

Algèbre linéaire avancée II

Friedrich Eisenbrand

11 mars 2017

Préface

Ceci sont mes notes du cours *Algèbre Linéaire Avancée II*. La qualité de ce texte dépend fortement de la participation des étudiants. Ces sources sont gérées sur *GitHub*, une plateforme importante de collaboration. Si vous trouvez des fautes, des erreurs typographique, ou même des démonstrations plus élégantes, ou des exemples qui vous aident à comprendre la matière, je vous invite à créer une *branch* des fichiers en question, où dedans vous éditez le texte. Après, vous publiez (*publish*) cette *branch* et vos collègues peuvent discuter vos modifications. Si vous êtes satisfait avec vos modifications, vous me demandez, avec un *Pull Request*, d'accepter vos modifications et finalement, le document peut être changé. Je me réjouis en avance de votre participation.

Contributions

Des corrections et modifications ont été implémentées par :

- Orane Jecker
- Natalia Karaskova
- Dylan Samuelian
- Aziz Benmosbah
- Djian Post
- Robin Mamie
- Alfonso Cevallos
- Kévin Jorand
- Charles Dufour
- Christoph Hunkenschröder
- Adam Cierniak
- Mann-Tchi Dang
- Yasmine Bennis
- Corentin

Table des matières

| | | |
|----------|--|-----------|
| 1 | Formes bilinéaires | 5 |
| 1.1 | Orthogonalité | 7 |
| 1.2 | Matrices congruentes | 10 |
| 1.3 | Le théorème de Sylvester | 13 |
| 1.4 | Le cas réel, défini positif | 16 |
| 1.5 | La méthode des moindres carrés | 22 |
| 1.6 | Formes linéaires, bilinéaires et l'espace dual | 24 |
| 1.7 | Formes sesquilinéaires et produits hermitiens | 27 |
| 1.8 | Espaces hermitiens | 31 |
| 2 | Le théorème spectral et la décomposition en valeurs singulières | 33 |
| 2.1 | Les endomorphismes auto-adjoints | 33 |
| 2.2 | Formes quadratiques réelles et matrices symétriques réelles | 38 |
| 2.3 | Encore les systèmes d'équations | 45 |
| 2.4 | Le meilleur sous-espace approximatif | 46 |
| 3 | Systèmes différentiels linéaires | 51 |
| 3.1 | L'exponentielle d'une matrice | 55 |
| 3.2 | Polynômes | 57 |
| 3.3 | La forme normale de Jordan | 61 |

1 Formes bilinéaires

Définition 1.1. Soit V un espace vectoriel sur un corps K . Une *forme bilinéaire* sur V est une correspondance qui à tout couple (v, w) d'éléments de V associe un scalaire, noté $\langle v, w \rangle \in K$, satisfaisant aux deux propriétés suivantes :

BL 1 Si u, v et w sont des éléments de V , et $\alpha \in K$ est un scalaire,

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \quad \text{et} \quad \langle u, \alpha \cdot w \rangle = \alpha \cdot \langle u, w \rangle.$$

BL 2 Si u, v et w sont des éléments de V , et $\alpha \in K$ est un scalaire,

$$\langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle \quad \text{et} \quad \langle \alpha \cdot u, w \rangle = \alpha \cdot \langle u, w \rangle.$$

La forme bilinéaire est dite *symétrique* si pour tout $v, w \in V$

$$\langle v, w \rangle = \langle w, v \rangle.$$

On dit que la forme bilinéaire est *non dégénérée à gauche* (respectivement à droite) si la condition suivante est vérifiée :

Si $v \in V$, et si $\langle v, w \rangle = 0$ pour tout $w \in V$, alors $v = 0$

ou de manière équivalente :

Si $v \in V$, $v \neq 0$, alors il existe $w \in V$ tel que $\langle v, w \rangle \neq 0$.

Si la forme bilinéaire est non dégénérée à gauche et à droite, on dit qu'elle est *non dégénérée*.

Exemple 1.1. Soit $V = K^n$, l'application

$$\begin{aligned} \langle \rangle : V \times V &\longrightarrow K \\ (u, v) &\longmapsto \sum_{i=1}^n u_i v_i \end{aligned}$$

est une forme bilinéaire. Vérifions (BL 1). Pour tous $u, v, w \in K^n$ et $\alpha \in K$:

$$\begin{aligned} \langle u, v + w \rangle &= \sum_{i=1}^n u_i (v_i + w_i) \\ &= \sum_{i=1}^n (u_i v_i + u_i w_i) \\ &= \sum_{i=1}^n u_i v_i + \sum_{i=1}^n u_i w_i \\ &= \langle u, v \rangle + \langle u, w \rangle \end{aligned}$$

1 Formes bilinéaires

et

$$\langle u, \alpha \cdot w \rangle = \sum_{i=1}^n u_i \alpha w_i = \alpha \sum_{i=1}^n u_i w_i = \alpha \langle u, w \rangle.$$

On appelle cette forme bilinéaire la *forme bilinéaire standard* de K^n . On vérifie aussi très facilement que la forme bilinéaire standard est symétrique et non dégénérée.

Exemple 1.2. Soit V l'espace des fonctions continues à valeurs réelles, définies sur l'intervalle $[0, 2 \cdot \pi]$. Si $f, g \in V$ on pose

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x) dx.$$

Clairement, \langle, \rangle est une forme bilinéaire symétrique sur V non dégénérée.

Exercice 1.1. Montrer que les formes bilinéaires des exemples 1.1 et 1.2 sont non dégénérées.

Soit V un espace vectoriel de dimension finie et $B = \{v_1, \dots, v_n\}$ une base de V . Pour une forme bilinéaire $f : V \times V \longrightarrow K$ et $x = \sum_i \alpha_i v_i$ et $y = \sum_j \beta_j v_j$ on a

$$\begin{aligned} f(x, y) &= f\left(\sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \beta_j v_j\right) \\ &= \sum_{i=1}^n \alpha_i f\left(v_i, \sum_{j=1}^n \beta_j v_j\right) \\ &= \sum_{i,j=1}^n \alpha_i \beta_j f(v_i, v_j) \end{aligned}$$

alors pour la matrice $A_B^f \in K^{n \times n}$, ayant comme composantes $f(v_i, v_j)$, on a

$$f(x, y) = [x]_B^T A_B^f [y]_B.$$

Exercice 1.2. Soit V de dimension finie et B une base de V . Deux formes bilinéaires $f, g : V \times V \longrightarrow K$ sont différentes si et seulement si $A_B^f \neq A_B^g$.

Pour mémoire, pour deux bases B, B' et étant donné $[x]_{B'}$, on trouve les coordonnées de x dans la base B , $[x]_B$, à l'aide de la matrice de changement de base $P_{B'B}$ comme

$$[x]_B = P_{B'B} [x]_{B'}.$$

Cette formule nous montre que

$$A_{B'}^f = P_{B'B}^T A_B^f P_{B'B}. \quad (1.1)$$

Exercice 1.3. Soit V un K -espace vectoriel de dimension finie et B une base de V . Une forme bilinéaire $f : V \times V \longrightarrow K$ est symétrique si et seulement si A_B^f est symétrique.

Proposition 1.1. Soit V un K -espace vectoriel de dimension finie, $B = \{b_1, \dots, b_n\}$ une base de V et $f : V \times V \longrightarrow K$ une forme bilinéaire. Les conditions suivantes sont équivalentes.

- i) $\text{rang}(A_B^f) = n$
- ii) f est non dégénérée à gauche, i.e. si $v \in V$, et si $\langle v, w \rangle = 0$ pour tout $w \in V$, alors $v = 0$
- iii) f est non dégénérée à droite, i.e. si $v \in V$, et si $\langle w, v \rangle = 0$ pour tout $w \in V$, alors $v = 0$

Démonstration. Nous montrons **i)** et **ii)** sont équivalentes. De la même manière, on démontre aussi que **i)** et **iii)** sont équivalentes.

i) \Rightarrow ii) : Supposons que $\text{rang}(A_B^f) = n$ et soit $v \in V$, $v \neq 0$. Pour $w \in V$ on a

$$f(v, w) = [v]_B^T A_B^f [w]_B.$$

Dès que $[v]_B \neq 0$, on a $[v]_B^T A_B^f \neq 0^T$ (car $\text{noyau}(A_B^f) = \{0\} \iff \text{rang}(A_B^f) = n$). Supposons que la i -ème composante de $[v]_B^T A_B^f$ n'est pas égale à 0. Alors $[v]_B^T A_B^f e_i \neq 0$ où toutes les composantes de e_i sont 0 sauf la i -ème composante, qui est égale à 1. Alors $f(v, b_i) \neq 0$. Donc f est non dégénérée à gauche.

ii) \Rightarrow i) : Si f est non dégénérée à gauche, alors $x^T A_B^f \neq 0$ pour tout $x \in K^n$ tel que $x \neq 0$ (sinon, on aurait trouvé un x tel que $x^T A_B^f y = 0$ pour tout $y \in K^n$). Ceci implique que les lignes de A_B^f sont linéairement indépendantes. Alors $\text{rang}(A_B^f) = n$. \square

1.1 Orthogonalité

Pour ce paragraphe **1.1**, s'il n'est pas spécifié autrement, V est toujours un espace vectoriel sur K muni d'une forme bilinéaire symétrique $\langle \cdot, \cdot \rangle$.

Définition 1.2. Deux éléments $u, v \in V$ sont *orthogonaux* ou *perpendiculaires* si $\langle u, v \rangle = 0$, et l'on écrit $v \perp u$.

Proposition 1.2. Soit $E \subseteq V$ une partie de V , alors $E^\perp = \{v \in V : v \perp e \text{ pour tout } e \in E\}$ est un sous-espace vectoriel de V .

Démonstration. Pour mémoire : $\emptyset \neq W \subseteq V$ est un sous-espace si les conditions suivantes sont vérifiées.

- i) Si $u, v \in W$ on a $u + v \in W$.
- ii) Si $c \in K$ et $u \in W$ on a $c \cdot u \in W$.

1 Formes bilinéaires

Si $u, v \in E^\perp$ alors pour tout $e \in E$

$$\langle e, u + v \rangle = \langle e, u \rangle + \langle e, -v \rangle = 0 + 0 = 0, \langle e, u + v \rangle = \langle e, u \rangle + \langle e, v \rangle = 0 + 0 = 0,$$

et pour $c \in K$

$$\langle e, c \cdot v \rangle = c \langle e, v \rangle = c \cdot 0 = 0.$$

□

Exercice 1.4. Soit $E \subseteq V$ et E^* le sous-espace de V engendré par les éléments de E . Montrer $E^\perp = E^{*\perp}$.

Exemple 1.3. Soient K un corps et $(a_{ij}) \in K^{m \times n}$ une matrice à m lignes et n colonnes. Le système homogène linéaire

$$AX = 0, \tag{1.2}$$

peut s'écrire sous la forme

$$\langle A_1, X \rangle = 0, \dots, \langle A_m, X \rangle = 0,$$

où les A_i sont les vecteurs lignes de la matrice A et $\langle \cdot, \cdot \rangle$ dénote la forme bilinéaire standard de K^n . Soit W le sous-espace de K^n engendré par les A_i et U le sous-espace de K^n des solutions du système (1.2). Alors on a $U = W^\perp$ et $\dim(W^\perp) = \dim(U) = n - \text{rang}(A) = \dim(\text{noyau}(A))$.

Définition 1.3. La caractéristique d'un anneau (unitaire) R , $\text{Char}(R)$ est l'ordre de 1_R comme élément du groupe abélien $(R, +)$. En d'autres mots, c'est le nombre

$$\min_{k \in \mathbb{N}_+} \underbrace{1 + \dots + 1}_{k \text{ fois}} = 0$$

Si cet ordre est infini, la caractéristique de R est 0.

Notation. Pour $n \in \mathbb{N}_+$ l'anneau des classes des restes est dénoté comme $\mathbb{Z}/n\mathbb{Z}$ ou plus brièvement \mathbb{Z}_n (parfois aussi noté \mathbb{F}_n). Ceci est un corps si et seulement si n est un nombre premier.

Exemple 1.4. Soit $n \in \mathbb{N}_+$. Alors la caractéristique de \mathbb{Z}_n est n . La caractéristique de \mathbb{Q}, \mathbb{R} et \mathbb{C} est zéro.

Lemme 1.3. Soit $\text{Char}(K) \neq 2$. Si $\langle u, u \rangle = 0$ pour tout $u \in V$ alors

$$\langle u, v \rangle = 0 \text{ pour tous } u, v \in V$$

On dit que la forme bilinéaire symétrique $\langle \cdot, \cdot \rangle$ est nulle.

Démonstration. Soient $u, v \in V$. On peut écrire

$$2 \cdot \langle u, v \rangle = \langle u + v, u + v \rangle - \langle u, u \rangle - \langle v, v \rangle$$

et comme $2 \neq 0$ on a $\langle u, v \rangle = 0$.

□

Définition 1.4. Une base $\{v_1, \dots, v_n\}$ de l'espace vectoriel V est une *base orthogonale* si $\langle v_i, v_j \rangle = 0$ pour $i \neq j$.

Théorème 1.4. Soit $\text{Char}(K) \neq 2$ et supposons que V est de dimension finie. Alors V possède une base orthogonale.

Démonstration. On montre le théorème par induction. Si $\dim(V) = 1$ alors toute base contient seulement un élément et alors est orthogonale.

Soit $\dim(V) > 1$. Si $\langle u, u \rangle = 0$ pour tout u , le lemme 1.3 implique que la forme bilinéaire symétrique est nulle et toute base de V est orthogonale. Autrement, soit $u \in V$ tel que $\langle u, u \rangle \neq 0$ et soit $V_1 = \text{span}\{u\}$. Pour $x \in V$ le vecteur

$$x - \langle x, u \rangle / \langle u, u \rangle \cdot u \in V_1^\perp$$

et alors $V = V_1 + V_1^\perp$. Cette somme est directe parce que chaque élément de $V_1 \cap V_1^\perp$ s'écrit comme $\beta \cdot u$ pour $\beta \in K$. Et $\langle u, \beta u \rangle = \beta \langle u, u \rangle = 0$ implique $\beta = 0$.

Alors $\dim(V_1^\perp) < \dim(V)$, et par induction, V_1^\perp possède une base orthogonale $\{v_2, \dots, v_n\}$. Alors $\{v_1, \dots, v_n\}$ est une base orthogonale de V . \square

Exercices

1. Soit K un corps. Si la caractéristique de K est différente de zéro, alors elle est un nombre premier.
2. Soit K un corps fini. Montrer que $|K| = q^\ell$ pour un nombre premier q et un nombre naturel $\ell \in \mathbb{N}$. *Indication : K est un espace vectoriel de dimension finie sur \mathbb{Z}_q pour un q premier.*
3. On considère les vecteurs

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \text{ et } v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{Z}_2^4.$$

Est-ce que $\text{span}\{v_1, v_2, v_3\}$ possède une base orthogonale par rapport à la forme bilinéaire symétrique standard de l'exemple 1.1 ?

4. En considérant la forme bilinéaire symétrique standard de l'exemple 1.1, trouver une base orthogonale du sous-espace de \mathbb{Z}_3^4 engendré par

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \text{ et } v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{Z}_3^4.$$

1.2 Matrices congruentes

Pour ce paragraphe 1.2, s'il n'est pas spécifié autrement, V est toujours un espace vectoriel sur K muni d'une forme bilinéaire symétrique $\langle \cdot, \cdot \rangle$.

Définition 1.5. Deux matrices $A, B \in K^{n \times n}$ sont dites *congruentes* s'il existe une matrice $P \in K^{n \times n}$ inversible telle que

$$A = P^T B P.$$

Nous écrivons $A \cong B$.

Exemple 1.5. Si V est de dimension finie et B, B' sont deux bases, (1.1) montrer que $A_B^{(\cdot)} \cong A_{B'}^{(\cdot)}$.

Lemme 1.5. La relation \cong est une relation d'équivalence.

Démonstration. Voir exercice. □

Le relation entre \cong et le concept de l'orthogonalité est précisée dans le lemme suivant.

Lemme 1.6. Soit $B = \{v_1, \dots, v_n\}$ une base de V . V possède une base orthogonale si et seulement s'il existe une matrice diagonale D telle que $A_B^{(\cdot)} \cong D$.

Démonstration. Soit $B = \{v_1, \dots, v_n\}$ une base de V et soit $B' = \{u_1, \dots, u_n\}$ une base orthogonale. Nous avons

$$\langle u_i, u_j \rangle = [u_i]_B^T A_B^{(\cdot)} [u_j]_B \quad (1.3)$$

égal à zéro si $i \neq j$. Pour la matrice de passage $P = P_{B'B} \in K^{n \times n}$ dont les colonnes sont les vecteurs $[u_1]_B, \dots, [u_n]_B$, on a alors que

$$P^T A_B^{(\cdot)} P$$

est une matrice diagonale. Aussi P est inversible, dès que B' est une base de V .

Soit $P \in K^{n \times n}$ inversible telle que

$$P^T A P = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}.$$

La base orthogonale est $B' = \{w_1, \dots, w_n\}$ donnée par

$$[w_j]_B = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix},$$

donc $w_j = \sum_{i=1}^n p_{ij} v_i$. □

Corollaire 1.7. Soit K un corps de caractéristique différente de 2. Toute matrice symétrique $A \in K^{n \times n}$ est congruente à une matrice diagonale.

Démonstration. Ceci est un corollaire du lemme 1.4 et du théorème 1.6 parce que K^n muni de la forme bilinéaire symétrique $\langle u, v \rangle = u^T A v$ possède une base orthogonale. \square

Maintenant soit $K = \mathbb{R}$ et $A \in \mathbb{R}^{n \times n}$ symétrique. Le Corollaire 1.7 implique qu'il existe une matrice inversible $P \in \mathbb{R}^{n \times n}$ tel que $P^T A P = D$ où D est une matrice diagonale. Si on échange deux colonnes de P et note P' la nouvelle matrice obtenue, alors $P'^T A P' = D'$, où D' est obtenue de D en échangeant les éléments diagonaux correspondants. Alors on peut trouver une matrice inversible $P \in \mathbb{R}^{n \times n}$ telle que

$$P^T A P = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}. \quad (1.4)$$

où les c_i sont ordonnés de sorte que $c_1, \dots, c_r > 0$, $c_{r+1}, \dots, c_s < 0$ et $c_{s+1}, \dots, c_n = 0$. En multipliant les premières $s + r$ colonnes et lignes de P par $1/\sqrt{|c_i|}$ on obtient en fait une factorisation (1.4) telle que $c_1, \dots, c_r = 1$, $c_{r+1}, \dots, c_s = -1$ et $c_{s+1}, \dots, c_n = 0$.

Alors on trouve $P \in \mathbb{R}^{n \times n}$ inversible telle que

$$P^T A P = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}. \quad (1.5)$$

Définition 1.6. Pour un espace vectoriel sur \mathbb{R} de dimension finie, on appelle une base B de V telle que $A_B^{(\cdot, \cdot)}$ a la forme décrite en (1.5) une *base de Sylvester*.

Algorithme 1.1. Cet algorithme trouve une matrice diagonale congruente à la matrice symétrique $A \in K^{n \times n}$ où K est un corps tel que $\text{Char}(K) \neq 2$. L'algorithme procède en n itérations. Après la $(i - 1)$ -ème itération, $i \geq 1$, (aussi après la 0-ème itération) l'algorithme a transformé A en une matrice congruente

$$\begin{pmatrix} c_1 & & & & & \\ & c_2 & & & & \\ & & \ddots & & & \\ & & & c_{i-1} & & \\ & & & & b_{i,i} & \dots & b_{i,n} \\ & & & & \vdots & & \vdots \\ & & & & b_{n,i} & \dots & b_{n,n} \end{pmatrix} \quad (1.6)$$

1 Formes bilinéaires

où les composantes des premières $(i - 1)$ lignes et colonnes sont zéro sauf pour la composante sur la diagonale de la matrice.

Pour $1 \leq i \leq n$, la i -ème itération est comme suit.

- Si la i -ème ligne est zéro, l'algorithme procède avec la $(i + 1)$ -ème itération.
- Autrement, si $b_{ii} = 0$: Soit $j \in \{i + 1, \dots, n\}$ tel que $b_{ij} \neq 0$. Si $b_{jj} \neq 0$ on échange la i -ème ligne et la j -ème ligne et après la i -ème colonne et la j -ème colonne. Autrement on additionne la ligne j sur la ligne i et on additionne la colonne j sur la colonne i .
- Étant donné que $\text{Char}(K) \neq 2$, on a alors maintenant $b_{ii} \neq 0$.
- Pour chaque $j \in \{i + 1, \dots, n\}$: on additionne $-b_{ij}/b_{ii}$ fois la i -ème ligne sur la j -ème ligne et on additionne $-b_{ij}/b_{ii}$ fois la i -ème colonne sur la j -ème colonne.

Exemple 1.6. Soit V un espace vectoriel sur \mathbb{Q} de dimension 3 muni d'une forme bilinéaire symétrique. Soit $B = \{v_1, v_2, v_3\}$ une base de V et

$$A_B^{\langle \cdot \rangle} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 4 \\ 2 & 4 & 0 \end{pmatrix}$$

Le but est de trouver une base orthogonale de V .

En utilisant notre algorithme on trouve

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

tel que

$$P^T \cdot A_B^{\langle \cdot \rangle} \cdot P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -\frac{28}{3} \end{pmatrix}.$$

Alors $B' = \{v_1, v_2, -2v_1 - (4/3)v_2 + v_3\}$ est une base orthogonale de V .

Exercices

1. Montrer que \cong est une relation d'équivalence sur l'ensemble des matrices $K^{n \times n}$.
2. Est-ce que la matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{Z}_2^{3 \times 3}$$

est congruente à une matrice diagonale? *Renseignement : voir l'exercice 3. de la section 1.1.*

3. Soit V un espace vectoriel sur un corps K de dimension finie muni d'une forme bilinéaire symétrique $\langle \cdot \rangle$. Soit $B = \{v_1, \dots, v_n\}$ une base de V . Montrer que $A_B^{\langle \cdot \rangle} \in K^{n \times n}$ est congruente à une matrice diagonale si et seulement si V possède une base orthogonale.

4. Soit K un corps de caractéristique 2 et soit V un espace vectoriel sur K de dimension finie muni d'une forme bilinéaire symétrique. Soit

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

- a) Soit $\dim(V) = 2$. Montrer que V ne possède pas de base orthogonale si et seulement s'il existe une base B de V telle que $A_B^{(\cdot)} = C$.
 b) Soit $\dim(V) = n$. Montrer que V ne possède pas de base orthogonale si et seulement s'il existe une base B de V telle que

$$A_B^{(\cdot)} = \begin{pmatrix} d_1 & & & & & \\ & d_2 & & & & \\ & & \ddots & & & \\ & & & d_k & & \\ & & & & C & \\ & & & & & \ddots \\ & & & & & & C \end{pmatrix}$$

et $d_1, \dots, d_k = 0$, et le nombre de C n'est pas égal à zéro.

5. Modifier l'algorithme 1.1 tel qu'il soit aussi correct pour des corps de caractéristique 2. Soit l'algorithme découvre que la matrice symétrique $A \in K^{n \times n}$ n'est pas congruente à une matrice diagonale, soit l'algorithme calcule une matrice diagonale congruente à A .
 6. Comment peut-on déterminer si un espace vectoriel de dimension finie muni d'une forme bilinéaire symétrique possède une base orthogonale ? Décrire très brièvement une méthode.
 7. Déterminer l'indice de nullité et l'indice de positivité des formes bilinéaires symétriques définies par les matrices suivantes

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

8. Soit V un espace euclidien de dimension n . Montrer que V possède une base B telle que pour tout $x, y \in V$

$$\langle x, y \rangle = [x]_B \cdot [y]_B,$$

où $[x]_B \cdot [y]_B$ dénote la forme bilinéaire standard de \mathbb{R}^n entre $[x]_B$ et $[y]_B$.

1.3 Le théorème de Sylvester

Soit V un espace vectoriel de dimension finie sur un corps K muni d'une forme bilinéaire symétrique. Nous avons vu (théorème 1.4) que V possède une base orthogonale.

1 Formes bilinéaires

Supposons que cette base est $B = \{v_1, \dots, v_n\}$ et considérons $x = \sum_i \alpha_i v_i \in V$ et $y = \sum_i \beta_i v_i \in V$. La forme bilinéaire s'écrit

$$\begin{aligned} \langle x, y \rangle &= \sum_{i,j} \alpha_i \beta_j \langle v_i, v_j \rangle \\ &= \sum_i \alpha_i \beta_i \langle v_i, v_i \rangle \\ &= [x]_B^T \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix} [y]_B \end{aligned}$$

où $c_i = \langle v_i, v_i \rangle$ pour tout i . Si $K = \mathbb{R}$ on peut ordonner la base afin d'avoir $c_1, \dots, c_r > 0$, $c_{r+1}, \dots, c_s < 0$ et $c_{s+1}, \dots, c_n = 0$. Nous allons maintenant démontrer, que les nombres r et s sont invariants par rapport à la base B de V .

Définition 1.7. Le sous espace $V_0 = \{v \in V : \langle v, x \rangle = 0 \text{ pour tout } x \in V\}$ est appelé l'espace de nullité de la forme bilinéaire symétrique $\langle \cdot \rangle$.

Théorème 1.8. Soit V un espace vectoriel de dimension finie sur un corps K de caractéristique $\neq 2$ et soit V muni d'une forme bilinéaire symétrique. Soit $B = \{v_1, \dots, v_n\}$ une base orthogonale de V . La dimension $\dim(V_0)$ est égale au nombre d'index i tel que $\langle v_i, v_i \rangle = 0$.

Démonstration. Nous utilisons la notation d'en-dessus et écrivons

$$\langle v, x \rangle = [v]_B^T \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix} [x]_B.$$

Cette expression est égale à zéro pour tout $x \in V$ si et seulement si $([v]_B)_i = 0$ pour tout i tel que $c_i \neq 0$. Ceci démontre que $\{v_i : \langle v_i, v_i \rangle = 0\}$ est une base de l'espace de nullité. \square

Théorème 1.9 (Théorème de Sylvester). Soit V un espace vectoriel de dimension finie sur \mathbb{R} muni d'une forme bilinéaire symétrique. Il existe un nombre entier $r \geq 0$ tel que, pour chaque base orthogonale $B = \{v_1, \dots, v_n\}$ de V , exactement r des indices i satisfont $\langle v_i, v_i \rangle > 0$.

Démonstration. Soient $\{v_1, \dots, v_n\}$ et $\{w_1, \dots, w_n\}$ des bases orthogonales de V ordonnées telles que $\langle v_i, v_i \rangle > 0$ si $1 \leq i \leq r$, $\langle v_i, v_i \rangle < 0$ si $r+1 \leq i \leq s$ et $\langle v_i, v_i \rangle = 0$ si $s+1 \leq i \leq n$. De même $\langle w_i, w_i \rangle > 0$ si $1 \leq i \leq r'$, $\langle w_i, w_i \rangle < 0$ si $r'+1 \leq i \leq s'$ et $\langle w_i, w_i \rangle = 0$ si $s'+1 \leq i \leq n$.

On démontre que $v_1, \dots, v_r, w_{r'+1}, \dots, w_n$ est linéairement indépendant. Ça implique que $r+n-r' \leq n$ et alors $r \leq r'$. Parce que l'argument est symétrique on peut conclure que $r = r'$.

1.3 Le théorème de Sylvester

Si $v_1, \dots, v_r, w_{r'+1}, \dots, w_n$ est linéairement dépendant, il existe des scalaires x_1, \dots, x_r et $y_{r'+1}, \dots, y_n$ respectivement non tous égaux à zéro tels que

$$x_1 v_1 + \dots + x_r v_r = y_{r'+1} w_{r'+1} + \dots + y_n w_n$$

et ça implique, dès que les v_i et respectivement les w_i sont orthogonaux,

$$x_1^2 \langle v_1, v_1 \rangle + \dots + x_r^2 \langle v_r, v_r \rangle = y_{r'+1}^2 \langle w_{r'+1}, w_{r'+1} \rangle + \dots + y_n^2 \langle w_n, w_n \rangle$$

Les $\langle v_i, v_i \rangle$ à gauche sont strictement positifs. Les $\langle w_i, w_i \rangle$ à droite sont non positifs. Alors $x_1 = 0, \dots, x_r = 0$ et dès que les w_i sont linéairement indépendants, on a $y_{r'+1} = 0, \dots, y_n = 0$. \square

Définition 1.8. L'entier r du théorème de Sylvester est appelé l'*indice de positivité* de la forme bilinéaire symétrique.

Exemple 1.7. Trouver une base de Sylvester et les indices de nullité et de positivité de la matrice

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

On utilise des transformations élémentaires sur les colonnes et les mêmes sur les lignes tour à tour en alternant.

Les transformations élémentaires sur les colonnes sont représentées par

$$P_1 = \begin{bmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

et transforment la matrice A en

$$\begin{pmatrix} 2 & 4 & 6 \\ 4 & 4 & 3 \\ 6 & 3 & 1 \end{pmatrix} \cdot \begin{bmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 4 & -4 & -9 \\ 6 & -9 & -17 \end{bmatrix}.$$

Alors

$$P_1^T \cdot A \cdot P = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & -9 \\ 0 & -9 & -17 \end{bmatrix}$$

Avec

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{9}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

on obtient

$$P_2^T P_1^T A P_1 P_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \frac{13}{4} \end{bmatrix}$$

1 Formes bilinéaires

L'indice de nullité est zéro et l'indice de positivité est 2. Le produit $P_1 \cdot P_2$ est égal à

$$P_1 \cdot P_2 = \begin{bmatrix} 1 & -2 & \frac{3}{2} \\ 0 & 1 & -\frac{9}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

Les colonnes sont une base de Sylvester. En divisant chaque colonne et chaque ligne par $\sqrt{2}$, $\sqrt{4}$ et $\sqrt{13/4}$ respectivement, et en échangeant les deux dernières colonnes et deux dernières lignes, on obtient une transformation P telle que $P^T A P = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$.

Exercices

1. Soit V un espace vectoriel de dimension finie sur \mathbb{R} et soit $\langle \cdot \rangle$ une forme bilinéaire sur V . Montrer que V admet une décomposition en somme directe

$$V_0 \oplus V^+ \oplus V^-$$

où V_0 est l'espace de nullité et V^+ et V^- sont des sous-espaces tels que

$$\langle v, v \rangle > 0 \text{ pour tout } v \in V^+ \setminus \{0\}$$

et

$$\langle v, v \rangle < 0 \text{ pour tout } v \in V^- \setminus \{0\}.$$

1.4 Le cas réel, défini positif

Définition 1.9. Soit V un espace vectoriel sur \mathbb{R} muni d'une forme bilinéaire symétrique. La forme bilinéaire symétrique est définie positive si $\langle v, v \rangle \geq 0$ pour tout $v \in V$, et si $\langle v, v \rangle > 0$ lorsque $v \neq 0$. Une forme bilinéaire symétrique définie positive est un *produit scalaire*.

Exemple 1.8. Soit $V = \mathbb{R}^n$. La forme bilinéaire symétrique

$$\langle u, v \rangle = \sum_{i=1}^n u_i v_i$$

est un produit scalaire, appelé le *produit scalaire ordinaire*. Aussi, la forme bilinéaire de l'exemple 1.2 est un produit scalaire.

Définition 1.10. Soit $\langle \cdot \rangle$ un produit scalaire. La *longueur* ou la *norme* d'un élément $v \in V$ est le nombre

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Un élément $v \in V$ est un *vecteur unitaire* si $\|v\| = 1$.

Pour le reste de ce paragraphe 1.4, s'il n'est pas spécifié autrement, V est toujours un espace vectoriel sur \mathbb{R} muni d'un produit scalaire.

Proposition 1.10. *Pour $v \in V$ et $\alpha \in \mathbb{R}$ on a*

$$\|\alpha v\| = |\alpha| \|v\|.$$

Démonstration.

$$\begin{aligned} \|\alpha v\| &= \sqrt{\langle \alpha v, \alpha v \rangle} \\ &= \sqrt{\alpha^2 \langle v, v \rangle} \\ &= |\alpha| \|v\|. \end{aligned}$$

□

Proposition 1.11 (Théorème de Pythagore). *Si v et w sont perpendiculaires*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Démonstration.

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v + w \rangle + \langle w, v + w \rangle \\ &= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \|w\|^2 \end{aligned}$$

□

Proposition 1.12 (Règle du parallélogramme). *Pour tous v et w , on a*

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Soit V un espace vectoriel sur un corps K muni d'une forme bilinéaire symétrique \langle, \rangle . Si w est un élément de V tel que $\langle w, w \rangle \neq 0$, pour tout $v \in V$, il existe un élément unique $\alpha \in K$ tel que $\langle w, v - \alpha w \rangle = 0$.

En fait,

$$\langle w, v - \alpha w \rangle = \langle w, v \rangle - \alpha \langle w, w \rangle.$$

Alors $\langle w, v - \alpha w \rangle = 0$ si et seulement si $\alpha = \langle v, w \rangle / \langle w, w \rangle$.

Définition 1.11. Soit V un espace vectoriel sur un corps K , muni d'un produit scalaire. Soit $w \in V \setminus \{0\}$ tel que $\langle w, w \rangle > 0$. Pour $v \in V$, soit $\alpha = \langle v, w \rangle / \langle w, w \rangle$. Le nombre α est la *composante* de v sur w , ou le *coefficient de Fourier* de v relativement à w . Le vecteur αw s'appelle la *projection* de v sur w .

1 Formes bilinéaires

Exemple 1.9. Soit V l'espace vectoriel de l'exemple 1.2 et $f(x) = \sin kx$, où $k \in \mathbb{N}_{>0}$. Alors

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_0^{2\pi} \sin^2 kx \, dx} = \sqrt{\pi}$$

Si g est une fonction quelconque, continue sur $[0, 2\pi]$, le coefficient de Fourier de g relativement à f est

$$\langle f, g \rangle / \langle f, f \rangle = \frac{1}{\pi} \int_0^{2\pi} g(x) \sin kx \, dx.$$

Théorème 1.13 (Inégalité de Cauchy-Schwarz). *Pour tous $v, w \in V$, on a*

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

Démonstration. Si $w = 0$, les deux termes de cette inégalité sont nuls et elle devient évidente. Supposons maintenant que w est un vecteur unitaire. Si $\alpha = \langle v, w \rangle$ est la composante de v sur w , $v - \alpha w$ est perpendiculaire à w , donc aussi à αw . D'après le théorème de Pythagore, on trouve

$$\begin{aligned} \|v\|^2 &= \|v - \alpha w\|^2 + \|\alpha w\|^2 \\ &= \|v - \alpha w\|^2 + \alpha^2, \end{aligned}$$

par conséquent $\alpha^2 \leq \|v\|^2$, si bien que $|\alpha| \leq \|v\|$.

Enfin, si $w \neq 0$, alors $w/\|w\|$ est unitaire. Donc par ce que nous venons de voir,

$$|\langle v, w/\|w\| \rangle| \leq \|v\|.$$

Cela implique

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

□

On appelle un espace vectoriel sur \mathbb{R} muni d'un produit scalaire un *espace euclidien*.

Théorème 1.14 (Inégalité triangulaire). *Si $v, w \in V$.*

$$\|v + w\| \leq \|v\| + \|w\|.$$

Démonstration.

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2, \end{aligned}$$

en recourant à l'inégalité de Cauchy-Schwarz.

□

Lemme 1.15. Soit V un espace vectoriel sur K muni d'un produit scalaire et soient v_1, \dots, v_n des éléments de V , deux à deux orthogonaux, tels que $\langle v_i, v_i \rangle \neq 0$ pour tout i . Soit $\alpha_i = \langle v, v_i \rangle / \langle v_i, v_i \rangle$ la composante de v sur v_i , alors le vecteur

$$v - \alpha_1 v_1 - \dots - \alpha_n v_n \text{ où } \alpha_i \in K$$

est perpendiculaire à tous v_1, \dots, v_n si et seulement si α_i est la composante de v sur v_i , i.e. $\alpha_i = \langle v, v_i \rangle / \langle v_i, v_i \rangle$ pour tout i .

Démonstration. Pour le vérifier, il suffit d'en faire le produit scalaire avec v_j pour tout j . Tous les termes $\langle v_i, v_j \rangle$ donnent zéro si $i \neq j$. Le reste

$$\langle v, v_j \rangle - \alpha_j \langle v_j, v_j \rangle$$

s'annule si et seulement si $\alpha_j = \langle v, v_j \rangle / \langle v_j, v_j \rangle$. \square

Notation. Soient V un espace vectoriel et $v_1, \dots, v_n \in V$. Le sous-espace engendré par v_1, \dots, v_n est dénoté par $\text{span}\{v_1, \dots, v_n\}$.

Théorème 1.16 (Le procédé d'orthogonalisation de Gram-Schmidt). Soient V un espace euclidien et $\{v_1, \dots, v_n\} \subseteq V$ un ensemble libre. Il existe un ensemble libre orthogonal $\{u_1, \dots, u_n\}$ de V tel que pour tout i , $\{v_1, \dots, v_i\}$ et $\{u_1, \dots, u_i\}$ engendrent le même sous-espace de V .

Démonstration. On montre le théorème par induction. On met $u_1 = v_1$ et on suppose qu'on a construit $\{u_1, \dots, u_{i-1}\}$ pour $i \geq 2$. L'ensemble $\{u_1, \dots, u_{i-1}, v_i\}$ est libre et une base du sous-espace engendré par $\{v_1, \dots, v_i\}$. On met

$$u_i = v_i - \alpha_{1,i} u_1 - \dots - \alpha_{i-1,i} u_{i-1}$$

où les $\alpha_{j,i}$ sont les composantes de v_i sur u_j . Comme ça

$$\begin{aligned} \text{span}\{u_1, \dots, u_i\} &= \text{span}\{u_1, \dots, u_{i-1}, v_i\} \\ &= \text{span}\{v_1, \dots, v_i\}. \end{aligned}$$

Surtout $\{u_1, \dots, u_i\}$ est un ensemble orthogonal. \square

Exercice 1.5. Est-ce qu'il faut vraiment supposer que le produit scalaire $\langle . \rangle$ soit réel et défini positif et sur \mathbb{R} pour ce procédé? Peux-tu imaginer une condition plus faible et satisfaite par le produit scalaire qui permet le procédé de Gram-Schmidt?

Définition 1.12. Une base $\{u_1, \dots, u_n\}$ d'un espace euclidien est *orthonormale* si elle est orthogonale et se compose de vecteurs tous unitaires.

Corollaire 1.17. Soit V un espace euclidien de dimension finie. Supposons $V \neq \{0\}$. V possède alors une base orthonormale.

Démonstration. Soient $\{v_1, \dots, v_n\}$ une base de V et $\{u_1, \dots, u_n\}$ le résultat du procédé Gram-Schmidt appliqué à $\{v_1, \dots, v_n\}$. Alors $\{u_1/\|u_1\|, \dots, u_n/\|u_n\|\}$ est une base orthonormale de V . \square

Exemple 1.10. Trouver une base orthonormale de l'espace vectoriel engendré par

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

Notons A, B et C les vecteurs. Soit $A' = A$ et

$$B' = B - \frac{A' \cdot B}{A' \cdot A'} \cdot A'$$

On trouve

$$B' = \frac{1}{3} \begin{pmatrix} 4 \\ -5 \\ 0 \\ 1 \end{pmatrix}$$

On calcule

$$C' = C - \frac{A' \cdot C}{A' \cdot A'} \cdot A' - \frac{B' \cdot C}{B' \cdot B'} \cdot B'$$

et on trouve

$$C' = \frac{1}{7} \begin{pmatrix} -4 \\ -2 \\ -1 \\ 6 \end{pmatrix}$$

La base orthonormale est

$$A'/\|A'\| = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, B'/\|B'\| = \frac{1}{\sqrt{42}} \begin{pmatrix} 4 \\ -5 \\ 0 \\ 1 \end{pmatrix} \text{ et } C'/\|C'\| = \frac{1}{\sqrt{57}} \begin{pmatrix} -4 \\ -2 \\ -1 \\ 6 \end{pmatrix}$$

Corollaire 1.18. Soit $A \in \mathbb{R}^{m \times n}$ une matrice de rang (colonne) plein. On peut factoriser A comme

$$A = A' \cdot R$$

où les colonnes de $A' \in \mathbb{R}^{m \times n}$ sont deux à deux orthonormales et $R \in \mathbb{R}^{n \times n}$ est une matrice triangulaire supérieure dont les valeurs diagonales sont positives.

Exemple 1.11. Trouver une factorisation Q, R du Corollaire 1.18 de la matrice

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

On trouve

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & \frac{4}{3} & -\frac{4}{7} \\ 1 & -\frac{5}{3} & -\frac{2}{7} \\ 0 & 0 & -1 \\ 1 & \frac{1}{3} & \frac{6}{7} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & 1 \\ 0 & 1 & \frac{3}{7} \\ 0 & 0 & 1 \end{bmatrix}$$

et alors

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2\sqrt{42}}{21} & -\frac{4}{\sqrt{105}} \\ \frac{1}{\sqrt{3}} & -\frac{5}{\sqrt{42}} & -\frac{2}{\sqrt{105}} \\ 0 & 0 & -\frac{\sqrt{105}}{15} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{42}} & \frac{2\sqrt{105}}{35} \end{bmatrix} \begin{bmatrix} \sqrt{3} & -\frac{\sqrt{3}}{3} & \sqrt{3} \\ 0 & \frac{\sqrt{42}}{3} & \frac{\sqrt{42}}{7} \\ 0 & 0 & \frac{\sqrt{105}}{7} \end{bmatrix}$$

Théorème 1.19 (Inégalité de Bessel). *Si v_1, \dots, v_n sont des vecteurs unitaires deux à deux orthogonaux et si $\alpha_i = \langle v, v_i \rangle / \langle v_i, v_i \rangle$ sont les coefficients de Fourier de v relativement à v_i alors*

$$\sum_{i=1}^n \alpha_i^2 \leq \|v\|^2.$$

Démonstration.

$$\begin{aligned} 0 &\leq \left\langle v - \sum_{i=1}^n \alpha_i v_i, v - \sum_{i=1}^n \alpha_i v_i \right\rangle \\ &= \langle v, v \rangle - 2 \cdot \sum \alpha_i \langle v, v_i \rangle + \sum \alpha_i^2 \\ &= \langle v, v \rangle - \sum \alpha_i^2 \end{aligned}$$

□

Exercices

1. Soit V un espace vectoriel muni d'un produit scalaire $\langle \cdot \rangle$. Soit $U \subseteq V$ un sous-espace. Montrer que $\langle \cdot \rangle$ restreint à U est un produit scalaire du sous-espace U .
2. Soient V un espace vectoriel muni d'un produit scalaire $\langle \cdot \rangle$ et $\{v_1, \dots, v_n\} \subseteq V$ un ensemble de vecteurs deux à deux orthogonaux.
 - a) Montrer que $\{v_1, \dots, v_n\}$ est un ensemble libre si pour tout i , $\langle v_i, v_i \rangle \neq 0$.
 - b) Donner un contre-exemple ou une démonstration de la réciproque.
3. Considérant l'exemple 1.2, montrer que l'ensemble

$$\{1, \sin x, \cos x, \sin(2x), \cos(2x), \sin(3x), \cos(3x), \dots\}$$

est un ensemble de vecteurs deux à deux orthogonaux.

4. Trouver la factorisation $Q \cdot R$ du Corollaire 1.18 de la matrice

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

Trouver la factorisation de la matrice $n \times n$

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

5. Trouver un produit scalaire de \mathbb{R}^n tel qu'il existe des vecteurs $u, v \in \mathbb{R}^n$ avec $\langle u, u \rangle < 0$ et $\langle v, v \rangle > 0$.
6. Soit V un espace vectoriel sur \mathbb{R} muni d'un produit scalaire. S'il existe des vecteurs $u, v \in V$ tels que $\langle u, u \rangle < 0$ et $\langle v, v \rangle > 0$, il existe un vecteur $w \neq 0$ tel que $\langle w, w \rangle = 0$.
7. Montrer que l'inégalité de Bessel (Théorème 1.19) est une égalité si v est dans le sous-espace engendré par les v_1, \dots, v_n .
8. On considère l'espace euclidien des fonctions continues sur l'intervalle $[0, 1]$ muni du produit scalaire

$$\langle f, g \rangle = \int_0^1 f(x)g(x) dx.$$

- i) Soit V le sous-espace engendré par $f(x) = x$ et $g(x) = x^2$. Trouver une base orthonormale de V .
- ii) Soit V le sous-espace engendré par $\{1, x, x^2\}$. Trouver une base orthonormale de V .
9. Soient V un espace euclidien, $\{u_1, \dots, u_n\}$ un ensemble orthonormal et $f, g \in \text{span}\{u_1, \dots, u_n\}$. Montrer l'identité de Parseval

$$\langle f, g \rangle = \sum_i \langle f, u_i \rangle \langle u_i, g \rangle.$$

1.5 La méthode des moindres carrées

Soient $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$ et supposons que le système des équations linéaires

$$Ax = b \tag{1.7}$$

n'a pas de solution. Dans beaucoup d'applications, on cherche un $x \in \mathbb{R}^n$ tel que la distance entre Ax et b est *minimale*. On aimerait alors résoudre le problème d'optimisation suivant

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|. \tag{1.8}$$

Pour le reste de ce paragraphe 1.5, s'il n'est pas spécifié autrement, V est toujours un espace euclidien.

Théorème 1.20. Soient v_1, \dots, v_n des vecteurs deux à deux orthogonaux et tels que $\|v_i\| > 0$ pour tout i . Soit v un élément de V et soit $\alpha_i = \langle v, v_i \rangle / \langle v_i, v_i \rangle$ la composante de v sur v_i . Pour $a_1, \dots, a_n \in \mathbb{R}$ alors

$$\left\| v - \sum_{i=1}^n \alpha_i v_i \right\| \leq \left\| v - \sum_{i=1}^n a_i v_i \right\|.$$

De plus, l'inégalité au-dessus est une égalité si et seulement si $a_i = \alpha_i$ pour tout i . Alors $\sum_{i=1}^n \alpha_i v_i$ est l'unique meilleure approximation de v par un vecteur du sous-espace engendré par les v_1, \dots, v_n .

Démonstration.

$$\begin{aligned} \left\| v - \sum_{i=1}^n a_i v_i \right\|^2 &= \left\| v - \sum_{i=1}^n \alpha_i v_i - \sum_{i=1}^n (a_i - \alpha_i) v_i \right\|^2 \\ &= \left\| v - \sum_{i=1}^n \alpha_i v_i \right\|^2 + \left\| \sum_{i=1}^n (a_i - \alpha_i) v_i \right\|^2 \end{aligned}$$

en utilisant le lemme 1.15 et le théorème de Pythagore. □

Maintenant nous pouvons décrire un *algorithme* pour résoudre le problème suivant.

Soient $v, v_1, \dots, v_n \in V$, trouver $u \in \text{span}\{v_1, \dots, v_n\}$ tel que la distance

$$\|v - u\|$$

est minimale.

Algorithme 1.2.

- i) Trouver une base orthonormale u_1, \dots, u_k du sous-espace $\text{span}\{v_1, \dots, v_n\}$ avec le procédé de Gram-Schmidt.
- ii) Retourner $u = \sum_{i=1}^k \langle v, u_i \rangle u_i$.

Théorème 1.21. Soient $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$. Les solutions du système

$$A^T A x = A^T b \tag{1.9}$$

sont les solutions optimales du problème (1.8).

1 Formes bilinéaires

Démonstration. Soit $\{a_1^*, \dots, a_k^*\}$ une base orthonormale du sous-espace $\text{Col}(A) = \{Ax : x \in \mathbb{R}^n\}$ engendré par les colonnes de A . Le théorème 1.20 implique que les solutions du problème (1.8) sont les solutions du système

$$Ax = \sum_i \langle b, a_i^* \rangle \cdot a_i^*.$$

Le lemme 1.15 implique que

$$b - \sum_i \langle b, a_i^* \rangle \cdot a_i^*$$

est perpendiculaire à tout a_i^* et dès que les a_i^* engendrent $\text{Col}(A)$ on a

$$A^T(Ax - b) = 0$$

pour toute solution optimale x de (1.8).

Maintenant, soit x une solution du système (1.9). Alors $Ax - b$ est perpendiculaire à tous les a_i^* . Le seul vecteur $v \in \text{span}\{a_1^*, \dots, a_k^*\} = \text{Col}(A)$ tel que $\langle b - v, v \rangle = 0$ est $v = \sum_i \langle a_i^*, b \rangle \cdot a_i^*$. Ceci démontre le théorème. \square

Exemple 1.12. Trouver une solution de moindres carrés sur les données

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \text{ et } b = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix} \text{ et } A^T b = \begin{pmatrix} 19 \\ 11 \end{pmatrix}.$$

La solution du système

$$\begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}.$$

est $x^* = (1, 2)^T$.

1.6 Formes linéaires, bilinéaires et l'espace dual

Soient V un espace vectoriel sur un corps K et V^* l'ensemble des applications linéaires de V dans K , où on considère K comme espace vectoriel de dimension 1 sur lui-même. Clairement, V^* est un espace vectoriel lui-même.

Définition 1.13. L'ensemble des applications linéaires de V dans K est noté V^* et, muni de l'addition et de la multiplication scalaire usuelles, est appelé l'*espace dual* de V . Les éléments de V^* sont appelés *formes linéaires*.

Exemple 1.13. Si $\langle \cdot, \cdot \rangle$ est un produit scalaire de V et $v \in V$, alors $f(x) = \langle x, v \rangle \in V^*$ est une forme linéaire.

Si V est de dimension finie et si $B = \{v_1, \dots, v_n\}$ est une base de V , l'image d'un vecteur $x = \sum_i \alpha_i v_i$ par une forme linéaire f est

$$\begin{aligned} f(x) &= f\left(\sum_i \alpha_i v_i\right) \\ &= \sum_i \alpha_i f(v_i) \\ &= (f(v_1), \dots, f(v_n))[x]_B, \end{aligned}$$

où $[x]_B = (\alpha_1, \dots, \alpha_n)^T$ sont les coordonnées de x dans la base B .

Lemme 1.22. *Supposons que V est de dimension finie et $\{v_1, \dots, v_n\}$ est une base de V . La fonction $\phi_j: V \rightarrow K$*

$$\phi_j\left(\sum_i \alpha_i v_i\right) = \alpha_j$$

est une forme linéaire.

Démonstration. Immédiate. □

Théorème 1.23. *Les formes linéaires $\{\phi_j \in V^*: j = 1, \dots, n\}$ du lemme 1.22 précédent forment une base de V^* .*

Démonstration. Si, pour $\beta_i \in K$, on a $\sum_i \beta_i \phi_i = 0$, alors

$$0 = \left(\sum_i \beta_i \phi_i\right)(v_j) = \beta_j,$$

c'est-à-dire les ϕ_j sont linéairement indépendantes. Les ϕ_j engendrent V^* puisque pour $f \in V^*$, $f = \sum_i f(v_i) \phi_i$. □

Définition 1.14. La base $\{\phi_1, \dots, \phi_n\}$ est la *base duale* de la base $\{v_1, \dots, v_n\}$.

Définition 1.15. Soit V un espace vectoriel sur un corps K . Une application

$$f: V \times V \rightarrow K$$

est une forme bilinéaire si pour tout $v \in V$, les applications $g, h: V \rightarrow K$ telles que $g(x) = f(v, x)$ et $h(x) = f(x, v)$ sont des formes linéaires.

Exemple 1.14. Un produit scalaire $\langle \cdot, \cdot \rangle$ est une forme bilinéaire symétrique.

Lemme 1.24. *Soit V un espace vectoriel sur un corps K de dimension finie muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Le produit scalaire est non dégénéré si et seulement si la matrice $A_B^{\langle \cdot, \cdot \rangle}$ est de rang plein pour toute base B .*

1 Formes bilinéaires

Démonstration. Soit n la dimension de V . Si le produit scalaire est dégénéré, alors il existe $0 \neq v \in V$ tel que $\langle v, u \rangle = 0$ pour tout $u \in V$. Alors

$$[v]_B^T A_B^\diamond x = 0 \text{ pour tout } x \in K^n,$$

alors $[v]_B^T A_B^\diamond = 0$ qui démontre $\text{rang}(A_B^\diamond) < n$.

Si $\text{rang}(A) < n$ alors il existe $0 \neq v \in V$ tel que $[v]_B^T A_B^\diamond = 0$. Alors, $\langle v, u \rangle = 0$ pour tout $u \in V$ qui démontre que $\langle . \rangle$ est dégénéré. □

Lemme 1.25. *Soit V un espace vectoriel sur le corps K de dimension finie muni d'un produit scalaire non dégénéré et soit $f: V \rightarrow K$ une forme linéaire. Il existe un $v \in V$ tel que $f(x) = \langle v, x \rangle$ pour tout $x \in V$.*

Démonstration. Soient $B = \{v_1, \dots, v_n\}$ une base de V et $A_B^\diamond \in K^{n \times n}$ la matrice associée au produit scalaire et à la base B . Soit $a \in K^n$ tel que $f(x) = a^T [x]_B$ pour tout $x \in V$. Dès que A_B^\diamond est de rang plein, alors il existe $v \in V$ tel que $[v]_B^T A_B^\diamond = a^T$. Ceci revient à résoudre un système d'équations linéaires (cf semestre 1) et comme la matrice A_B^\diamond est de rang plein, on a l'existence (et même l'unicité) d'une solution, i.e., $[v]_B^T A_B^\diamond = a^T$. Ainsi $f(x) = \langle v, x \rangle$ pour tout $x \in V$. □

Théorème 1.26 (Supplémentaire orthogonal). *Soient V un espace vectoriel de dimension finie sur corps K et W un sous-espace de V . Soit $\langle . \rangle$ un produit scalaire tel que, si restreint sur $W \times W$, il est non dégénéré. Alors $V = W \oplus W^\perp$.*

Démonstration. Pour un élément $u \in W \cap W^\perp$ on a $\langle u, w \rangle = 0$ pour tout $w \in W$. Dès que $\langle . \rangle$ est non dégénéré sur W on a $u = 0$, alors $W \cap W^\perp = \{0\}$.

Il reste à démontrer que $V = W + W^\perp$. Pour $v \in V$ le lemme 1.25 implique qu'il existe un $w_0 \in W$ tel que pour tout $u \in W$, $\langle u, v \rangle = \langle u, w_0 \rangle$ et ça démontre $v - w_0 \in W^\perp$ et alors $v \in W + W^\perp$. □

Exercices

1. Soient V un espace vectoriel de dimension finie, $f: V \rightarrow K$ une forme linéaire et B, B' des bases de V . Soit

$$f(x) = a^T [x]_B$$

où $a \in K^n$. Décrire $f(x)$ en termes de $P_{B'B}$ et $[x]_{B'}$.

2. Soient V un espace vectoriel de dimension finie, $f: V \times V \rightarrow K$ une forme bilinéaire et B, B' des bases de V . Soit

$$f(x, y) = [y]_B^T A_B^f [x]_B.$$

Décrire $f(x, y)$ en termes de $P_{B'B}$, $[x]_{B'}$, et $[y]_{B'}$.

3. On considère les vecteurs

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \text{ et } v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{Z}_2^4$$

et le produit scalaire standard. Trouver une base du $\text{span}\{v_1, v_2\}^\perp$. Est-ce que $\mathbb{Z}_2^4 = \text{span}\{v_1, v_2\} \oplus \text{span}\{v_1, v_2\}^\perp$?

4. Soit $V \subseteq \mathbb{R}[x]$ l'espace euclidien des polynômes de degré au plus n muni du produit scalaire $\langle p, q \rangle = \int_0^1 p(x)q(x) dx$. Décrire la matrice $A_B^\langle \cdot \rangle$ pour $B = \{1, x, \dots, x^n\}$.
5. Soit V un espace vectoriel sur un corps K et soient $f, g \in V^* \setminus \{0\}$ linéairement indépendants. Montrer que

$$\ker f \cap \ker g$$

est de dimension $n - 2$.

6. Soit V un espace vectoriel de dimension finie sur un corps K et soit $\langle \cdot \rangle$ un produit scalaire. Exprimez $(W_1 + W_2)^\perp$ et $(W_1 \cap W_2)^\perp$ en fonction de W_1^\perp et W_2^\perp .
7. Donner un exemple d'un espace vectoriel V muni d'un produit scalaire dégénéré et d'un sous-espace $W \subseteq V$ tel que V n'est pas la somme directe de W et W^\perp .

1.7 Formes sesquilineaires et produits hermitiens

Maintenant nous considérons le cas $K = \mathbb{C}$. Il faut un peu modifier la définition d'un produit scalaire pour obtenir des résultats similaires à ceux des sections précédentes. Le carré de la longueur d'un vecteur

$$v = \begin{pmatrix} a_1 + i \cdot b_1 \\ \vdots \\ a_n + i \cdot b_n \end{pmatrix} \in \mathbb{C}^n$$

où $a_i, b_i \in \mathbb{R}$, est égal à

$$\sum_i (a_i^2 + b_i^2) = \sum_i (a_i + ib_i) \cdot (a_i - ib_i) = \sum_i v_i \cdot \bar{v}_i,$$

où $v_i = a_i + i \cdot b_i$ et \bar{v}_i est la conjugaison de v_i . Ceci suggère la définition suivante.

Définition 1.16. Soit V un espace vectoriel sur un corps \mathbb{C} . Un *produit hermitien* sur V est une correspondance qui, à tout couple (v, w) d'éléments de V associe un nombre complexe, noté $\langle v, w \rangle$, satisfaisant les propriétés suivantes :

PH 1 On a $\langle v, w \rangle = \overline{\langle w, v \rangle}$ pour tout $v, w \in V$.

PH 2 Si u, v et w sont des éléments de V ,

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \text{ et } \langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle$$

1 Formes bilinéaires

PH 3 Si $x \in \mathbb{C}$ et $u, v \in V$,

$$\langle x \cdot u, v \rangle = x \langle u, v \rangle \text{ et } \langle u, x \cdot v \rangle = \bar{x} \cdot \langle u, v \rangle.$$

On dit que le produit hermitien est *non dégénéré* si, de plus, la condition suivante est vérifiée :

Si $v \in V$ et si $\langle v, w \rangle = 0$ pour tout $w \in V$, alors $v = 0$.

Pour tout $v \in V$ $\langle v, v \rangle \in \mathbb{R}$, puisque $\langle v, v \rangle = \overline{\langle v, v \rangle}$, on dit que le produit hermitien est *défini positif* si $\langle v, v \rangle \geq 0$ pour tout $v \in V$ et si $\langle v, v \rangle > 0$ lorsque $v \neq 0$.

Exemple 1.15. Le produit *hermitien standard* de \mathbb{C}^n

$$\langle u, v \rangle = \sum_i u_i \bar{v}_i$$

satisfait les condition PH 1-3 et est défini positif.

Les notions d'*orthogonalité*, de *perpendicularité*, de *base orthogonale* et de *supplémentaire orthogonal* sont définies comme avant, de même que les notions de *coefficients de Fourier* et la *projection de v sur w* .

Exemple 1.16. Soit V l'espace vectoriel des fonctions $f: \mathbb{R} \rightarrow \mathbb{C}$ continues sur l'intervalle $[0, 2\pi]$. Pour $f, g \in V$ on pose

$$\langle f, g \rangle = \int_0^{2\pi} f(x) \overline{g(x)} dx.$$

C'est un produit hermitien défini positif. Les fonctions $f_n(x) = e^{inx}$ pour $n \in \mathbb{Z}$ sont orthogonales dès que

$$f_n(x) \overline{f_m(x)} = e^{i(n-m)x} = \cos((n-m)x) + i \cdot \sin((n-m)x)$$

et alors

$$\langle f_n, f_n \rangle = \int_0^{2\pi} 1 dx = 2\pi$$

et pour $n \neq m$

$$\langle f_n, f_m \rangle = \int_0^{2\pi} \cos((n-m)x) dx + i \cdot \int_0^{2\pi} \sin((n-m)x) dx = 0.$$

Pour $f \in V$ la composante de f sur f_n , ou le coefficient de Fourier de f relativement à f_n , est

$$\frac{\langle f, f_n \rangle}{\langle f_n, f_n \rangle} = \frac{1}{2\pi} \cdot \int_0^{2\pi} f(x) e^{-inx} dx.$$

Définition 1.17. Une application $f: V \times V \rightarrow \mathbb{C}$ satisfaisant

PH 2 et PH 3 est appelé une *forme sesquilinéaire*.

1.7 Formes sesquilinéaires et produits hermitiens

Soient V un espace vectoriel sur \mathbb{C} de dimension finie et $f : V \times V \rightarrow \mathbb{C}$ une forme sesquilinéaire. Pour une base $B = \{v_1, \dots, v_n\}$ de V et $x = \sum_i \alpha_i v_i$ et $y = \sum_i \beta_i v_i$ on a

$$\langle x, y \rangle = \sum_{ij} \alpha_i \overline{\beta_j} f(v_i, v_j)$$

et avec la matrice $A_B^f = (f(v_i, v_j))_{1 \leq i, j \leq n}$ alors

$$\langle x, y \rangle = [x]_B^T A_B^f \overline{[y]_B} \quad (1.10)$$

où pour un vecteur $v \in \mathbb{C}^n$ le vecteur \bar{v} est tel que $\bar{v}_i = \overline{v_i}$ pour tout i . Pour une matrice $A \in \mathbb{C}^{m \times n}$, $\bar{A} \in \mathbb{C}^{m \times n}$ est la matrice telle que $\bar{A}_{ij} = \overline{A_{ij}}$ pour tout i, j .

Définition 1.18. Une matrice $A \in \mathbb{C}^{n \times n}$ est appelée *hermitienne* si on a

$$A = \bar{A}^T.$$

Proposition 1.27. Soit V un espace vectoriel sur \mathbb{C} de dimension finie et soit B une base de V . Une forme sesquilinéaire f est un produit hermitien si et seulement si A_B^f est hermitienne.

Définition 1.19. Deux matrices $A, B \in \mathbb{C}^{n \times n}$ sont *congruentes complexes* s'il existe une matrice inversible $P \in \mathbb{C}^{n \times n}$ telle que $A = P^T \cdot B \cdot \bar{P}$. Nous écrivons $A \cong_{\mathbb{C}} B$.

$\cong_{\mathbb{C}}$ est aussi une relation d'équivalence. On peut aussi modifier l'algorithme 1.1 tel qu'il calcule une matrice diagonale congruente complexe par rapport à une matrice hermitienne $A \in \mathbb{C}^{n \times n}$, voir l'exercice 6. Alors on a le théorème suivant.

Théorème 1.28. Soit V un espace vectoriel sur \mathbb{C} de dimension finie, muni d'un produit hermitien. Alors V possède une base orthogonale.

Démonstration. Soit $B = \{v_1, \dots, v_n\}$ une base de V . Pour $x, y \in V$ on a

$$\langle x, y \rangle = [x]_B^T A_B^{\langle \cdot, \cdot \rangle} \overline{[y]_B}.$$

Soit $P \in \mathbb{C}^{n \times n}$ inversible telle que

$$P^T A \bar{P} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}.$$

La base orthogonale est w_1, \dots, w_n dont $[w_j]_B$ est la j -ème colonne de P ,

$$[w_j]_B = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix},$$

alors $w_j = \sum_{i=1}^n p_{ij} w_i$.

□

Exemple 1.17. On considère la matrice hermitienne

$$A = \begin{bmatrix} 0 & -i & 3+4i \\ i & -2 & 12 \\ 3-4i & 12 & 5 \end{bmatrix}$$

et le but est de trouver une matrice inversible $P \in \mathbb{C}^{3 \times 3}$ telle que

$$P^T \cdot A \cdot \overline{P}$$

est une matrice diagonale. Nous échangeons la première et la deuxième colonne ainsi que la première et la deuxième ligne et obtenons

$$\begin{bmatrix} -2 & i & 12 \\ -i & 0 & 3+4i \\ 12 & 3-4i & 5 \end{bmatrix}.$$

Après on transforme

$$\begin{bmatrix} 1 & 0 & 0 \\ -0.5i & 1 & 0 \\ 6 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 & i & 12 \\ -i & 0 & 3+4i \\ 12 & 3-4i & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0.5i & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.5 & 3-2i \\ 0 & 3+2i & 77 \end{bmatrix}$$

La prochaine transformation est

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6-4i & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.5 & 3-2i \\ 0 & 3+2i & 77 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -6+4i \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 51 \end{bmatrix}.$$

Pour

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -0.5i & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -6-4i \\ 0 & 0 & 1 \end{bmatrix}$$

on obtient

$$P^T \cdot A \cdot \overline{P} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 51 \end{bmatrix}.$$

Exercices

1. Soit V un espace vectoriel sur \mathbb{C} et $f: V \times V \longrightarrow \mathbb{C}$ une application satisfaisant les axiomes

i) On a $f(v, w) = \overline{f(w, v)}$ pour tout $v, w \in V$.

ii) Si u, v et w sont des éléments de V ,

$$f(u, v + w) = f(u, v) + f(u, w)$$

iii) Si $x \in \mathbb{C}$ et $u, v \in V$,

$$f(x \cdot u, v) = x f(u, v).$$

Montrer que f est un produit hermitien.

2. Soit V un espace vectoriel sur \mathbb{C} de dimension finie et $f : V \times V \rightarrow \mathbb{C}$ une forme sesquilinéaire. Pour une base $B = \{v_1, \dots, v_n\}$ de V et $x = \sum_i \alpha_i v_i$ et $y = \sum_i \beta_i v_i$ montrer en détail que

$$\langle x, y \rangle = [x]_B^T A_B^f \overline{[y]_B}$$

avec la matrice $A_B^f = (f(v_i, v_j))_{1 \leq i, j \leq n}$. Indiquez l'application des axiomes PH 2) et PH 3) dans les pas correspondants.

3. Démontrer la proposition 1.27.
4. Soit V un espace vectoriel sur \mathbb{C} de dimension n et soit $\langle \cdot, \cdot \rangle$ un produit hermitien. Montrer que $\langle \cdot, \cdot \rangle$ est non dégénéré si et seulement si $\text{rang}(A_B^{\langle \cdot, \cdot \rangle}) = n$ pour chaque base B de V .
5. Montrer que $\cong_{\mathbb{C}}$ est une relation d'équivalence.
6. Modifier l'algorithme 1.1 afin qu'il calcule une matrice diagonale congruente complexe par rapport à une matrice hermitienne $A \in \mathbb{C}^{n \times n}$.
7. Soient V un espace vectoriel sur \mathbb{C} et $\dim(V) = 3$ et $B = \{v_1, v_2, v_3\}$ une base de V . Avec les matrices $A_i \in \mathbb{C}^{3 \times 3}$ décrites en bas et les applications $f_i(x, y) = [x]_B^T A_i [y]_B$, cocher ce qui s'applique.

| | A_1 | A_2 | A_3 |
|----------------------|--------------------------|--------------------------|--------------------------|
| forme sesquilinéaire | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| produit hermitien | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

$$A_1 = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 2 & 1+i & 3 \\ 1 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 2 & 1+2 \cdot i & 3-i \\ 1-2 \cdot i & 0 & 2-i \\ 3-i & 2+i & 0 \end{pmatrix}.$$

1.8 Espaces hermitiens

Pour le reste de ce paragraphe, si pas spécifié autrement, V est toujours un espace vectoriel sur \mathbb{C} muni d'un produit hermitien défini positif, alors V est un espace hermitien.

Définition 1.20. Soit $\langle \cdot, \cdot \rangle$ un produit hermitien défini positif. La *longueur* ou la *norme* d'un élément $v \in V$ est le nombre

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Un élément $v \in V$ est un *vecteur unitaire* si $\|v\| = 1$.

1 Formes bilinéaires

Aussi, l'inégalité de Cauchy-Schwarz est démontrée comme avant :

$$|\langle u, v \rangle| \leq \|u\| \|v\|. \quad (1.11)$$

Les propriétés suivantes sont facilement vérifiées.

- i) Pour tout $v \in V$, $\|v\| \geq 0$ et $\|v\| = 0$ si et seulement si $v = 0$.
- ii) Pour $\alpha \in \mathbb{C}$ et $v \in V$ on a $\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$.
- iii) Pour chaque $u, v \in V$ $\|u + v\| \leq \|u\| + \|v\|$.

Aussi, nous avons le théorème de Pythagore, l'inégalité de Bessel et la règle du parallélogramme. L'équivalent du procédé de Gram-Schmidt pour les espaces hermitiens est comme suit.

Théorème 1.29 (Le procédé d'orthogonalisation de Gram-Schmidt). *Soit V un espace hermitien et $\{v_1, \dots, v_n\} \subseteq V$ un ensemble libre. Il existe un ensemble libre orthogonal $\{u_1, \dots, u_n\}$ de V tel que pour tout i , $\{v_1, \dots, v_i\}$ et $\{u_1, \dots, u_i\}$ engendrent le même sous-espace de V .*

Comme avant, une base orthonormale est une base orthogonale consistant de vecteurs unitaires et le procédé de Gram-Schmidt nous donne le corollaire suivant.

Corollaire 1.30. *Soit V un espace hermitien de dimension finie. Supposons $V \neq \{0\}$. V possède alors une base orthonormale.*

Exercices

- 1. Montrer l'inégalité de Cauchy-Schwarz.
- 2. Montrer l'inégalité triangulaire [iii](#)).
- 3. Montrer qu'un espace hermitien de dimension finie possède une base B telle que $\langle x, y \rangle = [x]_B^T \cdot [y]_B$, où \cdot est le produit hermitien standard.

2 Le théorème spectral et la décomposition en valeurs singulières

Dans ce chapitre, nous allons étudier les espaces euclidiens et hermitiens d'une manière plus profonde. Lorsque l'on parle de \mathbb{C} ou de \mathbb{R} , nous allons utiliser la lettre \mathbb{K} pour dénoter \mathbb{C} ou \mathbb{R} . On va rappeler quelques notions importantes du cours du premier semestre. Un *endomorphisme* est une application linéaire $f: V \rightarrow V$. Si V est un espace vectoriel de dimension finie et si $B = \{v_1, \dots, v_n\}$ est une base de V , on a

$$f(x) = \phi_B^{-1}(A_B \phi_B(x)),$$

où ϕ_B est l'isomorphisme $\phi_B: V \rightarrow K^n$, $\phi_B(x) = [x]_B$ sont les coordonnées de x par rapport à la base B . On a le diagramme suivant

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \downarrow \phi_B & & \downarrow \phi_B \\ K^n & \xrightarrow{A_B \cdot x} & K^n \end{array}$$

Les colonnes de la matrice A_B sont les coordonnées de $f(v_1), \dots, f(v_n)$. Une matrice $A \in K^{n \times n}$ est *diagonalisable* s'il existe une matrice inversible $P \in K^{n \times n}$ telle que $P^{-1} \cdot A \cdot P$ est une matrice diagonale.

2.1 Les endomorphismes auto-adjoints

Dans ce paragraphe 2.1, V est toujours un espace euclidien ou un espace hermitien de dimension finie.

Définition 2.1. Un endomorphisme F est *auto-adjoint* si

$$\langle F(v), w \rangle = \langle v, F(w) \rangle \text{ pour tous } v, w \in V.$$

Théorème 2.1. Soient $B = \{v_1, \dots, v_n\}$ une base orthonormale de V et F un endomorphisme. Alors F est auto-adjoint si et seulement si sa matrice A_B dans la base B est symétrique ($\mathbb{K} = \mathbb{R}$) ou hermitienne ($\mathbb{K} = \mathbb{C}$).

2 Le théorème spectral et la décomposition en valeurs singulières

Démonstration. On traite seulement le cas $\mathbb{K} = \mathbb{C}$. Le cas $\mathbb{K} = \mathbb{R}$ est démontré d'une manière analogue. Nous avons, où \cdot dénote le produit hermitien standard,

$$\langle F(v), w \rangle = (A_B[v]_B) \cdot [w]_B = [v]_B^T A_B^T \overline{[w]_B},$$

et

$$\langle v, F(W) \rangle = [v]_B^T \overline{A_B[w]_B}.$$

Alors $\overline{A_B} = A_B^T$. □

Lemme 2.2. Soit $A \in \mathbb{C}^{n \times n}$ une matrice hermitienne. Les valeurs propres de A sont réelles.

Démonstration. Soient $\lambda \in \mathbb{C}$ une valeur propre et $v \neq 0$ son vecteur propre. Alors

$$\lambda v^T \bar{v} = v^T A^T \bar{v} = v^T \overline{A} v = \bar{\lambda} v^T \bar{v}.$$

□

Corollaire 2.3. Soit F un endomorphisme auto-adjoint, alors toutes ses valeurs propres sont réelles.

Démonstration. Soit $B = \{v_1, \dots, v_n\}$ une base orthonormale. Les valeurs propres de F sont les valeurs propres de la matrice hermitienne A_B . □

Corollaire 2.4. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ (hermitienne $A \in \mathbb{C}^{n \times n}$) a une valeur propre réelle.

Démonstration. Le polynôme caractéristique $p(x) = \det(A - x \cdot I_n)$ a une racine complexe selon le théorème fondamental de l'algèbre. Les valeurs propres de A sont les racines de $p(x)$. Mais toutes ces racines sont réelles selon le corollaire 2.3. □

Lemme 2.5. Soient F un endomorphisme auto-adjoint et $u, v \neq 0$ deux vecteurs propres dont leurs valeurs propres sont différentes, alors $\langle u, v \rangle = 0$.

Démonstration. Soient $\lambda \neq \gamma$ les valeurs propres correspondant aux vecteurs propres $u, v \neq 0$ respectivement. Puisque $\lambda, \gamma \in \mathbb{R}$ on a

$$\lambda \langle u, v \rangle = \langle F(u), v \rangle = \langle u, F(v) \rangle = \gamma \langle u, v \rangle$$

et alors $\langle u, v \rangle = [u]_B \cdot [v]_B = 0$, où \cdot dénote le produit scalaire/hermitien standard. □

Définition 2.2. Une matrice inversible $U \in \mathbb{R}^{n \times n}$ est *orthogonale* si $U^{-1} = U^T$. Une matrice inversible $U \in \mathbb{C}^{n \times n}$ est *unitaire* si $U^{-1} = \overline{U}^T$.

Si U est orthogonale (unitaire), les colonnes de U sont une base orthonormale de \mathbb{R}^n (\mathbb{C}^n), où l'orthonormalité est entendue au sens du produit scalaire (hermitien) standard.

Notation. Nous allons écrire A^* pour dénoter \overline{A}^T pour une matrice A .

Théorème 2.6 (Théorème spectral). *Soit $A \in \mathbb{K}^{n \times n}$ une matrice symétrique (hermitienne), alors A est diagonalisable avec une matrice orthogonale (unitaire) $P \in \mathbb{K}^{n \times n}$ telle que*

$$P^* \cdot A \cdot P = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad (2.1)$$

où $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ sont les valeurs propres de A .

Démonstration. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique. Le cas où $A \in \mathbb{C}^{n \times n}$ est hermitienne est un exercice.

Le théorème est démontré par induction. Si $n = 1$, l'assertion est triviale.

Supposons le théorème vrai jusqu'à $n - 1 \in \mathbb{N}, n \geq 2$. Montrons le pour n .

Soit $\lambda_1 \in \mathbb{R}$ une valeur propre de A et $v \neq 0 \in \mathbb{R}^n$ un vecteur propre correspondant à λ_1 . Avec la méthode de Gram-Schmidt, on peut trouver une base orthonormale $\{v, u_2, \dots, u_n\}$ dans \mathbb{R}^n . Soit $U \in \mathbb{R}^{n \times n-1}$ la matrice dont les colonnes sont u_2, \dots, u_n . On considère la matrice

$$U^T \cdot A \cdot U \in \mathbb{R}^{n-1 \times n-1}.$$

Cette matrice est symétrique. En effet :

$$(U^T \cdot A \cdot U)^T = U^T \cdot A^T \cdot (U^T)^T = U^T \cdot A \cdot U,$$

car A est symétrique. Par hypothèse d'induction, cette matrice peut être diagonalisée avec une matrice orthogonale $K \in \mathbb{R}^{n-1 \times n-1}$, alors

$$K^T \cdot U^T \cdot A \cdot U \cdot K = \begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Maintenant, soit $P \in \mathbb{R}^{n \times n}$ la matrice

$$P = (v, U K) \in \mathbb{R}^{n \times n}$$

La matrice P est orthogonale puisque

$$P^T P = \begin{pmatrix} v^T v & v^T U K \\ (U K)^T v & (U K)^T (U K) \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Et

$$P^T A P = \begin{pmatrix} v^T A \\ K^T U^T A \end{pmatrix} (v \ U K) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

□

Corollaire 2.7. Soit V un espace euclidien (hermitien) de dimension fini et F un endomorphisme auto-adjoint. Alors V possède une base $\{v_1, \dots, v_n\}$ orthonormale de vecteurs propre de F .

Démonstration. Soit $B = \{u_1, \dots, u_n\}$ une base de V tel que $\langle x, y \rangle = [x]_B \cdot [y]_B$ où \cdot dénote le produit hermitien standard (voir chapitre 1.8, exercice 3) et soit A_F la matrice symétrique (hermitienne) telle que $F(x) = \phi_B^{-1}(A_F[x]_B)$. Selon théorème 2.6 A_F est diagonalisable avec une matrice orthogonale (unitaire) P

$$P^* \cdot A_F \cdot P = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Soient p_1, \dots, p_n les colonnes de P . La base orthonormale de vecteurs propres de F est $\{v_1, \dots, v_n\}$ où $v_i = \phi_B^{-1}(p_i)$. \square

Cours du 12. Avril 2016

Comment peut-on calculer la diagonalisation (2.1) ? Voici un procédé pour diagonaliser une matrice symétrique (hermitienne) $A \in \mathbb{K}^{n \times n}$.

- i) Trouver les racines $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ du polynôme caractéristique

$$p(x) = \det(A - x \cdot I).$$

- ii) Pour tout $j \in \{1, \dots, k\}$:

- a) Trouver une base $b_1^{(j)}, \dots, b_{d_j}^{(j)}$ du noyau de la matrice $A - \lambda_j I$, par exemple avec l'algorithme de Gauss.
- b) Trouver une base orthonormale $p_1^{(j)}, \dots, p_{d_j}^{(j)}$ du $\text{span}\{b_1^{(j)}, \dots, b_{d_j}^{(j)}\}$, par exemple avec le procédé de Gram-Schmidt.

- iii) $P = \left(p_1^{(1)}, \dots, p_{d_1}^{(1)}, \dots, p_1^{(k)}, \dots, p_{d_k}^{(k)} \right)$

Exemple 2.1. Soit V un espace euclidien de dimension 3 et soit F un endomorphisme auto-adjoint de V . Soit $B = \{v_1, v_2, v_3\}$ une base telle que

$$A_B = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

Trouver une base orthonormale qui se compose des vecteurs propres.

Le polynôme caractéristique de A_B est

$$p(x) = \det(A_B - x I) = -x^3 + 12x^2 - 21x - 98 = -(x - 7)^2(x + 2)$$

On trouve les bases des espaces propres

$$\lambda_1 = 7 : b_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, b_2^{(1)} = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \lambda_2 = -2 : b_1^{(2)} = \begin{pmatrix} -1 \\ -1/2 \\ 1 \end{pmatrix}$$

Les vecteurs $b_1^{(1)}$ et $b_2^{(1)}$ ne sont pas orthogonaux. Le procédé de Gram-Schmidt produit $b_2^{(1)*} = (-1/4, 1, 1/4)^T$. Les vecteurs $b_1^{(1)}, b_2^{(1)*}, b_1^{(2)}$ sont une base orthogonale de vecteurs propres. Maintenant il reste à les normaliser et on obtient

$$p_1^{(1)} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{bmatrix}, p_2^{(1)} = \begin{bmatrix} -\frac{\sqrt{2}}{6} \\ \frac{2\sqrt{2}}{3} \\ \frac{\sqrt{2}}{6} \end{bmatrix}, p_1^{(2)} = \begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Alors $\left\{ \frac{\sqrt{2}}{2}v_1 + \frac{\sqrt{2}}{2}v_3, -\frac{\sqrt{2}}{6}v_1 + \frac{2\sqrt{2}}{3}v_2 + \frac{\sqrt{2}}{6}v_3, -\frac{2}{3}v_1 - \frac{1}{3}v_2 + \frac{2}{3}v_3 \right\}$ est une base ortho-normale de vecteurs propre de F .

Le lemme 2.2 et le corollaire 2.4 démontrent qu'une matrice symétrique réelle possède une valeur propre réelle. Cette démonstration passe par les nombres complexes et utilise le théorème fondamental de l'algèbre. Pour le cas où $A = A^T \in \mathbb{R}^{n \times n}$, nous allons maintenant démontrer l'assertion du corollaire 2.4 d'une manière géométrique. L'ensemble

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$$

est appelé la n -sphère.

Définition 2.3. Une *forme quadratique* est une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = x^T A x$ où $A \in \mathbb{R}^{n \times n}$ est une matrice symétrique.

En effet, si $B \in \mathbb{R}^{n \times n}$ n'est pas symétrique, la matrice $A = 1/2(B^T + B)$ est symétrique et $x^T B x = 1/2(x^T B^T x + x^T B x) = x^T A x$. Alors aussi la fonction $g(x) = x^T B x$ est une forme quadratique.

Une forme quadratique est un polynôme de degré 2 et une fonction continue. Puisque S^{n-1} est compact et $f(x)$ est continue, $f(x)$ possède un maximum sur S^{n-1} . Nous sommes prêts à démontrer le lemme d'une manière géométrique.

Lemme 2.8. Soient $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et $v \in S^{n-1}$ le maximum de la fonction $f(x) = x^T A x$ sur S^{n-1} . On a $Av = \lambda v$ pour un $\lambda \in \mathbb{R}$. En particulier, A possède une valeur propre réelle.

Démonstration. Supposons qu'il n'existe pas de $\lambda \in \mathbb{R}$ tel que $Av = \lambda v$. Alors, en particulier, $Av \neq 0$ et on peut écrire

$$Av = \alpha v + \beta w$$

où $w \in S^{n-1}$, $w \perp v$ et $\beta \neq 0$. Pour $x \in [-1, 1]$ on a

$$\sqrt{(1-x^2)}v + xw \in S^{n-1}.$$

2 Le théorème spectral et la décomposition en valeurs singulières

On voit facilement que $\|\sqrt{(1-x^2)}v + xw\| = 1$ en utilisant le fait que $v \perp w$, et $v, w \in S^{n-1}$. Nous considérons la fonction $g: [-1, 1] \rightarrow \mathbb{R}$

$$g(x) = \left(\sqrt{(1-x^2)}v + xw \right)^T A \left(\sqrt{(1-x^2)}v + xw \right).$$

Notons que $g(0) = f(v)$, donc si v maximise f sur la n -sphère, en particulière $x = 0$ doit maximiser $g(x)$ dans l'intervalle $[-1, 1]$. Si on démontre que $g'(0) \neq 0$, nous avons déduit une contradiction et la démonstration est faite.

Comme $w^T Av = v^T Aw$, clairement

$$g(x) = (1-x^2)v^T Av + (2 \cdot \sqrt{(1-x^2)} \cdot x) w^T Av + x^2 w^T Aw.$$

Ceci démontre que $g'(0) = 2 \cdot w^T Av = 2 \cdot \beta \neq 0$.

□

Exercices

1. Soit $z = x + iy \in \mathbb{C}^n$, où $x, y \in \mathbb{R}^n$. Montrer que x et y sont linéairement indépendants sur \mathbb{R} si et seulement si z et \bar{z} sont linéairement indépendants sur \mathbb{C} .

2.2 Formes quadratiques réelles et matrices symétriques réelles

Définition 2.4. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est

- définie positive, si $x^T Ax > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$
- définie négative, si $x^T Ax < 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$
- semi-définie positive, si $x^T Ax \geq 0$ pour tout $x \in \mathbb{R}^n$
- semi-définie négative, si $x^T Ax \leq 0$ pour tout $x \in \mathbb{R}^n$.

La forme quadratique $x^T Ax$ correspondante est appelée définie positive, définie négative, semi-définie positive ou semi-définie négative, en accord avec A .

Théorème 2.9. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est

1. définie positive, si toutes ses valeurs propres sont strictement positives.
2. définie négative, si toutes ses valeurs propres sont strictement négatives.
3. semi-définie positive, si toutes ses valeurs propres sont positives (ou zéro).
4. semi-définie négative, si toutes ses valeurs propres sont négatives (ou zéro).

Démonstration. D'après le théorème 2.6 il existe une matrice orthogonale $U \in \mathbb{R}^{n \times n}$ telle que

$$A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T. \quad (2.2)$$

2.2 Formes quadratiques réelles et matrices symétriques réelles

où les λ_i sont les valeurs propres de A . Les colonnes u_1, \dots, u_n de U forment une base orthonormale de \mathbb{R}^n de vecteurs propres de A . Soit $x \in \mathbb{R}^n$, alors $x = \sum_i \alpha_i u_i$ et

$$x^T A x = \sum_i (\alpha_i^2) \lambda_i.$$

D'ici l'assertion suit directement. \square

Le *k-mineur principal* d'une matrice A est le déterminant de la matrice qui est construite en choisissant les premières k lignes et colonnes de A ; c'est-à-dire $\det(B_k)$ où $B_k \in \mathbb{R}^{k \times k}$ telle que $b_{ij} = a_{ij}$, $1 \leq i, j \leq k$. Soit $K = \{l_1, \dots, l_k\} \subseteq \{1, \dots, n\}$ où $l_1 < l_2 < \dots < l_k$. La matrice $B_K \in \mathbb{R}^{k \times k}$ est définie par $b_{ij} = a_{l_i l_j}$, pour $1 \leq i, j \leq k$. Un *k-mineur symétrique* de A est le déterminant d'une matrice B_K ; c'est-à-dire $\det(B_K)$.

Théorème 2.10. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique.

- a) A est définie positive si et seulement si tous ses mineurs principaux sont strictement positifs.
- b) A est semi-définie positive si et seulement si tous ses mineurs symétriques sont non négatifs.

Démonstration. On démontre a), tandis que b) est un exercice. Si A est définie positive, alors les matrices B_k , $1 \leq k \leq n$, sont symétriques et définies positives. Leurs valeurs propres sont toutes positives. Selon le théorème 2.6, $\det(B_k)$ est le produit des valeurs propres de B_k , alors $\det(B_k) > 0$.

Supposons maintenant que $\det(B_k) > 0$ pour tout $k \in \{1, \dots, n\}$. Nous appliquons notre algorithme 1.1. Avant la i -ème itération, la matrice A est de la forme

$$P^T A P = \begin{pmatrix} c_1 & & & & & \\ & c_2 & & & & \\ & & \ddots & & & \\ & & & c_{i-1} & & \\ & & & & b_{i,i} & \dots & b_{i,n} \\ & & & & \vdots & & \vdots \\ & & & & b_{n,i} & \dots & b_{n,n} \end{pmatrix}.$$

On observe que $b_{ii} \neq 0$. Parce que, si $b_{ii} = 0$ pour la première fois, nous n'avons jamais échangé de colonnes ou de lignes avant et $0 = c_1 \cdots c_{i-1} \cdot b_{ii} = \det(B_i)$, et c'est une contradiction. Alors il n'est jamais nécessaire d'échanger des lignes et colonnes et notre algorithme trouve une matrice R , triangulaire supérieure, dont les éléments diagonaux sont tous égaux à 1, et telle que

$$R^T \cdot A \cdot R = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}$$

On observe que $\det(B_k) = c_1 \cdots c_k$, alors toutes les c_i sont positives. Alors A est définie positive. \square

2 Le théorème spectral et la décomposition en valeurs singulières

Théorème 2.11. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et $f(x) = x^T A x$ la forme quadratique correspondante à A . On a

$$\max_{x \in S^{n-1}} f(x) = \lambda_1 \text{ et } \min_{x \in S^{n-1}} f(x) = \lambda_n \quad (2.3)$$

où λ_1 et λ_n sont les valeurs propres maximale et minimale de A respectivement.

Démonstration. Nous utilisons la factorisation

$$A = P^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P$$

où $P \in \mathbb{R}^{n \times n}$ est une matrice orthogonale dont les colonnes sont p_1, \dots, p_n . Si $x = \sum_i \alpha_i p_i$, alors

$$\|x\|^2 = \sum_i \alpha_i^2 \text{ et } x^T A x = \sum_i (\lambda_i \alpha_i^2)$$

et si $\|x\|^2 = 1$,

$$\lambda_n = \lambda_n \sum_i \alpha_i^2 \leq \sum_i (\lambda_i \alpha_i^2) \leq \lambda_1 \sum_i \alpha_i^2 = \lambda_1.$$

Ça démontre que p_1 et p_n sont les solutions optimales des problèmes d'optimisation (2.3). \square

Définition 2.5. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique. Pour $x \in \mathbb{R}^n \setminus \{0\}$ le *quotient Rayleigh-Ritz* est

$$R_A(x) = \frac{x^T A x}{x^T x}.$$

Pour $x \in \mathbb{R}^n \setminus \{0\}$ $x/\|x\| \in S^{n-1}$ et $R_A(x) = (x/\|x\|)^T \cdot A(x/\|x\|)$.

Théorème 2.12 (Théorème Min-Max). Soit $A \in \mathbb{R}^n$ une matrice symétrique avec les valeurs propres $\lambda_1 \geq \dots \geq \lambda_n$. Si U dénote un sous-espace de \mathbb{R}^n alors

$$\lambda_k = \max_{\dim(U)=k} \min_{x \in U \setminus \{0\}} R_A(x) \quad (2.4)$$

et

$$\lambda_k = \min_{\dim(U)=n-k+1} \max_{x \in U \setminus \{0\}} R_A(x) \quad (2.5)$$

Démonstration. Nous démontrons (2.4). La partie (2.5) est un exercice. Soit $\{u_1, \dots, u_n\}$ une base orthonormale de vecteurs propres associés à $\lambda_1 \leq \dots \leq \lambda_n$ respectivement. On fixe un entier k , et un espace U de dimension k . Clairement $\text{span}\{u_k, \dots, u_n\} \cap U \supsetneq \{0\}$. Alors il existe un vecteur $0 \neq x = \sum_{i=k}^n \alpha_i u_i \in U$. Clairement $R_A(x) \leq \lambda_k$. Pour $U = \text{span}\{u_1, \dots, u_k\}$, $\min_{x \in U \setminus \{0\}} R_A(x) = \lambda_k$. Ensemble ça démontre (2.4). \square

Exercices

1. Montrer la partie **b)** du théorème 2.10.
2. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est définie négative, si et seulement si $\det(B_k) \neq 0$ et $\det(B_k) = (-1)^k |\det(B_k)|$ pour tout k .
3. Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est semi-définie négative, si et seulement si $\det(B_K) = (-1)^{|K|} |\det(B_K)|$ pour tout $K \subseteq \{1, \dots, n\}$.
4. Montrer la partie (2.5) du théorème 2.12.
5. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique avec les valeurs propres $\lambda_1 \geq \dots \geq \lambda_n$. Soit B_K une matrice comme décrite en dessus où $|K| = k$ avec les valeurs propres $\mu_1 \geq \dots \geq \mu_{n-k}$. Pour $1 \leq i \leq k$, alors

$$\lambda_i \geq \mu_i \geq \lambda_{i+k}.$$

Cours du 19. Avril 2016

Définition 2.6. Une matrice hermitienne $A \in \mathbb{C}^{n \times n}$ est

- définie positive, si $x^T A \bar{x} > 0$ pour tout $x \in \mathbb{C}^n \setminus \{0\}$,
- définie négative, si $x^T A \bar{x} < 0$ pour tout $x \in \mathbb{C}^n \setminus \{0\}$,
- semi-définie positive, si $x^T A \bar{x} \geq 0$ pour tout $x \in \mathbb{C}^n$,
- semi-définie négative, si $x^T A \bar{x} \leq 0$ pour tout $x \in \mathbb{C}^n$.

Le théorème 2.9 trouve son analogue comme suivant. La démonstration est un exercice.

Théorème 2.13. Une matrice hermitienne $A \in \mathbb{C}^{n \times n}$ est

1. définie positive, si toutes ses valeurs propres sont (strictement) positives.
2. définie négative, si toutes ses valeurs propres sont (strictement) négatives.
3. semi-définie positive, si toutes ses valeurs propres sont non-négatives (donc positives ou zéro).
4. semi-définie négative, si toutes ses valeurs propres sont non-positives (négatives ou zéro).

On commence avec un théorème qui décrit la décomposition en valeurs singulières et montre qu'elle existe.

Théorème 2.14. Une matrice $A \in \mathbb{C}^{m \times n}$ peut être décomposée comme

$$A = P \cdot D \cdot Q$$

où $P \in \mathbb{C}^{m \times m}$ et $Q \in \mathbb{C}^{n \times n}$ sont unitaires et $D \in \mathbb{R}_{\geq 0}^{m \times n}$ est une matrice diagonale. Si A est réelle, P et Q sont réelles.

2 Le théorème spectral et la décomposition en valeurs singulières

Démonstration. La matrice $A^* \cdot A$ est hermitienne et semi-définie positive dès que

$$x^* A^* A x = (Ax)^*(Ax) \geq 0.$$

Alors les valeurs propres de $A^* A$ sont non-négatives ($\lambda_i \geq 0$). Soient $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2 \geq 0$ les valeurs propres et soit $\{u_1, \dots, u_n\}$ une base orthonormale correspondante de vecteurs propres. La matrice $Q \in \mathbb{C}^{n \times n}$ est la matrice dont les lignes sont u_1^*, \dots, u_n^* .

Soit $r \in \mathbb{N}_0$ tel que $\sigma_r > 0$ et $\sigma_{r+1} = 0$; on a $\sigma_1 \geq \dots \geq \sigma_r \geq 0 = \sigma_{r+1} = \dots = 0 = \sigma_n$. Nous construisons les vecteurs

$$v_i = A u_i / \sigma_i, \quad 1 \leq i \leq r.$$

Les v_i sont orthonormaux, parce que

$$\|v_i\|^2 = (A u_i)^* A u_i / \sigma_i^2 = 1$$

et pour $1 \leq i \neq j \leq r$,

$$v_i^* v_j = u_i^* u_j = 0.$$

Avec le procédé de Gram-Schmidt nous complétons les v_i tel que $\{v_1, \dots, v_m\}$ est une base orthonormale de \mathbb{C}^m . Les colonnes de la matrices P sont alors v_1, \dots, v_m dans cet ordre. La matrice $D \in \mathbb{C}^{m \times n}$ est la matrice diagonale dont les r premières composantes sur la diagonale sont $\sigma_1, \dots, \sigma_r$ dans cet ordre. Avec ces matrices P, D et Q nous avons

$$A = P \cdot D \cdot Q$$

ou, équivalent

$$P^* \cdot A \cdot Q^* = D,$$

Nous montrons ça en détail. Nous avons

$$(P^* \cdot A \cdot Q^*)_{ij} = v_i^* A u_j$$

et c'est égal à zéro si $j \geq r+1$, parce que dans ce cas $A u_j = 0$ dès que $u_j^* A^* A u_j = 0$. Si $j \leq r$ et si $i > r$ alors v_i est perpendiculaire à $A u_j$ et $v_i^* A u_j = 0$.

Et si $i, j \leq r$, alors

$$u_i^* A^* A u_j / \sigma_i = u_i^* u_j \sigma_j^2 / \sigma_i = \begin{cases} \sigma_i & \text{si } i = j \\ 0 & \text{autrement.} \end{cases}$$

□

Définition 2.7. En suivant la notation du théorème 2.14, les nombres $\sigma_1, \dots, \sigma_r$ sont les *valeurs singulières* de A . La factorisation $A = P \cdot D \cdot Q$ est une *décomposition en valeurs singulières*.

Exemple 2.2. Trouver une décomposition en valeurs singulières de

$$A = \begin{pmatrix} 0 & -1.6 & 0.6 \\ 0 & 1.2 & 0.8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On commence avec

$$A^* \cdot A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On obtient $\sigma_1 = 2, \sigma_2 = 1$ et $\sigma_3 = 0$. Les valeurs singulières sont les $\sigma_i > 0$, i.e., $\sigma_1 = 2$ et $\sigma_2 = 1$. On calcule les vecteurs propres correspondant à $\sigma_1 = 2, \sigma_2 = 1$ et $\sigma_3 = 0$. La matrice Q est

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

et

$$v_1 = \frac{1}{2} \cdot A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.8 \\ 0.6 \\ 0 \\ 0 \end{pmatrix}, v_2 = A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.8 \\ 0 \\ 0 \end{pmatrix}.$$

On complète avec $v_3 = e_3$ et $v_4 = e_4$, alors

$$P = \begin{pmatrix} -0.8 & 0.6 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

et

$$\begin{pmatrix} -0.8 & 0.6 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = A.$$

Définition 2.8. La *pseudo inverse* d'une matrice

$$D = \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$$

2 Le théorème spectral et la décomposition en valeurs singulières

où $\sigma_i \in \mathbb{R}_{>0}$ est

$$D^+ = \begin{pmatrix} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix} \in \mathbb{R}^{n \times m}$$

Toutes les composantes qui ne sont pas décrites sont zéro. La *pseudo inverse* d'une matrice $A \in \mathbb{C}^{m \times n}$ avec une décomposition en valeurs singulières $A = P \cdot D \cdot Q$ est

$$A^+ = Q^* D^+ P^*.$$

Exemple 2.3. La pseudo inverse de la matrice A d'exemple 2.2 est

$$A^+ = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -0.8 & 0.6 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Pourquoi est-ce que nous parlons de la pseudo inverse ? Parce qu'elle est unique.

Théorème 2.15. Soit K un corps et $A \in K^{m \times n}$, alors il existe au plus une seule matrice $X \in K^{n \times m}$ telle que les quatre conditions de Penrose sont satisfaites :

- i) $AXA = A$
- ii) $(AX)^* = AX$
- iii) $XAX = X$
- iv) $(XA)^* = XA$.

Démonstration. Soient X et Y deux matrices satisfaisant i-iv. Alors

$$\begin{aligned} X &= XAX \\ &= XAYAX \\ &= XAYAYAYAX \\ &= (XA)^*(YA)^*Y(AY)^*(AX)^* \\ &= A^*X^*A^*Y^*YY^*A^*X^*A^* \\ &= (AXA)^*Y^*YY^*(AXA)^* \\ &= A^*Y^*YY^*A^* \\ &= (YA)^*Y(AY)^* \\ &= YAYAY \\ &= YAY \\ &= Y. \end{aligned}$$

□

Théorème 2.16. La pseudo inverse d'une matrice $A \in \mathbb{C}^{m \times n}$ satisfait les conditions *i-iv*.

Démonstration. Soit $A = PDQ$ une décomposition en valeurs singulières et $A^+ = Q^*D^+P^*$. Il est facile de voir que D^+ satisfait les conditions *i-iv* relatives à D . Les conditions sont aussi vite montrées pour A et A^+ . Par exemple *i* est montrée comme suivant.

$$\begin{aligned} AA^+A &= PDQQ^*D^+P^*PDQ \\ &= PDD^+DQ \\ &= PDQ \\ &= A. \end{aligned}$$

Il est un exercice de vérifier les conditions *ii-iv*. □

2.3 Encore les systèmes d'équations

Nous considérons encore une fois un système

$$Ax = b, \tag{2.6}$$

où $A \in \mathbb{C}^{m \times n}$ et $b \in \mathbb{C}^m$.

Définition 2.9. La *solution minimale* de (2.6) est la solution du problème des moindres carrés

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|^2$$

correspondant avec norme $\|x\|$ minimale.

Théorème 2.17. La *solution minimale* de (2.6) est

$$x = A^+b,$$

où A^+ est la pseudo inverse de A .

Démonstration. On a

$$\begin{aligned} \min_{x \in \mathbb{C}^n} \|Ax - b\| &= \min_{x \in \mathbb{C}^n} \|PDQx - b\| = \min_{x \in \mathbb{C}^n} \|P^*(PDQx - b)\| \\ &= \min_{x \in \mathbb{C}^n} \|DQx - P^*b\| = \min_{y \in \mathbb{C}^n} \|Dy - P^*b\| \\ &= \min_{y \in \mathbb{C}^n} \|Dy - c\| \end{aligned}$$

où $c = P^*b$. Les solutions optimales sont les $y \in \mathbb{C}^n$ tel que $y_i = c_i/\sigma_i$ pour $1 \leq i \leq r$ et $y_{r+1} \dots y_n$ sont arbitraires. Parmi ces vecteurs, la solution où $y_{r+1} = \dots = y_n = 0$ est celle de norme minimale. Elle est donnée par

$$y = D^+c.$$

2 Le théorème spectral et la décomposition en valeurs singulières

La solution minimale du (2.6) est alors

$$x = Q^*y = Q^*D^+P^*b = A^+b.$$

□

Exemple 2.4. Trouver la solution minimale du système

$$\begin{pmatrix} 0 & -1.6 & 0.6 \\ 0 & 1.2 & 0.8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 5 \\ 7 \\ 3 \\ -2 \end{pmatrix}.$$

La pseudo-inverse de la matrice ci-dessus est

$$A^+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -0.4 & 0.3 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0 \end{pmatrix}$$

et

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ -0.4 & 0.3 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 7 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.1 \\ 8.6 \end{pmatrix}.$$

Cours du 3. Mai 2016

2.4 Le meilleur sous-espace approximatif

Nous nous occupons du problème suivant. Soient $a_1, \dots, a_m \in \mathbb{R}^n$ des vecteurs et $1 \leq k \leq n$, trouver un sous-espace $H \subseteq \mathbb{R}^n$ de dimension k tel que

$$\sum_i d(a_i, H)^2$$

soit minimale. Ici $d(a_i, H)$ est la *distance* de a_i à H . Si $H = \text{span}\{u_1, \dots, u_k\}$ où $\{u_1, \dots, u_k\}$ est une base orthonormale de H , alors $a_i = \sum_{j=1}^k \langle a_i, u_j \rangle u_j + d_i$ où $d_i = a_i - \sum_{j=1}^k \langle a_i, u_j \rangle u_j$ est orthogonal à u_1, \dots, u_k et alors à H . Avec le théorème de Pythagore (Proposition 1.11) alors

$$d(a_i, H)^2 + \sum_{j=1}^k \langle a_i, u_j \rangle^2 = \|a_i\|^2.$$

2.4 Le meilleur sous-espace approximatif

Le sous-espace H de dimension k qui minimise $\sum_i d(a_i, H)^2$ est alors celui qui maximise

$$\sum_i \sum_{j=1}^k \langle a_i, u_j \rangle^2 = \sum_{j=1}^k \|Au_j\|^2 = \sum_{j=1}^k u_j^T A^T A u_j.$$

Pour $k = 1$, nous connaissons déjà une manière de résoudre ce problème. Il faut résoudre

$$\max_{u \in S^{n-1}} u^T A^T A u$$

est maximal. La matrice $A^T A$ est symétrique, alors on peut la factoriser comme

$$A^T A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \quad (2.7)$$

où $U = (u_1, \dots, u_n) \in \mathbb{R}^{n \times n}$ est orthogonale. Nous pouvons supposer que les λ_i sont ordonnés comme $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. Les valeurs propres sont non négatives dès que $A^T A$ est semi-définie positive. Selon Théorème 2.11 la solution est $H = \text{span}\{u_1\}$.

La généralisation suivante du Théorème 2.11 est un exercice.

Théorème 2.18. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et $f(x) = x^T A x$ la forme quadratique correspondante à A . Soit

$$A = U \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T$$

une factorisation de A telle que $U = (u_1, \dots, u_n) \in \mathbb{R}^{n \times n}$ est orthogonale et $\lambda_1 \geq \dots \geq \lambda_n$. Pour $1 \leq \ell < n$ on a

$$\max_{\substack{x \in S^{n-1} \\ x \perp u_1, \dots, x \perp u_\ell}} f(x) = \lambda_{\ell+1} = \min_{\substack{x \in S^{n-1} \\ x \perp u_{\ell+2}, \dots, x \perp u_n}} f(x) \quad (2.8)$$

et $u_{\ell+1}$ est une solution optimale.

Maintenant nous pouvons résoudre le problème central

$$\min_{\substack{H \subseteq \mathbb{R}^n \\ \dim(H)=k}} \sum_{i=1}^m d(a_i, H)^2. \quad (2.9)$$

Théorème 2.19. Soient $a_1, \dots, a_m \in \mathbb{R}^n$ $A = (a_1, \dots, a_m)^T$ et u_1, \dots, u_k les premières colonnes de la matrice orthogonale $U \in \mathbb{R}^{n \times n}$ de la factorisation (2.7). Le sous-espace $H = \text{span}\{u_1, \dots, u_k\}$ est une solution du problème (2.9).

2 Le théorème spectral et la décomposition en valeurs singulières

Démonstration. Pour $k = 1$ nous avons déjà montré l'assertion. Soit $k \geq 2$ et $W \trianglelefteq \mathbb{R}^n$ une solution optimale du problème (2.9) et soit w_1, \dots, w_k une base orthonormale de W . Nous pouvons supposer $w_k \perp \text{span}\{v_1, \dots, v_{k-1}\}$, (voir Exercice 6).

Par induction, nous avons

$$\sum_{j=1}^{k-1} w_j^T A^T A w_j \leq \sum_{j=1}^{k-1} u_j^T A^T A u_j.$$

Dès que

$$\max_{\substack{x \in S^{n-1} \\ x \perp \text{span}\{v_1, \dots, v_{k-1}\}}} x^T A^T A x$$

est atteint à v_k nous avons

$$w_k^T A^T A w_k \leq v_k^T A^T A v_k$$

et alors

$$\sum_{j=1}^k w_j^T A^T A w_j \leq \sum_{j=1}^k u_j^T A^T A u_j.$$

□

Définition 2.10. Soit $A \in \mathbb{R}^{m \times n}$. La *norme Frobenius* de A est le nombre

$$\|A\|_F = \sqrt{\sum_{ij} a_{ij}^2}.$$

Définition 2.11. Soit $A \in K^{n \times n}$. La *trace* de A est la somme de ses coefficients diagonaux, $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$.

Lemme 2.20. Pour $A, B \in K^{n \times n}$ $\text{Tr}(AB) = \text{Tr}(BA)$.

Lemme 2.21. Pour $A \in \mathbb{R}^{m \times n}$, $\|A\|_F^2 = \sum_{i=1}^r \sigma_i^2$ où $\sigma_1, \dots, \sigma_r$ sont les valeurs singulières de A .

Démonstration. On a $\|A\|_F^2 = \text{Tr}(A^T A) = \text{Tr}(U \cdot \text{diag}(\sigma_1^2, \dots, \sigma_n^2) \cdot U^T)$ où $U \in \mathbb{R}^{n \times n}$ est orthogonale. Alors

$$\|A\|_F^2 = \text{Tr}(\text{diag}(\sigma_1^2, \dots, \sigma_n^2)) = \sum_{i=1}^r \sigma_i^2$$

□

Maintenant nous allons résoudre le problème suivant. Étant donné $A \in \mathbb{R}^{m \times n}$ et $k \in \mathbb{N}$, trouver une matrice $B \in \mathbb{R}^{m \times n}$ de $\text{rang}(B) \leq k$ tel que

$$\|A - B\|_F$$

soit minimale.

Si $A = P \cdot \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \cdot Q$ est une décomposition en valeurs singulières où les colonnes de P sont v_1, \dots, v_m et les lignes de Q sont u_1^T, \dots, u_n^T on peut écrire

$$A = \sum_{i=1}^r \sigma_i v_i u_i^T \quad (2.10)$$

et on dénote la somme des premiers k termes comme

$$A_k = \sum_{i=1}^k \sigma_i v_i u_i^T$$

Le rang de A_k est au plus k .

Lemme 2.22. *Les lignes de A_k sont les projections des lignes de A dans le sous-espace $V_k = \text{span}\{u_1, \dots, u_k\}$.*

Démonstration. Soit a^T une ligne de A . La projection de a dans le sous-espace $\text{span}\{u_1, \dots, u_k\}$ est

$$\sum_{i=1}^k a^T u_i \cdot u_i^T.$$

Alors les projections des lignes de A dans le sous-espace V_k sont données par $A u_i u_i^T = \sum_{i=1}^k \sigma_i u_i v_i^T = A_k$. \square

Théorème 2.23. *Pour une matrice $B \in \mathbb{R}^{m \times n}$ de rang plus petit ou égal à k on a*

$$\|A - A_k\|_F \leq \|A - B\|_F.$$

Démonstration. On dénote les lignes de A par a_1^T, \dots, a_m^T et soit B une matrice de rang au plus k . Les lignes de B sont dénotées comme b_1^T, \dots, b_m^T . Soit $H = \text{span}\{b_1, \dots, b_m\}$. La dimension de H est $\text{rang}(B) \leq k$. On a

$$\|A - B\|_F^2 = \sum_{i=1}^m \|a_i - b_i\|^2 \geq \sum_{i=1}^m d(a_i, H)^2.$$

Soit $\tilde{H} = \text{span}\{u_1, \dots, u_k\}$. Nous avons démontré que

2 Le théorème spectral et la décomposition en valeurs singulières

- i) \tilde{H} est le meilleur sous-espace approximatif des lignes de A alors $\sum_{i=1}^m d(a_i, H)^2 \geq \sum_{i=1}^m d(a_i, \tilde{H})^2$ et
- ii) Les lignes de A_k sont les projections des lignes de A dans \tilde{H} .

En dénotant les lignes de A_k $\tilde{a}_1^T, \dots, \tilde{a}_m^T$, alors

$$\|A - B\|_F^2 \geq \sum_{i=1}^m d(a_i, \tilde{H})^2 = \|a_i - \tilde{a}_i\|^2 = \|A - A_k\|_F^2.$$

□

Exercices

1. Est-ce que la décomposition en valeurs singulières est unique ? Est-ce que les valeurs singulières sont uniques ?
2. Dans la démonstration du théorème 2.14, montrer que $\text{rang}(A) = r$.
3. Démontrer que la pseudo-inverse satisfait les conditions ii-iv.
4. Si $Ax = b$ a plusieurs solutions, il existe une solution unique avec une norme minimale.
5. Montrer Théorème 2.18.
6. Soient $G, H \subseteq \mathbb{R}^n$ des sous-espaces de \mathbb{R}^n et $k = \dim(G) > \dim(H)$. Montrer que G possède une base orthonormale w_1, \dots, w_k telle que $w_k \perp H$.

3 Systèmes différentiels linéaires

Cours du 17. Mai 2016

On considère le système suivant

$$\begin{aligned} \mathbf{x}'_1(t) &= a_{11}\mathbf{x}_1(t) + \cdots + a_{1n}\mathbf{x}_n(t) \\ \mathbf{x}'_2(t) &= a_{21}\mathbf{x}_1(t) + \cdots + a_{2n}\mathbf{x}_n(t) \\ &\vdots \\ \mathbf{x}'_n(t) &= a_{n1}\mathbf{x}_1(t) + \cdots + a_{nn}\mathbf{x}_n(t) \end{aligned} \tag{3.1}$$

où les $a_{ij} \in \mathbb{R}$. En notation de vecteur matrice on peut écrire ça comme

$$\mathbf{x}' = A \mathbf{x}$$

où

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

On cherche des fonctions dérivables $\mathbf{x}_i : \mathbb{R} \rightarrow \mathbb{R}$ qui, ensembles constituent \mathbf{x} et qui satisfont (3.1). Un tel \mathbf{x} est une *solution* du système (3.1).

Exemple 3.1. Considérons l'équation différentielle $\mathbf{x}'(t) = \mathbf{x}(t)$. Une solution est $\mathbf{x}(t) = e^t$. Une autre solution est $\mathbf{x}(t) = 2 \cdot e^t$. Si on spécifie la *condition initiale* $\mathbf{x}(0) = 1$, alors $\mathbf{x}(t) = e^t$ est la solution qui satisfait cette condition initiale. Autrement, si on spécifie $\mathbf{x}(0) = \alpha$, alors $\mathbf{x}(t) = \alpha \cdot e^t$ est une solution qui satisfait la condition initiale.

Considérons $\mathbf{x}'(t) = -\mathbf{x}(t)$, une solution est $\mathbf{x}(t) = e^{-t}$. C'est aussi une solution qui respecte la condition initiale $\mathbf{x}(0) = 1$.

Essayons d'abord de résoudre le système comme suivant $\mathbf{x}(t) = e^{\lambda t} \mathbf{v}$ où $\mathbf{v} \in \mathbb{R}^n$ est un vecteur constant. Dans ce cas $\mathbf{x}' = A\mathbf{x}$ se récrit comme $\lambda e^{\lambda t} \mathbf{v} = e^{\lambda t} A \mathbf{v}$. Nous avons démontré le lemme suivant.

Lemme 3.1. Si $\lambda \in \mathbb{R}$ est une valeur propre de A et si $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ est un vecteur propre correspondant, alors $\mathbf{x}(t) = e^{\lambda t} \mathbf{v}$ est une solution du système (3.1) pour les conditions initiales $\mathbf{x}(0) = \mathbf{v}$.

Le théorème suivant est démontré en cours *analyse 2*.

Théorème 3.2 (Cours d'analyse II). Étant données les conditions initiales $\mathbf{x}(0)$ il existe une solution \mathbf{x} unique du système (3.1).

3 Systèmes différentiels linéaires

Nous sommes concernés avec le problème de *trouver* la solution \mathbf{x} explicitement. On commence avec une observation qui est un exercice simple.

Lemme 3.3. *L'ensemble $\mathcal{X} = \{\mathbf{x} : \mathbf{x} \text{ est une solution du système (3.1)}\}$ est un espace vectoriel sur \mathbb{R} .*

Est-ce que c'est possible de donner une base de \mathcal{X} explicitement ? Dans le cas où A est diagonalisable comme

$$A = P \cdot \text{diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1}$$

où $P \in \mathbb{R}^{n \times n}$ est inversible et les $\lambda_i \in \mathbb{R}$ sont particulièrement agréables comme c'est décrit dans le lemme suivant.

Théorème 3.4. *Si \mathbb{R}^n possède une base $\{v_1, \dots, v_n\} \subseteq \mathbb{R}^n$ de vecteurs propres de A telle que $Av_i = \lambda_i v_i$, alors*

$$\mathbf{x}^{(i)}(t) = e^{\lambda_i t} \cdot v_i, \quad i = 1, \dots, n$$

est une base de \mathcal{X} .

Démonstration. Montrons d'abord que les $\mathbf{x}^{(i)}$ sont linéairement indépendants. Supposons que $\sum_i \alpha_i \mathbf{x}^{(i)} = 0$. C'est à dire que les n fonctions qui sont les composantes de $\sum_i \alpha_i \mathbf{x}^{(i)}$ sont toutes la fonction $f(x) = 0$. Dès que $e^{\lambda_i 0} = 1$, ça implique que

$$0 = \sum_i \alpha_i v_i e^{\lambda_i 0} = \sum_i \alpha_i v_i.$$

Mais les v_i sont linéairement indépendants. Alors $\alpha_i = 0$ pour tout i ce qui démontre que les $\mathbf{x}^{(i)}$ sont linéairement indépendants.

Maintenant soit $\mathbf{y} \in \mathcal{X}$ et soient $\alpha_i \in \mathbb{R}$ tels que

$$\mathbf{y}(0) = \sum_i \alpha_i v_i.$$

Alors $\mathbf{x} = \sum_i \alpha_i \mathbf{x}^{(i)} \in \mathcal{X}$ et dès que $\mathbf{x}(0) = \mathbf{y}(0)$, le Théorème 3.2 implique $\mathbf{x} = \mathbf{y}$. C'est à dire que les $\mathbf{x}^{(i)}$ engendrent \mathcal{X} , alors $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\}$ est une base de \mathcal{X} . \square

Est-ce qu'on peut aussi trouver une solution dans la cas où A est diagonalisable dans les nombres complexes, donc si

$$A = P \cdot \text{diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1}$$

où $P \in \mathbb{C}^{n \times n}$ est inversible et les $\lambda_i \in \mathbb{C}$? Pour discuter de ça, il faut d'abord définir, ce qu'est une solution complexe du système (3.1). Toute fonction $f : \mathbb{R} \rightarrow \mathbb{C}$ s'écrit comme

$$f(x) = f_{\Re}(x) + i \cdot f_{\Im}(x)$$

où $f_{\Re}(x), f_{\Im}(x)$ sont des fonctions de $\mathbb{R} \rightarrow \mathbb{R}$. Si f_{\Re} et f_{\Im} sont dérivables, on dit que $f(x)$ est dérivable et on définit

$$f'(x) = f'_{\Re}(x) + i \cdot f'_{\Im}(x).$$

Si $\mathbf{x}_1, \dots, \mathbf{x}_n: \mathbb{R} \rightarrow \mathbb{C}$ sont dérivables, comme avant

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$$

est une *solution complexe* du système (3.1) si $\mathbf{x}' = A\mathbf{x}$. Et comme avant, on peut noter le lemme suivant, en se rappelant que $e^{a+ib} = e^a(\cos b + i \cdot \sin b)$.

Lemme 3.5. *Si $\lambda \in \mathbb{C}$ est une valeur propre de A et si $v \in \mathbb{C}^n \setminus \{0\}$ est un vecteur propre correspondant, alors $\mathbf{x}(t) = e^{\lambda t}v$ est une solution du système (3.1) pour les conditions initiales $\mathbf{x}(0) = v$.*

Démonstration. On écrit

$$\mathbf{x}' = \lambda e^{\lambda t}v = e^{\lambda t}Av = A\mathbf{x}.$$

□

Lemme 3.6. *Étant donnée une solution complexe $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ du système (3.1), alors \mathbf{x}_{\Re} et \mathbf{x}_{\Im} sont des solutions réelles.*

Démonstration. Dès que $\mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ est une solution, on a

$$\mathbf{x}'_{\Re} + i\mathbf{x}'_{\Im} = \mathbf{x}' = A\mathbf{x} = A\mathbf{x}_{\Re} + A\mathbf{x}_{\Im}.$$

Dès que A est réelle on voit $\mathbf{x}'_{\Re} = A\mathbf{x}_{\Re}$ et $\mathbf{x}'_{\Im} = A\mathbf{x}_{\Im}$.

□

Supposons alors que $A \in \mathbb{R}^{n \times n}$ est diagonalisable. Et soit $\{v_1, \dots, v_n\}$ une base de \mathbb{C}^n de vecteurs propres associés à $\lambda_1, \dots, \lambda_n$ respectivement. Si $v_i = u_i + i \cdot w_i$ où $u_i, w_i \in \mathbb{R}^n$, les $u_1, \dots, u_n, v_1, \dots, v_n$ engendrent \mathbb{R}^n , voir exercice 3. Comme nous avons noté

$$\mathbf{x}^{(j)} = e^{\lambda_j t}v_j$$

sont des solutions complexes du système (3.1).

Aussi, on peut supposer que la base et les valeurs propres sont tels que les vecteurs/valeurs propres complexes viennent en paires conjuguées complexes. Plus précisément

$$v_{2j+1} = \overline{v_{2j}} \text{ et } \lambda_{2j+1} = \overline{\lambda_{2j}} \text{ pour } 1 \leq j \leq n/2 \quad (3.2)$$

et

$$v_j \in \mathbb{R}^n, \lambda_j \in \mathbb{R} \text{ pour } j > 2k + 1. \quad (3.3)$$

3 Systèmes différentiels linéaires

Considérons maintenant une solution impliquée par $v = u + iw$ $\lambda = a + ib$.

$$\begin{aligned}\mathbf{x} &= e^{at} (\cos(bt) + i \sin(bt)) (u + iw) \\ &= e^{at} (\cos(bt)u - \sin(bt)w) + ie^{at} (\sin(bt)u + \cos(bt)w).\end{aligned}$$

Ça nous donne alors ces deux solutions réelles

$$\begin{aligned}\mathbf{x}^{(1)} &= e^{at} (\cos(bt)u - \sin(bt)w) \\ \mathbf{x}^{(2)} &= e^{at} (\sin(bt)u + \cos(bt)w).\end{aligned}$$

Remarque 3.7. Les solutions réelles impliquées par v et λ sont les mêmes que les solutions réelles impliquées par \bar{v} et $\bar{\lambda}$.

Nous pouvons alors noter une marche à suivre pour résoudre le système (3.1) étant donné $\mathbf{x}(0)$ si A est diagonalisable.

1. Trouver une base de vecteurs propres v_1, \dots, v_n de A ordonnée comme dans (3.2) et (3.3).
2. Pour chaque paire v_j, λ_j , $1 \leq j \leq k$ trouver les deux solutions réelles dénotées comme $\mathbf{x}^{(2j)}$ et $\mathbf{x}^{(2j+1)}$.
3. Pour chaque paire réelle v_j, λ_j $n \geq j > 2k + 1$, trouver la solution $\mathbf{x}^{(j)}$.
4. Trouver la combinaison linéaire

$$\mathbf{x}(0) = \sum_j \alpha_j \mathbf{x}^{(j)}(0)$$

5. La solution est

$$\mathbf{x} = \sum_j \alpha_j \mathbf{x}^{(j)}$$

Exemple 3.2. Résoudre le système $\mathbf{x}' = A\mathbf{x}$ où

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \text{ et } \mathbf{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

On trouve que $\lambda_1 = 1 + 2i$ et $\lambda_2 = 1 - 2i$ sont les valeurs propres de A et

$$v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ et } v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

sont les vecteurs propres correspondants. Les deux solutions impliquées par v_1 sont

$$\begin{aligned}\mathbf{x}^{(1)} &= e^t \left(\cos(2t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \sin(2t) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ \mathbf{x}^{(2)} &= e^t \left(\sin(2t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \cos(2t) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).\end{aligned}$$

La solution qu'on cherche est

$$\mathbf{x} = \begin{pmatrix} e^t \sin(2t) + e^t \cos(2t) \\ -e^t \sin(2t) + e^t \cos(2t) \end{pmatrix}.$$

Exercices

1. Montrer Lemme 3.3.
2. Une fonction $f : \mathbb{C} \longrightarrow \mathbb{C}$ est *holomorphe* en $z_0 \in \mathbb{C}$ si

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existe. Soit f holomorphe sur \mathbb{C} et $g = f|_{\mathbb{R}}$ la fonction f réduite à \mathbb{R} . Montrer

- i) $g(x) = g_{\Re}(x) + i \cdot g_{\Im}(x)$ est dérivable au sens de notre définition, particulièrement $g_{\Re}(x)$ et $g_{\Im}(x)$ sont dérivables.
 - ii) $f'_{|\mathbb{R}}(x) = g'_{\Re}(x) + i \cdot g'_{\Im}(x)$.
3. Soit $\{u_1 + i \cdot w_1, \dots, u_n + i \cdot w_n\}$ une base de \mathbb{C}^n où $u_i, w_i \in \mathbb{R}^n$ pour tout i . Montrer que $\text{span}\{u_i, w_i : 1 \leq i \leq n\} = \mathbb{R}^n$.

3.1 L'exponentielle d'une matrice

Cours du 24. Mai 2016

Définition 3.1. Pour $A \in \mathbb{C}^{n \times n}$ on définit

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$$

On rappelle la définition d'une série intégrable

$$\sum_{j=0}^{\infty} a_j z^j,$$

où les coefficients $a_j \in \mathbb{C}$, et qui converge sur un *disque* de rayon ρ . C'est à dire que, si $|z| < \rho$ la série converge et la fonction $f : \{x \in \mathbb{C} : |x| < \rho\} \rightarrow \mathbb{C}$ définie par $f(x) = \sum_{j=0}^{\infty} a_j x^j$ est *holomorphe* avec dérivée $f'(x) = \sum_{j=0}^{\infty} j a_j x^{j-1}$. Une série intégrable importante est la série

$$e^x = \sum_{j=0}^{\infty} \frac{1}{j!} x^j,$$

qui définit la fonction holomorphe $\exp : \mathbb{C} \longrightarrow \mathbb{C}$

$$e^{a+ib} = e^a (\cos b + i \sin b).$$

On va maintenant généraliser la définition de la *norme Frobenius* pour les matrices complexes. Pour $A \in \mathbb{C}^{m \times n}$,

$$\|A\|_F = \sqrt{\sum_{ij} |a_{ij}|^2}.$$

3 Systèmes différentiels linéaires

Lemme 3.8. Pour $A, B \in \mathbb{C}^{m \times n}$ on a

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

Démonstration. Soient $a_1^T, \dots, a_n^T \in \mathbb{C}^m$ les lignes de A et $\bar{b}_1, \dots, \bar{b}_n \in \mathbb{C}^m$ les colonnes de B . Avec Cauchy-Schwarz

$$|(AB)_{ij}|^2 = (a_i^T \bar{b}_j)(\bar{a}_i^T b_j) \leq \|a_i\|^2 \|b_j\|^2$$

et donc

$$\|AB\|_F^2 = \sum_{ij} |(AB)_{ij}|^2 \leq \sum_i \|a_i\|^2 \cdot \sum_j \|b_j\|^2 = \|A\|_F^2 \cdot \|B\|_F^2.$$

□

Lemme 3.9. La série e^A converge.

Démonstration.

$$\begin{aligned} \left\| \sum_{k=m}^{\infty} \frac{1}{k!} A^k \right\|_F &\leq \sum_{k=m}^{\infty} \frac{1}{k!} \|A^k\|_F \\ &\leq \sum_{k=m}^{\infty} \frac{1}{k!} \|A\|_F^k. \end{aligned}$$

Alors

$$\lim_{m \rightarrow \infty} \left\| \sum_{k=m}^{\infty} \frac{1}{k!} A^k \right\|_F = 0.$$

□

Nous avons montré que $e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ converge pour tout $t \in \mathbb{R}$. Plus précisément chaque composante $\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ est une série intégrable avec un rayon de convergence ∞ . Alors, nous pouvons dériver les éléments pour obtenir

$$\frac{d}{dt} e^{At} = A e^{At}. \quad (3.4)$$

Théorème 3.10. La solution du problème initial $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = v$ est

$$\mathbf{x}(t) = e^{At} v.$$

Démonstration. Soit $\mathbf{x}(t) = e^{At} v$. Alors $\mathbf{x}'(t) = A e^{At} v = A \mathbf{x}(t)$. Plutôt $\mathbf{x}(0) = v$. □

Définition 3.2. Une matrice N est *nilpotente* s'il existe un $k \in \mathbb{N}$ tel que $N^k = 0$.

Nous allons montrer ce théorème dans le prochain cours.

Théorème 3.11. *Chaque matrice $A \in \mathbb{C}^{n \times n}$ peut être factorisée comme*

$$A = P(\text{diag}(\lambda_1, \dots, \lambda_n) + N)P^{-1}$$

où $N \in \mathbb{C}^{n \times n}$ est nilpotente, $P \in \mathbb{C}^{n \times n}$ est inversible, $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ sont les valeurs propres de A et $\text{diag}(\lambda_1, \dots, \lambda_n)$ et N commutent.

Lemme 3.12. *Pour $A, B \in \mathbb{C}^{n \times n}$, si $A \cdot B = B \cdot A$ on a $e^{A+B} = e^A e^B$.*

Comment peut-on maintenant résoudre le problème initial $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = v$ explicitement? Nous savons que cette solution est $\mathbf{x} = e^{tA} \cdot v$ et nous savons que c'est une solution réelle pour $A \in \mathbb{R}^{m \times n}$. Mais les premiers termes s'écrivent comme

$$\sum_{i=0}^m t^i A^i = P \left(\sum_{i=0}^m t^i \text{diag}(\lambda_1, \dots, \lambda_n) + t^i N \right) P^{-1}$$

où nous avons utilisé le théorème 3.11. Dès que N et $\text{diag}(\lambda_1, \dots, \lambda_n)$ commutent, la solution *réelle* que l'on cherche est

$$\begin{aligned} \mathbf{x} &= P e^{t \text{diag}(\lambda_1, \dots, \lambda_n)} e^{tN} P^{-1} v \\ &= P \left(\text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \cdot \sum_{j=0}^{k-1} t^j N^j / j! \right) P^{-1}, \end{aligned}$$

où $k \in \mathbb{N}$ est tel que $N^k = 0$.

3.2 Polynômes

Soit K un corps. On dénote l'anneau des polynômes de K par $K[x]$. Un élément de $K[x]$ s'écrit comme

$$p(x) = a_0 + a_1 x + \dots + a_n x^n$$

où les *coefficients* $a_i \in K$. Le *degré* de $p(x) \neq 0$ est

$$\deg(p) = \max\{i : a_i \neq 0\}$$

et $\deg(0) = -\infty$. Si $p \neq 0$, le coefficient $a_{\deg(p)}$ est le *coefficient dominant* de p . Un polynôme de degré zéro est une *constante*.

La formule de multiplication de deux polynômes $f(x) = a_0 + a_1 x + \dots + a_n x^n$ et $g(x) = b_0 + \dots + b_m x^m$ est

$$f(x) \cdot g(x) = \sum_{i=0}^{m+n} \left(\sum_{k+l=i} a_k b_l \right) x^i \quad (3.5)$$

Théorème 3.13. *Pour $f, g \in K[x]$, $\deg(f \cdot g) = \deg(f) + \deg(g)$.*

Démonstration. La formule (3.5) révèle que $\deg(f \cdot g) \leq \deg(f) + \deg(g)$. Soient $f(x) = a_0 + \dots + a_n x^n$ et $g(x) = b_0 + \dots + b_m x^m$ tels que $a_n, b_m \neq 0$. Le coefficient de x^{n+m} est $a_n \cdot b_m \neq 0$. \square

Définition 3.3. Un polynôme $f(x) \in K[x]$ tel que $\deg(f) \geq 1$ est *irréductible* si

$$f(x) = g(x) \cdot h(x)$$

implique $\deg(g) \cdot \deg(h) = 0$, alors un des facteurs est une constante.

La division avec reste est l'opération suivante.

Théorème 3.14. Soient $f, g \in K[x]$ et $\deg(g) > 0$. Il existe $q, r \in K[x]$ tels que

$$f(x) = q(x)g(x) + r(x)$$

et $\deg(r) < \deg(g)$.

Démonstration. La preuve se fait par induction sur $\deg(f)$. Si $\deg(f) < \deg(g)$, alors on pose $q = 0$ et $r = f$.

Soit alors $\deg(f) = n \geq \deg(g) = m$ et

$$f(x) = a_0 + \cdots + a_n x^n \text{ et } g(x) = b_0 + \cdots + b_m x^m$$

où a_n et b_m sont les coefficients dominants de f et g respectivement. Clairement

$$\deg \left(f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \right) < \deg(f(x))$$

et par induction

$$f(x) - \frac{a_n}{b_m} x^{n-m} g(x) = q(x)g(x) + r(x)$$

tel que $\deg(r(x)) < \deg(g(x))$. On obtient alors

$$f(x) = (q(x) + \frac{a_n}{b_m} x^{n-m})g(x) + r(x).$$

□

Définition 3.4. Pour $f(x) = a_0 + \cdots + a_n x^n \in K[x]$ et $\alpha \in K$, l'évaluation $f(\alpha)$ est $a_0 + a_1 \alpha + \cdots + a_n \alpha^n \in K$.

Définition 3.5. Soit $f(x) \in \mathbb{K}[x] \setminus \{0\}$. Un $\alpha \in K$ tel que $f(\alpha) = 0$ est une *racine* de $f(x)$.

Définition 3.6. Un polynôme $q(x)$ *divise* un autre polynôme $f(x)$ s'il existe un polynôme $g(x)$ tel que $f(x) = g(x) \cdot q(x)$. On dit que $q(x)$ est un diviseur de $f(x)$ et on écrit $q(x) \mid f(x)$.

Théorème 3.15. Soit $f(x)$ un polynôme et $\alpha \in K$, alors α est une racine de f si et seulement si $(x - \alpha) \mid f(x)$.

Démonstration. Si $f(x) = q(x) \cdot (x - \alpha)$, alors $f(\alpha) = 0$.

Dans l'autre sens, si f est une constante, $f(\alpha) = 0$ implique que $f = 0$ et $(x - \alpha)$ divise f .

Si f n'est pas une constante, il existe $q(x)$ et $r(x)$ tels que

$$f(x) = q(x) \cdot (x - \alpha) + r(x)$$

avec $\deg(r) \leq 0$. Alors $f(\alpha) = 0$ implique $r = 0$. \square

Définition 3.7. Un diviseur commun de $a(x) \in K[x]$ et $b(x) \in K[x]$ est un diviseur de $a(x)$ et $b(x)$. Un diviseur commun le plus grand de $a(x)$ et $b(x)$ est un diviseur commun de $a(x)$ et $b(x)$ tel que tous les autres diviseurs communs de $a(x)$ et $b(x)$ le divisent. On dénote les plus grands diviseurs communs de a et b par $\text{pgdc}(a, b)$ (ou, en anglais, $\text{gcd}(a, b)$, greatest common divisor).

Théorème 3.16. Soient $a(x), b(x)$ deux polynômes, tels que $\deg(a) + \deg(b) > 0$. Un polynôme

$$d(x) = g(x)a(x) + h(x)b(x) \neq 0 \tag{3.6}$$

où $g, h \in K[x]$ de degré minimale est un plus grand diviseur commun de a et b .

Démonstration. On montre qu'un tel $d(x)$ est un diviseur commun de a et b en procédant par l'absurde. Supposons que d ne divise pas a . Alors il existe q et r tels que

$$a = q \cdot d + r$$

et $\deg(r) < \deg(d)$. Alors

$$r = a - q \cdot d = (1 - gq)a - hqb$$

est un polynôme de la forme (3.6) avec un degré strictement plus petit que celui de d .

Il est clair que tous les diviseurs communs de a et b divisent b . \square

Théorème 3.17. Soit $p(x)$ irréductible et supposons que $p(x) \mid f(x) \cdot g(x)$, alors $p(x) \mid f(x)$ ou $p(x) \mid g(x)$.

Démonstration. Si $p(x)$ ne divise $f(x)$ ni $g(x)$ alors $1 = f(x)h_1(x) + p(x)h_2(x)$ et $1 = g(x)h_3(x) + p(x)h_4(x)$ alors $\text{gcd}(p(x), f(x)g(x)) = 1$. \square

Théorème 3.18. Un polynôme $f(x) \in K[x]$ a une factorisation

$$f(x) = a^* \prod_j p_j(x)$$

où $a^* \in K$ et les $p_j(x)$ sont irréductibles avec coefficient dominant 1. Cette factorisation est unique sauf pour des permutations des p_j .

Définition 3.8. Soit V un espace vectoriel sur un corps K , $A : V \rightarrow V$ un endomorphisme et $f(x) = a_0 + \cdots + a_n x^n \in K[x]$. L'évaluation de f sur A est l'endomorphisme $f(A) : V \rightarrow V$

$$f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 \cdot I_n \in K^{n \times n}.$$

Définition 3.9. Soit $A : V \rightarrow V$ un endomorphisme et $W \subseteq V$ un sous-espace de V . On dit que W est *invariant sous A* si $A(x) \in W$ pour tout $x \in W$.

Lemme 3.19. Soit $f(x) \in K[x]$ et $A : V \rightarrow V$ un endomorphisme, alors $\ker(f(A))$ est invariant sous A .

Démonstration. Si $v \in \ker(f(A))$ on trouve que $f(A)Av = Af(A)v = 0$. Alors, $Av \in \ker(f(A))$. \square

Théorème 3.20. Soit $f(x) = f_1(x) \cdot f_2(x)$ et $A : V \rightarrow V$ un endomorphisme tel que

$$i) \deg(f_1) \cdot \deg(f_2) \neq 0,$$

$$ii) \gcd(f_1, f_2) = 1$$

alors $\ker(f(A)) = \ker(f_1(A)) \oplus \ker(f_2(A))$.

Démonstration. Dès que $\gcd(f_1, f_2) = 1$ il existe $g_1(x), g_2(x)$ tels que

$$1 = g_1(x)f_1(x) + g_2(x)f_2(x)$$

et alors

$$g_1(A) \cdot f_1(A) + g_2(A)f_2(A) = I. \quad (3.7)$$

Pour $v \in \ker(f(A))$ alors

$$g_1(A) \cdot f_1(A) \cdot v + g_2(A)f_2(A) \cdot v = v.$$

Mais $g_1(A) \cdot f_1(A) \cdot v \in \ker(f_2(A))$ dès que

$$f_2(A) \cdot g_1(A) \cdot f_1(A) \cdot v = g_1(A) \cdot f_1(A)f_2(A) \cdot v = g_1(A)f(A)v = 0$$

et d'une manière similaire on voit que $g_2(A)f_2(A) \cdot v \in \ker(f_1(A))$. Il reste à démontrer que la somme est directe.

Soit alors $v \in \ker(f_1(A)) \cap \ker(f_2(A))$. L'équation (3.7) montre

$$v = g_1(A) \cdot f_1(A) v + g_2(A)f_2(A) v = 0,$$

qui démontre que la somme est directe. \square

Exercices

1. Montrer que $K[x]$ est un anneau avec $1_{K[x]} = 1_K$.
2. Montrer que $a(x) \in K[x]$ et $b(x) \in K[x]$ $\deg(a) + \deg(b) > 0$ possèdent exactement un diviseur commun le plus grand avec coefficient principal égal à 1_K .

3.3 La forme normale de Jordan

Définition 3.10. Un *bloc Jordan* est une matrice de la forme

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

où les éléments non décrits sont zéro.

Une matrice $A \in \mathbb{C}^{n \times n}$ est en *forme normale de Jordan* si A est en forme bloc diagonale, où tout les blocs sur la diagonale sont des blocs Jordan, alors une matrice de la forme

$$A = \begin{pmatrix} B_1 & & \\ & B_2 & \\ & & \ddots \\ & & & B_k \end{pmatrix}$$

où les matrices $B_j \in \mathbb{C}^{n_j \times n_j}$ sont des blocs Jordan.

Notre but est de montrer le théorème suivant.

Théorème 3.21. Soit $A \in \mathbb{C}^{n \times n}$, alors il existe des matrices $P, J \in \mathbb{C}^{n \times n}$ telles que J est en forme normale de Jordan, P est inversible et

$$A = P^{-1} J P.$$

Définition 3.11. Soit $V = \mathbb{C}^n$. Le *décalage* est l'application linéaire

$$U \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}.$$

Le décalage plus une constante est une application linéaire

$$U + \lambda \cdot I.$$

3 Systèmes différentiels linéaires

C'est facile à voir que la matrice représentant le décalage plus λ est un seul bloc Jordan

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

Lemme 3.22. Soit V un espace vectoriel de dimension finie sur \mathbb{C} et soit $T: V \longrightarrow V$ une application linéaire. Alors V est la somme directe de sous-espaces $V = V_1 \oplus \cdots \oplus V_K$ tels que

i) $T(V_i) \subseteq V_i$ pour tout i et

ii) $T|_{V_i}: V_i \longrightarrow V_i$ est de la forme $N_i + \lambda I$ où N_i est nilpotente.

Démonstration. Des que l'espace des applications linéaires sur V est un espace vectoriel sur \mathbb{C} de dimension finie il existe un $k \in \mathbb{N}$ tel que

$$I, T, T^2, \dots, T^k$$

sont linéairement dépendants. Alors il existe un polynôme $p(x) \in \mathbb{C}[x] \setminus \{0\}$ tel que $p(T) = 0$. Le théorème fondamental d'algèbre implique que

$$p(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$$

avec des λ_i différentes. Le diviseur le plus grand de $(x - \lambda_i)^{m_i}$ et $p(x)/(x - \lambda_i)^{m_i}$ est 1 pour $i \neq j$. En utilisant théorème 3.20 en $k - 1$ étapes alors

$$V = \ker p(T) = \ker(T - \lambda_1 I)^{m_1} \oplus \cdots \oplus \ker(T - \lambda_k I)^{m_k}$$

et avec $V_i = \ker(T - \lambda_i I)^{m_i}$ on a $V = V_1 \oplus \cdots \oplus V_K$ et i). □

Remarque 3.23. Lemme 3.22 démontre qu'il existe une base

$$\mathcal{B} = b_1^1, \dots, b_{\ell_1}^1, b_1^2, \dots, b_{\ell_2}^2, \dots, b_1^k, \dots, b_{\ell_k}^k$$

où $b_1^i, \dots, b_{\ell_i}^i$ est une base de V_i telle que la matrice $A_{\mathcal{B}}^T$ de T par rapport à la base \mathcal{B} est une matrice bloc diagonale

$$A_{\mathcal{B}}^T = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}$$

et les matrices $B_i \in \mathbb{C}^{n_i \times n_i}$ sont de la forme $B_i = N_i + \lambda_i I$ où les N_i sont nilpotentes.

Rappel : Si $\phi_{\mathcal{B}}$ est l'isomorphisme $\phi_{\mathcal{B}}: V \longrightarrow \mathbb{C}^n$, où $\phi_{\mathcal{B}}(x) = [x]_{\mathcal{B}}$ sont les coordonnées de x par rapport à la base \mathcal{B} , on a le diagramme suivant

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \downarrow \phi_{\mathcal{B}} & & \downarrow \phi_{\mathcal{B}} \\ \mathbb{C}^n & \xrightarrow{A_{\mathcal{B}}^T \cdot x} & \mathbb{C}^n \end{array}$$

Il est clair, qu'il faut s'occuper maintenant des applications linéaires

$$T|_{V_i}: V_i \longrightarrow V_i$$

qui sont de la forme $N + \lambda I$ pour une application nilpotente N . Le théorème suivant s'occupe des applications linéaires nilpotentes. La matrice de λI est toujours λI pour chaque base. Il est alors clair que le théorème suivant démontre le théorème 3.21.

Théorème 3.24. *Soit V un espace vectoriel sur \mathbb{C} de dimension finie et $N: V \longrightarrow V$ une application linéaire nilpotente. Alors V possède une base \mathcal{B} de la forme*

$$x_1, Nx_1, \dots, N^{m_1-1}x_1, x_2, Nx_2, \dots, N^{m_2-1}x_2, \dots, x_k, Nx_k, \dots, N^{m_k-1}x_k$$

telle que $N^{m_i}x_i = 0$ pour tout i .

Remarque 3.25. Si on inverse l'ordre de la base \mathcal{B} alors si on liste les éléments de droite à gauche on obtient une base \mathcal{B}' et la matrice $A_{\mathcal{B}'}^N$ de l'application N a la forme

$$A_{\mathcal{B}'}^N = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_k \end{pmatrix}$$

en forme normale de Jordan, où

$$J_i = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \in \mathbb{C}^{m_i \times m_i}.$$

Par conséquence $N + \lambda I$ est représentée par

$$A_{\mathcal{B}'}^{N+\lambda I} = A_{\mathcal{B}'}^N + \lambda I_n$$

en forme normale de Jordan.

Démonstration du Théorème 3.24. Pour $x \in V \setminus \{0\}$ on appelle

$$m_x = \min\{i: N^i x = 0\}$$

la *durée de vie* de x . La séquence

$$x, Nx, \dots, N^{m_x-1}x$$

est l'*orbite* de x sous N .

En concaténant les orbites des éléments d'une base, nous obtiendrons un ensemble de vecteurs qui engendre V . Supposons alors que nous avons un ensemble x_1, \dots, x_ℓ dont les orbites

$$x_1, Nx_1, \dots, N^{m_1-1}x_1, \dots, x_\ell, Nx_\ell, \dots, N^{m_\ell-1}x_\ell \quad (3.8)$$

engendrent V . Ici m_i est la durée de vie de x_i . Si (3.8) est linéairement dépendant nous allons remplacer un x_i par un vecteur y tel que

3 Systèmes différentiels linéaires

- i) Les orbites de $x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n$ engendrent ensemble V aussi,
- ii) la somme des durées de vie de $x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n$ est strictement plus petite que la somme des durées de vie de x_1, \dots, x_ℓ .

Comme ça nous avons montré le théorème parce que un tel procédé doit se terminer.

Dès que l'ensemble (3.8) est linéairement dépendant il existe une combinaison linéaire non triviale de (3.8) qui est égale à 0

$$0 = \beta_0^1 x_1 + \beta_1^1 N x_1 + \dots + \beta_{m_1-1}^1 N^{m_1-1} x_1 + \dots + \beta_0^\ell x_\ell + \beta_1^\ell N x_\ell + \dots + \beta_{m_\ell-1}^\ell N^{m_\ell-1} x_\ell$$

Maintenant, nous allons appliquer l'application N k -fois où k est le dernier moment, où pas tous les termes

$$\beta_i^j N^{k+i} x$$

sont égaux à zéro. Comme ça nous avons trouvé un sous-ensemble $J \subseteq \{1, \dots, k\}$ et des $\gamma_j \neq 0$ tels que

$$\sum_{j \in J} \gamma_j N^{m_j-1} x_j = 0.$$

Soit $m = \min_{j \in J} m_j - 1$ et soit $i \in J$ un index où le minimum est atteint. Alors

$$0 = N^m \sum_{j \in J} \gamma_j N^{m_j-1-m} x_j$$

et $\gamma_i x_i$ est un terme dans la somme en dessus.

Maintenant on remplace x_i avec $y = \sum_{j \in J} \gamma_j N^{m_j-1-m} x_j$. Il est facile de voir que les orbites de

$$x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n$$

engendrent encore V . Et la durée de vie de y est au plus $m < m_i$. On a alors démontré le théorème. \square

Exercices

1. Montrer que les orbites de $x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n$ engendrent encore V . (Voir démonstration du théorème 3.24).