inégalité

$$\sup_{||X||\neq 0} \frac{||AX||}{||X||} \leq |||A|||.$$

□

Théorème 6.2.8. *Soit* $||\cdot||$ *une norme sur* $M_{n,1}(\mathbb{C})$.

1. *Sa norme subordonnée* $|||\cdot|||$ *sur* $M_n(\mathbb{C})$ *est une norme matricielle.*
2. *Cette norme vérifie la propriété suivante.*

$$\forall A \in M_n(\mathbb{C}) \ , \forall X \in M_{n,1}(\mathbb{C}) \ , \ ||AX|| \leq |||A||| \cdot ||X||$$

Démonstration. Montrons d'abord que si $|||A||| = 0$, alors $A = 0$. Pour tout $i \in [[1, n]]$, on note A_i la i-ème colonne de A. Si $|||A||| = 0$, alors pour tout $X \in M_{n,1}(\mathbb{C})$, $AX = 0$. Or pour tout $i \in [[1, n]]$, $A_i = Ae_i$, donc $A_i = 0$. Toutes les colonnes de A sont nulles, donc $A = 0$.

On montrera en exercice que pour tout $\lambda \in \mathbb{C}$ et tout $A \in M_n(\mathbb{C})$, $|||\lambda A||| = |\lambda| \cdot |||A|||$.

Soient $\lambda \in \mathbb{C}$ et $A \in M_n(\mathbb{C})$. Montrons que $|||\lambda A||| = |\lambda| \cdot |||A|||$ par double inégalité. C'est vrai si $\lambda = 0$. Supposons donc λ non nul. Pour tout X de norme 1,

$$|\lambda| \cdot ||AX|| = ||\lambda AX|| \leq |||\lambda A|||$$

donc $||AX|| \leq \dfrac{1}{|\lambda|}|||\lambda A|||$. Comme c'est vrai pour tout X de norme 1, $|||A||| \leq \dfrac{1}{|\lambda|}|||\lambda A|||$ donc

$$|\lambda| \cdot |||A||| \leq |||\lambda A|||$$

En appliquant cette inégalité à $\dfrac{1}{\lambda}$ et λA, on obtient l'autre inégalité :

$$\frac{1}{|\lambda|} \cdot |||\lambda A||| \leq |||\frac{1}{\lambda}\lambda A||| = |||A|||$$

donc $|||\lambda A||| \leq |\lambda| \cdot |||A|||$.

Montrons l'inégalité triangulaire pour l'addition. Soit $X \in M_{n,1}(\mathbb{C})$ de norme 1.

$$\begin{aligned} ||(A+B)X|| &\leq ||AX|| + ||BX|| \\ &\leq |||A||| + |||B||| \end{aligned}$$

Donc $\sup\limits_{||X||=1} ||(A+B)X|| \leq |||A||| + |||B|||$, c'est-à-dire

$$|||A+B||| \leq |||A||| + |||B|||.$$

Montrons maintenant l'inégalité triangulaire pour la multiplication. Soient A et B dans $M_n(\mathbb{C})$. Montrons que $|||AB||| \leq |||A|||\cdot|||B|||$. Soit $X \in M_{n,1}(\mathbb{C})$ tel que $X \neq 0$ Supposons d'abord que $BX \neq 0$.

$$\begin{aligned}\frac{||ABX||}{||X||} &= \frac{||ABX||}{||BX||}\frac{||BX||}{||X||} \\ &\leq |||A||| \cdot |||B|||\end{aligned}$$

Si $BX = 0$, alors $ABX = 0$ et on a encore l'inégalité

$$\frac{||ABX||}{||X||} \leq |||A||| \cdot |||B|||$$

On obtient donc

$$|||AB||| = \sup_{X \neq 0} \frac{||ABX||}{||X||} \leq |||A||| \cdot |||B|||.$$

La propriété 2 du théorème est une conséquence directe de la définition de $||| \cdot |||$. □

Définition 6.2.9. *Soient $|| \cdot ||$ une norme sur $M_{n,1}(\mathbb{C})$ et $\mathcal{N}$ une norme matricielle sur $M_n(\mathbb{C})$. On dit que ces normes sont compatibles si pour tout $A \in M_n(\mathbb{C})$ et tout $X \in M_{n,1}(\mathbb{C})$,*

$$||AX|| \leq \mathcal{N}(A)||X||.$$

On note $|||\cdot|||_1$, $|||\cdot|||_2$ et $|||\cdot|||_\infty$ les normes respectivement subordonnées à $|| \cdot ||_1$, $|| \cdot ||_2$ et $|| \cdot ||_\infty$.

Théorème 6.2.10. *Soit $A = (a_{i,j}) \in M_n(\mathbb{C})$.*

$$|||A|||_\infty = \max_{i \in [[1,n]]} \sum_{j=1}^{n} |a_{i,j}| \quad , \quad |||A|||_1 = \max_{j \in [[1,n]]} \sum_{i=1}^{n} |a_{i,j}|.$$

Démonstration. Voir les TD. □

Corollaire 6.2.11. $|||A|||_1 = |||{}^tA|||_\infty$.

Remarque. On a vu au premier chapitre que

$$\rho(A) \leq \max_{i \in [[1,n]]} \sum_{j=1}^{n} |a_{i,j}| \quad \text{et que} \quad \rho(A) \leq \max_{j \in [[1,n]]} \sum_{i=1}^{n} |a_{i,j}|.$$

Proposition 6.2.12. *Si $\mathcal{N}$ est une norme matricielle, alors $\rho(A) \leq \mathcal{N}(A)$.*

Démonstration. Voir les TD. □

Théorème 6.2.13. *Pour tout $A \in M_n(\mathbb{C})$, $|||A|||_2 = \sqrt{\rho({}^t\overline{A}A)}$.*

Démonstration. Voir les TD. □

Corollaire 6.2.14. *1. Pour tout A dans $M_n(\mathbb{C})$, $|||A|||_2 = |||{}^t\overline{A}|||_2$.*

2. Si A est hermitienne, alors $|||A|||_2 = \rho(A)$. En particulier, cette égalité est vraie si A est symétrique réelle.

Démonstration. On sait que si $M, N \in M_n(\mathbb{C})$, Les matrices MN et NM ont même polynôme caractéristique. En appliquant ce fait à A et ${}^t\overline{A}$, on obtient : $\rho({}^t\overline{A}A) = \rho(A{}^t\overline{A})$, donc $|||A|||_2 = |||{}^t\overline{A}|||_2$. Si A est hermitienne, ${}^t\overline{A} = A$ donc ${}^t\overline{A}A = A^2$. On conclut en remarquant que $\rho(A^2) = \rho(A)^2$ (voir le lemme 6.2.15 suivant). □

On a utilisé le fait que $\rho(A^2) = \rho(A)^2$. Ici, la matrice A est diagonalisable, mais c'est vrai aussi dans le cas général.

Lemme 6.2.15. *Pour tout $A \in M_n(\mathbb{C})$ et tout entier positif k,*

$$\rho(A^k) = \rho(A)^k.$$

Démonstration. Il existe $P \in \mathrm{GL}_n(\mathbb{C})$ tel que $T = P^{-1}AP$ est triangulaire supérieure. On note $(d_1, \ldots, d_n)$ la diagonale de T. Alors $\mathrm{Spec}(A) = \{d_1, \ldots, d_n\}$. La diagonale de T^k est $(d_1^k, \ldots, d_n^k)$. Or, $A = PTP^{-1}$, donc $A^k = PT^kP^{-1}$. Ainsi, $\mathrm{Spec}(A^k) = \{d_1^k, \ldots, d_n^k\}$. On en déduit le résultat. □

6.3. La suite $(A^k)_{k\in\mathbb{N}}$

Soit E un $\mathbb{C}$-espace vectoriel. Soit N une norme définie sur E. Soit (u_k) une suite d'éléments de E. On dit que (u_k) est convergente vers $u \in E$ pour la norme N si

$$\lim_{k\to+\infty} N(u_k - u) = 0.$$

Si E est de dimension finie, toutes les normes sur E sont équivalentes. Rappelons que dans ce cas, les notions de limites et de continuité sont indépendantes de la norme utilisée.

Lemme 6.3.1. *Soit $Q \in GL_n(\mathbb{C})$. Pour tout $v \in M_{n,1}(\mathbb{C})$, on pose*

$$\mathcal{N}(v) = ||Q^{-1}v||_\infty.$$

Alors $\mathcal{N}$ est une norme sur $M_{n,1}(\mathbb{C})$. Soit N' sa norme subordonnée. Alors pour tout $A \in M_n(\mathbb{C})$,

$$\mathcal{N}'(A) = |||Q^{-1}AQ|||_\infty.$$

Démonstration. Exercice. □

Lemme 6.3.2. *Soit $A \in M_n(\mathbb{C})$ et soit $\varepsilon > 0$. Il existe une norme $||\cdot||$ sur $M_{n,1}(\mathbb{C})$ dont la norme subordonnée $|||\cdot|||$ sur $M_n(K)$ vérifie*

$$|||A||| \leq \rho(A) + \varepsilon.$$

Démonstration. On triangularise A. Il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que

$$T = P^{-1}AP = \begin{pmatrix} \lambda_1 & t_{1,2} & \dots & t_{1,n} \\ & \ddots & \ddots & \vdots \\ & 0 & \ddots & t_{n-1,n} \\ & & & \lambda_n \end{pmatrix} \in \mathcal{T}_{sup,n}(\mathbb{C}).$$

Soient δ un réel positif, $D = \mathrm{Diag}(1, \delta, \dots, \delta^{n-1})$ et $T_\delta = D^{-1}TD = (PD)^{-1}A(PD)$. Alors

$$T_\delta = D^{-1}TD = \begin{pmatrix} \lambda_1 & \delta t_{1,2} & \dots & \delta^{n-1} t_{1,n} \\ & \ddots & \ddots & \vdots \\ & 0 & \ddots & \delta t_{n-1,n} \\ & & & \lambda_n \end{pmatrix}$$

En effet, la multiplication à gauche par D^{-1} multiplie chaque ligne i de T par δ^{1-i} et la multiplication à gauche par D multiplie chaque colonne j de T par δ^{j-1}. Donc le coefficient (i,j) de $D^{-1}TD$ est égal à $T_{i,j}\delta^{1-i+j-1} = T_{i,j}\delta^{j-i}$.

Fixons $\delta > 0$ tel que pour tout $i \in [[1, n]]$,

$$\sum_{j=i+1}^{n} |\delta^{j-i} t_{i,j}| < \varepsilon.$$

Soit $\mathcal{N}$ la norme sur $M_{n,1}(\mathbb{C})$ définie par

$$\mathcal{N}(v) = ||(PD)^{-1}v||_\infty.$$

Soit $\mathcal{N}'$ sa norme subordonnée. D'après le lemme 6.3.1,

$$\mathcal{N}'(A) = |||(PD)^{-1}A(PD)|||_\infty = |||T_\delta|||_\infty.$$

Alors

$$\begin{aligned} \mathcal{N}'(A) &= \max_{i\in[[1,n]]} \sum_{j=1}^{n} |(T_\delta)_{i,j}| \\ &= \max_{i\in[[1,n]]} \left(|\lambda_i| + \sum_{j=i+1}^{n} |\delta^{j-i} t_{i,j}| \right) \\ &\leq \max_{i\in[[1,n]]} |\lambda_i| + \varepsilon \\ &\leq \rho(A) + \varepsilon. \end{aligned}$$

□

Théorème 6.3.3. *Soit $A \in M_n(\mathbb{C})$. Les propriétés suivantes sont équivalentes.*

(i) *La suite $(A^k)_{k\geq 0}$ converge vers* 0.

(ii) *Pour tout $v \in M_{n,1}(\mathbb{C})$, la suite $(A^k v)_{k\geq 0}$ converge vers* 0.

(iii) $\rho(A) < 1$.

(iv) *Il existe une norme subordonnée $\mathcal{N}$ pour laquelle $\mathcal{N}(A) < 1$.*

(v) *Il existe une norme matricielle $\mathcal{N}$ pour laquelle $\mathcal{N}(A) < 1$.*

Démonstration. (i) $\Longrightarrow$ (ii). L'application f de $M_n(\mathbb{C})$ dans $M_{n,1}(\mathbb{C})$ qui à A associe Av est linéaire. Elle est donc continue puisque les espaces vectoriels considérés sont de dimension finie. Donc si $\lim A^k = 0$, alors $\lim f(A^k) = 0$, c'est-à-dire $\lim A^k v = 0$.

(ii) $\Longrightarrow$ (iii). On suppose (ii) vérifiée. Soient $\lambda \in \mathrm{Spec}(A)$ et v un vecteur propre associé à λ. Alors $Av = \lambda v$ et pour tout $k \in \mathbb{N}$, $A^k v = \lambda^k v$. Comme $\lim_{k\to+\infty} A^k v = 0$, c'est que $\lim_{k\to+\infty} \lambda^k = 0$, donc $|\lambda| < 1$.

(iii) $\Longrightarrow$ (iv). On suppose que $\rho(A) < 1$. Soit $\varepsilon \in]0, 1-\rho(A)[$. Alors $\varepsilon > 0$ et $\rho(A) + \varepsilon < 1$. D'après le lemme 6.3.2, il existe une norme subordonnée $\mathcal{N}$ telle que $\mathcal{N}(A) \leq \rho(A) + \varepsilon < 1$.

(iv) $\Longrightarrow$ (v). C'est clair.

(v) $\Longrightarrow$ (i). Soit $\mathcal{N}$ une norme matricielle telle que $\mathcal{N}(A) < 1$. Pour tout entier $k > 0$, $\mathcal{N}(A^k) \leq \mathcal{N}(A)^k$. Comme $0 < \mathcal{N}(A) < 1$, $\lim_{k\to+\infty} \mathcal{N}(A)^k = 0$ donc $\lim_{k\to+\infty} \mathcal{N}(A^k) = 0$, c'est-à-dire $\lim_{k\to+\infty} A^k = 0$. □

Corollaire 6.3.4. *Soit $||\cdot||$ une norme matricielle de $M_n(\mathbb{C})$. Alors pour toute matrice A,*

$$\lim_{k\to+\infty} ||A^k||^{1/k} = \rho(A).$$

Démonstration. Pour tout A, $\rho(A) \leq ||A||$, donc

$$\rho(A)^k = \rho(A^k) \leq ||A^k||.$$

On en déduit que

$$\rho(A) \leq ||A^k||^{1/k}.$$

Montrons que pour tout $\varepsilon > 0$, il existe C tel que si $k \geq C$, $||A^k||^{1/k} \leq \rho(A) + \varepsilon$. Soit $B = \dfrac{A}{\rho(A)+\varepsilon}$. Alors $\rho(B) < 1$ et

$$\lim_{k\to+\infty} B^k = 0.$$

Donc il existe C tel que si $k \geq C$, $||B^k|| \leq 1$, c'est-à-dire

$$\frac{||A^k||}{(\rho(A)+\varepsilon)^k} \leq 1$$

donc

$$||A^k|| \leq (\rho(A)+\varepsilon)^k,$$

ce qui donne

$$||A^k||^{1/k} \leq \rho(A)+\varepsilon.$$

□

6.4. Conditionnement

Soit $\mathcal{N}$ une norme matricielle sur $M_n(\mathbb{C})$. Soit $A \in \mathrm{GL}_n(\mathbb{C})$. On appelle conditionnement de A relativement à $\mathcal{N}$ le nombre réel

$$\mathrm{cond}(A) = \mathcal{N}(A)\mathcal{N}(A^{-1}).$$

Proposition 6.4.1. *Si $\mathcal{N}$ est une norme subordonnée, alors cond$(A) \geq 1$ pour tout $A \in GL_n(\mathbb{C})$.*

Démonstration. Exercice : montrer d'abord que dans ce cas, $\mathcal{N}(I_n) = 1$, puis utiliser le fait que $AA^{-1} = I_n$. □

Le conditionnement permet de majorer l'erreur d'un résultat provenant d'erreurs dans les données, dues par exemple à des arrondis ou la précision de mesures.

Par exemple, si l'on cherche à résoudre l'équation $AX = B$, où l'on ne connaît pas la valeur exacte de B, mais seulement une approximation. Alors on n'aura accès qu'à une approximation de la solution. Il importe de pouvoir majorer l'erreur commise en fonction de l'erreur sur B. C'est l'objet de la proposition suivante.

Proposition 6.4.2. *On considère une norme $||\cdot||$ de $M_{n,1}(\mathbb{C})$, sa norme subordonnée $|||\cdot|||$ de $M_n(\mathbb{C})$ et cond le conditionnement relatif à cette norme. Soient $A \in GL_n(\mathbb{C})$, $B \in M_{n,1}(\mathbb{C})$ et $\Delta B \in M_{n,1}(\mathbb{C})$. Soient enfin X et $X + \Delta X \in M_{n,1}(\mathbb{C})$ tels que $AX = B$ et $A(X+\Delta X) = B + \Delta B$ (donc $A\Delta X = \Delta B$). Alors*

$$\frac{||\Delta X||}{||X||} \leq \mathit{cond}(A)\frac{||\Delta B||}{||B||}$$

Démonstration. Exercice. □

Chapitre 7

Méthodes itératives

7.1. Le principe

On s'intéresse à l'équation $AX = B$, où $A = (a_{i,i}) =\in \mathrm{GL}_n(\mathbb{C})$ et $B \in M_{n,1}(\mathbb{C})$. On cherche à construire une suite $(X^{(k)})$ de vecteurs de $M_{n,1}(\mathbb{C})$ qui converge vers la solution. Pour cela, on décompose A en une somme $A = M + N$ où M est facile à inverser. On définit alors une suite $X^{(k)}$ par la donnée d'un vecteur initial $X^{(0)} \in M_{n,1}(\mathbb{C})$ et d'une relation de récurrence

$$X^{(k+1)} = -M^{-1}NX^{(k)} + M^{-1}B.$$

Alors si $(X^{(k)})$ converge vers X, $X = -M^{-1}NX + M^{-1}B$, donc $AX = B$.

On pose $K = -M^{-1}N$ et $C = M^{-1}B$. La relation de récurrence devient

$$X^{(k+1)} = KX^{(k)} + C$$

Lemme 7.1.1. *Les deux assertions suivantes sont équivalentes.*

(i) $\rho(K) < 1$.

(ii) $I_n - K$ *est inversible et pour tout* $X^{(0)}$*, la suite* $(X^{(k)})$ *converge vers* $(I_n - K)^{-1}C$ *(c'est-à-dire* $A^{-1}B$*).*

Démonstration. (i) $\Longrightarrow$ (ii) On suppose que $\rho(K) < 1$. Alors $1 \notin \mathrm{Spec}(K)$ donc $I_n - K$ est inversible. Grâce au théorème 6.3.3, le fait que $\rho(K) < 1$ implique qu'il existe une norme $\mathcal{N}$ sur $M_{n,1}(\mathbb{C})$ dont la norme subordonnée $\mathcal{N}'$ sur $M_n(\mathbb{C})$ vérifie

$$\mathcal{N}'(K) < 1$$

On a

$$X^{(k+1)} - X^{(k)} = K(X^{(k)} - X^{(k-1)}).$$

Par récurrence, on voit que

$$X^{(k+1)} - X^{(k)} = K^k(X^{(1)} - X^{(0)}).$$

Ainsi,

$$\begin{aligned}\mathcal{N}(X^{(k+1)} - X^{(k)}) &\leq \mathcal{N}'(K^k)\mathcal{N}(X^{(1)} - X^{(0)})\\ &\leq \mathcal{N}'(K)^k\mathcal{N}(X^{(1)} - X^{(0)}).\end{aligned}$$

Ainsi, si l'on pose $\alpha = \mathcal{N}'(K)$, $\alpha < 1$ et

$$\mathcal{N}(X^{(k+1)} - X^{(k)}) \leq \alpha^k\mathcal{N}(X^{(1)} - X^{(0)}).$$

Soient $k \geq 0$ et $l > 0$,

$$\begin{aligned}\mathcal{N}\left(X^{(k+l)} - X^{(k)}\right) &\leq \sum_{i=0}^{l-1}\mathcal{N}\left(X^{(k+i+1)} - X^{(k+i)}\right)\\ &\leq \sum_{i=0}^{l-1}\alpha^{k+i}\mathcal{N}\left(X^{(1)} - X^{(0)}\right)\\ &\leq \mathcal{N}\left(X^{(1)} - X^{(0)}\right)\frac{\alpha^k}{1-\alpha}\end{aligned}$$

Or

$$\lim_{k\to+\infty}\mathcal{N}\left(X^{(1)} - X^{(0)}\right)\frac{\alpha^k}{1-\alpha} = 0$$

Donc pour tout $\varepsilon > 0$, il existe $N > 0$ tel que si $k > N$ et si $l \geq 0$, alors $\mathcal{N}(X^{(k+l)} - X^{(k)}) < \varepsilon$. La suite $(X^{(k)})$ est donc une suite de Cauchy dans un espace complet. Cela montre qu'elle converge vers une limite L qui vérifie $L = KL + C$, c'est-à-dire : $L = (I_n - K)^{-1}C$.

(ii) $\Longrightarrow$ (i) Soit $Y^{(k)} = (I_n - K)X^{(k)} - C$. Alors $(Y^{(k)})$ converge vers 0.

$$\begin{aligned}Y^{(k+1)} &= (I_n - K)X^{(k+1)} - C\\ &= (I_n - K)(KX^{(k)} + C) - C\\ &= K((I_n - K)X^{(k)} - C)\\ &= KY^{(k)}\end{aligned}$$

Soient λ une valeur propre de K et Y une valeur propre associée. Choisissons comme valeur initiale $X^{(0)} = (I_n - K)^{-1}(Y + C)$ pour la suite $(X^{(k)}$. Alors $Y^{(0)} = Y$ et pour tout k,

$$Y^{(k)} = K^kY = \lambda^kY.$$

Comme $\lim_{k\to} Y^{(k)} = 0$, c'est que $\lim_{k\to}\lambda^k = 0$ donc $|\lambda| < 1$. On en déduit que $\rho(K) < 1$. □

7.2. Notations.

On décompose la matrice A sous la forme suivante.

$$A = L + D + U$$

où L est triangulaire inférieure à diagonale nulle, U est triangulaire supérieure à diagonale nulle et où D est diagonale. Autrement dit :

$$L_{i,j} = \begin{cases} a_{i,j} & \text{si } i > j \\ 0 & \text{sinon} \end{cases} \qquad D_{i,j} = \begin{cases} a_{i,j} & \text{si } i = j \\ 0 & \text{sinon} \end{cases} \qquad U_{i,j} = \begin{cases} a_{i,j} & \text{si } i < j \\ 0 & \text{sinon} \end{cases}$$

Quitte faire des permutations sur les lignes, on peut supposer que $a_{i,i} \neq 0$ pour tout $i \in [[1, n]]$, pour la raison suivante.

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

Comme $\det A \neq 0$, il existe $\sigma \in S_n$ tel que

$$\prod_{i=1}^{n} a_{\sigma(i),i} \neq 0$$

Si l'on applique aux lignes de A la permutation σ, on obtient une nouvelle matrice telle que $a_{i,i} \neq 0$ pour tout $i \in [[1, n]]$.

On suppose donc que D est inversible.

7.3. La méthode de Jacobi

On choisit la décomposition $A = M + N$ où $M = D$ et $N = L + U$. Alors la relation de récurrence à considérer est

$$\begin{aligned} X^{(k+1)} &= -D^{-1}(L + U)X^{(k)} + D^{-1}B \\ &= (I_n - D^{-1}A)X^{(k)} + D^{-1}B \end{aligned}$$

Si l'on note $X_i^{(k)}$ le i-ème coefficient de $X^{(k)}$, cela donne

$$X_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \sum_{j \neq i} a_{i,j} X_j^{(k)} \right). \tag{7.1}$$

La proposition suivante est une conséquence directe du lemme 7.1.1.

Proposition 7.3.1. *La suite $(X^{(k)})$ converge vers la solution X du système pour tout choix de $X^{(0)}$ si et seulement si*

$$\rho(I - D^{-1}A) < 1$$

On dit alors que la méthode de Jacobi converge pour tout vecteur initial.

Corollaire 7.3.2. *Si A est à diagonale strictement dominante, la méthode de Jacobi converge pour tout vecteur initial.*

Démonstration.

$$I - D^{-1}A = \begin{pmatrix} 0 & \frac{a_{1,2}}{a_{1,1}} & \cdots & \frac{a_{1,n}}{a_{1,1}} \\ \frac{a_{2,1}}{a_{2,2}} & 0 & & \frac{a_{2,n}}{a_{2,2}} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{a_{n,1}}{a_{n,n}} & \cdots & \frac{a_{n,n-1}}{a_{n,n}} & 0 \end{pmatrix}$$

Comme A est à diagonale strictement dominante, pour tout k,

$$\sum_{j \neq k} \frac{|a_{k,j}|}{|a_{k,k}|} < 1$$

donc d'après la proposition 2.2.2 (conséquence du théorème de Gershgorin-Hadamard), $\rho(I - D^{-1}A) < 1$. □

7.4. La méthode de Gauss-Seidel

Elle peut être vue comme un avatar de la méthode de Jacobi : on reprend l'égalité (7.1) avec la transformation suivante : on calcule les coordonnées $X_i^{(k+1)}$ pour i allant de 1 à n, en changeant les $X_j^{(k)}$ par $X_j^{(k+1)}$ quand celui-ci a déjà été calculé, c'est-à-dire quand $j < i$.

$$X_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \sum_{j<i} a_{i,j} X_j^{(k+1)} - \sum_{j>i} a_{i,j} X_j^{(k)} \right). \tag{7.2}$$

Donnons une écriture matricielle de cette récurrence. L'égalite 7.2 s'écrit matriciellement

$$X^{(k+1)} = D^{-1} \left(B - LX^{(k+1)} - UX^{(k)} \right)$$

ce qui donne

$$X^{(k+1)} = -(D+L)^{-1} UX^{(k)} + (D+L)^{-1} B$$

Donc ici, $M = D + L$ et $N = U$.

Proposition 7.4.1. *La suite $(X^{(k)})$ converge vers la solution X du système pour tout choix de $X^{(0)}$ si et seulement si*

$$\rho((D+L)^{-1}U) < 1$$

On dit alors que la méthode de Gauss-Seidel converge pour tout vecteur initial.

Corollaire 7.4.2. *Si A est à diagonale strictement dominante, la méthode de Gauss-Seidel converge pour tout vecteur initial.*

Démonstration. Soient $L' = D^{-1}L$ et $U' = D^{-1}U$. Alors

$$\begin{aligned}(D+L)^{-1}U &= (D(I+D^{-1}L))^{-1}D(D^{-1}U)\\ &= (I+D^{-1}L)^{-1}D^{-1}D(D^{-1}U)\\ &= (I+L')^{-1}U'\end{aligned}$$

On calcule le polynôme caractéristique de $(I+L')^{-1}U'$.

$$\begin{aligned}\chi_{(I+L')^{-1}U'} &= \det\left(XI-(I+L')^{-1}U'\right)\\ &= \det(I+L')\det\left(XI-(I+L')^{-1}U'\right)\end{aligned}$$

puisque $\det(I+L') = 1$. On obtient donc

$$\chi_{(I+L')^{-1}U'} = \det(X(I+L')-U')$$

Soit $\lambda \in \mathrm{Spec}((I+L')^{-1}U')$. Par ce qui précède, $\det(\lambda(I+L')-U') = 0$, c'est-à-dire

$$\begin{vmatrix} \lambda & -\dfrac{a_{1,2}}{a_{1,1}} & \dots & -\dfrac{a_{1,n}}{a_{1,1}} \\ \lambda\dfrac{a_{2,1}}{a_{2,2}} & \lambda & & -\dfrac{a_{2,n}}{a_{2,2}} \\ \vdots & \ddots & \ddots & \vdots \\ \lambda\dfrac{a_{n,1}}{a_{n,n}} & \dots & \lambda\dfrac{a_{n,n-1}}{a_{n,n}} & \lambda \end{vmatrix} = 0$$

Comme A est à diagonale strictement dominante, pour tout k,

$$\sum_{j\neq k}\left|\frac{a_{k,j}}{a_{k,k}}\right| < 1$$

Donc pour tout k

$$\sum_{j\neq k}|\lambda|\left|\frac{a_{k,j}}{a_{k,k}}\right| < |\lambda|$$

ce qui s'écrit aussi

$$\sum_{j<k} |\lambda| \left|\frac{a_{k,j}}{a_{k,k}}\right| + \sum_{j>k} |\lambda| \left|\frac{a_{k,j}}{a_{k,k}}\right| < |\lambda| \tag{7.3}$$

Supposons par l'absurde que $|\lambda| \geq 1$. En utilisant cette inégalité dans les termes de la somme de l'inégalité (7.3) correspondant aux indices où $j > k$, on obtient

$$\sum_{j<k} |\lambda| \left|\frac{a_{k,j}}{a_{k,k}}\right| + \sum_{j>k} \left|\frac{a_{k,j}}{a_{k,k}}\right| \leq \sum_{j<k} |\lambda| \left|\frac{a_{k,j}}{a_{k,k}}\right| + \sum_{j>k} |\lambda| \left|\frac{a_{k,j}}{a_{k,k}}\right| < |\lambda|$$

Cela signifie que $\lambda(I + L') - U'$ est à diagonale strictement dominante, donc son déterminant ne peut pas être nul. On en déduit que $|\lambda| < 1$. □

7.5. La relaxation

La relaxation utilise une méthode itérative comme la méthode de Jacobi ou celle de Gauss Seidel pour en définir une nouvelle. On choisit donc ici l'une de ces deux méthode que l'on appelle méthode de base.

On fixe un paramètre réel $\omega > 0$. Admettons que $X^{(k)}$ ait déjà été calculé, ainsi que X_j^{k+1} pour $j < i$. Soit alors $\overline{X_i^{(k+1)}}$ la i-ème coordonnée que l'on obtiendrait en utilisant la méthode de base choisie (Jacobi ou Gauss-Seidel). On définit alors $X_i^{(k+1)}$ par

$$X_i^{(k+1)} = \omega \overline{X_i^{(k+1)}} + (1-\omega) X_i^{(k)}.$$

7.5.1 Relaxation de Jacobi

Si la méthode de base est la méthode de Jacobi, on utilise la récurrence (7.1) et on obtient la nouvelle relation

$$X_i^{(k+1)} = \frac{\omega}{a_{i,i}} \left(b_i - \sum_{j \neq i} a_{i,j} X_j^{(k)} \right) + (1-\omega) X_i^{(k)}. \tag{7.4}$$

Voyons l'écriture matricielle. La récurrence précédente s'écrit

$$X^{(k+1)} = \omega D^{-1}(B - (L+U)X^{(k)}) + (1-\omega) X^{(k)}$$

ce qui donne

$$X^{(k+1)} = (I - \omega D^{-1} A) X^{(k)} + \omega D^{-1} B$$

Proposition 7.5.1. *La suite $(X^{(k)})$ converge vers la solution X du système pour tout choix de $X^{(0)}$ si et seulement si*

$$\rho(I - \omega D^{-1}A) < 1$$

On dit alors que la méthode de relaxation basée sur Jacobi converge pour tout vecteur initial.

Corollaire 7.5.2. *Si A est à diagonale strictement dominante et si $\omega \in]0,1[$, alors la méthode de relaxation basée sur Jacobi de paramètre ω converge pour tout vecteur initial.*

Démonstration. Exercice. □

7.5.2 Relaxation de Gauss-Seidel

Si la méthode de base est la méthode de Gauss-Seidel, on utilise la récurrence (7.2) et on obtient

$$X_i^{(k+1)} = \frac{\omega}{a_{i,i}}\left(b_i - \sum_{j<i} a_{i,j}X_j^{(k+1)} - \sum_{j>i} a_{i,j}X_j^{(k)}\right) + (1-\omega)X_i^{(k)}. \qquad (7.5)$$

Voyons en la traduction matricielle. L'égalité (7.5) donne

$$a_{i,i}X_i^{(k+1)} + \omega\sum_{j<i} a_{i,j}X_j^{(k+1)} = \omega b_i - \omega\sum_{j>i} a_{i,j}X_j^{(k)} + a_{i,i}(1-\omega)X_i^{(k)}$$

Cela s'écrit

$$(D+\omega L)X^{(k+1)} = \omega B + ((1-\omega)D - \omega U)X^{(k)}$$

soit

$$X^{(k+1)} = (D+\omega L)^{-1}((1-\omega)D - \omega U)X^{(k)} + \omega(D+\omega L)^{-1}B$$

Proposition 7.5.3. *La suite $(X^{(k)})$ converge vers la solution X du système pour tout choix de $X^{(0)}$ si et seulement si*

$$\rho\left((D+\omega L)^{-1}((1-\omega)D - \omega U)\right) < 1$$

On dit alors que la méthode de relaxation basée sur Gauss-Seidel de paramètre ω converge pour tout vecteur initial.

Corollaire 7.5.4.

1. *Pour que la méthode de relaxation basée sur Gauss-Seidel converge pour tout vecteur initial, il faut que $0 < \omega < 2$.*
2. *Si A est à diagonale strictement dominante et si $\omega \in]0,1]$, cette méthode converge pour tout vecteur initial.*

Démonstration. Exercice. □

Chapitre 8

Factorisations de matrices

8.1. Factorisation QR

Notation. On rappelle que

$$\mathcal{T}^{++}_{sup,n}(\mathbb{R}) = \{T \in \mathcal{T}_{sup,n}(\mathbb{R}) \ : \forall i \in [[1,n]] \ T_{i,i} > O\}$$

On note $\mathrm{Sym}_n^{++}(\mathbb{R})$ l'ensemble des matrices de $\mathrm{Sym}_n(\mathbb{R})$ qui sont définies positives.

Théorème 8.1.1. *Pour tout $A \in GL_n(\mathbb{R})$, il existe $(Q,R) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_{sup,n}(\mathbb{R})$ unique tel que $A = QR$.*

Démonstration. Voir les TD (on considère l'orthonormalisée de Gram-Schmidt de la base de $M_n(\mathbb{R})$ formée par les colonnes de A). □

8.2. Factorisation de Cholesky

Notation. On note $\mathrm{Sym}_n^{++}(\mathbb{R})$ l'ensemble des matrices de $\mathrm{Sym}_n(\mathbb{R})$ qui sont définies positives.

Théorème 8.2.1. *Pour tout $S \in Sym_n^{++}(\mathbb{R})$, il existe $R \in \mathcal{T}^{++}_{sup,n}(\mathbb{R})$ unique telle que $S = {}^tRR$.*

Démonstration. On considère le produit scalaire $(X|Y) = {}^tXSY$ sur $M_{n,1}(\mathbb{R})$. Soit ε_i l'orthonormalisée de Gram-Schmidt de (e_i). Soit R la matrice de passage de (ε_i) à (e_i). La matrice de Gram de (e_i) est S. Celle de (ε_i) est I, donc $S = {}^tRIR = {}^tRR$.

L'unicité est laissée en exercice. □

8.3. Méthodes de calcul

8.3.1 Orthonormalisation de Gram-Schmidt

Pour calculer la décomposition de Cholesky, on peut donc utiliser le procédé d'orthonormalisation de Gram-Schmidt comme dans la preuve ci-dessus. Cela se fait en $n^3/3 + O(n^2)$ opérations dans $\mathbb{R}$.

8.3.2 Coefficients indéterminés

On peut aussi utiliser une méthode par substitution de coefficients indéterminées.

Exemple. Soit $S = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$. On cherche $R = \begin{pmatrix} r_{1,1} & r_{1,2} \\ 0 & r_{2,2} \end{pmatrix}$ telle que $S = {}^tRR$. On calcule donc ${}^tRR = \begin{pmatrix} r_{1,1}^2 & r_{1,1}r_{1,2} \\ 0 & r_{1,2}^2 + r_{2,2}^2 \end{pmatrix}$. On obtient donc $r_{1,1} = 2$, $r_{1,2} = -1/2$ et $r_{2,2} = \sqrt{4 - r_{1,2}^2} = \frac{\sqrt{15}}{2}$. Finalement, $R = \begin{pmatrix} 2 & -1/2 \\ 0 & \sqrt{15}/2 \end{pmatrix}$.

Plus généralement, on obtient la décomposition de Cholesky de S par la proposition suivante.

Proposition 8.3.1. *Soit $S = (s_{i,j}) \in Sym_n^{++}(\mathbb{R})$ et $S = {}^tRR$ sa décomposition de Cholesky. On pose $R = (r_{i,j})$. Bien sûr, si $i > j$, $r_{i,j} = 0$. On suppose que les lignes $1, \ldots, i-1$ ont été calculées, c'est-à-dire que pour tout $(k, j) \in [[1, i-1]] \times [[1, n]]$ sont connus.*

1. $r_{i,i} = \sqrt{s_{i,i} - \sum_{k=1}^{i-1} r_{k,i}^2}$.
2. *Pour tout $j > i$,* $r_{i,j} = \frac{1}{r_{i,i}} \left(s_{i,j} - \sum_{k=1}^{i-1} r_{k,i} r_{k,j} \right)$.

Démonstration. Exercice. Utiliser : $s_{i,j} = \sum_{k=1}^{n} r_{k,i} r_{k,j} = \sum_{k=1}^{i} r_{k,i} r_{k,j}$ (pour $i \leq j$). □

Cela mène à un algorithme qui permet de calculer la décomposition de Cholesky en au plus $n^3/3 + O(n^2)$ opérations dans $\mathbb{R}$.

8.3.3 Variante

On peut chercher une décomposition de la forme $S = {}^tUDU$ où $U \in \mathcal{T}^1_{sup,n}(\mathbb{R})$ et où D est diagonale à coefficients diagonaux strictement positifs. Cela évite le calcul de racines carrées en cours de l'algorithme. De plus, cela donne parfois un résultat si S n'est pas définie positive. Dans ce cas, les coefficients diagonaux de D ne sont pas strictement positifs.

Si l'on pose $U = (u_{i,j})$ et $D = \operatorname{diag}((d_i))$, on obtient alors pour i allant de 1 à n :

1. $d_i = s_{i,i} - \displaystyle\sum_{k=1}^{i-1} d_k u_{k,i}^2$
2. pour $j > i$: $u_{i,j} = \dfrac{1}{d_i}\left(s_{i,j} - \displaystyle\sum_{k=1}^{i-1} d_k u_{k,i} u_{k,j}\right)$

Exercice. Décomposer $S = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$ sous cette forme en utilisant cette méthode. En déduire la décomposition de Cholesky $S = {}^tRR$ de S.

8.3.4 Méthode de Gauss

On considère la forme quadratique associée à S

$$q(x_1, \ldots, x_n) = \sum_{i,j} s_{i,j} x_i x_j$$

On veut trouver un changement de variable $Y = UX$ tel que

$$q(x_1, \ldots, x_n) = \sum_i d_i y_i^2$$

On va voir qu'on peut le faire pour toute forme quadratique. Celle-ci sera définie positive si et seulement si $d_i > 0$ pour tout $i \in [[1, n]]$.

Soit $D = \operatorname{diag}(d_1, \ldots, d_n)$. Alors pour tout $X \in M_{n,1}(\mathbb{R})$

$${}^tXSX = {}^tYDY = {}^tX{}^tUDUX$$

donc

$$S = {}^tUDU$$

Exemple. Soit $S = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$. Soit q la forme quadratique de $\mathbb{R}^2$ muni de sa base canonique que S définit.

$$\begin{aligned} q(x) &= 4x_1^2 - 2x_1x_2 + 4x_2^2 \\ &= 4\left((x_1 - \frac{1}{4}x_2)^2 - \frac{1}{16}x_2^2\right) + 4x_2^2 \\ &= 4\left(x_1 - \frac{1}{4}x_2\right)^2 + \frac{15}{4}x_2^2 \end{aligned}$$

Si $U = \begin{pmatrix} 1 & -\frac{1}{4} \\ 0 & 1 \end{pmatrix}$ et $D = \operatorname{diag}\left(4, \frac{15}{4}\right)$ alors $S = {}^tUDU$. Ainsi, si

$$R = \begin{pmatrix} 2 & 0 \\ 0 & \frac{\sqrt{15}}{2} \end{pmatrix} U = \begin{pmatrix} 2 & -\frac{1}{2} \\ 0 & \frac{\sqrt{15}}{2} \end{pmatrix}$$

on obtient la décomposition de Cholesky $S = {}^tRR$.

Cas général. Supposons d'abord que S est définie positive. Alors $s_{1,1} = q(e_1) > 0$. On utilise l'égalité

$$s_{1,1}x_1^2 + 2\sum_{j>1} s_{i,j}x_1x_j = s_{1,1}\left(\left(x_1 + \sum_{j>1}\frac{s_{1,j}}{s_{1,1}}x_j\right)^2 - \left(\sum_{j>1}\frac{s_{1,j}}{s_{1,1}}x_j\right)^2\right)$$

Ainsi,

$$q'(x_2, \ldots, x_n) = q(x_1, \ldots, x_n) - s_{1,1}\left(x_1 + \sum_{j>1}\frac{s_{1,j}}{s_{1,1}}x_j\right)^2$$

est une forme quadratique qui ne dépend plus de x_1. On itère le procédé.

Si S n'est pas définie positive, il se peut que $s_{1,1} = 0$. Alors la méthode ci-dessus ne fonctionne pas. S'il existe $i \geq 1$ tel que $s_{i,i} \neq 0$, la méthode ci-dessus peut s'appliquer. Sinon, et si $q \neq 0$, il existe i, j tels que $1 \leq i < j$ et $s_{i,j} \neq 0$. On utilise alors l'égalité

$$x_ix_j = \left(\frac{x_i + x_j}{2}\right)^2 - \left(\frac{x_i - x_j}{2}\right)^2$$

On fait le changement de variable

$$x'_i = \frac{x_i + x_j}{2} \quad , \quad x'_j = \frac{x_i - x_j}{2}$$

On obtient une nouvelle forme quadratique q'' où le coefficient $s''_{i,i}$ de $x_i'^2$ est non nul. La méthode ci-dessus s'applique donc à q''. On peut ensuite faire le changement de variable inverse

$$x_i = x'_i + x'_j \quad , \quad x_j = x'_i - x'_j$$

Proposition 8.3.2. *La matrice U obtenue par cet algorithme est inversible. Si S est définie positive, $U \in \mathcal{T}^1_{sup,n}(\mathbb{R})$.*

Exemple. Soit $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Soit q la forme quadratique de $\mathbb{R}^3$ qu'elle définit de façon canonique.

$$q(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$$

On fait le changement de variable

$$x'_1 = \frac{x_1 + x_2}{2} \quad , \quad x'_2 = \frac{x_1 - x_2}{2}$$

donc

$$x_1 = x'_1 + x'_2 \quad , \quad x_2 = x'_1 - x'_2$$

Alors

$$\begin{aligned} q(x_1, x_2, x_3) &= 2(x_1'^2 - x_2'^2) + 2(x'_1 + x'_2)x_3 + 2(x'_1 - x'_2)x_3 \\ &= 2x_1'^2 + 4x'_1x_3 - 2x_2'^2 \\ &= 2(x'_1 + x_3)^2 - 2x_3^2 - 2x_2'^2 \end{aligned}$$

On fait alors le changement de variables inverse.

$$\begin{aligned} q(x_1, x_2, x_3) &= 2\left(\frac{x_1 + x_2}{2} + x_3\right)^2 - 2x_3^2 - 2\left(\frac{x_1 - x_2}{2}\right)^2 \\ &= \frac{1}{2}(x_1 - x_2 + 2x_3)^2 - \frac{1}{2}(x_1 - x_2)^2 - 2x_3^2 \end{aligned}$$

On trouve

$$S = {}^tUDU$$

où $U = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $D = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -2 \end{pmatrix}$

8.4. Problème des moindres carrés

8.4.1 Motivation

Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$ inconnue, dont on connaît les valeurs en certains points $f(t_1), \dots, f(t_m)$.

On cherche à approcher f par une combinaison linéaire de fonctions $g_1, \dots, g_n$, où $n \leq m$.

Si par exemple on veut approcher f par un polynôme de degré inférieur ou égal à $n-1$, on prend $g_i = X^{i-1}$.

On cherche donc $x_1, \dots, x_n$ tels que $\sum_{j=1}^n x_j \begin{pmatrix} g_j(t_1) \\ \vdots \\ g_j(t_m) \end{pmatrix}$ soit "proche" de $\begin{pmatrix} f(t_1) \\ \vdots \\ f(t_m) \end{pmatrix}$, c'est-à-dire tels que AX soit "proche" de B, où l'on a posé $A = (g_j(t_i))$ et $B = (f(t_i))$.

8.4.2 Résolution

En général, l'équation $AX = B$ n'aura pas de solutions. Résoudre ce problème au sens des moindres carrés consiste à chercher $X \in M_n(\mathbb{R})$ tel que $||AX - B||_2$ soit minimale.

Théorème 8.4.1. $||AX - B||_2$ *est minimal si et seulement si* ${}^tAAX = {}^tAB$.

Démonstration. Voir les TD. □

On est donc ramenés à résoudre l'équation

$${}^tAAX = {}^tAB,$$

appelée équation normale du problème des moindres carrés associé à A et B.

Or, ${}^tAA \in \mathrm{Sym}_n(\mathbb{R})$. De plus, on a le résultat suivant.

Proposition 8.4.2. *Si* $rgA = n$, *alors* ${}^tAA \in Sym_n^{++}(\mathbb{R})$.

Si $\mathrm{rg}A = n$, ce qui en pratique arrive très souvent, il y a unicité de la solution au problème des moindres carrés. On peut le résoudre en trouvant $R \in \mathcal{T}_{sup,n}^{++}(\mathbb{R})$ telle que ${}^tAA = {}^tRR$, puis on résout ${}^tRR = {}^tAB$, ce qui se fait en résolvant des systèmes triangulaires.

8.4.3 Instabilité numérique

Soit pour $\varepsilon > 0$

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

On vérifie que $\mathrm{rg}(A) = 3$, mais un calcul mené avec une précision insuffisante peut conduire à un arrondi de ${}^tAA$ qui est de rang 1. La méthode utilisant l'équation normale peut donc rencontrer des problèmes d'instabilité numérique.

Il existe d'autres méthodes pour ce problème. L'une d'entre elles passe par une généralisation de la décomposition QR aux matrices rectangulaires en utilisant une méthode appelée méthode de Householder. Nous ne le ferons pas ici.