Algèbre linéaire avancée II

Friedrich Eisenbrand

6 mai 2026

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 typographiques, 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
— Lucas Gehrt
— Jooyoung Kim
— Léo Navarro
— Arthur Serres
— Matthieu Haeberle
— Chady Bensaid
— Belarbi Zied
— Sébastien de Morsier
— Cedric Lehr
— Orso Renucci
— Anselm Albrecher
— Beatrice Serrurier
— Pablo Habib
— Vittorio Sandri

— Youssef Jamali

Le deuxième chapitre sur les valeurs propres est basé, en partie, sur les notes du cours de Daniel Kressner.

Table des matières

1 Rappel des notions fondamentales

On se rappelle de la notion d'un espace vectoriel V sur un corps K. Pendant notre cours, on travaille presque exclusivement avec des espaces vectoriels de dimension fini, $\dim(V) = n \in \mathbb{N} = \{0, 1, 2, ...\}$. Dans ce cas il existe une base finie

$$B = \{v_1, \ldots, v_n\}.$$

On considère cette base B comme ordonnée. Alors il y a un premier élément de cet ensemble B un deuxième etc.

Tout élément $v \in V$ possède une représentation unique comme *combinaison linéaire*

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n,$$

où $\alpha_i \in K$ pour $i = 1, \ldots, n$. Le vecteur

$$[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$$

est appelé les *coordonnées* de v *relatif à la base* B. Si B' est une autre base de V on peut obtenir les *coordonnées* de v relatif à la base B' à l'aide de la *matrice de changement de base* $P_{BB'} \in K^{n \times n}$. Pour tout $v \in V$ on a

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

Exemple 1.1. Supposons $\dim(V) = 3$ et $B = \{v_1, v_2, v_3\}$ et $B' = \{v_1', v_2', v_3'\}$ sont deux bases de V. Les v_i' sont représentées comme combinaisons linéaires des v_i comme suivant.

$$\begin{aligned} v_1' &= 2v_1 + v_2 - v_3 \\ v_2' &= v_1 + 3v_2 + v_3 \\ v_3' &= v_1 + v_2 - v_3. \end{aligned}$$

La matrice de changement de base $P_{B'B}$ est

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

On vérifie par exemple

$$[v_1']_{B'} = e_1 \text{ et } [v_1']_B = P_{B'B} \cdot [v_1,]_{B'} = P_{B'B} \cdot e_1.$$

Ici, pour $i = 1, \ldots, n$, $e_i \in K^n$ est le vecteur dont la i-ième composante est 1 et les autres sont tous égaux à 0. La matrice de changement de base $P_{BB'}$ est

$$P_{BB'} = P_{B'B}^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} \\ -1 & \frac{3}{4} & -\frac{5}{4} \end{pmatrix}.$$

1.1 Applications linéaires et matrices

Soient V et W deux espaces vectoriels sur un corps K. Une *application linéaire* est une fonction $f\colon V \longrightarrow W$ tel que

(l.1) Pour tous $u, v \in V$ on a $f(u+v) = f(u) + f(v)$.

(l.2) Pour tout $u \in V$ et $\alpha \in K$ on a $f(\alpha u) = \alpha f(u)$.

Si $B_V = \{v_1, \ldots, v_n\}$ est une base de V et $B_W = \{w_1, \ldots, w_m\}$ est une base de W et si

$$f(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m$$

alors pour tout $v \in V$ on vérifie facilement

$$[f(v)]_{B_W} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} [v]_{B_V}. \tag{1.1}$$

Exemple 1.2. Soient V et W deux espaces vectoriels sur $\mathbb{Z}_3$, $B_V = \{v_1, v_2, v_3, v_4\}$ et $B_W = \{u_1, u_2, u_3, u_4\}$ bases de V et W respectivement. Soit $f\colon V \longrightarrow W$ une application linéaire et les images des v_i comme suivant.

$$\begin{aligned} f(v_1) &= 1 \cdot u_1 + 2 \cdot u_2 + 2 \cdot u_3 + 2 \cdot u_4 \\ f(v_2) &= 2 \cdot u_1 + 1 \cdot u_2 + 1 \cdot u_3 + 1 \cdot u_4 \\ f(v_3) &= 0 \cdot u_1 + 2 \cdot u_2 + 0 \cdot u_3 + 2 \cdot u_4 \\ f(v_4) &= 1 \cdot u_1 + 2 \cdot u_2 + 2 \cdot u_3 + 2 \cdot u_4 \\ f(v_5) &= 2 \cdot u_1 + 2 \cdot u_2 + 1 \cdot u_3 + 2 \cdot u_4 \end{aligned}$$

Alors la matrice

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 2 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 2 & 2 \end{pmatrix} \in \mathbb{Z}_3^{4\times 5}$$

nous donne pour tout $v \in V$

$$[f(v)]_{B_W} = A \cdot [v]_{B_V}.$$

Par exemple on observe

$$\begin{aligned} f(v_1 + 2v_3 + 1v_5) &= 1 \cdot (1 \cdot u_1 + 2 \cdot u_2 + 2 \cdot u_3 + 2 \cdot u_4) \\ &\quad +2 \cdot (0 \cdot u_1 + 2 \cdot u_2 + 0 \cdot u_3 + 2 \cdot u_4) \\ &\quad +1 \cdot (2 \cdot u_1 + 2 \cdot u_2 + 1 \cdot u_3 + 2 \cdot u_4) \\ &= 0 \cdot u_1 + 2 \cdot u_2 + 0 \cdot u_3 + 2 \cdot u_4 \end{aligned}$$

et

$$\begin{aligned} A \cdot [v_1 + 2v_3 + 1v_5]_{B_V} &= A \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix} \\ &= [f(v)]_{B_W} \end{aligned}$$

1.2 L'élimination de Gauss

Rappelons l'élimination de Gauss. Étant donné une matrice $A \in K^{m \times n}$, cette algorithme transforme A en A' en *forme échelonée*. C'est à dire que :

i) Si A' possède r lignes non nuls ($\neq 0^T$), ces lignes sont les premiers r lignes de A'.

ii) Il existe une suite $1 \leq p_1 < p_2 < \cdots < p_k \leq n$ les *indices pivot* t.q. pour tout $i \in \{1, \ldots, r\}$, $a'_{i,p_i} \neq 0$ et pour tout $1 \leq j < p_i$ on a $a'_{i,p_i} = 0$.

On peut facilement prouver que l'algorithme est correcte. Tout d'abord, l'algorithme n'effectue que des opérations élémentaires sur les lignes de la forme

i) échanger deux lignes

ii) additionner un multiple (scalaire) d'une ligne à une **autre** ligne.

L'opération i), si les lignes i et j sont échangées, correspond à la multiplication

$$A' := T_{i,j} \cdot A',$$

où $T_{i,j} \in K^{m \times m}$ est obtenue à partir de la matrice identité $I \in K^{m \times m}$ en échangeant la i-ème et la j-ème ligne.

L'opération ii), si on additionne α fois la ligne i sur la ligne j, correspond à la multiplication

$$A' := L_{i,j}(\alpha) \cdot A',$$

où $L_{i,j}(\alpha) \in K^{m \times m}$ est obtenue à partir de la matrice identité $I \in K^{m \times m}$ en additionnant α *fois* la i-ème sur la j-ème ligne de I.

Algorithme 1 Élimination de Gauss

Entrée : $A \in K^{m \times n}$
Sortie : $A' \in K^{m \times n}$ sous forme échelonnée telle qu'il existe une matrice inversible $Q \in K^{m \times m}$ vérifiant $Q \cdot A = A'$.

$A' := A$
$i := 1$
tant que $(i \leq m)$
    trouver le *plus petit* $1 \leq j \leq n$ tel qu'il existe $k \geq i$ avec $a'_{kj} \neq 0$
    Si un tel élément n'existe pas, alors **arrêter**
    échanger les lignes i et k dans A'
    pour $k = i + 1, \ldots, m$
        soustraire (a'_{kj}/a'_{ij}) fois la ligne i de la ligne k dans A'
    $i := i + 1$
retourner : A'

Exemple 1.3. En $\mathbb{Q}^{4\times 4}$ on a

$$T_{1,3} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad L_{3,2}(-5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La matrice résultante A' peut être obtenue via

$$A' = Q \cdot A$$

avec une matrice $Q \in K^{m \times m}$ non singulière. En fait, si les opérations sont effectuées en multipliant A' avec les matrices élémentaires $P_1, \ldots, P_\mu$ à gauche, Q est le produit

$$Q = P_\mu \cdots P_1.$$

Par ailleurs, nous avons l'invariant suivant.

> Après chaque itération de la boucle **tant que**, la matrice H obtenue à partir des j premières colonnes de A' est sous forme échelonnée et les lignes $i, i + 1, \ldots, m$ de H sont entièrement nulles, voir la figure 1.1.

Combien d'opérations arithmétiques l'élimination de Gauss effectue-t-elle ? Soustraire un multiple d'une ligne à une autre ligne dans A' peut se faire en $O(n)$ opérations arithmétique du corps K. Ainsi, le nombre total d'opérations effectuées dans la boucle **tant que** est $O(m \cdot n)$. Il y a $O(m)$ itérations de la boucle **tant que**. Au total, cela montre que l'élimination de Gauss effectue $O(m^2 \cdot n)$ opérations.

Exemple 1.4 (Exemple 1.2 continué). On laisse l'algorithme de Gauss s'exécuter sur

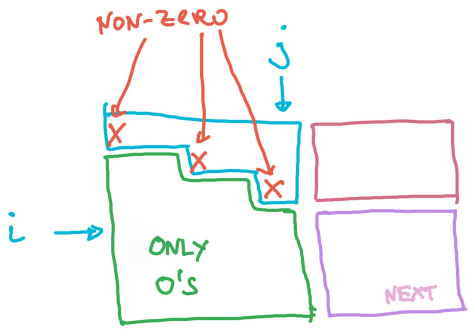

FIGURE 1.1 – Élimination de Gauss : la matrice A' avant la i-ème itération de la boucle **tant que**.

l'entrée

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 2 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 2 & 2 \end{pmatrix} \in \mathbb{Z}_3^{4\times 5},$$

et en calcule la matrice $Q \in \mathbb{Z}_3^{4\times 4}$ telle que $Q \cdot A$ est en forme échelonnée. Au début $A' := A$ et $Q := I_4$. On effectue les opérations suivantes sur A' et Q :

$\text{ligne}(2) := \text{ligne}(2) - 2 \cdot \text{ligne}(1)$

$\text{ligne}(3) := \text{ligne}(3) - 2 \cdot \text{ligne}(1)$

$\text{ligne}(4) := \text{ligne}(4) - 2 \cdot \text{ligne}(1)$.

Après nous avons

$$A' = \begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 \end{pmatrix} \quad \text{et } Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Enfin, après l'opération suivante sur A' et Q

$\text{ligne}(4) := \text{ligne}(4) - 1 \cdot \text{ligne}(2)$

on obtient

$$A' = \begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{et } Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}.$$

La matrice A' est en forme échelonnée et on vérifie

$$Q \cdot A = A'.$$

1.3 L'image et le noyaux d'une application linéaire

Soient encore V et W deux espaces vectoriels sur K de dimension fini et $f\colon V \longrightarrow W$ une application linéaire. On se rappelle de la définition de l'*image* de f

$$\operatorname{im}(f) = \{f(v)\colon v \in V\} \subseteq W.$$

L'image $\operatorname{im}(f)$ est un sous espace vectoriel de W. Le noyaux de f est défini comme

$$\ker(f) = \{v \in V\colon f(v) = 0\} \subseteq V.$$

le noyaux $\ker(f)$ est un sous espace vectoriel de V. On se rappelle du théorème suivant du premier semestre [Mic26].

Théorème 1.1. *Soient V et W deux espaces vectoriels sur K de dimension fini et $f\colon V \longrightarrow W$ une application linéaire. Alors*

$$\dim(V) = \dim(\operatorname{im}(f)) + \dim(\ker(f)).$$

Comment calculer des bases des espaces vectoriels $\operatorname{im}(f)$ et $\ker(f)$? Si $A' \in K^{m\times n}$ est en forme échelonnée avec des indices de pivot $1 \leq p_1 < p_2 < \cdots < p_k \leq n$ et si les colonnes de A' sont $a'_1, \ldots, a'_n \in K^m$, alors les vecteurs

$$a'_{p_1}, \ldots, a'_{p_k} \text{ sont linéairement indépendants.} \tag{1.2}$$

Aussi pour tout $j \in \{1, \ldots, n\} \setminus \{p_1, \cdots, p_k\}$, le vecteur $a'_j \in K^m$ est une combinaison linéaire des colonnes $a'_{p_1}, \ldots, a'_{p_\ell}$ où ℓ est maximal tq. $j > p_\ell$.

$$a'_j = \alpha_1^{(j)} a'_{p_1} + \cdots + \alpha_\ell^{(j)} a'_{p_\ell}.$$

Alors tous vecteurs

$$e_j - (\alpha_1^{(j)} e_{p_1} + \cdots + \alpha_\ell^{(j)} e_{p_\ell}) \quad j \in \{1, \ldots, n\} \setminus \{p_1, \cdots, p_k\} \tag{1.3}$$

sont dans le noyaux de A' et alors dans le noyaux de A. Cet ensemble de vecteurs est libre. Parce que le nombre total de vecteurs (1.2) et (1.3) est n, alors (1.2) est une base de l'image de A' et (1.3) est une base du noyau de A' (aussi de A respectivement).

Exemple 1.5 (Exemple 1.2 continué). On continue avec la matrice

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

Alors $\{e_1, e_3\}$ est une base de l'image de A' (et de A) et

$$\{1 \cdot u_1 + 2 \cdot u_2 + 2 \cdot u_3 + 2 \cdot u_4, 0 \cdot u_1 + 2 \cdot u_2 + 0 \cdot u_3 + 2 \cdot u_4\}$$

est une base de $\mathrm{im}(f)$. L'ensemble

$$\{e_2 + e_1, e_4 + 2e_1, e_5 + e_3 + e_1\}$$

est une base de $\ker(A') = \ker(A)$ et alors

$$\{v_2 + v_1, v_4 + 2v_1, v_5 + v_3 + v_1\}$$

est une base de noyau $\ker(f)$.

1.3.1 Endomorphismes, matrices semblables

Un *endomorphisme* est une application linéaire

$$f : V \longrightarrow V,$$

où V est un espace vectoriel sur K. Dans le cas où V est de dimension $n \in \mathbb{N}_+$ [1] et B est une base V on dénote la matrice décrite en (1.1) $A_B \in K^{n\times n}$. Pour tour $v \in V$ on a

$$[f(v)]_B = A_B^f \cdot [v]_B.$$

Pour une base B' et la matrice changement de base $P_{B'B}$ et son inverse $P_{B'B}^{-1} = P_{BB'}$, alors

$$\begin{aligned} [f(v)]_{B'} &= P_{B'B}^{-1} A_B P_{B'B} \cdot [v]_{B'} \\ A'_B &= P_{B'B}^{-1} A_B P_{B'B}. \end{aligned} \tag{1.4}$$

Définition 1.1. Deux matrices $A, B \in K^{n\times n}$ sont *semblables*, s'il existe une matrice inversible $P \in K^{n\times n}$ tel que $A = P^{-1} \cdot B \cdot P$. On écrit $A \approx B$.

L'équation (1.4) montre que, pour B et B' deux bases de V, les matrices A_B et $A_{B'}$ d'un endomorphisme $f : V \longrightarrow V$ sont semblables. La relation $\approx$ est une relation d'équivalence [Mic26].

Le thème principal de ce cours est la *diagonalisation*. Une matrice $A \in K^{n\times n}$ est diagonalisable si A est similaire à une matrice diagonale

$$A \approx \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

En endomorphisme f est diagonalisable si A_B est diagonalisable ce qui est équivalent au fait que V possède une base de vecteurs propres de f. Quand est-ce que un endomorphisme est diagonalisable ? La réponse à cette question révèle des relations profondes entre des concepts fondamentaux de l'algèbre, tels que les polynômes, les plus grands communs diviseurs et leur factorisation. Un résultat qui illustre ces liens est le suivant (Theorème 8.18).

1. $\mathbb{N}_+ = \{1, 2, 3, \dots\}$

Théorème 1.2. *Une matrice $A \in K^{n\times n}$ est diagonalisable, si et seulement si le polynôme minimal $p(x) \in K[x]$ de A factorise en termes linéaires distinctes, c.à.d.*

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

où $\lambda_1, \ldots, \lambda_k \in K$ sont les valeurs propres distinctes de A.

Pour y parvenir, nous commençons par un traitement approfondi des sujets relatifs aux polynômes.

Exercices

1. Complétez les omissions marquées avec $(*)$ pour que l'algorithme de Gauss calcule aussi la matrice de transition $Q \in K^{m\times m}$ t.q. $Q \cdot A = A'$ en forme échelonnée.

   Entrée : $A \in K^{m\times n}$
   Sortie : $A' \in K^{m\times n}$ sous forme échelonnée et $Q \in K^{m\times m}$ inversible
   vérifiant $Q \cdot A = A'$.
   $A' := A$
   $Q := I_m$
   $i := 1$
   tant que $= (i \leq m)$
   trouver le *plus petit* $1 \leq j \leq n$ tel qu'il existe $k \geq i$ avec $a'_{kj} \neq 0$
   Si un tel élément n'existe pas, alors **arrêter**
   échanger les lignes i et k dans A'
   $(*)$
   pour $k = i + 1, \ldots, m$
   soustraire (a'_{kj}/a'_{ij}) fois la ligne i de la ligne k dans A'
   $(*)$
   $i := i + 1$
   retourner : A' et Q

2. a) Quel est l'inverse de la matrice $T_{i,j} \in K^{n\times n}$, $i, j \in \{1, \ldots, n\}$, $i \neq j$?
   b) Quel est l'inverse de la matrice $L_{i,j}(\alpha) \in K^{n\times n}$, $i, j \in \{1, \ldots, n\}$, $i \neq j$ et $\alpha \in K$?
   c) Décrire en mots l'effet de la multiplication de ces matrices à *droite*, c.à.d. étant donnée $A \in K^{m\times n}$ décrire

   $$A \cdot T_{i,j} \quad \text{et} \quad A \cdot L_{i,j}(\alpha).$$

3. a) Une matrice $N \in K^{n\times n}$ est appelée *nilpotente* s'il existe un $k \in \mathbb{N}_+$ t.q. $N^k = 0$. Montrez que pour une telle matrice nilpotente

   $$(I + N)^{-1} = I - N + +N^2 - N^3 + \cdots + (-1)^{k-1} N^{k-1}.$$

   b) Basé sur 3a), donner une formule pour l'inverse R^{-1} dune matrice $R \in K^{n\times n}$ triangulaire supérieure avec 1 sur la diagonale.

c) Trouver une formule pour l'inverse R^{-1} dune matrice $R \in K^{n\times n}$ triangulaire supérieure avec des éléments diagonales non nuls.

4. Montrer que tout matrice $A \in K^{m\times n}$ peut être factorisée comme

$$U \cdot A \cdot V = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix},$$

où I_k est le rang de A et $U \in K^{m\times m}$ et $V \in K^{n\times n}$ sont des matrices inversibles.
Indication : Commencer avec la forme échelonnée.

5. Soient V et W deux espaces vectoriels sur $\mathbb{Z}_3$, $B_V = \{v_1, v_2, v_3, v_4\}$ et $B_W = \{u_1, u_2, u_3, u_4\}$ bases de V et W respectivement. Soit $f : V \longrightarrow W$ une application linéaire et les images des v_i comme suivant.

$$\begin{aligned} f(v_1) &= 1\cdot u_1 + 0\cdot u_2 + 2\cdot u_3 + 2\cdot u_4 \\ f(v_2) &= 2\cdot u_1 + 1\cdot u_2 + 2\cdot u_3 + 1\cdot u_4 \\ f(v_3) &= 0\cdot u_1 + 2\cdot u_2 + 0\cdot u_3 + 2\cdot u_4 \\ f(v_4) &= 1\cdot u_1 + 2\cdot u_2 + 1\cdot u_3 + 2\cdot u_4 \\ f(v_5) &= 2\cdot u_1 + 2\cdot u_2 + 1\cdot u_3 + 1\cdot u_4. \end{aligned}$$

Calculer bases du $\ker(f)$ et $\mathrm{im}(f)$.

2 Polynômes

2.1 Notions fondamentales

Soit R un anneau. On se souvient que cela signifie que R est un ensemble, muni des opérations binaires $+ : R \times R \to R$ et $\cdot : R \times R \to R$ telles que :

(R1) $a + b = b + a$ pour tout $a, b \in R$.

(R2) $a + (b + c) = (a + b) + c$, pour tout $a, b, c \in R$.

(R3) Il existe un élément $0 \in R$ tel que $0 + a = a$ pour chaque $a \in R$.

(R4) Pour chaque $a \in R$ il existe un élément $-a \in R$ tel que $a + (-a) = 0$.

(R5) $a(bc) = (ab)c$, pour tout $a, b, c \in R$.

(R6) Il existe un élément $1 \in R$ tel que $a \cdot 1 = 1 \cdot a = a$ pour chaque $a \in R$.

(R7) $a(b + c) = ab + ac$ et $(b + c)a = ba + ca$ pour tout $a, b, c \in R$.

Si, de plus, on a

(R8) $ab = ba$ pour tous $a, b \in R$.

alors l'anneau R est appelé *anneau commutatif*. Le *centre* de R est l'ensemble $Z(R) = \{r \in R : ra = ar \text{ pour tout } a \in R\}$. Un élément $a \neq 0$ de R est un *diviseur de zéro* s'il existe un $b \neq 0$ tel que $ab = 0$ ou $ba = 0$. Un anneau commutatif sans diviseurs de zéro et $1 \neq 0$ est appelé *anneau intègre*.

Exercice 2.1. Soit R un anneau, alors l'élément 1 est unique.

Exemple 2.1. i) Les nombres entiers $\mathbb{Z}$ avec l'addition et la multiplication standard forment un anneau commutatif sans diviseurs de zéro. Aussi $1 \neq 0$. $\mathbb{Z}$ est donc un anneau intègre.

ii) $5 \cdot \mathbb{Z} = \{5z : z \in \mathbb{Z}\}$ avec l'addition et la multiplication standard n'est pas un anneau. L'axiome (R6) n'est pas satisfait.

iii) L'ensemble des matrices $\mathbb{Z}^{2\times 2}$ avec l'addition et multiplication des matrices est un anneau non commutatif avec des diviseurs de zéro.

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

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

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

Un élément $r \in R$ est *inversible* s'il existe un élément $r^{-1} \in R$ tel que

$$r \cdot r^{-1} = r^{-1} \cdot r = 1.$$

On dénote l'ensemble des éléments inversibles par R^*. On se rappelle que $(R^*, \cdot)$ est un groupe, le *groupe des éléments inversibles*. Un anneau intègre tel que $R \setminus \{0\} = R^*$ est un *corps*.

Exercice 2.2. Soit R un anneau et $r \in R^*$. Alors r n'est pas un diviseur de zéro.

Exercice 2.3. Soit R un anneau et $R^{n\times n}$ l'anneau des matrices $n \times n$ sur R. Montrer que le centre de $R^{n\times n}$ est $Z(R^{n\times n}) = \{aI_n \colon a \in Z(R)\}$.

Le théorème suivant est trouvé dans les notes de cours du premier semestre [Mic26].

Théorème 2.1. *Soit R un anneau et $S \subseteq R$. Si les deux conditions suivantes sont vérifiées*

a) $1 \in S$

b) Pour $s, t \in S$ alors $s \cdot t \in S$ et $s - t \in S$.

alors S est aussi un anneau.

Le théorème suivant est démontré dans le cours *anneaux et corps*. Il nous donne une manière formelle d'introduire le concept d'une *indéterminée* (ou *variable*) qui n'est autre que ledit élément $x \in S \setminus R$.

Théorème 2.2. *Soit R un anneau, alors il existe un anneau $R' \supseteq R$ (R est un sous-anneau de R') satisfaisant $1_{R'} = 1_R$ et qui possède un élément $x \in R \setminus R$ tel que*

(i) $ax = xa$ pour chaque $a \in R$.

(ii) Si $a_0 + a_1x + \cdots + a_nx^n = 0$ et $a_i \in R$ pour tout i, alors $a_i = 0$ pour tout i.

Définition 2.1. Un polynôme sur R est une expression de la forme $p(x) = a_0 + a_1x + \cdots + a_nx^n$, où $n \in \mathbb{N}$, $a_i \in R$ pour tout i. L'élément a_i est le i-ème *coefficient* de $p(x)$. L'ensemble des polynômes sur R est dénoté par $R[x]$.

Exemple 2.2. i) $p(x) = 3 + x^2 + 5x^4 \in \mathbb{Z}[x]$.

ii) $p(x) = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 3 & 3 \\ 2 & 1 \end{smallmatrix}\right)x^3 \in \mathbb{Z}^{2\times 2}[x]$.

Proposition 2.3. *Deux polynômes*

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots \quad \text{et} \quad q(x) = b_0 + b_1x + b_2x^2 + \cdots \tag{2.1}$$

sont égaux *si et seulement si $a_i = b_i$ pour tout $i \in \mathbb{N}$. Dans ce cas, on écrit $p(x) = q(x)$.*

Exercice 2.4. Démontrez la proposition 2.3.

Théorème 2.4. *$R[x]$ est un anneau. Si R est commutatif, alors $R[x]$ est commutatif. Si R est anneau sans diviseur de zéro alors $R[x]$ est un anneau sans diviseur de zéro.*

Démonstration. Pour montrer que $R[x]$ est un anneau on invoque Theorème 2.1. Clairement $1 \in R[x]$ ce qui implique a). Pour b) il faut montrer que pour deux polynômes $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ et $g(x) = b_0 + b_1 x + \cdots + b_m x^m$ sur R, on a :

i) $f - g \in R[x]$ et

ii) $f \cdot g \in R[x]$.

La somme s'écrit comme

$$f(x) + g(x) = \sum_{i=0}^{\max\{m,n\}} (a_i + b_i)\, x^i \in R[x], \tag{2.2}$$

où $a_i = 0$ pour $i > n$ et $b_i = 0$ pour $i > m$. Alors on a $f - g \in R[x]$.

Le produit de f et g s'écrit comme

$$p(x) \cdot q(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \cdots . \tag{2.3}$$

Bref, on a la formule

$$f(x) \cdot g(x) = \sum_{i=0}^{m+n} \Big(\sum_{j+k=i} a_j b_k \Big)\, x^i \in R[x]. \tag{2.4}$$

C'est à dire que $R[x]$ est stable pour l'opération $\cdot$ de S. Alors, $R[x]$ est un sous-anneau de S et donc $R[x]$ est un anneau. La formule (2.4) implique que $R[x]$ est commutatif, si R est commutatif.

Finalement, si $f(x) = a_0 + a_1 x + \cdots \neq 0$ et $g(x) = b_0 + b_1 x + \cdots \neq 0$ sont deux polynômes non-nuls et si R est un anneau sans diviseur de zéro, il faut montrer que $f(x) \cdot g(x) \neq 0$. Soit $n = \max\{i\colon a_i \neq 0\}$ et $m = \max\{i\colon b_i \neq 0\}$. Le coefficient de x^{m+n} du polynôme $f \cdot g$ est $a_n \cdot b_m$. Ce coefficient n'est pas nul dès que R est un anneau sans diviseur de zéro. □

Exemple 2.3 (Multiplication de polynômes).

$$\begin{aligned} f(x) &= 3x^3 + x + 2 \in \mathbb{Z}[x] \\ g(x) &= 2x^4 + 2x^2 + 1 \in \mathbb{Z}[x] \\ f(x) \cdot g(x) &= 6x^7 + 8x^5 + 4x^4 + 5x^3 + 4x^2 + x + 2 \in \mathbb{Z}[x] \end{aligned}$$

Définition 2.2. Le *degré* de $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots \neq 0$ est

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

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*.

Exemple 2.4. Soit $f(x) = 2 + 3x + 5x^3 \in \mathbb{Z}[x]$, alors $\deg(f) = 3$ et le coefficient dominant de f est 5.

Théorème 2.5. *Soit R un anneau, $f, g \in R[x] \setminus \{0\}$ tel que le coefficient dominant de f ou de g n'est pas un diviseur de zéro. Alors,* $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Démonstration. Soient $f(x) = a_0 + \cdots + a_n x^n$ et $g(x) = b_0 + \cdots b_m x^m$ tels que $a_n, b_m \neq 0$. Le coefficient de x^{n+m} est $a_n \cdot b_m \neq 0$. Les coefficients de x^k, $k > n + m$ sont tous nuls. □

Un polynôme $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in R[x]$ induit une application $f_p : R \longrightarrow R$, $f_p(r) = a_0 + a_1 r + \cdots + a_n r^n$. Nous écrivons aussi $p(r)$ pour $f_p(r)$ et on parle de l'*évaluation* de p sur r.

Exemple 2.5. Soit $p(x) = x \in R[x]$ et $q(x) = (x + a) \in R[x]$, alors $p(x)q(x) = x^2 + ax$. Quand est-ce que l'on a $p(r)q(r) = (p \cdot q)(r)$? C'est le cas si et seulement si

$$r^2 + ra = r^2 + ar$$

c'est-à-dire si et seulement si $ra = ar$. On attire l'attention sur le fait que cet exemple traite de l'évaluation en $r \in R$ de polynômes. Ici, $p(r)q(r)$ dénote l'évaluation de p et q en r *puis* d'en multiplier les résultats. Tandis que $(p \cdot q)(r)$ représente la multiplication de p et q *puis* on évalue ce produit en r.

Théorème 2.6. *Soit R un anneau et $\alpha \in Z(R)$ un élément du centre de R. L'application*

$$\begin{array}{rccc} \Phi\colon & R[x] & \to & R \\ & f(x) & \mapsto & f(\alpha) \end{array}$$

est un morphisme d'anneaux surjectif.

Exercice 2.5. Démontrer le théorème 2.6.

2.2 Interpolation

Deux polynômes différents p et q peuvent induire la même application f_p et f_q, même s'ils sont des polynômes sur un corps K.

Exemple 2.6. Soit $K = \mathbb{Z}_2$, $p(x) = x + x^2$ et $q(x) = x^2 + x^3$ deux polynômes sur K. Il est clair que $p \neq q$. Mais $f_p : K \to K$ et $f_q : K \to K$ sont les mêmes applications car pour tout $x \in \mathbb{Z}_2$ on a $p(x) = q(x) = 0$.

Par contre, si K est un corps infini, deux polynômes différents induisent deux applications différentes. Nous allons voir les détails de ce fait maintenant.

Théorème 2.7. *Soit K un corps, $r_0, \ldots, r_n \in K$ des éléments distincts (c.-à-d. $r_i \neq r_j$ pour $i \neq j$) et $y_0, \ldots, y_n \in K$. Il existe exactement un seul polynôme $f(x) \in K[x]$ de degré au plus n tel que*

$$f(r_i) = y_i \text{ pour tout } i \in \{0, \ldots, n\}.$$

Démonstration. Un polynôme $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ satisfait $f(r_i) = y_i$ pour tout i si et seulement si le coefficients $a_0, \ldots, a_n$ satisfont

$$\begin{pmatrix} 1 & r_0 & \cdots & r_0^n \\ 1 & r_1 & \cdots & r_1^n \\ & & \cdots & \\ 1 & r_n & \cdots & r_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} \tag{2.5}$$

La matrice $V_{r_0,\ldots,r_n} \in K^{(n+1)\times(n+1)}$ à gauche de (2.5) est connue sous le nom de *matrice de Vandermonde* des éléments $r_0, \ldots, r_n$. Le théorème sera prouvé une fois que nous aurons démontré que $\det(V_{r_0,\ldots,r_n}) \neq 0$.

On démontre $\det(V_{r_0,\ldots,r_n}) \neq 0$ par récurrence sur n. Pour $n = 0$, le déterminant est 1. Pour $n > 0$, on soustrait r_0 fois la colonne n de la colonne $n+1$. Après, on soustrait r_0 fois la colonne $n-1$ de la colonne n etc. Le résultat est la matrice

$$\begin{pmatrix} 1 & 0 & \ldots\ldots\ldots\ldots & 0 \\ 1 & r_1 - r_0 & r_1(r_1 - r_0) & \cdots & r_1^{n-1}(r_1 - r_0) \\ & & & & \cdots \\ 1 & r_n - r_0 & r_n(r_n - r_0) & \cdots & r_n^{n-1}(r_n - r_0) \end{pmatrix} \tag{2.6}$$

Le déterminant de la matrice (2.6) est celle de $V_{r_0,\ldots,r_n}$. Développement du déterminant le long de la première ligne donne le déterminant de la matrice

$$\begin{pmatrix} r_1 - r_0 & r_1(r_1 - r_0) \cdots & r_1^{n-1}(r_1 - r_0) \\ & & \cdots \\ r_n - r_0 & r_n(r_n - r_0) \cdots & r_n^{n-1}(r_n - r_0) \end{pmatrix} \tag{2.7}$$

Alors

$$\det(V_{r_0,\ldots,r_n}) = (r_n - r_0) \cdots (r_1 - r_0) \det(V_{r_1,\ldots,r_n})$$

Comme les r_i sont distincts, le produit $(r_n - r_0) \cdots (r_1 - r_0)$ n'est pas zéro. Par l'hypothèse de récurrence, $\det(V_{r_1,\ldots,r_n}) \neq 0$ et donc $\det(V_{r_0,r_1,\ldots,r_n}) \neq 0$. □

Le théorème 2.7 ne contredit pas l'exemple 2.6 car le corps $\mathbb{Z}_2$ n'admet que 2 éléments. Il existe bel et bien un unique polynôme $p \in \mathbb{Z}_2[x]$ de degré inférieur ou égal à 1 tel que $p(0) = p(1) = 0$. Une énumération des polynômes possibles ou la résolution du système linéaire donne $p(x) = 0$.

Exercice 2.6. Montrer que le déterminant de $V_{r_0,\ldots,r_n}$ est

$$\det(V_{r_0,\ldots,r_n}) = \prod_{0 \leq i < j \leq n} (r_j - r_i).$$

Exemple 2.7. On cherche le polynôme $f(x) \in \mathbb{Z}_5[x]$ de degré au plus 3 tel que $f(0) = 2$, $f(1) = 2$, $f(2) = 2$ et $f(3) = 3$.

En trouvant l'unique solution du système

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & 4 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 3 \end{pmatrix},$$

on obtient

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}.$$

Ainsi, $f(x) = 2 + 2x + 2x^2 + x^3$ est le polynôme recherché.

Corollaire 2.8. *Soit K un corps infini. Deux polynômes $p(x), q(x) \in K[x]$ sont égaux, si et seulement si les applications f_p et f_q sont les mêmes.*

Démonstration. Si $p(x) = q(x)$, alors $p(r) = q(r)$ pour tout $r \in K$ et donc $f_p = f_q$. D'autre part, si $f_p = f_q$, on peut supposer qu'un des deux polynômes n'est pas le polynôme 0 (s'ils sont tous les deux nuls, alors $p = q$). Comme $p(r) = q(r)$ pour tout $r \in K$ et le corps K est infini, les polynômes prennent les mêmes valeurs sur $\max\{\deg(p), \deg(q)\} + 1$ éléments distincts de K. Il en résulte que $p(x) = q(x)$ par le Theorème 2.7. □

2.3 Division avec reste

Soit R un anneau et K un corps pour ce chapitre. La *division avec reste* est l'opération suivante.

Théorème 2.9. *Soient $f, g \in R[x]$, $g \neq 0$ et le coefficient dominant de g un élément inversible de R. Il existe $q, r \in R[x]$ uniques tels que*

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

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

Démonstration. Si $\deg(g) = 0$, alors $g(x) = b_0 \in R^*$. Dans ce cas, on met $q(x) = f(x) \cdot b_0^{-1}$ et $r(x) = 0$.

Autrement, soit $\deg(g) \geq 1$. La preuve se fait par récurrence 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

$$\deg\left(f(x) - a_n b_m^{-1} x^{n-m} g(x)\right) < \deg(f(x))$$

et par hypothèse de récurrence il existe $q_0, r \in R[x]$ tel que

$$f(x) - a_n {b_m}^{-1} x^{n-m} g(x) = q_0(x)g(x) + r(x)$$

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

$$\begin{aligned} f(x) &= \left(q_0(x) + a_n {b_m}^{-1} x^{n-m}\right) g(x) + r(x) \\ &= = q(x)g(x) + r(x), \end{aligned}$$

où $q(x) = q_0(x) + a_n {b_m}^{-1} x^{n-m}$. Supposons maintenant qu'il existe deux autres polynômes $q'(x) \neq q(x)$ et $r'(x) \neq r(x)$ tels que

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

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

$$r(x) - r'(x) = (q'(x) - q(x)) \cdot g(x).$$

Par le théorème 2.5

$$\deg(r - r') = \deg(q' - q) + \deg(g) \geq \deg(g),$$

ce qui contredit le fait que $\deg(r) < \deg(g)$ et $\deg(r') < \deg(g)$. □

Exemple 2.8. La division avec reste du polynôme $x^5 + 2x^2 + 1$ par $2x^3 + x + 1$ de $\mathbb{Z}_3[x]$ donne

$$x^5 + 2x^2 + 1 = (2x^2 + 2)(2x^3 + x + 1) + (x + 2).$$

On peut formuler la division avec reste comme un procédé récursif comme suivant. On suppose ici $\deg(g(x)) > 0$.

Algorithme 2 Division avec reste

```
1:  procedure DIV(f(x), g(x) )
2:      Soient n = deg(f) et m = deg(g)
3:      if n < m then
4:          return 0, f(x)
5:      else
6:          Soient a, b ∈ R coefficients dominants de f et g respectivement
7:          (q_0(x), r(x)) ← DIV(f(x) − (a/b)x^{n−m} g(x), g(x))
8:          return (a/b)x^{n−m} + q_o(x), r(x))
9:      end if
10: end procedure
```

Le Théorème suivant est une variante du Théorème 2.9 dans laquelle le *quotient* $q(x)$ est multiplié à droite, c.à.d. on exige

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

Théorème 2.10. *Soient $f, g \in R[x]$, $g \neq 0$ et le coefficient dominant de g un élément inversible de R. Il existe $q, r \in R[x]$ uniques tels que*

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

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

Exercice 2.7. Modifier la démonstration du Théorème 2.9 pour obtenir une démonstration du Théorème 2.10.

Exemple 2.9. C'est exemple montre que les quotients et les restes peuvent être divers dans les cas où le quotient est multiplié à droite et à gauche. Soit $\mathbb{Z}^{2\times 2}$ l'anneau des matrices 2×2 sur les entiers et

$$\begin{aligned} f(x) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot x^4 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{Z}^{2\times 2}[x] \\ g(x) &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot x^2 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{Z}^{2\times 2}[x] \end{aligned}$$

deux polynômes. Division avec reste (Théorème 2.9) nous donne $f(x) = q(x) \cdot g(x) + r(x)$ avec

$$q(x) = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} x^2 + \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix} \text{ et } r(x) = \begin{pmatrix} 2 & -2 \\ 0 & 1 \end{pmatrix}$$

Si le quotient $q(x)$ doit être multiplié à $g(x)$ à droite (Théorème 2.10), on obtient $f(x) = g(x) \cdot q(x) + r(x)$ avec

$$q(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x^2 + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \text{ et } r(x) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

2.4 Divisibilité et racines

Définition 2.3. Soit R un anneau commutatif. Un polynôme $q(x) \in R[x]$ *divise* un polynôme $f(x) \in R[x]$ s'il existe un polynôme $g(x) \in R[x]$ tel que $f(x) = g(x) \cdot q(x)$. On dit que $q(x)$ est un *diviseur* de $f(x)$ et on écrit $q(x) \mid f(x)$.

Exemple 2.10. Soient $q(x) = x^2 + 1 \in \mathbb{Z}_2[x]$ et $f(x) = x^3 + x^2 - x + 1 \in \mathbb{Z}_2[x]$. On a $f(x) = q(x)(x + 1)$ et donc $q(x) \mid f(x)$.

Définition 2.4. Soit R un anneau commutatif et $p(x) \in R[x] \setminus \{0\}$. Un $\alpha \in R$ tel que $f_p(\alpha) = 0$ est une *racine* de $f(x)$.

Exemple 2.11. Soit $p(x) = x^4 + x^3 + x^2 + x + 1 \in \mathbb{Z}_5(x)$, alors $\alpha = 1$ est une racine de $p(x)$.

Théorème 2.11 (Théorème fondamental de l'algèbre). *Tout polynôme $p(x) \in \mathbb{C}[x] \setminus \{0\}$ non constant admet au moins une racine complexe.*

Théorème 2.12. *Soit R un anneau commutatif et $f(x) \in R[x] \setminus \{0\}$ un polynôme et $\alpha \in R$, 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

$$\begin{aligned} f(\alpha) &= q(\alpha) \cdot (\alpha - \alpha) \\ &= 0. \end{aligned}$$

Pour l'autre sens, 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 $r = 0$. □

Définition 2.5. Soit R un anneau commutatif. La *multiplicité* d'une racine α de $p(x) \in R[x] \setminus \{0\}$ est le plus grand entier $i \geq 1$ tel que $(x-\alpha)^i \mid p(x)$.

Exemple 2.12 (Suite de l'exemple 1.10). Le polynôme $p(x) = x^4 + x^3 + x^2 + x + 1 \in \mathbb{Z}_5$ est divisible par $x - 1$, et

$$p(x) = (x^3 + 2x^2 + 3x + 4)(x - 1).$$

De plus, 1 est aussi une racine de $x^3 + 2x^2 + 3x + 4$. En fait

$$x^4 + x^3 + x^2 + x + 1 = (x-1)^4,$$

alors la multiplicité de la racine 1 de $p(x)$ est 4 parce que $(x-1)^4$ divise p mais pas $(x-1)^5$, en tant que polynôme de degré 5 (théorème 2.5).

2.5 Le plus grand diviseur commun

Soit K un corps dans ce chapitre.

Théorème 2.13. *Soient $f(x)$ et $g(x)$ deux polynômes sur K non tous deux nuls et soit*

$$I = \{u \cdot f + v \cdot g \colon u, v \in K[x]\}.$$

Il existe un polynôme $d(x) \in K[x]$ tel que

$$I = \{h \cdot d \colon h \in K[x]\}. \tag{2.8}$$

Démonstration. Remarquons que I contient des polynômes non nuls (notamment f ou g). Soit $d \in I \setminus \{0\}$ de degré minimal et $u', v' \in K[x]$ tels que

$$u' \cdot f + v' \cdot g = d.$$

Soit $u \cdot f + v \cdot g \in I$. La division avec reste donne $u \cdot f + v \cdot g = qd + r$, avec $\deg(r) < \deg(d)$. Alors

$$r = (u - qu') \cdot f + (v - qv') \cdot g \in I$$

et par minimalité de $d \in I \setminus \{0\}$, $r = 0$. Ainsi il existe un $h \in K[x]$ tel que $h \cdot d = u \cdot f + v \cdot g$. Il est clair que $h \cdot d \in I$ pour tous les $h \in K[x]$ et l'assertion est démontrée. □

Définition 2.6. Un polynôme $f(x) \in K[x] \setminus \{0\}$ dont le coefficient dominant est 1 est appelé *polynôme unitaire*.

Définition 2.7. Soient $f, g \in K[x]$ non tous deux nuls. Un *diviseur commun* de f et g est un diviseur de f et g.

Théorème 2.14. *Soient f, g et d comme dans le théorème 2.13.*

i) d est un diviseur commun de f et g.

ii) Chaque diviseur commun de f et g est un diviseur de d.

Démonstration. L'assertion i) suit du fait que $f, g \in I$ et de (2.8). Soient $u, v \in K[x]$ tels que $d = u \cdot f + v \cdot g$ et soit w un diviseur commun de f et g. Alors il existe $f', g' \in K[x]$ tels que $f = f'w$ et $g = g'w$. Par conséquent

$$d = (u \cdot f' + v \cdot g')w,$$

ce que montre que $w \mid d$ et ii). □

Proposition 2.15. *Soient $f, g \in K[x]$ non tous les deux nul et soit $d \in K[x]$ un polynôme satisfaisant les conditions i) et ii) du Théorème 2.14. Le polynôme d est unique à multiplication près par une constante $\alpha \neq 0$. Si d est unitaire, alors d est unique.*

Démonstration. Soient d et d' deux polynômes satisfaisant i) et ii). C.à.d. d et d' sont des diviseurs commun de f et g et chaque diviseur commun de f et g divise d et d'. Les conditions impliquent $d \mid d'$ et $d' \mid d$. Alors il existe $z, z' \in K[x]$ tel que $d = d'z'$ et $d' = dz$. Par suite, $d = d \cdot z \cdot z'$. Le theorème 2.5 ($d, d' \neq 0$) implique que $z, z' \in K$.

Si d et d' sont unitaires, alors $z = z' = 1$, ce que démontre 2.15). □

Définition 2.8. L'unique polynôme unitaire $d \in K[x]$ satisfaisant (2.8) est appelé le *plus grand commun diviseur* de f et g. Il est noté $\gcd(f, g)$ ou $\mathrm{pgcd}(f, g)$.

2.6 L'algorithme d'Euclide

Pour calculer le plus grand diviseur commun de $f(x)$ et $g(x)$ on peut utiliser l'algorithme d'Euclide. Soient $f_0, f_1 \in K[x]$ et pas tous les deux nuls et $\deg(f_0) \geq \deg(f_1)$. Si $f_1 = 0$, alors

$$\gcd(f_0, f_1) = a^{-1} f_0,$$

où a est le coefficient dominant de f_0. Autrement, on applique la division avec reste

$$f_0 = q_1 f_1 + f_2,$$

où $q_1, f_2 \in K[x]$ et $\deg(f_2) < \deg(f_1)$. Un polynôme $d \in K[x]$ est un diviseur commun de f_0 et f_1 si et seulement si d est un diviseur commun de f_1 et f_2. L'algorithme d'Euclide est le procédé de calculer la suite $f_0, f_1, f_2, \ldots, f_{k-1}, f_k \in K[x]$ où $\deg(f_{k-1}) \geq 0$, $f_k = 0$ et

$$f_{i-1} = q_i f_i + f_{i+1}$$

est le résultat de la division avec reste de f_{i-1} par f_i. Le procédé se termine toujours car la suite des degrés $\{\deg(f_i)\}$ est entière et strictement décroissante (méthode de descente infinie de Fermat). Le dernier reste non nul f_{k-1} est un multiple constant de $\gcd(f_0, f_1)$: il suffit de diviser par le coefficient dominant pour le rendre unitaire.

Exemple 2.13. On calcule le plus grand diviseur commun de $f_0 = 4x^6+x^4+2x^2+2 \in \mathbb{Z}_5[x]$ et $f_1 = 3x^4 + x^3 + 2x^2 + 2x + 2 \in \mathbb{Z}_5[x]$.

$$q_1 = 3x^2 + 4x + 2,\ f_2 = 4x^3 + 4x^2 + 3x + 3,$$

$$q_2 = 2x + 2,\ f_3 = 3x^2 + 1,$$

$$q_3 = 3x + 3,\ f_4 = 0.$$

Alors tout diviseur commun de f_0 et f_1 divise $f_3 = 3x^2 + 1$ et f_3 est aussi un diviseur commun. On divise par 3 (ou multiplie par 2) et on obtient $\gcd(f_0, f_1) = x^2 + 2$.

Le calcul des suites f_i et q_i donne aussi une représentation $\gcd(f_0, f_1) = u \cdot f_0 + v \cdot f_1$, $u, v \in K[x]$. En effet

$$\begin{pmatrix} f_i \\ f_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \begin{pmatrix} f_{i-1} \\ f_i \end{pmatrix}$$

et alors

$$\begin{pmatrix} f_{k-1} \\ f_k \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_{k-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \end{pmatrix}.$$

Exemple 2.14. On continue l'exemple 2.13.

$$\begin{pmatrix} 3x+3 & x^3+4x^2+2x \\ x^2+2x+2 & 2x^4+4x^2+3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2x+2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 3x+3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 2x^2+x+3 \end{pmatrix}$$

et

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

$$\begin{pmatrix} x^2+2 \\ 0 \end{pmatrix} = \begin{pmatrix} x+1 & 2x^3+3x^2+4x \\ 2x^2+4x+4 & 4x^4+3x^2+1 \end{pmatrix} \begin{pmatrix} 4x^6+x^4+2x^2+2 \\ 3x^4+x^3+2x^2+2x+2 \end{pmatrix}$$

Alors

$$x^2 + 2 = \gcd(f_0, f_1) = (x+1)f_0 + (2x^3 + 3x^2 + 4x)f_1.$$

2.7 Factorisation en irréductibles

On fixe un corps K dans ce chapitre.

Définition 2.9. Un polynôme $p(x) \in K[x] \setminus \{0\}$ est *irréductible*, si

i) $\deg(p) \geq 1$ et

ii) si $p(x) = f(x)g(x)$ alors $\deg(f) = 0$ ou $\deg(g) = 0$.

Exemple 2.15.

1. Chaque polynôme linéaire $p(x) = ax + b \in K[x]$, $a \in K \setminus \{0\}$ est irréductible. En effet, si $p(x) = f(x)g(x)$ et $\deg(f) > 0$ et $\deg(g) > 0$, alors le théorème 2.5 implique que $\deg(p) > 1$.
2. $x^2 + 1 \in \mathbb{R}[x]$ est irréductible. Autrement, il existe un polynôme linéaire $f(x) = x - \alpha \in \mathbb{R}[x]$ qui divise f ce qui implique que α est une racine de x^2+1. Cependant, aucun nombre α réel ne satisfait $\alpha^2 = -1$.

Clairement, tout polynôme $f(x) \in K[x] \setminus \{0\}$ peut être factorisé comme

$$f(x) = a \cdot \prod_i p_i(x), \tag{2.9}$$

dont les p_i sont irréductibles et unitaires et $a \in K$. On va voir maintenant que cette factorisation est unique.

Théorème 2.16. *Soit $p(x) \in K[x]$ irréductible et supposons que $p(x)$ divise un produit $f_1(x) \cdots f_k(x)$ de polynômes non nul. Alors $p(x)$ divise un polynôme $f_i(x)$.*

Démonstration. Par récurrence il suffit de démontrer l'assertion pour $k = 2$. Ainsi, supposons $p \mid fg$, $f, g \in K[x] \setminus \{0\}$. Si p ne divise pas f, alors $\gcd(p, f) = 1$ car les seuls diviseurs de p sont des multiples constants de 1 et lui-même. Soient donc $u, v \in K[x]$ t.q. $up + vf = 1$. Alors $upg + vfg = g$, et donc $p \mid g$. □

Théorème 2.17. *La factorisation (2.9) est unique à l'ordre près des p_i.*

Démonstration. Pour une factorisation $f(x) = a \prod_i q_i(x)$, où les q_i sont irréductibles et unitaires on utilise le théorème 2.16 pour déduire qu'il existe j tel que $p_1 \mid q_j$. Comme p_1 et q_j sont irréductibles et unitaires, $p_1 = q_j$. En divisant par p_1, l'assertion suit par récurrence. □

Corollaire 2.18. *Soient $f(x) \in K[x] \setminus \{0\}$ et $\alpha_1, \ldots, \alpha_\ell$ des racines de f de multiplicité $k_1, \ldots, k_\ell$ respectivement. Alors il existe $g(x) \in K[x]$ tel que*

$$f(x) = g(x) \cdot \prod_{i=1}^{\ell} (x - \alpha_i)^{k_i}.$$

Exercices

1. Démontre le Corollaire 2.18.
2. Soit K un corps fini, $n = |K|$.
   a) Chaque $a \in K^*$ satisfait $a^{n-1} = 1$. *Indice : Thm. de Lagrange*
   b) Le polynôme $f(x) = x^n - x + 1 \in K[x]$ ne possède pas de racines.
   c) Conclure qu'il existent des polynômes irréductibles en $K[x]$ de degré aux moins 2. *Indice : Factorisation de $f(x)$ en irréductibles.*

2.8 Le polynôme minimal d'une matrice

Soit $A \in K^{n\times n}$ une matrice et

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x]$$

un polynôme sur K. *L'évaluation* de $f(x)$ en A est la matrice

$$f(A) = a_0 I_n + a_1 A + \cdots + a_n A^n \in K[x].$$

Exercice 2.8. Pour $f(x), g(x) \in K[x]$ on a

i) $(f+g)(A) = f(A) + g(A)$,

ii) $(f \cdot g)(A) = f(A) \cdot g(A)$ et

iii) $1(A) = I_n$.

En autres mots, l'application

$$\begin{array}{rccc} \phi & : & K[x] & \longrightarrow & K^{n\times n} \\ & & f(x) & \longmapsto & f(A) \end{array}$$

est un homomorphisme d'anneaux.

Un polynôme $f(x) \in K[x]$ *annule* A, si $f(A)$ est la matrice zéro

$$f(A) = 0 \in K^{n\times n}.$$

Théorème 2.19. *Soit $p(x) \in K[x] \setminus \{0\}$ un polynôme de degré minimal parmi les polynômes qui annulent A. Alors $p(x)$ divise chaque polynôme $f(x) \in K[x]$ tel que $f(A) = 0$. Le polynôme p est unique à multiplication prés par une constante $\alpha \neq 0$. Si p est unitaire, alors p est unique.*

Démonstration. Supposons que $f(x) \in K[x]$ annule A

$$f(A) = 0.$$

On effectue division avec reste

$$f = q \cdot p + r \text{ où } \deg(r) < \deg(p).$$

Parce que $r(A) = 0$ on a forcement $r = 0$. (p est de degré minimal parmi les polynômes évaluant à zéro). Alors $p \mid f$.

Comme dans la démonstration de la Proposition 2.15 on montre que p est unique à multiplication prés par une constante $\alpha \neq 0$. Alors si p est unitaire, p est unique. □

Définition 2.10. Le polynôme *unitaire* $p(x) \in K[x] \setminus \{0\}$ du Théorème 8.16 est appelé le *polynôme minimal* de $A \in K^{n \times n}$.

Comment trouver le polynôme minimal ?

Soit $A \in K^{n \times n}$. Le théorème de Cayley-Hamilton (Théorème 3.11) affirme que $f_A(A) = 0$, où $f_A \in K[x]$ est le *polynôme caractéristique* de A. Le degré de f_A est n. Une façon de calculer le polynôme minimal consiste à chercher le plus petit $d \in \mathbb{N}_+$, $d \leq n$ tel que

$$I = A^0, A, A^1, \cdots, A^d$$

soient linéairement dépendants et de trouver $a_0, a_1, \ldots, a_d$ tel que

$$a_0 I + a_1 A + a_2 A^2 + \cdots + a_d A^d = 0. \tag{2.10}$$

Le polynôme minimal de A est $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. Ce d aussi bien les coefficients $a_0, \ldots, a_d$ de $p(x)$ se laissent calculé par l'algorithme de Gauss sur la matrice

$$B = (c_0, \ldots, c_n) \in K^{n^2 \times (n+1)} \tag{2.11}$$

dont la colonne $c_k \in K^{n^2}$ représente la matrice $A^k \in K^{n \times n}$ pour $k = 0, \ldots, n$. C'est à dire, la $n \cdot i + j$-ème composante de c_k est

$$(A^k)_{ij}. \tag{2.12}$$

En fait, le premier vecteur (j minimal) dans la liste de vecteurs de la base du $\ker(B)$, comme décrit en Chapitre 1.3 (1.3) nous donne le degré et les coefficients du polynôme minimal.

La complexité de l'élimination de Gauss avec m' lignes et n' colonnes est

$$\Omega(m' \cdot n' \cdot \min\{m', n'\})$$

opérations arithmétique du corps K. Dans notre cas, cela équivaut à $\Omega(n^4)$. Cependant, il existe un algorithme de calcul du polynôme minimal qui effectue $O(n^3)$ opérations arithmétiques du corps K. Cet algorithme est aléatoire, c'est à dire, que cet algorithme fais des choix basé sur des expériments de probabilité. À la fin, l'algorithme retourne un polynôme $p(x) \in K[x]$. Ce polynôme est *probablement* le polynôme minimal de $A \in K^{n \times n}$. En fait, la probabilité de l'évènement

$$p(x) \text{ n'est pas le polynôme minimal de } A$$

est au plus $1/2^{100}$, un nombre rationnel plus petit que le réciproque du nombre de particules élémentaires de l'univers.

Soit $v \in K^n \setminus \{0\}$. On montre de manière similaire à la démonstration du Théorème 2.19 la propriété suivante.

Théorème 2.20. *Soit $p_v(x) \in K[x] \setminus \{0\}$ de degré minimal tel que*

$$p_v(A)\, v = 0 \tag{2.13}$$

soit vérifiée. Ce polynôme $p_v(x)$ divise tout polynôme $h(x) \in K[x]$ tel que $h(A)v = 0$. Le polynôme p_v est unique à multiplication prés par une constante $\alpha \neq 0$. Si p est unitaire, alors p est unique.

Le polynôme $p_v(x)$ est appelé l'*A-annulateur de v*.

Le démonstration est comme celle du Théorème 2.19. Soit $h(x) \in K[x]$ un polynôme tel que $h(A)v = 0$. Si $p_v(x)$ ne divise pas $h(x)$, alors, par division euclidienne comme ci-dessus, on obtient

$$r(A) = h(A) - q(A)p_v(A) \in K^{n\times n}$$

et par conséquent $r(A)\, v = 0$, ce qui contredit la minimalité du degré de $p_A^v(x)$. Alors $p_v(x)$ divise tout polynôme $h(x) \in K[x]$ tel que $h(A)v = 0$. Par conséquence p_v est unique à multiplication prés par une constante $\alpha \neq 0$. Alors si p est unitaire, p est unique.

Algorithme aléatoire

L'algorithme est le suivant. Fixons un ensemble fini $S \subseteq K$ avec $|S| = M$ et choisissons $v \in K^n$ aléatoirement en sélectionnant chaque composante de v uniformément dans S. Ensuite, calculons $p_v(x)$ en trouvant le plus petit d tel que

$$I\, v, A\, v, A^2\, v, \cdots, A^i\, v \quad \text{soient linéairement dépendants.}$$

On retourne le polynôme $p_v(x)$.

Cette fois on effectue l'Algorithme de Gauss sur la matrice

$$C = \left(I\, v, A\, v, A^2\, v, \cdots, A^n\, v\right) \in K^{n\times(n+1)}. \tag{2.14}$$

Le calcul de $I\, v, A\, v, A^2\, v, \cdots, A^n\, v$ nécessite $O(n^3)$ opérations dans le corps. Ce sont des produits matrice-vecteur. Aussi l'algorithme de Gauss effectue $O(n^3)$ opérations.

Probabilité d'échec

Théorème 2.21. *La probabilité de l'événement $p_v(x) \neq p(x)$ est au plus $n/|S|$.*

Démonstration. Soit $p(x) = \ell_1(x)^{e_1} \cdots \ell_k(x)^{e_k}$ la factorisation unique de $p(x)$ en facteurs irréductibles distincts. On remarque que $k \leq \deg(p(x)) \leq n$. Si $p_v(x)$ n'est pas égal à $p(x)$, alors il doit exister un facteur irréductible, disons ℓ_1, tel que $(p(x)/\ell_1)(A)v = 0$ soit vérifié. Considérons maintenant ℓ_1. Puisque $(p(x)/\ell_1)(A) \neq 0$, on a clairement

$$P\left[(m_A/\ell_1)(A)v = 0\right] \leq 1/|S|.$$

La borne annoncée découle alors de l'application de l'inégalité de Boole (union bound). □

Si l'ensemble S est de cardinalité $|S| = 2n$, la probabilité de l'événement $p_v(x) \neq p(x)$ est au plus $1/2$. Avec une telle choix de S, si l'algorithme est répété k-fois, la probabilité de l'évènement

$$P[\text{ aucun } p_v \text{ est égal à } p] \leq 2^{-k}.$$

Pour $k = 100$, par exemple, cette probabilité est inférieure aux réciproque des particules élémentaires de l'univers et en peut argumenter que cette probabilité es négligeable.

Algorithme 3 Polynôme minimal

Entrée : $A \in K^{n\times n}$, $k \in \mathbb{N}_+$, $S \subseteq K$ t.q. $|S| = 2n$
Sortie : $p(x) \in K[x]$
Répète100 fois :
    Choisir $v \in S^n$ aléatoirement
    Calculer $p_v(x)$
retour : Polynôme $p_v(x)$ de degré maximal.

Théorème 2.22. *Soit $A \in K^{n\times n}$ et $k \in n$. L'algorithme 2.8 effectue $O(n^3)$ opérations de corps. La probabilité de l'évènement que le polynôme retourné n'est pas le polynôme minimal de A est au plus 2^{-100}.*

Exercices

1. Soit $A \in K^{n\times n}$.
   i) Montrer que l'ensemble $I = \{f(x) \in K[x] \mid f(A) = 0\}$ est un idéal de $K[x]$ et qu'il est engendré par le polynôme minimal de T.
   ii) Soit $v \in V$. Montrer que l'ensemble $I = \{f(x) \in K[x] \mid f(A)v = 0\}$ est un idéal de $K[x]$ et que celui-ci est engendré par $p_v(x)$, le A-annulateur de v.
   iii) Montrer que $p(x)$ est le plus petit commun multiple (des) polynômes $p_v(x)$, $v \in V$.
2. Soit $A \in K^{n\times n}$ la matrice

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{n-1} \end{pmatrix}$$

   i) Quel est le polynôme minimal $p_{e_1}(x) \in K[x]$ du vecteur $e_1 \in K^n$?
   ii) Quel est le polynôme minimal $p(x) \in K[x]$ de la matrice A ? *Indice : Thm. de Cayleigh-Hamilton et Théorème 2.20.*

2.9 Construction formelle de l'anneau des polynômes

Dans cette section, nous fournirons une preuve du Théorème 2.2. Soit R un anneau. Une *suite* est une fonction $\sigma : \mathbb{N}_0 \longrightarrow R$. Nous écrirons σ comme

$$\sigma = [a_0, a_1, a_2, \dots),$$

où $a_i = \sigma(i)$. Nous désignons l'ensemble de toutes les suites par S. Nous identifions chaque élément $a \in R$ comme l'élément

$$[a, 0, 0, \dots) \in S.$$

Ensuite, nous définissons les opérations d'addition et de multiplication sur S. L'addition est l'opération naturelle suivante

$$[a_0, a_1, a_2, \dots) + [b_0, b_1, b_2, \dots) = [a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots).$$

Il est simple de montrer que $(S, +)$ est un groupe abélien, c'est-à-dire que (R1)-(R4) tiennent. Par exemple (R1) est montrée comme suit. Soient $[a_0, a_1, a_2, \dots), [b_0, b_1, b_2, \dots) \in S$. On a

$$\begin{aligned} [a_0, a_1, a_2, \dots) + [b_0, b_1, b_2, \dots) &= [a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots) && (2.15)\\ &= [b_0 + a_0, b_1 + a_1, b_2 + a_2, \dots) && (2.16)\\ &= [b_0, b_1, b_2, \dots) + [a_0, a_1, a_2, \dots). && (2.17) \end{aligned}$$

Ici, (2.16) est à cause de (R1) dans l'anneau R.

L'opération de multiplication est la suivante. Étant donné deux suites $[a_0, a_1, a_2, \dots) \in S$ et $[b_0, b_1, b_2, \dots) \in S$, leur produit est

$$[a_0, a_1, a_2, \dots) \cdot [b_0, b_1, b_2, \dots) = [c_0, c_1, c_2, \dots),$$

où

$$c_i = \sum_{j+k=i} a_j b_k.$$

Il est simple de montrer que les conditions (R5) et (R7) sont satisfaites. L'élément unité $1 \in S$ est la suite

$$[1, 0, 0, \dots).$$

Cela montre que S est un anneau. L'élément $x \in S$ spécifié dans le Théorème 2.2 est la suite

$$x = [0, 1, 0, \dots).$$

Clairement, pour chaque $a \in R$ on a

$$[a, 0, 0, \dots) \cdot x = x \cdot [a, 0, 0, \dots),$$

ce qui est la condition i du Théorème 2.2. Nous observons en outre que

$$x = [0, 1, 0, 0, \dots), \quad x^2 = [0, 0, 1, 0, \dots), \quad x^3 = [0, 0, 0, 1, 0 \dots) \quad \cdots$$

De cela, on peut conclure que, si

$$\mathbf{a}_0 + \mathbf{a}_1 x + \cdots + \mathbf{a}_n x^n = 0,$$

chaque $\mathbf{a}_i$ est de la forme $[a_i, 0, 0, \cdots)$ avec $a_i \in R$, alors chaque a_i doit être nul. Ceci conclut la preuve du Théorème 2.2.

3 Valeurs propres

3.1 Valeurs propres et vecteurs propres

Définition 3.1. Soit V un espace vectoriel sur un corps K et $f: V \longrightarrow V$ un endomorphisme. Un *vecteur propre* de f associé à la *valeur propre* $\lambda \in K$ est un vecteur $v \neq 0$ de V tel que $f(v) = \lambda v$.

Exemple 3.1. Soit $f: V \longrightarrow V$, l'endomorphisme $f(v) = 0$ pour tous $v \in V$. Alors tous $0 \neq v \in V$ est un vecteur propre associé à $\lambda = 0$.

Lemme 3.1. *Soit $B = \{v_1, \ldots, v_n\}$ une base de V et $A \in K^{n\times n}$ la matrice de l'endomorphisme $f: V \longrightarrow V$ relatif à B. La matrice A est une matrice diagonale, c'est à dire A est de la forme*

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

si et seulement si v_i est un vecteur propre associé à la valeur propre λ_i pour tout $i = 1, \ldots, n$.

Démonstration. Pour $v \in V$ soit $[v]_B \in K^n$ le vecteur des coordonnées de v relatif à B. L'application $\phi: V \longrightarrow K^n$, $\phi(x) = [x]_B$ est un isomorphisme. On a $[f(v_i)]_B = A\,[v_i]_B$ pour $i = 1, \ldots, n$. Supposons que $\{v_1, \ldots, v_n\}$ est une base de vecteurs propres. Des que $[v_i]_B = e_i$, et $f(v_i) = \lambda_i v_i$ alors

$$\lambda_i \cdot e_i = A\, e_i, \text{ pour } i \in \{1, \ldots, n\},$$

c.à.d. que A est une matrice diagonale.
La direction d'inverse est analogue. □

Définition 3.2. Un endomorphisme $f: V \longrightarrow V$ pour lequel existe une base de V composée de vecteurs propres est *diagonalisable*.

Définition 3.3. Soit $A \in K^{n\times n}$ une matrice. Un *vecteur propre* de A associé à la *valeur propre* $\lambda \in K$ est un vecteur propre de l'endormophisme $f(x) = Ax$ de K^n.

Exemple 3.2. 1. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Alors

— $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_1 = 1$,

— $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_2 = 0$,

— $v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ n'est pas un vecteur propre.

2. Soit $A = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \in \mathbb{R}^{2\times 2}$ pour $\phi \in \mathbb{R}$.

— Si $\phi \neq k\pi$, $k \in \mathbb{N}$, alors A n'a pas de valeur propre (réelle).

— Si $\phi = (2k+1)\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ a une valeur propre $\lambda = -1$ et tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.

— Si $\phi = 2k\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ a une valeur propre $\lambda = 1$ et encore tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.

On va voir que si on considère A comme une matrice complexe, alors on a toujours les valeurs propres $\cos\phi + \mathrm{i}\sin\phi$ et $\cos\phi - \mathrm{i}\sin\phi$.

Lemme 3.2. *Un vecteur $v \in V \setminus \{0\}$ est un vecteur propre de $f\colon V \longrightarrow V$ associé à la valeur propre $\lambda \in K$ si et seulement si $v \in \ker(f - \lambda \cdot \mathrm{Id})$.*

Rappel : L'endomophisme $\mathrm{Id}\colon V \longrightarrow V$ défini comme $\mathrm{Id}(v) = v$ pour tous $v \in V$ est appelé l'*identité*.

Définition 3.4. Soit λ une valeur propre de l'endomorphisme $f\colon V \longrightarrow V$. Le sous espace E_λ de V, défini comme

$$E_\lambda = \ker(f - \lambda \cdot \mathrm{Id})$$

est l'espace propre de f associé à λ. La dimension de E_λ est la multiplicité géométrique de λ.

Lemme 3.3. *Soient $v_1, \ldots, v_r \in V$ des vecteurs propres, associés aux valeurs propres $\lambda_1, \ldots, \lambda_r$ distinctes (c'est à dire $\lambda_i \neq \lambda_j$ pour $i \neq j$), alors $\{v_1, \ldots, v_r\}$ est un ensemble libre.*

Démonstration. Supposons que le théorème soit faux et soit $r \geq 1$ minimal, tel qu'il existent des vecteurs propres $v_1, \ldots, v_r \in V$ associés aux valeurs propres $\lambda_1, \ldots, \lambda_r$ qui sont linéairement dépendants. Des que $v_i \neq 0$ alors $r > 1$. Considérons une combinaison linéaire non triviale

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = 0. \tag{3.1}$$

Puisque (3.1) est un contre exemple minimal, on a $\alpha_i \neq 0$ pour tous i. Nous pouvons supposer que $\lambda_r \neq 0$. Autrement, on réarrange (3.1).

Si on applique f à l'expression (3.1) on obtient

$$\lambda_1\alpha_1 v_1 + \cdots + \lambda_r\alpha_r v_r = 0$$

et en divisant par λ_r

$$(\lambda_1/\lambda_r)\alpha_1 v_1 + \cdots + \alpha_r v_r = 0. \tag{3.2}$$

On soustrait (3.2) de (3.1) et on obtient

$$(1-\lambda_1/\lambda_n)\alpha_1 v_1 + \cdots + (1-\lambda_{r-1}/\lambda_r)\alpha_{r-1}v_{r-1}$$

Ceci est en contradiction avec la minimalité de r. □

Corollaire 3.4. *Soit $f\colon V \longrightarrow V$ un endomorphisme d'un espace vectoriel V sur K de dimension $n \in \mathbb{N}$ et soient $\lambda_1, \ldots, \lambda_r$ les valeurs propres différentes de f et soient $n_1, \ldots, n_r$ leurs multiplicités géométriques respectives. Soient $B_i = \{v_1^{(i)}, \ldots, v_{n_i}^{(i)}\}$ des bases de E_{λ_i} respectivement, pour $i = 1, \ldots, r$. Alors*

$$\{v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \cdots, v_1^{(r)}, \ldots, v_{n_r}^{(r)}\}$$

est un ensemble libre. L'application f est diagonalisable si et seulement si

$$n_1 + \cdots + n_r = n.$$

Démonstration. Soit la combinaison linéaire

$$\sum_{i=1}^{r}\sum_{j=1}^{n_i} \alpha_{ij} v_j^{(i)} = 0.$$

Remarquons que les vecteurs $\sum_{j=1}^{n_i} \alpha_{ij} v_j^{(i)}$ appartiennent à E_{λ_i} pour tout i respectivement. Le lemme 3.3 garantit donc que tous ses vecteurs soient nuls. Par suite, $\sum_{j=1}^{n_i} \alpha_{ij} v_j^{(i)}$ et les α_{ij} sont tous égaux à zéro car les $v_1^{(i)}, \ldots, v_{n_i}^{(i)}$ sont linéairement indépendants. Ça démontre que

$$\{v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \cdots, v_1^{(r)}, \ldots, v_{n_r}^{(r)}\}$$

est un ensemble libre. En plus, si $n_1 + \cdots + n_r = n$, f est diagonalisable par définition car l'ensemble forme une base de K^n.

À l'inverse, si f est diagonalisable, et si m_i dénote le nombre vecteurs propres en E_{λ_i} dans la base consistant de vecteurs propres, alors $m_i \leq n_i$, et on a

$$n = m_1 + \cdots + m_r \leq n_1 + \cdots + n_r \leq n,$$

et donc $n_1 + \cdots + n_r = n$. □

Voici une marche à suivre afin de déterminer si $f\colon V \longrightarrow V$ est diagonalisable ou non.

1. Déterminer les différentes $\lambda_1, \ldots, \lambda_r \in K$ tel que $\ker(f - \lambda \operatorname{Id}) \neq \{0\}$
2. Pour chaque λ_i calculer une base $\{v_1^{(i)}, \ldots, v_{n_i}^{(i)}\}$ de E_{λ_i}.
3. f est diagonalisable si et seulement si $n_1 + \cdots + n_r = n$.

Exercices

1. Une matrice $A \in K^{n\times n}$ est appelée *diagonalisable*, si endomorphisme $\phi\colon K^n \longrightarrow K^n$ défini comme $\phi(x) = Ax$ est diagonalisable. Démontrer que A est diagonalisable, si et seulement s'il existe $U \in K^{n\times n}$ inversible tel que $U^{-1}AU$ est une matrice diagonale.

3.2 Le polynôme caractéristique

Durant ce chapitre nous allons étudier les endomorphismes $f\colon V \longrightarrow V$ d'un espace vectoriel de dimension fini $n \in \mathbb{N}$. Si $B = \{v_1, \ldots, v_n\}$ est une base de V, on a

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

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

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

Les colonnes de la matrice A_B sont les coordonnées de $f(v_1), \ldots, f(v_n)$ dans la base B. Si B' est une autre base de V on a

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

où $P_{BB'}$ est la matrice de changement de base de B en B'. Comme on a

$$[f(v)]_{B'} = A_{B'}[v]_{B'} = A_{B'}P_{BB'}[v]_B$$

et

$$[f(v)]_{B'} = P_{BB'}[f(v)]_B,$$

on trouve

$$[f(v)]_B = P_{BB'}^{-1}A_{B'}P_{BB'}[v]_B \quad \text{pour tous } v \in V.$$

Et ça implique

$$A_B = P_{BB'}^{-1}A_{B'}P_{BB'} \tag{3.3}$$

En particulier,

$$\det(A_{B'}) = \det(A_B)$$

ce qui laisse nous définir le *déterminant d'un endomorphisme* f comme $\det(f) = \det(A_B)$.

Clairement, λ est une valeur propre de f si et seulement si λ est une valeur propre de A_B et c'est le cas si et seulement si

$$\det(A_B - \lambda I_n) = 0. \tag{3.4}$$

Rappelons la formule de Leibniz pour le déterminant d'une matrice $B \in K^{n\times n}$

$$\det(B) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} b_{i\pi(i)} \tag{3.5}$$

et si on regroupe les puissances de λ, on a

$$\det(A_B - \lambda I_n) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 \tag{3.6}$$

où $a_n, \ldots, a_0 \in K$.

Définition 3.5. Le polynôme $\det(A_B - \lambda I_n) \in K[\lambda]$ est le *polynôme caractéristique* de f.

Remarquons que $\det(A_B) = \det(A - 0 \cdot I_n)$, d'où $a_0 = \det(A_B)$. L'expression (3.6) est un polynôme avec indéterminée λ et comme polynôme formel, est défini par la formule de Leibniz

$$p_A(\lambda) = \det(A - \lambda I_n) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} (A - \lambda I_n)_{i,\pi(i)}.$$

En tant que somme des polynômes $\operatorname{sgn}(\pi)\prod_{i=1}^{n}(A - \lambda I_n)_{i,\pi(i)}$, son degré est au plus n. Considérons la permutation triviale $\pi = \mathrm{Id}$ donnant le produit de degré n

$$\operatorname{sgn}(\mathrm{Id}) \prod_{i=1}^{n} (A - \lambda I_n)_{i\,\mathrm{Id}(i)} = \prod_{i=1}^{n} (A_{ii} - \lambda),$$

et en remarquons que toutes les autres permutations aboutissent à un produit au degré inférieur à $n-2$. Cela signifie en particulier que $a_n = (-1)^n$, et donc que le polynôme caractéristique est de degré n.

Lemme 3.5. *Soit $p_A(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n$ le polynôme caractéristique de la matrice $A \in K^{n\times n}$. Alors, $a_0 = \det(A)$ et $a_n = (-1)^n$.*

Corollaire 3.6. *Soit $V \neq \{0\}$ un espace vectoriel de dimension fini sur $K = \mathbb{C}$, et $f : V \to V$ un endomorphisme. Alors f possède une valeur propre.*

Démonstration. Soit $f(\lambda) \in \mathbb{C}[\lambda]$ le polynôme caractéristique de f et n la dimension de V. Le degré de f est égal à $n \geq 1$, et donc $f(x)$ possède une racine $\lambda^* \in \mathbb{C}$ (théorème fondamental de l'algèbre). Cette racine λ^* est une valeur propre de f. □

Remarque 3.7. Pour deux bases B et B', comme on a $A_B = P_{BB'}^{-1} A_{B'} P_{BB'}$, alors

$$\begin{aligned} \det(A_B - \lambda I_n) &= \det(P_{BB'}^{-1} A_{B'} P_{BB'} - \lambda P_{BB'}^{-1} I_n P_{BB'}) \\ &= \det(P_{BB'}^{-1}) \det(A_{B'} - \lambda I_n) \det(P_{BB'}) \\ &= \det(A_{B'} - \lambda I_n). \end{aligned}$$

La définition 3.5 ne dépend ainsi pas de la base choisie et a donc un sens.

Définition 3.6. Soit $\lambda \in K$ une valeur propre de l'endomorphisme $f : V \longrightarrow V$. La *multiplicité algébrique* de λ est la multiplicité de λ comme racine de $\det(f - x\,\mathrm{Id}) \in K[x]$.

Proposition 3.8. *Soit $f\colon V \to V$ un endomorphisme et soit $\lambda \in K$ une valeur propre de f. La multiplicité géométrique de λ est au plus la multiplicité algébrique de λ.*

Démonstration. Soit m la multiplicité géométrique de λ et soit $\{v_1, \ldots, v_m\}$ une base de E_λ. On la complète en une base

$$B = \{v_1, \ldots, v_m, w_1, \ldots, w_{n-m}\}$$

de V.

La matrice A_B de l'endomorphisme f dans la base B est alors de la forme

$$A_B = \begin{pmatrix} \lambda I_m & C \\ 0 & D \end{pmatrix}$$

où $C \in K^{m \times n-m}$ et $D \in K^{(n-m)\times(n-m)}$. En effet, on a par définition $A_B[v_i]_B = [f(v_i)]_B$, et par conséquent $A_B e_i = \lambda e_i$, où e_i est le i-ème vecteur canonique de dimension n.

Lorsqu'on développe le déterminant d'une matrice en blocs comme A_B grâce à la formule de Leibniz, seules les permutations envoyant $\{1, \ldots, m\}$ et $\{m+1, \ldots, n\}$ sur eux-même donnent un produit non nul. On peut alors diviser la somme en deux pour obtenir que le déterminant est exactement le produit des déterminants des blocs diagonaux. Le polynôme caractéristique $p(x) \in K[x]$ de f est alors

$$\begin{aligned} p(x) &= \det \begin{pmatrix} (\lambda - x) I_m & C \\ 0 & D - x I_{n-m} \end{pmatrix} \\ &= (\lambda - x)^m \det (D - x I_{n-m}) . \end{aligned}$$

La multiplicité algébrique de λ est donc au moins m. □

Théorème 3.9 (Théorème de diagonalisation). *Soit V un espace vectoriel sur K de dimension n, $f : V \longrightarrow V$ un endomorphisme et $\lambda_1, \ldots, \lambda_r \in K$ les valeurs propres distinctes de f. Alors f est diagonalisable si et seulement si*

i) le polynôme caractéristique $p_f(x)$ de f décompose en facteurs linéaires, c'est-à-dire,

$$p_f(x) = (-1)^n \prod_{i=1}^{r} (x - \lambda_i)^{a_i}$$

où a_i est la multiplicité algébrique de $\lambda_i \in K$ pour tous i.

ii) $\dim(E_{\lambda_i}) = a_i$, *pour tous $i = 1, \ldots, r$. C'est à dire, les multiplicités algébriques et géométriques sont les mêmes.*

Démonstration. Supposons f diagonalisable. Soit B une base composée de vecteurs propres de f et A la matrice de f associée à la base B. Le lemme 3.1 implique que A est diagonale et alors $p_f(x) = \det(A - x\,\mathrm{Id}) = (-1)^n \prod_{i=1}^{r}(x - \lambda_i)^{a_i}$. La dimension de E_{λ_i} est celle du noyau $\ker(A - \lambda_i I_n)$. Clairement $\dim(\ker(A - \lambda_i I_n)) = a_i$, et on a alors montré i) et ii).

Supposons maintenant que i) et ii) tiennent. Soient g_i les multiplicités géométriques des valeurs propres λ_i, $i = 1, \ldots, r$. Comme on a

$$\deg((-1)^n \prod_{i=1}^{r}(x - \lambda_i)^{a_i}) = n,$$

alors $g_1 + \cdots + g_r = n$ et f est diagonalisable grâce au corollaire 3.4. □

Exemple 3.3. Le polynôme caractéristique de la matrice

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

est $(x-1)^2$. La multiplicité géométrique de $\lambda = 1$ est 1 et la multiplicité algébrique est 2. La matrice n'est pas diagonalisable.

Exemple 3.4. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ donnée par

$$f(x) = Ax, \text{ où } A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

Pour la base canonique $B = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$, on a $A_B = A$. Le polynôme caractéristique de f est

$$p(x) = -x^3 + x^2 + x - 1 = -(x-1)^2(x+1).$$

Les valeurs propres de f sont $\lambda_1 = 1$ et $\lambda_2 = -1$ et

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \text{ et } \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

sont des bases de E_{λ_1} et E_{λ_2} respectivement. Alors f est diagonalisable et pour la base

$$B' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

on a

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

et

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

On peut vérifier qu'on a bien

$$A_B = P_{BB'}^{-1} A_{B'} P_{BB'}. \tag{3.7}$$

Exercices

1. Donner un exemple d'un corps fini K et deux polynomes $p(x) \neq q(x) \in K[x]$ tel que $f_p = f_q$.

3.3 Spectre de matrices semblables

Définition 3.7. L'ensemble des valeurs propres d'une matrice $A \in K^{n\times n}$ (resp. d'un endomorphisme $f\colon V \longrightarrow V$) est appelé le *spectre* de A (resp. de f), noté $\operatorname{spec}(A)$ (resp. $\operatorname{spec}(f)$).

Théorème 3.10. *Soit $A \in K^{n\times n}$ une matrice et $P \in K^{n\times n}$ une matrice inversible.*

i) Le spectre de A et celui de $P^{-1}AP$ sont les mêmes.

ii) $v \in K^n$ est un vecteur propre de A si et seulement si $P^{-1}v$ est un vecteur propre de $P^{-1}AP$.

iii) Les polynômes caractéristiques $p_A(x)$ et $p_{P^{-1}AP}(x)$ sont identiques.

3.4 Théorème de Hamilton-Cayley

Soit $A \in K^{n\times n}$ et $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x] \setminus \{0\}$ un polynôme. On peut évaluer le polynôme en la matrice A comme suit :

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

Maintenant, soit $p_A(x)$ le polynôme caractéristique de A et v un vecteur propre de A associé à la valeur propre λ. On voit

$$p_A(A) \cdot v = a_0 v + a_1 \lambda v + \cdots + a_n \lambda^n v = p_A(\lambda) v = 0 \cdot v = 0.$$

Dans le cas où A est diagonalisable, il existe une base de vecteurs propres $\{v_1, \ldots, v_n\}$. On a alors $p_A(A) \cdot v_i = 0$ pour tous i, et donc $p_A(A) = 0$.

Théorème 3.11 (Hamilton-Cayley). *Soit $A \in K^{n\times n}$ et $p_A(\lambda)$ le polynôme caractéristique de A, alors*

$$p_A(A) = 0.$$

Démonstration. On écrit

$$\det(A - \lambda I_n) I_n = \operatorname{cof}(A - \lambda I_n)^T (A - \lambda I_n),$$

où $\operatorname{cof}(A - \lambda I_n)$ est la comatrice de $(A - \lambda I_n)$.

En regroupant les coefficients de λ^i dans $\operatorname{cof}(A - \lambda I_n)^T$ on obtient

$$\operatorname{cof}(A - \lambda I_n)^T = \sum_{i=0}^{n-1} \lambda^i B_i$$

avec certaines matrices $B_i \in K^{n \times n}$. Alors

$$a_0 I_n + a_1 \lambda I_n + \cdots + a_n \lambda^n I_n = B_0 A + \sum_{i=1}^{n-1} \lambda^i (B_i A - B_{i-1}) - \lambda^n B_{n-1},$$

où $p_A(\lambda) = a_0 + \cdots + a_n \lambda^n$. Ceci implique

$$\begin{array}{rcl} a_0 I_n & = & B_0 A \\ a_i I_n & = & B_i A - B_{i-1} \text{ pour } i \in \{1, \ldots, n-1\} \\ a_n I_n & = & -B_{n-1} \end{array} \tag{3.8}$$

ce que sont des équations de matrices en $K^{n \times n}$. Si on multiplie les matrices indicées par i à droite par A^i et qu'on somme les équations, on obtient $p_A(A)$ à gauche du signe d'égalité. À droite, on obtient une somme téléscopique égale à la matrice nulle. □

Exemple 3.5. Le polynôme caractéristique de $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ est $p_A(t) = (1-t)(2-t)$. On a

$$p_A(A) = (I_n - A)(2I_n - A) = 0$$

Pour la matrices $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, on a bien sûr que $p_A(A) = 0$ pour $p_A(t) = (2-t)^2$. Cependant, il existe un polynôme unitaire de degré strictement inférieur tel que $q(A) = 0$, à savoir $q(t) = t - 2$.

Définition 3.8. Le polynôme unitaire de degré minimal parmi ceux qui annulent A est appelé *polynôme minimal* de A.

Les résultats suivants donnent des utilisations typiques du théorème 3.11

Corollaire 3.12. *Soit $A \in K^{n \times n}$.*

(i) Toute puissance A^k avec $k \in \mathbb{N}$ peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \ldots, A^{n-1}$.

(ii) Si A est inversible, alors l'inverse A^{-1} peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \ldots, A^{n-1}$.

Démonstration. (i). Trivialement, l'assertion est vraie pour $k = 0, 1, \ldots, n-1$. On montre le cas $k = n$. Par le théorème 3.11 :

$$0 = p_A(A) = \alpha_0 I + \alpha_1 A \cdots + \alpha_{n-1} A^{n-1} + A^n \quad \Rightarrow \quad A^n = -\alpha_0 I - \alpha_1 A \cdots - \alpha_{n-1} A^{n-1}.$$

De façon similaire, on montre le cas $k > n$ par récurrence, utilisant $0 = A^{k-n} p_A(A)$.

(ii). Si A est inversible alors $\alpha_0 = \det(A)$ est inversible. De $0 = p_A(A)$ on obtient que

$$I = -\frac{\alpha_1}{\alpha_0} A \cdots - \frac{\alpha_{n-1}}{\alpha_0} A^{n-1} - \frac{1}{\alpha_0} A^n = A\Big(- \frac{\alpha_1}{\alpha_0} I \cdots - \frac{\alpha_{n-1}}{\alpha_0} A^{n-2} - \frac{1}{\alpha_0} A^{n-1}\Big)$$

et donc $A^{-1} = -\frac{\alpha_1}{\alpha_0} I \cdots - \frac{\alpha_{n-1}}{\alpha_0} A^{n-2} - \frac{1}{\alpha_0} A^{n-1}$. □

4 Formes bilinéaires

Dans le Chapitre 3 nous avons vu le concept de similarité de deux matrices $A, B \in K^{n\times n}$. Les matrices A et B sont semblables, s'il existe une matrice inversible $P \in K^{n\times n}$ tel que

$$A = P^{-1} \cdot B \cdot P.$$

Le Théorème 3.9 explique, quand une matrice $A \in K^{n\times n}$ est semblable à une matrice diagonale. C'est la cas si et seulement si le polynôme caractéristique de A décompose en facteurs linéaires et le multiplicités algébriques et géométriques de ces racines sont les mêmes.

Dans ce chapitre on ce concentre sur une autre relation d'équivalence. 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 dans ce cas $A \cong B$ et on se demande quand est-ce que une matrice $A \in K^{n\times n}$ est congruente à une matrice diagonale. $A \cong \operatorname{diag}(a_1, \ldots, a_n)$. Dans ce cas, on a certainement

$$A^T = A,$$

c.à.d. que A est une matrice *symétrique*.

Considérons un exemple. Soit $A \in \mathbb{Z}_3^{n\times n}$ la matrice

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

L'effet de la multiplication de A avec

$$P_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

à droite est que la nouvelle première colonne de A est la somme des deux colonnes. L'effet de la multiplication de A avec P^T à gauche est que la nouvelle première ligne de A est la somme des deux lignes. Alors on obtient

$$P_1^T A P_1 = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

D'ici on peut additionner la première colonne de A sur la deuxième et additionner la première ligne de A sur la deuxième et on obtient avec

$$P_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$(P_1P_2)^T A P_1 P_2 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Maintenant, si $A \in \mathbb{Z}_2^{n\times n}$ est encore la matrice

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

on observe que

$$P_1^T A P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

En fait, on peut montrer que, sur $\mathbb{Z}_2$, la matrice A n'est pas congruente à une matrice diagonale.

Dans ce chapitre, nous allons voir que toute matrice symétrique $A \in K^{n\times n}$ est congruente à une matrice diagonale si $1_K + 1_K \neq 0$.

4.1 Formes bililnéaires : Définition et propriétés de base

Définition 4.1. 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,

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

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

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

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 $v = 0$.

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

Exemple 4.1. Soit $V = K^n$, l'application

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

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

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

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 4.2. Soit $V = \mathbb{R}^3$ et

$$\langle u, v\rangle = u^T \begin{pmatrix} 1 & 0 & 1\\ 2 & 0 & 2\\ 1 & 1 & 0\end{pmatrix} v \text{ pour } u, v \in \mathbb{R}^3,$$

est une forme bilinéaire non-symmetrique, dégénérée à droite et à gauche.

Exemple 4.3. 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 4.1. Montrer que les formes bilinéaires des exemples 4.1 et 4.3 sont non dégénérées.

La matrice d'une forme bilinéaire relatif à une base

Soit V un espace vectoriel de dimension finie et $B = \{v_1, \ldots, v_n\}$ 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

$$f(x,y) = [x]_B^T \, A_B^f \, [y]_B. \tag{4.1}$$

Proposition 4.1. *Soit $B \in K^{n\times n}$ tel que pour tous $x, y \in V$*

$$f(x,y) = [x]_B^T \, B \, [y]_B,$$

alors $B = A_B^f$.

Démonstration. C'est immédiat, parce que pour tous $i, j \in \{1, \ldots, n\}$

$$f(v_i, v_j) = (b)_{ij}.$$

□

Exemple 4.4. Soit $V = \{p(x) : p \in \mathbb{R}[x], \deg(p) \leq 2\}$ l'espace vectoriel des polynômes réelles de degré au plus 2 et $B = \{1, x, x^2\}$ une base de V et $f(p,q) = \int_0^1 p(x) \cdot q(x) dx$. Il est facile de vérifier que f est une forme bilinéaire sur V. La matrice A_B^f est

$$A_B^f = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}.$$

Pour $p(x) = 2 + 3\,x - 5\,x^2$ et $q(x) = 2\,x + 3\,x^2$ on obtient

$$\int_0^1 f(x)p(x)dx = \begin{pmatrix} 2 & 3 & -5 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$$

Changement de base

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

$$A_{B'}^f = P_{B'B}^T \, A_B^f \, P_{B'B}. \tag{4.2}$$

Définition 4.2. 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 dans ce cas $A \cong B$.

Exemple 4.5. Si V est de dimension finie et B, B' sont deux bases de V, la relation (4.2) montre que $A_B^{\langle , \rangle} \cong A_{B'}^{\langle , \rangle}$.

Lemme 4.2. *La relation $\cong$ est une relation d'équivalence.*

Démonstration. Voir exercice. □

Exercice 4.2. 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 4.3. *Soit V un K-espace vectoriel de dimension finie, $B = \{b_1, \ldots, b_n\}$ 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 a 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 a 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 $\text{rang}(A_B^f) = n$. □

4.2 Orthogonalité

Pour ce paragraphe 4.2, s'il n'est pas spécifié autrement, V est toujours un espace vectoriel sur K muni d'une forme bilinéaire symétrique $\langle , \rangle$.

Définition 4.3. Deux éléments $u, v \in V$ sont *orthogonaux* ou *perpendiculaires* si $\langle u, v\rangle = 0$, et l'on écrit $u \perp v$.

Proposition 4.4. *Soit $E \subseteq V$ une partie de V, alors $E^\perp = \{v \in V : v \perp e$ 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 $c \cdot u \in W$.

Des que $0 \in E^\perp$, on a que $E^\perp \neq \emptyset\varnothing$. Si $u, v \in E^\perp$ alors pour tout $e \in E$

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

et pour $c \in K$

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

□

Exercice 4.3. Soit $E \subseteq V$ et E^* le sous-espace de V engendré par les éléments de E. Montrer $E^\perp = E^{*\perp}$.

Exemple 4.6. Soient K un corps et $(a_{ij}) \in K^{m \times n}$ une matrice à m lignes et n colonnes. Le système homogène linéaire

$$A\,X = 0, \tag{4.3}$$

peut s'écrire sous la forme

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

où les A_i sont les vecteurs lignes de la matrice A et $\langle , \rangle$ 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 (4.3). Alors on a $U = W^\perp$ et $\dim(W^\perp) = \dim(U) = n - \text{rang}(A) = \dim(\text{noyau}(A))$.

Définition 4.4. 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

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

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 4.7. 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 4.5. *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 $\langle , \rangle$ *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 4.5. Une base $\{v_1, \ldots, v_n\}$ de l'espace vectoriel V est une *base orthogonale* si $\langle v_i, v_j \rangle = 0$ pour $i \neq j$.

Remarque 4.6. Pour une forme bilinéaire symétrique $\langle , \rangle$ et une base $B = \{v_1, \ldots, v_n\}$. On se rappelle que

$$\langle v_i, v_j \rangle = (A_B^{\langle , \rangle})_{ij}$$

Alors B est une base orthogonale, si et seulement si, $A_B^{\langle , \rangle}$ est une matrice diagonale.

Théorème 4.7. *Soit* $\mathrm{Char}(K) \neq 2$ *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 4.5 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 $V_1 = \mathrm{span}\{u\}$. 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 $\{v_2, \ldots, v_n\}$. Alors $\{u, v_2, \ldots, v_n\}$ est une base orthogonale de V. □

Exemple 4.8. Soit $V = \mathbb{Z}_5^3$ et $\langle , \rangle \colon \mathbb{Z}_5^3 \times \mathbb{Z}_5^3 \to \mathbb{Z}_5$ défini comme

$$\langle x, y \rangle = x^T A y,$$

où

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

Le but est de trouver une base orthogonale de $\mathbb{Z}_5^3$. On va trouver une matrice inversible $P \in \mathbb{Z}_5^{3 \times 3}$ tel que $P^T A P$ est une matrice diagonale. Si p_1, p_2, p_3 sont les colonnes de P, alors

$$\{p_1, p_2, p_3\}$$

est une base de $\mathbb{Z}_5^3$ et c'est une base orthogonale, des que

$$\langle p_i, p_j \rangle = 0 \text{ si } i \neq j,\ 1 \leq i, j \leq 3.$$

Nous allons additionner la 2-ème colonne de A sur la 1-ère colonne et la 2-ème ligne de A sur la 1-ère ligne de A. C'est à dire on calcule

$$P^T AP = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}$$

où

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

Après on va additionner $2\cdot$ la 1-ère colonne sur la deuxième, et l'opération correspondante de lignes et on obtient

$$P^T AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 4 \\ 0 & 4 & 0 \end{pmatrix}$$

où

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

Après on va additionner $4\cdot$ la 2-ère colonne sur la 3-ème, et l'opération correspondante de lignes et on obtient

$$P^T AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

où

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

Nous avons trouvé une base orthogonale

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

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 4.1 ?

4. En considérant le forme bilinéaire symétrique standard de l'exemple 4.1, trouver une base orthogonale du sous-espace de $\mathbb{Z}_3^4$ engendré par

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

4.3 Matrices congruentes à une matrice diagonale

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

Lemme 4.8. *Soit V un espace vectoriel de dimension finie et $B = \{v_1, \ldots, v_n\}$ une base quelconque. Alors V possède une base orthogonale si et seulement s'il existe une matrice diagonale D telle que $A_B^{\langle\cdot,\cdot\rangle} \cong D$.*

Démonstration. Si B' est une base orthogonale de V, alors $A_{B'}^{\langle\cdot,\cdot\rangle}$ est une matrice diagonale. Grâce à la relation (4.2), $A_B^{\langle\cdot,\cdot\rangle}$ est congruente à une matrice diagonale.

Si $A_B^{\langle\cdot,\cdot\rangle} \cong D$ où $D \in K^{n\times n}$ est une matrice diagonale, alors il existe une matrice $P \in K^{n\times n}$ inversible, telle que

$$P^T A_B^{\langle\cdot,\cdot\rangle} P = D.$$

La base $B' = \{w_1, \ldots, w_n\}$ donnée par les colonnes de P (en tant que coordonées dans la base B) :

$$[w_j]_B = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix} \iff w_j = \sum_{i=1}^{n} p_{ij} v_i, \quad \forall j = 1, \ldots, n,$$

est donc une base orthogonale. □

Corollaire 4.9. *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 4.7 et du théorème 4.8 parce que K^n muni de la forme bilinéaire symétrique $\langle u, v\rangle = u^T A v$ possède une base orthogonale. □

Cet algorithme trouve une matrice diagonale congruente à la matrice symétrique $A \in K^{n\times n}$ où K est un corps tel que $\mathrm{Char}(K) \neq 2$. 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

$$\begin{pmatrix} c_1 & & & & & & \\ & c_2 & & & & & \\ & & \ddots & & & & \\ & & & c_{i-1} & & & \\ & & & & b_{i,i} & \dots & b_{i,n} \\ & & & & \vdots & & \vdots \\ & & & & b_{n,i} & \dots & b_{n,n} \end{pmatrix} \tag{4.4}$$

où les composantes des premières $(i-1)$ lignes et colonnes sont nulles sauf éventuellement sur la diagonale.

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

- Soit k un indice tel que $k \geq i$ et $b_{kk} \neq 0$. On échange la i-ème ligne et la k-ème ligne puis la i-ème colonne et la k-ème colonne. Ceci permet (entre autres) d'échanger les coefficients b_{ii} et b_{kk} de la diagonale, s'assurant ainsi d'avoir un coefficient non nul.
- Si l'indice k de l'étape précédente n'existe pas (tous les coefficients diagonaux après c_{i-1} sont nuls), soit $j \in \{i+1, \dots, n\}$ un indice vérifiant $b_{ij} \neq 0$. On ajoute la j-ème ligne à la i-ème ligne puis la j-ème colonne à la i-ème colonne. Le i-ème coefficient de la diagonale devient alors $2b_{ij} + b_{jj} = 2b_{ij} \neq 0$.
- Si, à son tour, un tel indice j n'existe pas, on peut procéder à la $i+1$-ème itération car la matrice est déjà de la forme (4.4) (avec $i+1$ à la place de i).
- 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. Ceci permet d'annuler les coefficients à droite et sous le coefficient b_{ii}. On peut alors poursuivre à l'étape $i+1$.

Remarquons que chaque opération est faite à la fois sur les lignes et sur les colonnes. Ceci garantit que la matrice résultante reste symétrique. De plus, les opérations sont faites sur les lignes et colonnes d'indices $j \geq i$, laissant intacte la forme de la matrice (4.4).

4.4 Un algorithme

Maintenant, nous allons formaliser la procédé appliquée dans example 4.8.

Exemple 4.9. Soit V une espace vectoriel sur $\mathbb{Q}$ de dimension 3 muni d'une forme bilinéaire symétrique $\langle , \rangle$. Soit $B = \{v_1, v_2, v_3\}$ une base de V et

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

Le but est de trouver une base orthogonale de V.

En utilisant notre algorithme on trouve

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

telle que

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

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

Théorème 4.10 (Supplémentaire orthogonal). *Soient V un espace vectoriel de dimension finie sur corps K et W un sous-espace de V. Soit $\langle . \rangle$ une forme bilinéaire symétrique tel que, si restreint sur $W \times W$, elle est non dégénérée. Alors $V = W \oplus W^{\perp}$.*

Exercices

1. Soit
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{Z}_3^{2\times 2}.$$
Trouver une matrice $P \in \mathbb{Z}^{2\times 2}$ inversible tel que $P \cdot A \cdot P^T$ est une matrice diagonale. Est-ce que A est diagonalisable ?
2. Démontrer Théorème 4.10.
3. Montrer que $\cong$ est une relation d'équivalence sur l'ensemble des matrices $K^{n\times n}$.
4. 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 ? *Indice : voir l'exercice 3. de la section 4.2.*

5. Soit V un espace vectoriel sur un corps K de dimension finie muni d'une forme bilinéaire symétrique $\langle . \rangle$. Soit $B = \{v_1, \ldots, v_n\}$ une base de V. Montrer que $A_B^{\langle . \rangle} \in K^{n \times n}$ est congruente à une matrice diagonale si et seulement si V possède une base orthogonale.
6. Soit K un corps de caractéristique 2 et soit V un espace vectoriel sur K de dimension finie muni d'une forme bilinéaire symétrique non-nulle. 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 $A_B^{\langle . \rangle} = C$.
   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

$$A_B^{\langle . \rangle} = \begin{pmatrix} d_1 & & & & & & \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_k & & & \\ & & & & C & & \\ & & & & & \ddots & \\ & & & & & & C \end{pmatrix}$$

   et $d_1, \ldots, d_k = 0$, et le nombre de C n'est pas égal à zéro.
7. Modifier l'algorithme 4.4 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.
8. 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.
9. 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 $[y]_B$.

4.5 Le théorème de Sylvester

Soit V un espace vectoriel de dimension finie sur un corps K, $\mathrm{Char}(K) \neq 2$, muni d'une forme bilinéaire symétrique. Nous avons vu (théorème 4.7) que V possède une

base orthogonale. Supposons que cette base est $B = \{v_1, \ldots, v_n\}$ 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 $c_1, \ldots, c_r > 0$, $c_{r+1}, \ldots, c_s < 0$ et $c_{s+1}, \ldots, c_n = 0$.

Maintenant soit $K = \mathbb{R}$ et $A \in \mathbb{R}^{n \times n}$ symétrique. Le Corollaire 4.9 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

$$P^T A P = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}. \tag{4.5}$$

où les c_i sont ordonnés de sorte que $c_1, \ldots, c_r > 0$, $c_{r+1}, \ldots, c_s < 0$ et $c_{s+1}, \ldots, c_n = 0$. En multipliant les premières s colonnes de P par $1/\sqrt{|c_i|}$ on obtient en fait une factorisation (4.5) telle que $c_1, \ldots, c_r = 1$, $c_{r+1}, \ldots, c_s = -1$ et $c_{s+1}, \ldots, c_n = 0$.

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

$$P^T A P = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}. \tag{4.6}$$

Définition 4.6. Pour un espace vectoriel sur $\mathbb{R}$ de dimension finie, on appelle une base B de V telle que $A_B^{\langle , \rangle}$ a la forme décrite en (4.6) une *base de Sylvester*.

Nous allons maintenant démontrer, que les nombres r et s sont invariants par rapport au choix de la base B de V.

Définition 4.7. 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 $\langle . \rangle$.

Théorème 4.11. *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 $B = \{v_1, \ldots, v_n\}$ une base orthogonale de V. La dimension* $\dim(V_0)$ *est égale au nombre d'indices i tel que $\langle v_i, v_i\rangle = 0$.*

Démonstration. Nous utilisons la notation d'au-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. □

Définition 4.8. La dimension de l'espace de nullité $\dim(V_0)$ est appelé l'*indice de nullité* de la forme bilinéaire symétrique.

Théorème 4.12 (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 $B = \{v_1, \ldots, v_n\}$ de V, exactement r des indices i satisfont $\langle v_i, v_i\rangle > 0$.*

Démonstration. Soient $\{v_1, \ldots, v_n\}$ et $\{w_1, \ldots, w_n\}$ 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 $v_1, \ldots, v_r, w_{r'+1}, \ldots w_n$ 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 $r = r'$.

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

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

et ça implique, car les v_i et respectivement les w_i sont orthogonaux entre eux,

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

Les $\langle v_i, v_i\rangle$ à gauche sont strictement positifs. Les $\langle w_i, w_i\rangle$ à droite sont négatifs ou nuls. Il suit que $x_1 = 0, \ldots, x_r = 0$ et, comme les w_i sont linéairement indépendants, on a également $y_{r'+1} = 0, \ldots, y_n = 0$. □

Définition 4.9. L'entier r du théorème de Sylvester est appelé l'*indice de positivité* de la forme bilinéaire symétrique.

Exemple 4.10. Trouver une base de Sylvester de $\mathbb{R}^4$ et les indices de nullité et de positivité de la forme bilinéaire symétrique $x^T Ay$ où

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

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

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

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

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

En divisant les colonnes de P par $\sqrt{2}, \sqrt{4}$ et $\sqrt{13/4}$ respectivement, on obtient une transformation P telle que $P^T AP = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$.

Les colonnes de P sont une base de Sylvester.

Exercices

1. Démontrer, à l'aide des théorèmes 4.11 et 4.12, que l'indice de négativité (l'entier s de l'équation (4.5)) ne dépend lui aussi pas de la base choisie.
2. Déterminer l'indice de nullité et l'indice de positivité des formes bilinéaire symétriques définies par les matrices suivantes

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

3. Soit V un espace vectoriel de dimension finie sur $\mathbb{R}$ et soit $\langle . \rangle$ une forme bilinéaire symétrique 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

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

5 Produits scalaires et hermitiens

Définition 5.1. 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*.

Pour ce chapitre 5, sauf pour Chapitre 5.2, V est toujours un espace vectoriel sur $\mathbb{R}$ muni d'un produit scalaire. On appelle un espace vectoriel sur $\mathbb{R}$ muni d'un produit scalaire un *espace euclidien*.

Exemple 5.1. 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 4.3 est un produit scalaire.

Définition 5.2. Soit $\langle , \rangle$ 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 $\|v\| = 1$.

Proposition 5.1. *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 5.2 (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 5.3 (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'un produit scalaire $\langle,\rangle$. Si $w \neq 0$ est un élément de V, alors $\langle w, w\rangle > 0$. 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 5.3. Soit V un espace euclidien. Soit $w \in V \setminus \{0\}$. 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 $\alpha\, w$ s'appelle la *projection* de v sur w.

Exemple 5.2. Soit V l'espace vectoriel de l'exemple 4.3 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 5.4 (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 $w \neq 0$. Si $\alpha = \langle v, w\rangle / \langle w, w\rangle$ est la composante de v sur w, $v - \alpha w$ est perpendiculaire à w, donc aussi à $\alpha\, w$. D'après le théorème de Pythagore, on trouve

$$\begin{aligned}\|v\|^2 &= \|v - \alpha\, w\|^2 + \|\alpha\, w\|^2 \\ &\geq \alpha^2 \|w\|^2 \\ &= \langle v, w\rangle^2 / \|w\|^2.\end{aligned}$$

Cela implique

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

□

Théorème 5.5 (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 5.6. *Soit V un espace euclidien et soient $v_1, \ldots, v_n$ des éléments de V, deux à deux orthogonaux, tels que $v_i \neq 0$ pour tout i, et soit $a_1, \ldots, a_n \in \mathbb{R}$. Le vecteur*

$$v - a_1 v_1 - \cdots - a_n v_n$$

est perpendiculaire à tous les $v_1, \ldots, v_n$ si et seulement si a_i est la composante de v sur v_i, c'est-à-dire $a_i = \langle v, v_i\rangle / \langle v_i, v_i\rangle$ 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

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

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

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

Théorème 5.7 (Le procédé d'orthogonalisation de Gram-Schmidt). *Soient V un espace euclidien et $\{v_1, \ldots, v_n\} \subseteq V$ un ensemble libre. Il existe un ensemble libre orthogonal $\{u_1, \ldots, u_n\}$ de V tel que pour tout i, $\{v_1, \ldots, v_i\}$ et $\{u_1, \ldots, u_i\}$ 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 $\{u_1, \ldots, u_{i-1}\}$ pour $i \geq 2$. L'ensemble $\{u_1, \ldots, u_{i-1}, v_i\}$ est libre et une base du sous-espace engendré par $\{v_1, \ldots, v_i\}$. On met

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

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

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

Surtout $\{u_1, \ldots, u_i\}$ est un ensemble orthogonal. □

Exercice 5.1. 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 la forme bilinéaire symétrique qui permet le procédé de Gram-Schmidt ?

Définition 5.4. Une base $\{u_1, \ldots, u_n\}$ d'un espace euclidien est *orthonormale* si elle est orthogonale et se compose de vecteurs tous unitaires.

Corollaire 5.8. *Soit V un espace euclidien de dimension finie. Supposons $V \neq \{0\}$. V possède alors une base orthonormale.*

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

Exemple 5.3. 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

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

La base orthonormale est

$$A'/\|A'\| = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, B'/\|B'\| = \frac{1}{\sqrt{42}} \begin{pmatrix} 4 \\ -5 \\ 0 \\ 1 \end{pmatrix} \text{ et } C'/\|C'\| = \frac{1}{\sqrt{105}} \begin{pmatrix} -4 \\ -2 \\ -7 \\ 6 \end{pmatrix}$$

Corollaire 5.9. *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.

Démonstration. Comme $\text{rang}(A) = n$, les colonnes de A sont libres ; dès lors on peut appliquer le procédé de Gram-Schmidt à $\{a_1, \ldots, a_n\}$ où a_j désigne la j-ième colonne de A. On obtient alors une base orthogonale $B = \{a'_1, \ldots, a'_n\}$ avec la relation :

$$a_1 = a'_1 \quad , \quad a_j = \sum_{i=1}^{j-1} \alpha_{i,j} a'_i + a'_j$$

pour tout $j \in \{2, \ldots, n\}$, et où $\alpha_{i,j}$ est le coefficient de Fourier de a_j relativement à a'_i. Grâce à ce procédé, on a pu écrire a_j comme une combinaison linéaire de $\{a'_1, \ldots, a'_n\}$. On peut représenter cela avec un produit matrice-vecteur :

$$a_j = (a'_1 \ \ldots \ a'_n) \begin{pmatrix} \alpha_{1,j} \\ \alpha_{2,j} \\ \vdots \\ \alpha_{j-1,j} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

En posant

$$S = \begin{pmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \cdots & \alpha_{1,n-1} & \alpha_{1,n} \\ 0 & 1 & \alpha_{2,3} & \cdots & \alpha_{2,n-1} & \alpha_{2,n} \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & \alpha_{n-1,n} \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n},$$

on a, par les propriétés du produit matriciel,

$$A = A'S,$$

où $A' = (a'_1 \ldots a'_n)$. Il nous faut encore normaliser les colonnes de la matrice A'. Pour cela, on définit les deux matrices diagonales suivantes :

$$D = \begin{pmatrix} 1/\|a'_1\| & & \\ & \ddots & \\ & & 1/\|a'_n\| \end{pmatrix}, \quad D^{-1} = \begin{pmatrix} \|a'_1\| & & \\ & \ddots & \\ & & \|a'_n\| \end{pmatrix} \quad D, D^{-1} \in \mathbb{R}^{n\times n}$$

En posant $A^* = A'D$ et $R = D^{-1}S$ on obtient :

$$A = A^* R$$

où $A^* \in \mathbb{R}^{m\times n}$ et $R \in \mathbb{R}^{n\times n}$ sont des matrices qui vérifient les propriétés de l'énoncé. □

Exemple 5.4. Trouver une factorisation Q, R du Corollaire 5.9 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 5.10 (Inégalité de Bessel). *Si $v_1, \ldots, v_n$ sont des vecteurs unitaires deux à deux orthogonaux et si $\alpha_i = \langle v, v_i\rangle$ sont les coefficients de Fourier de v relativement à v_i alors*

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

Démonstration.

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

□

Exercices

1. Soit V un espace vectoriel sur $\mathbb{R}$ de dimension fini, muni d'une forme bilinéaire $\langle . \rangle$. Soit $B = \{b_1, \ldots, b_n\}$ une base orthogonale et $U = \text{span}\{b_i : i = 1, \ldots, n, \langle b_i, b_i\rangle > 0\}$. Montrer que $\langle . \rangle$ restreint à U est un produit scalaire du sous-espace U.

2. Soient V un espace vectoriel muni d'une forme bilinéaire symétrique $\langle . \rangle$ et $\{v_1, \ldots, v_n\} \subseteq V$ un ensemble de vecteurs deux à deux orthogonaux.
   a) Montrer que $\{v_1, \ldots, v_n\}$ 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 4.3, montrer que l'ensemble

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

   est un ensemble de vecteurs deux à deux orthogonaux.
4. Trouver la factorisation $Q \cdot R$ du Corollaire 5.9 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 une forme bilinéaire symétrique de $\mathbb{R}^n$ telle 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'une forme bilinéaire symétrique. 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 5.10) est une égalité si v est dans le sous-espace engendré par les $v_1, \ldots, v_n$.
8. On considère l'espace euclidien des fonctions continues sur l'intervalle $[0, 1]$ muni de la forme bilinéaire symétrique

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

   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, $\{u_1, \ldots, u_n\}$ un ensemble orthonormal et $f, g \in \text{span}\{u_1, \ldots, u_n\}$. Montrer l'*identité de Parseval*

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

5.1 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

$$Ax = b \tag{5.1}$$

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

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|. \tag{5.2}$$

Ici la norme $\|\cdot\|$ est par rapport au produit scalaire standard de $\mathbb{R}^n$.

Lemme 5.11. *Soit V un espace euclidien, $H \subseteq V$ un sous espace de V, $b \in V$ et $h \in H$ tel que $b - h \in H^\perp$, alors*

$$\|b - h'\| > \|b - h\|$$

pour tout $h' \in H$, $h \neq h'$. En particulière h est unique.

Démonstration. Soit $h' \neq h \in H$. Puisque $h - h' \in H$, Pythagore implique

$$\begin{aligned} \|b - h'\|^2 &= \|b - h + (h - h')\|^2 \\ &= \|b - h\|^2 + \|(h - h')\|^2 \\ &> \|b - h\|^2. \end{aligned}$$

□

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

Soient $b, a_1, \ldots, a_n \in V$, trouver $h \in H = \text{span}\{a_1, \ldots, a_n\}$ tel que la distance

$$\|b - h\|$$

est minimale.

---

i) Trouver une base orthonormale $\{u_1, \ldots, u_k\}$ du sous-espace $H = \text{span}\{a_1, \ldots, a_n\}$ avec le procédé de Gram-Schmidt.

ii) Retourner $h = \sum_{i=1}^k \langle b, u_i\rangle u_i$.

Pour $b \in \mathbb{R}^m$ et $a_1, \ldots, a_n \in \mathbb{R}^m$ étant les colonnes des $A \in \mathbb{R}^{m\times n}$ en (5.1), une solution $x^* \in \mathbb{R}^n$ de

$$Ax = h \tag{5.3}$$

est une solution optimale de (5.2). Une procédé plus simple est comme suivant.

Théorème 5.12. *Soient $A \in \mathbb{R}^{m\times n}$ et $b \in \mathbb{R}^m$. Les solutions du système*

$$A^T Ax = A^T b. \tag{5.4}$$

sont les solutions optimales du problème (5.2) *(où l'on considère la norme euclidienne sur $\mathbb{R}^m$, qui est la norme induite par le produit scalaire ordinaire)*

Démonstration. Soit $H = \text{span}\{a_1, \ldots, a_n\}$. Soit $x^* \in \mathbb{R}^n$ et $h = Ax^*$. Le vecteur $b - h$ est orthogonal a tout les colonnes de A (et alors à H) si et seulement si $0 = A^T(b - h) = A^T b - A^T Ax^*$.

L'assertion est donc une conséquence du Lemme 5.11. □

Remarque 5.13. Pour une norme $||\cdot||$ quelconque engendrée par un produit scalaire $\langle\cdot,\cdot\rangle$), une preuve similaire montre que les solutions du système

$$A^T F^{\langle\cdot,\cdot\rangle} Ax = A^T F^{\langle\cdot,\cdot\rangle} b$$

sont les solutions optimales du problème (5.2), où $F^{\langle\cdot,\cdot\rangle}$ est la matrice du produit scalaire $\langle\cdot,\cdot\rangle$ selon la base canonique (i.e. $(F^{\langle\cdot,\cdot\rangle})_{i,j} = \langle e_i, e_j\rangle$).

Exemple 5.5. Trouver une solution de moindre carrées 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 $x^* = (1, 2)^T$.

Exercices

1. Soient V un espace vectoriel de dimension finie, $f : V \longrightarrow K$ 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, $f : V \times V \longrightarrow K$ une forme bilinéaire et B, B' des bases de V. Soit
$$f(x, y) = [y]_B^T A_B^f [y]_B.$$
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 la forme bilinéaire standard. Trouver une base du $\operatorname{span}\{v_1, v_2\}^\perp$. Est-ce que $\mathbb{Z}_2^4 = \operatorname{span}\{v_1, v_2\} \oplus \operatorname{span}\{v_1, v_2\}^\perp$?

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 $A_B^{\langle\rangle}$ pour $B = \{1, x, \ldots, x^n\}$.
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 $\langle . \rangle$ une forme bilinéaire symétrique. 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 $W^\perp$.

5.2 Formes sesquilinéaires 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

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

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 5.5. Soit V un espace vectoriel sur un corps $\mathbb{C}$ et

$$\langle , \rangle : V \times V \longrightarrow \mathbb{C}$$

une correspondance qui à tout couple (v, w) d'éléments de V associe un nombre complexe, noté $\langle v, w\rangle$. En considérant 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,

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

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

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

on dit que $\langle,\rangle$ est

i) une *forme sesquilinéaire*, si $\langle,\rangle$ satisfait PH 2 et PH 3.

ii) une *forme hermitienne*, si $\langle,\rangle$ satisfait PH 1, PH 2 et PH 3.

iii) un *produit hermitien*, si $\langle,\rangle$ satisfait PH 1, PH 2 et PH 3 et

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

Une forme sesquilinéaire est *non dégénérée à gauche* si la condition suivante est vérifiée.

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

Remarque 5.14. Si $\langle,\rangle$ est une forme hermitienne, pour tout $v \in V$, on a $\langle v, v\rangle \in \mathbb{R}$ dès que $\langle v, v\rangle = \overline{\langle v, v\rangle}$ par PH 1. On dit que la forme hermitienne est *définie positif* si $\langle v, v\rangle > 0$ pour tout $v \in V \setminus \{0\}$. Alors un produit hermitien est une forme hermitienne définie positif.

Exemple 5.6. Le produit *hermitien standard* de $\mathbb{C}^n$

$$\langle u, v\rangle = \sum_i u_i \overline{v_i}$$

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. Aussi les *coefficients de Fourier* sont les mêmes.
Soit $w \in V$, $w \neq 0$ et $v \in V$. On cherche $\alpha \in \mathbb{C}$ satisfaisant

$$\langle w, v - \alpha w\rangle = 0.$$

On trouve $\overline{\alpha} = \langle w, v\rangle / \langle w, w\rangle$. Puisque $\langle w, w\rangle \in \mathbb{R}$ et $\overline{\langle w, v\rangle} = \langle v, w\rangle$, alors

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

Exemple 5.7. Soit V l'espace vectoriel des fonctions $f : \mathbb{R} \longrightarrow \mathbb{C}$ 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

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

Soient V un espace vectoriel sur $\mathbb{C}$ de dimension finie et $f : V \times V \longrightarrow \mathbb{C}$ une forme sesquilinéaire. Pour une base $B = \{v_1, \ldots, v_n\}$ 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

$$\langle x, y \rangle = [x]_B^T A_B^f \overline{[y]_B} \tag{5.5}$$

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

Définition 5.6. Une matrice $A \in \mathbb{C}^{n \times n}$ est appelée *hermitienne* si on a

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

Proposition 5.15. *Soit V un espace vectoriel sur $\mathbb{C}$ de dimension finie et soit B une base de V. Une forme sesquilinéaire f est une forme hermitienne si et seulement si A_B^f est hermitienne.*

Définition 5.7. 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 $A = P^T \cdot B \cdot \overline{P}$. Nous écrivons $A \cong_{\mathbb{C}} B$.

$\cong_{\mathbb{C}}$ est aussi une relation d'équivalence. On peut aussi modifier l'algorithme 4.4 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 5.16. *Soit V un espace vectoriel sur $\mathbb{C}$ de dimension finie, muni d'une forme hermitienne. Alors V possède une base orthogonale.*

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

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

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

$$P^T A_B^{\langle \cdot \rangle} \overline{P} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}.$$

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

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

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

□

Exemple 5.8. 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 $f\colon V \times V \longrightarrow \mathbb{C}$ 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 une forme hermitienne.

2. Soit V un espace vectoriel sur $\mathbb{C}$ de dimension finie et $f : V \times V \longrightarrow \mathbb{C}$ une forme sesquilinéaire. Pour une base $B = \{v_1, \ldots, v_n\}$ 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 5.15.

4. Soit V un espace vectoriel sur $\mathbb{C}$ de dimension n et soit $\langle . \rangle$ une forme sesquilinéaire. Montrer que $\langle . \rangle$ est non dégénéré si et seulement si $\operatorname{rang}(A_B^{\langle . \rangle}) = n$ pour chaque base B de V.

5. Montrer que $\cong_{\mathbb{C}}$ est une relation d'équivalence.

6. Modifier l'algorithme 4.4 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 $f_i(x,y) = [x]_B^T A_i \overline{[y]_B}$, cocher ce qui s'applique.

	A_1	A_2	A_3
forme sesquilinéaire			
forme hermitienne			

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

5.3 Espaces hermitiens

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

Définition 5.8. Soit $\langle , \rangle$ un produit hermitien. 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 $\|v\| = 1$.

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

$$|\langle u, v \rangle| \leq \|u\|\|v\|. \tag{5.6}$$

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 5.17 (Le procédé d'orthogonalisation de Gram-Schmidt). *Soit V un espace hermitien et $\{v_1, \ldots, v_n\} \subseteq V$ un ensemble libre. Il existe un ensemble libre orthogonal $\{u_1, \ldots, u_n\}$ de V tel que pour tout i, $\{v_1, \ldots, v_i\}$ et $\{u_1, \ldots, u_i\}$ 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 5.18. *Soit V un espace hermitien de dimension finie. V possède alors une base orthonormale.*

Exercices

1. (Composante de u sur v ; indépendance de la direction) Soient $u, v \in V$ tel que $\langle v, v \rangle \neq 0$. Montrer qu'il existe un seul $\alpha \in \mathbb{C}$ tel que

   $$\langle u - \alpha \cdot v, v \rangle = 0.$$

   Pour ce α on a

   $$\langle v, u - \alpha \cdot v \rangle = 0.$$

2. Montrer l'inégalité de Cauchy-Schwarz.
3. Montrer l'inégalité triangulaire iii).
4. Montrer qu'un espace hermitien de dimension finie possède une base B telle que $\langle x, y \rangle = [x]_B \cdot [y]_B$, où $\cdot$ est le produit hermitien standard.

6 Le théorème spectral et la décomposition en valeurs singulières

Pour $A \in \mathbb{C}^{m\times n}$ nous allons écrire $\overline{A}^T$ comme

$$A^* = \overline{A}^T.$$

Alors une matrice $A \in \mathbb{C}^{n\times n}$ est une matrice hermitienne si $A^* = A$. Aussi, une matrice réelle $A \in \mathbb{R}^{n\times n}$ est symétrique, si $A^* = A$. On observe pour $A \in \mathbb{C}^{m\times n}$ et $B \in \mathbb{C}^{n\times k}$

$$(AB)^* = B^*A^*.$$

6.1 Le Théorème spectral

Lemme 6.1. *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

$$\lambda\, v^T\overline{v} = v^TA^T\overline{v} = v^T\overline{Av} = \overline{\lambda}\, v^T\overline{v}.$$

□

Corollaire 6.2. *Une matrice hermitienne $A \in \mathbb{C}^{n\times n}$ possède une valeur propre réelle $\lambda \in \mathbb{R}$.*

Démonstration. Le polynôme caractéristique $p(x) = \det(A - x \cdot I_n)$ a une racine complexe selon le *théorème fondamental* de l'algèbre (Théorème 2.11). Les valeurs propres de A sont les racines de $p(x)$. Mais toutes ces racines sont réelles selon Lemme 6.1. □

Remarque 6.3. La même preuve, par le théorème fondamental de l'algèbre, montre en fait que le polynôme caractéristique de A est scindé sur $\mathbb{R}$, i.e. A possède n valeurs propres réelles (en comptant les multiplicités algébriques).

Lemme 6.4. *Soient $A \in \mathbb{C}^{n\times n}$ une matrice hermitienne ($A^* = A$) et $u, v \in \mathbb{C}^n \setminus \{0\}$ deux vecteurs propres dont leurs valeurs propres sont différentes, alors $u^*v = 0$.*

Démonstration. Soient $\lambda \neq \gamma \in \mathbb{R}$ les valeurs propres correspondant aux vecteurs propres $u, v \in \mathbb{C}^n \setminus \{0\}$ respectivement. On a

$$\lambda u^*v = (Au)^*v = u^*A^*v = u^*Av = \gamma u^*v$$

et alors $u^*v = 0$. □

Définition 6.1. 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.

Théorème 6.5 (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*

$$P^* \cdot A \cdot P = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \tag{6.1}$$

où $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 laissé en 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 $0 \neq v \in \mathbb{R}^n$ un vecteur propre unitaire correspondant à λ_1. Avec la méthode de Gram-Schmidt, on peut trouver une base $\{v, u_2, \ldots, u_n\}$ dans $\mathbb{R}^n$ qui est orthonormale par rapport au produit scalaire ordinaire. Soit $U \in \mathbb{R}^{n\times n-1}$ la matrice dont les colonnes sont $u_2, \ldots, u_n$. 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

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

car $v^T \cdot v = 1$, $U^T \cdot U = I_{n-1}$ et $v^T \cdot U = 0$. Et

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

□

Comment peut-on calculer la diagonalisation (6.1) ? Voici un procédé pour diagonaliser une matrice symétrique (hermitienne) $A \in \mathbb{K}^{n \times n}$.

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

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

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

   a) Trouver une base $b_1^{(j)}, \ldots, b_{d_j}^{(j)}$ du noyau de la matrice $A - \lambda_j I$, par exemple avec l'algorithme de Gauss.

   b) Trouver une base orthonormale $p_1^{(j)}, \ldots, p_{d_j}^{(j)}$ du span$\{b_1^{(j)}, \ldots, b_{d_j}^{(j)}\}$, par exemple avec le procédé de Gram-Schmidt.

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

Exemple 6.1. Soit

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

Trouver une base orthonormale qui se compose des vecteurs propres.

Le polynôme caractéristique de A est

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

On trouve les bases des espaces propres

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

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

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

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 $\overline{z}$ sont linéairement indépendants sur $\mathbb{C}$.

6.2 Formes quadratiques réelles et matrices symétriques réelles

Le lemme 6.1 et le corollaire 6.2 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 6.2 d'une manière géométrique. L'ensemble

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

est appelé la *n-sphère.*

Définition 6.2. Une *forme quadratique* est une fonction $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$, $f(x) = x^T Ax$ où $A \in \mathbb{R}^{n\times n}$ est une matrice symétrique.

En fait, si $B \in \mathbb{R}^{n\times n}$ n'est pas symétrique, la matrice $A = 1/2(B^T + B)$ est symétrique et $x^T Bx = 1/2(x^T B^T x + x^T Bx) = x^T Ax$. Alors la fonction $g(x) = x^T Bx$ est aussi 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 6.6. *Soient $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et $v \in S^{n-1}$ le maximum de la fonction $f(x) = x^T Ax$ 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

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

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

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

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

□

Définition 6.3. 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 6.7. *Une matrice symétrique $A \in \mathbb{R}^{n \times n}$ est*

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

Démonstration. D'après le théorème 6.5 il existe une matrice orthogonale $U \in \mathbb{R}^{n \times n}$ telle que

$$A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T. \tag{6.2}$$

où les λ_i sont les valeurs propres de A. Les colonnes $u_1, \ldots, u_n$ 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

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

D'ici l'assertion suit directement. □

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 $K = \{l_1, \ldots, l_k\} \subseteq \{1, \ldots, n\}$ où $l_1 < l_2 < \cdots < l_k$. 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 6.8. *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 positifs (ou zéro).

Ce théorème est connu sous le nom de "critère de Sylvestre".

Démonstration. On démontre a), tandis que b) est un exercice. Soit A une matrice définie positive, montrons que les matrices B_k, $1 \leq k \leq n$, sont également définies positives : soit $z_k \in \mathbb{R}^k$ non-nul, on complète ce vecteur en un vecteur $x_k \in \mathbb{R}^n$ en lui rajoutant $n-k$ zéros. On vérifie facilement que $z_k^T B_k z_k = x_k^T A x_k > 0$. Les valeurs propres de B_k sont donc toutes strictement positives en vertu du théorème précédent. En se servant alors du théorème 6.5, on obtient facilement que $\det(B_k)$ est le produit des valeurs propres de B_k, alors $\det(B_k) > 0$.

Supposons maintenant que $\det(B_k) > 0$ pour tout $k \in \{1, \ldots, n\}$. L'argument est par récurrence. Le cas $n = 1$ est trivial.

Soit $n > 1$. Les matrices B_k $k = 1, \ldots, n-1$ sont définies positives par récurrence. Si A elle même n'est pas définie positive, $\det(A) > 0$ implique qu'il existe au moins deux valeurs propres négatives, disons μ et λ dans la factorisation donnée par le théorème spectral, et deux vecteurs propres orthogonaux $u, v \in \mathbb{R}^n$ correspondants. Leurs dernières composantes sont pas égales à zéro, parce que A_{n-1} est définie positive. Alors il existe $\beta \neq 0$ tel que la dernière composante de $u + \beta v$ est égale a zéro. Mais

$$(u+\beta v)^T A(u+\beta v) = \mu + \beta^2 \lambda < 0,$$

ce qui est une contradiction au fait que A_{n-1} est définie positive. □

Exemple 6.2. Nous pouvons alors montrer qu'une matrice symétrique est définie positive de deux manières différentes d'après les deux théorèmes précédents : considérons la matrice suivante

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

• Son polynôme caractéristique est égal à $\det(A - XI_n) = -X^3 + 6X^2 - 10X + 4$ dont les racines sont 2, $2+\sqrt{2}$ et $2-\sqrt{2}$. Les valeurs propres de A sont toutes les trois strictement positives donc la matrice est définie positive.

• D'une autre façon, $\det(B_1) = 2 > 0$, $\det(B_2) = 3 > 0$ et $det(B_3) = det(A) = 4 > 0$ donc la matrice est définie positive.

Théorème 6.9. *Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et $f(x) = x^T Ax$ la forme quadratique correspondante à A. On a*

$$\max_{x \in S^{n-1}} f(x) = \lambda_1 \quad et \quad \min_{x \in S^{n-1}} f(x) = \lambda_n \tag{6.3}$$

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

Démonstration. Nous utilisons la factorisation

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

où $P \in \mathbb{R}^{n\times n}$ est une matrice orthogonale dont les colonnes sont $p_1, \ldots, p_n$. 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 (6.3) avec les valeurs optimales λ_1 et λ_n respectivement.

□

Définition 6.4. 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 6.10 (Théorème Min-Max). *Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique avec les valeurs propres $\lambda_1 \geq \cdots \geq \lambda_n$. Si U dénote un sous-espace de $\mathbb{R}^n$ alors*

$$\lambda_k = \max_{\dim(U)=k} \min_{x\in U\setminus\{0\}} R_A(x) \tag{6.4}$$

et

$$\lambda_k = \min_{\dim(U)=n-k+1} \max_{x\in U\setminus\{0\}} R_A(x) \tag{6.5}$$

Démonstration. Nous démontrons (6.4). La partie (6.5) est un exercice. Soit $\{u_1, \ldots, u_n\}$ une base orthonormale de vecteurs propres associés à $\lambda_1 \geq \ldots \geq \lambda_n$ respectivement. On fixe un entier k, et un espace U de dimension k. Clairement $\text{span}\{u_k, \ldots, u_n\} \cap U \supsetneq \{0\}$. 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 $U = \text{span}\{u_1, \ldots, u_k\}$, $\min_{x\in U\setminus\{0\}} R_A(x) = \lambda_k$. Ensemble ça démontre (6.4). □

Exercices

1. Une matrice réelle symétrique, différente de la matrice zéro, telle que toute composante sur la diagonale est zéro ne peut pas être semi définie positive, ni semi définie négative.
2. Une matrice réelle symétrique, différente de la matrice zéro, dont la diagonale est égale à zéro, possède un 2×2 mineur symétrique négatif.

3. Soit $L \in \mathbb{R}^{n\times n}$ une matrice de la forme

$$L = \left(\begin{array}{c|c} H & 0 \\ \hline C & I_{n-i} \end{array}\right)$$

où $H \in \mathbb{R}^{i\times i}$ et $i \geq 0$. Soit $Q \in \mathbb{R}^{n\times n}$ la matrice de la permutation (transposition) qui échange $\mu, \nu > i$. Montrer

$$Q \cdot L = \left(\begin{array}{c|c} H & 0 \\ \hline C' & I_{n-i} \end{array}\right) \cdot Q$$

où C' provient de C en échangeant les lignes $\mu - i$ et $\nu - i$.

4. En s'appuyant sur l'exercice 3) montrer l'assertion suivante. Si l'algorithme 4.4 a exécuté k-itérations et dans chacune de ces k itérations, il existe un $j \geq i$ tel que $b_{jj} \neq 0$ (avec la notation de la i-ème itération), alors il existe une matrice de permutation Q telle que le résultat de ces premières k itérations s'écrit

$$R^T Q^T A Q R,$$

où R est une matrice triangulaire supérieure dont les éléments diagonaux sont 1. En autres mots, il existe une permutation si appliquée aux lignes et colonnes, l'algorithme 4.4 n'échange pas de lignes et colonnes pendant ces premiers k itérations.

5. Soient $A \in \mathbb{R}^{n\times n}$ réelle symétrique et A' la matrice obtenue de A en échangeant les lignes i et j ($A' = Q^T A Q$ pour une matrice de permutation (transposition) Q). Montrer que pour tout $K \subseteq \{1, ..., n\}$, il existe $K' \subseteq \{1, ..., n\}$ tel que $\det(A_K) = \det(A'_{K'})$ (mineurs symétriques) et inversément. En d'autres termes, montrer que tous les mineurs symétriques de A se retrouvent dans A' et vice versa.

6. Montrer la partie b) du théorème 6.8.
Indication pour montrer $\Longleftarrow$: Il existe une matrice de permutation Q telle que l'algorithme 4.4 n'échange pas de colonnes et lignes si confronté avec $Q^T \cdot A \cdot Q$ comme input et toutes itérations sont telles que $b_{ii} \neq 0$ jusqu'à un point, où tout le reste de la matrice est 0. Il faut s'appuyer sur les exercices 2 et 4.

7. 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.

8. 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 $K \subseteq \{1, \ldots, n\}$.

9. Montrer la partie (6.5) du théorème 6.10.

Définition 6.5. Une matrice hermitienne $A \in \mathbb{C}^{n\times n}$ est
— définie positive, si $x^T A \overline{x} > 0$ pour tout $x \in \mathbb{C}^n \setminus \{0\}$,
— définie négative, si $x^T A \overline{x} < 0$ pour tout $x \in \mathbb{C}^n \setminus \{0\}$,
— semi-définie positive, si $x^T A \overline{x} \geq 0$ pour tout $x \in \mathbb{C}^n$,
— semi-définie négative, si $x^T A \overline{x} \leq 0$ pour tout $x \in \mathbb{C}^n$.

Le théorème 6.7 trouve son analogue comme suivant. La démonstration est un exercice.

Théorème 6.11. *Une matrice hermitienne $A \in \mathbb{C}^{n\times n}$ est*

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

6.3 La décomposition en valeurs singulières

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

Théorème 6.12. *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 $D \in \mathbb{R}^{m\times n}_{\geq 0}$ est une matrice diagonale. Si A est réelle, P et Q sont réelles.

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 $\sigma_1^2 \geq \sigma_2^2 \cdots \geq \sigma_n^2 \geq 0$ les valeurs propres où $\sigma_i \in \mathbb{R}_{\geq 0}$ pour tous $i \in \{1, \ldots, n\}$ et soit $\{u_1, \ldots, u_n\}$ une base orthonormale correspondante de vecteurs propres. La matrice $Q \in \mathbb{C}^{n\times n}$ est la matrice dont les lignes sont $u_1^*, \ldots, u_n^*$.

Soit $r \in \mathbb{N}_0$ le nombre de σ_i strictement positif,

$$r = \mid \{i \in \{1, \ldots, n\} \colon \sigma_i > 0, \} \mid .$$

Nous construisons les vecteurs

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

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 tels que $\{v_1, \ldots, v_m\}$ est une base orthonormale de $\mathbb{C}^m$. Les colonnes de la matrices P sont alors $v_1, \ldots, v_m$ 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 $\sigma_1, \ldots, \sigma_r$ dans cet ordre. Avec ces matrices P, D et Q nous avons

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

ou de manière équivalente

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

Nous montrons ça en détail. Nous avons

$$(P^* \cdot A \cdot Q^*)_{ij} = v_i^* A u_j \tag{6.6}$$

et c'est égal à zéro si $j > r$, parce que dans ce cas $Au_j = 0$ dès que $u_j^* A^* A u_j = 0$. Si $i > r$ (6.6) pas égale a zéro implique $Au_j \neq 0$ et alors $j \leq r$. Mais dans ce cas, par construction, v_i est orthogonal à Au_j/σ_j et (6.6) est néanmoins zéro.

Et si $1 \leq i, j \leq r$, 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 6.6. En suivant la notation du théorème 6.12, les nombres $\sigma_1, \ldots, \sigma_r$ sont les *valeurs singulières* de A. La factorisation $A = P \cdot D \cdot Q$ est une *décomposition en valeurs singulières.*

Exemple 6.3. 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 6.7. La *pseudo inverse* d'une matrice

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

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 6.4. La pseudo inverse de la matrice A d'exemple 6.3 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 6.13. *Soit $A \in \mathbb{C}^{m\times n}$, alors il existe au plus une seule matrice $X \in \mathbb{C}^{n\times m}$ 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 6.14. *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 suit :

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

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

6.4 Encore les systèmes d'équations

Nous considérons encore une fois un système

$$Ax = b, \tag{6.7}$$

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

Définition 6.8. La *solution minimale* de (6.7) est la solution du problème des moindres carrés

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

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

Théorème 6.15. *La solution minimale de* (6.7) *est*

$$x = A^+ b,$$

où A^+ est la pseudo inverse de A.

Démonstration. Tout d'abord on remarque que pour $B \in \mathbb{C}^{n\times n}$ unitaire et $y \in \mathbb{C}^n$, $\|y\|^2 = y^T\overline{y} = y^T B^T\overline{B}\overline{y} = \|By\|^2$. Ainsi 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 $c = P^*b$. Dès lors on peut facilement vérifier que y est une solution minimale de $Dy = c \Leftrightarrow Q^*y$ est une solution minimale de $Ax = b$. Mais comme les solutions optimales de $Dy = c$ sont les $y \in \mathbb{C}^n$ tels que $y_i = c_i/\sigma_i$ pour $1 \leq i \leq r$ et $y_{r+1} \ldots y_n$ sont arbitraires alors la solution où $y_{r+1} = \cdots = y_n = 0$ est celle de norme minimale. Elle est donnée par

$$y = D^+ c.$$

La solution minimale de (6.7) est alors

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

□

Exemple 6.5. 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

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

6.5 Le meilleur sous-espace approximatif

Nous nous occupons du problème suivant. Soient $a_1, \ldots, a_m \in \mathbb{R}^n$ 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 a H. Si $H = \operatorname{span}\{u_1, \ldots, u_k\}$ où $\{u_1, \ldots, u_k\}$ 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 $u_1, \ldots, u_k$ et alors à H. Avec le théorème de Pythagore (Proposition 5.2), on a

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

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

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

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

$$A^T A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T \tag{6.8}$$

où $U = (u_1, \ldots, u_n) \in \mathbb{R}^{n \times n}$ est orthogonale. Nous pouvons supposer que les λ_i sont ordonnés comme $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Les valeurs propres sont non négatives dès que $A^T A$ est semi-définie positive. Selon Théorème 6.9 la solution est $H = \operatorname{span}\{u_1\}$.

La généralisation suivante du Théorème 6.9 est un exercice.

Théorème 6.16. *Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et $f(x) = x^T Ax$ 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 $U = (u_1, \ldots, u_n) \in \mathbb{R}^{n\times n}$ est orthogonale et $\lambda_1 \geq \cdots \geq \lambda_n$. Pour $1 \leq \ell < n$ on a

$$\max_{\substack{x\in S^{n-1} \\ x\perp u_1,\ldots,x\perp u_\ell}} f(x) = \lambda_{\ell+1} = \min_{\substack{x\in S^{n-1} \\ x\perp u_{\ell+2},\ldots,x\perp u_n}} f(x) \tag{6.9}$$

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

Maintenant, nous pouvons résoudre le problème central

$$\min_{\substack{H\trianglelefteq\mathbb{R}^n \\ \dim(H)=k}} \sum_{i=1}^{m} d(a_i, H)^2. \tag{6.10}$$

Théorème 6.17. *Soient $a_1, \ldots, a_m \in \mathbb{R}^n$, $A = (a_1, \cdots, a_m)^T$ et $u_1, \ldots, u_k$ les premières colonnes de la matrice orthogonale $U \in \mathbb{R}^{n\times n}$ de la factorisation* (6.8)*. Le sous-espace $H = \operatorname{span}\{u_1, \ldots, u_k\}$ est une solution du problème* (6.10)*.*

Démonstration. Pour $k = 1$ nous avons déjà montré l'assertion. Soit $k \geq 2$ et $W \trianglelefteq \mathbb{R}^n$ un sous espace de dimension k et soit $w_1, \ldots, w_k$ une base orthonormale de W. Nous pouvons supposer $w_k \perp \operatorname{span}\{u_1, \ldots, u_{k-1}\}$, (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

$$\max_{\substack{x\in S^{n-1} \\ x\perp \operatorname{span}\{u_1,\ldots,u_{k-1}\}}} x^T A^T A x$$

est atteint à u_k nous avons

$$w_k^T A^T A w_k \leq u_k^T A^T A u_k$$

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 6.9. Soit $A \in \mathbb{R}^{m\times n}$.

1) On définit la *norme Frobenius* de A comme le nombre

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

2) Pour une norme vectorielle $\|\cdot\|$ définie positive sur $\mathbb{R}^n$, on définit la norme $|||\cdot|||$ de A, appelée *norme matricielle subordonnée* à $\|\cdot\|$, comme le nombre

$$|||A||| = \sup_{v \neq 0} \frac{\|Av\|}{\|v\|} = \sup_{\|v\|=1} \|Av\|.$$

Remarque 6.18. La norme Frobenius n'est pas une norme matricielle subordonnée.

Exemple 6.6. Soit $A \in \mathbb{R}^{m \times n}$. On considère la norme euclidienne standard (aussi appelée norme 2) sur $\mathbb{R}^n$:

$$\|v\|_2 = \sqrt{\sum_{i=1}^n (v_i)^2}.$$

Alors la norme matricielle subordonnée à $\|\cdot\|_2$ correspond à

$$|||A|||_2 = \sup_{v \neq 0} \frac{\|Av\|_2}{\|v\|_2} = \sup_{\|v\|_2=1} \|Av\|_2 = \sup_{\|v\|_2=1} \sqrt{v^T A^T A v} = \sqrt{\lambda_1},$$

où λ_1 est la plus grande valeur propre de $A^T A$ (ou de manière équivalente, $\sqrt{\lambda_1}$ est la plus grande valeur singulière de A).

Exercice 6.1. Soient $A \in \mathbb{R}^{m \times n}$, $\|\cdot\|$ une norme définie positive sur $\mathbb{R}^n$, $|||\cdot|||$ la norme matricielle subordonnée à $\|\cdot\|$. Alors pour tout $v \in \mathbb{R}^n$, on a :

$$\|Av\| \leq |||A||| \cdot \|v\|.$$

Définition 6.10. Soit $A \in K^{n \times n}$. La *trace* de A est la somme de ses coefficients diagonaux, $\mathrm{Tr}(A) = \sum_{i=1}^n a_{ii}$.

Lemme 6.19. *Pour $A, B \in K^{n \times n}$ $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.*

Lemme 6.20. *Pour $A \in \mathbb{R}^{m \times n}$, $\|A\|_F^2 = \sum_{i=1}^r \sigma_i^2$ où $\sigma_1, \ldots, \sigma_r$ sont les valeurs singulières de A.*

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

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

□

Maintenant nous allons résoudre le problème suivant. Étant donnés $A \in \mathbb{R}^{m\times n}$ 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 $A = P \cdot \text{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots 0) \cdot Q$ est une décomposition en valeurs singulières où les colonnes de P sont $v_1, \ldots, v_m$ et les lignes de Q sont $u_1^T, \ldots, u_n^T$ on peut écrire

$$A = \sum_{i=1}^{r} \sigma_i v_i u_i^T \tag{6.11}$$

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

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

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

Écrit comme ligne (c.à.d. matrice en $\mathbb{R}^{1\times n}$) c'est égal a

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

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

Théorème 6.22. *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 $a_1^T, \ldots, a_m^T$ et soit B une matrice de rang au plus k. Les lignes de B sont dénotées comme $b_1^T, \ldots, b_m^T$. Soit $H = span\{b_1, \ldots, b_m\}$. 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 $\tilde{H} = \text{span}\{u_1, \ldots, u_k\}$. Nous avons démontré que

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 par $\tilde{a}_1^T, \ldots, \widetilde{a_m}^T$, alors

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

□

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 6.12, 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 6.16.
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 $w_1, \ldots, w_k$ telle que $w_k \perp H$.

7 Systèmes différentiels linéaires

On considère le système différentiel suivant

$$\begin{array}{rcl} \mathbf{x}_1'(t) &=& a_{11}\mathbf{x}_1(t) + \cdots + a_{1n}\mathbf{x}_n(t) \\ \mathbf{x}_2'(t) &=& a_{21}\mathbf{x}_1(t) + \cdots + a_{2n}\mathbf{x}_n(t) \\ & \vdots & \\ \mathbf{x}_n'(t) &=& a_{n1}\mathbf{x}_1(t) + \cdots + a_{nn}\mathbf{x}_n(t) \end{array} \tag{7.1}$$

où les $a_{ij} \in \mathbb{R}$. En notation matricielle, on peut écrire le système comme

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

où

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad \mathbf{x} \in C^1(\mathbb{R}; \mathbb{R}^n).$$

On cherche des fonctions dérivables $\mathbf{x}_i : \mathbb{R} \longrightarrow \mathbb{R}$ qui, ensemble, constituent $\mathbf{x}$ et qui satisfont (7.1). Un tel $\mathbf{x}$ est une *solution* du système (7.1). Soit $v \in \mathbb{R}^n$. Une solution $\mathbf{x}$ du système (7.1) respecte les *conditions initiales* données par v, si $x_i(0) = v_i$ pour $i = 1, \ldots, n$. On écrit $\mathbf{x}(0) = v$.

Exemple 7.1. 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 l'unique solution qui satisfait cette condition initiale. Généralement, si on spécifie $\mathbf{x}(0) = \alpha$, alors $\mathbf{x}(t) = \alpha \cdot e^t$ est la 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 $\mathbf{x}(0) = 1$.

Exemple 7.2. Soit $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. On trouve deux solutions

$$\mathbf{x}_1(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

et

$$\mathbf{x}_2(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

et chaque combinaison linéaire des deux

$$\alpha \cdot \mathbf{x}_1(t) + \beta \cdot \mathbf{x}_2(t)$$

est une solution. Si on spécifie la condition initiale $\mathbf{x}(0) = \binom{1}{2}$ on cherche alors $\alpha, \beta \in \mathbb{R}$ tel que

$$\binom{1}{2} = \alpha \cdot x_1(0) + \beta \cdot x_2(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

On met $\alpha = 1$ et $\beta = 2$, alors $x_1(t) = \cos(t) - 2\sin(t)$ et $x_2(t) = \sin(t) + 2\cos(t)$ est une solution, respectant les conditions initiales $\mathbf{x}(0) = \binom{1}{2}$.

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

Théorème 7.1 (Cours d'analyse II). *Étant donné les* conditions initiales $\mathbf{x}(0)$ *il existe une unique solution* $\mathbf{x}$ *du système* (7.1).

Nous sommes concernés par le problème de *trouver* la solution $\mathbf{x}$ explicitement.

Exercices

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

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

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

2. Soit $\{u_1 + i \cdot w_1, \ldots, u_n + i \cdot w_n\}$ une base de $\mathbb{C}^n$ où $u_i, w_i \in \mathbb{R}^n$ pour tout i. Montrer que $\operatorname{span}\{u_i, w_i \colon 1 \leq i \leq n\} = \mathbb{R}^n$.

7.1 L'exponentielle d'une matrice

Définition 7.1. Pour $A \in \mathbb{C}^{n \times n}$ on définit

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

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 $f \colon \{x \in \mathbb{C} \colon |x| < \rho\} \to \mathbb{C}$ 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 $\exp : \mathbb{C} \longrightarrow \mathbb{C}$

$$e^{a+ib} = e^a(\cos b + i \sin b), \quad a, b \in \mathbb{R}.$$

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

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

Lemme 7.2. *Pour $A \in \mathbb{C}^{n\times m}$ et $B \in \mathbb{C}^{m\times n}$ on a*

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

Démonstration. Soient $a_1^T, \ldots, a_n^T \in \mathbb{C}^m$ les lignes de A et $\overline{b_1}, \ldots, \overline{b_n} \in \mathbb{C}^m$ les colonnes de B. Avec Cauchy-Schwarz

$$|(AB)_{ij}|^2 = (a_i^T \overline{b_j})(\overline{a_i}^T b_j) \leq \|a_i\|^2 \|b_j\|^2$$

et donc

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

□

Lemme 7.3. *Soit $A \in \mathbb{C}^{n\times n}$. La série*

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

converge.

Démonstration. Considérons la suite $(b_n)_{n\in\mathbb{N}} \in \mathbb{C}^{n\times n}$

$$b_n = \sum_{j=0}^{n} \frac{A^j}{j!}$$

On montre que cette suite et une suite Cauchy par rapport la norme Frobenius $\|\cdot\|_F$. Soient $m \geq n \in \mathbb{N}$, alors

$$\begin{aligned} \| b_m - b_n \|_F &= \| \sum_{j=n+1}^{m} \frac{A^j}{j!} \|_F \\ &\leq \sum_{j=n+1}^{m} \frac{\| A^j \|_F}{j!} \\ &\leq \sum_{j=n+1}^{m} \frac{\| A \|_F^j}{j!}. \end{aligned}$$

Le première inégalité au dessus est l'inégalité triangulaire *(Théorème 5.5)* et la deuxième est Lemme 7.2. Puisque la suite $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}$

$$a_n = \sum_{j=0}^{n} \frac{\| A \|_F^j}{j!}$$

es Cauchy *(Analyse 1)*, alors pour tout $\epsilon > 0$ il existe $N_\epsilon \in \mathbb{N}$ tel que pour tout $m, n \geq N_\epsilon$

$$|a_m - a_n| < \epsilon.$$

Alors, pour tous $m, n \geq N_\epsilon$, on a

$$\| b_m - b_n \|_F < \epsilon.$$

Ceci montre que (b_n) est Cauchy et alors que e^A converge. □

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 ∞. Nous pouvons donc dériver les termes de la somme pour obtenir

$$\frac{d}{dt} e^{At} = A e^{At}. \tag{7.2}$$

Théorème 7.4. *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 $\mathbf{x}'(t) = Ae^{At}v = A\mathbf{x}(t)$. Plutôt $\mathbf{x}(0) = v$. □

Exemple 7.3 (Exemple 7.2 continué). Soit encore $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et les conditions initiales données par $v = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$. On calcule

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

dans ce cas. Pour $k \in \mathbb{N}$

$$A^{2k} = (-1)^k I \quad \text{et} \quad A^{2k+1} = \begin{pmatrix} 0 & (-1)^{k+1} \\ (-1)^k & 0 \end{pmatrix}$$

Alors, on a

$$e^{tA} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} & \sum_{k=0}^{\infty} \frac{(-1)^{k+1} t^{2k+1}}{(2k+1)!} \\ \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} & \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \end{pmatrix}$$

On se souvient des formules

$$\cos(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \text{ et } \sin(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}$$

Alors

$$e^{tA} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$$

Pour $v_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ alors comme avant,

$$\mathbf{x} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \cos(t) - 2\sin(t) \\ 2\cos(t) + \sin(t) \end{pmatrix}$$

est une solution respectant les conditions initiales $v_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Définition 7.2. Soit K un corps. Une matrice $N \in K^{n\times n}$ est *nilpotente* s'il existe un $k \in \mathbb{N}$ tel que $N^k = 0$.

Nous allons montrer ce théorème dans les prochaines cours.

Théorème 7.5. *Chaque matrice $A \in \mathbb{C}^{n\times n}$ peut être factorisée comme*

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

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

On se souvient maintenant, comment montrer que $e^{x+y} = e^x e^y$ pour $x, y \in \mathbb{C}$. Pour $N \in \mathbb{N}$, le produit des termes $\sum_{k=0}^{N} x^k/k!$ et $\sum_{k=0}^{N} y^k/k!$ s'écrit comme

$$\left(\sum_{k=0}^{N} x^k/k!\right)\left(\sum_{k=0}^{N} y^k/k!\right) = \sum_{k=0}^{2N} \sum_{i+j=k} \frac{x^i}{i!}\frac{y^j}{j!} \tag{7.3}$$

$$= \sum_{k=0}^{2N} \sum_{j=0}^{k} \binom{k}{j} \frac{x^{k-j} \cdot y^j}{k!}. \tag{7.4}$$

En fait, c'est le cas pour $x, y \in R$, où R est un anneau et $x \cdot y = y \cdot x$, c.à.d. x et y commutent. Dans ce cas, on a aussi

$$\sum_{k=0}^{2N} \frac{(x+y)^j}{k!} = \sum_{k=0}^{2N} \sum_{j=0}^{k} \binom{k}{j} \frac{x^{k-j}y^j}{k!}.$$

Basé sur cette observation, le lemme suivant est un exercice.

Lemme 7.6. *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 $\mathbf{x}' = Ax$, $\mathbf{x}(0) = v$ 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

$$\sum_{i=0}^{m} t^i A^i = P \left(\sum_{i=0}^{m} (t \operatorname{diag}(\lambda_1, \ldots, \lambda_n) + tN)^i / i! \right) P^{-1}.$$

Maintenant,

$$\sum_{i=0}^{\infty} (t \operatorname{diag}(\lambda_1, \ldots, \lambda_n) + tN)^i / i! = e^{t\cdot \operatorname{diag}(\lambda_1,\ldots,\lambda_n)+tN}.$$

Puisque $t \cdot \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ et $t \cdot N$ commutent, le théorème 7.5 implique

$$e^{t\cdot \operatorname{diag}(\lambda_1,\ldots,\lambda_n)+tN} = e^{t\cdot \operatorname{diag}(\lambda_1,\ldots,\lambda_n)} \cdot e^{tN}$$

La solution *réelle* que l'on cherche est alors

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

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

Exemple 7.4 (Exemple 7.2 continué). La matrice

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

est diagonalisable sur $\mathbb{C}$. En fait avec

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

on a

$$A = P \cdot D \cdot P^{-1}.$$

Les premiers N termes de la série e^{tA}, sont alors

$$\begin{aligned} \sum_{k=0}^{N} t^k A^k / k! &= \sum_{k=0}^{N} t^k \left((P \cdot D^k P^{-1}) \right) / k! \\ &= P \left(\sum_{k=0}^{N} (t \cdot D)^k / k! \right) P^{-1} \\ &= P \cdot \begin{pmatrix} \sum_{k=0}^{N} (-i \cdot t)^k / k! & 0 \\ 0 & \sum_{k=0}^{N} (i \cdot t)^k / k! \end{pmatrix} \cdot P^{-1} \end{aligned}$$

Puisque pour tout $t \in \mathbb{R}$

$$\sum_{k=0}^{\infty}(-i\cdot t)^k/k! = \cos(t) - i\sin(t) \quad \text{et} \quad \sum_{k=0}^{\infty}(i\cdot t)^k/k! = \cos(t) + i\sin(t),$$

et

$$P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

on a alors

$$\begin{aligned} e^{t\cdot A} &= P\cdot D\cdot P^{-1} \\ &= \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \end{aligned}$$

La solution respectant les conditions initiales $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ est, comme avant

$$\mathbf{x} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \cos(t) - 2\sin(t) \\ 2\cos(t) + \sin(t) \end{pmatrix}.$$

8 La forme normale de Jordan

Dans ce chapitre, on va démontrer le Théorème 7.5. On va donner une forme normale spécifique pour chaque matrice complexe, le *forme normale de Jordan.* Dans ce chapitre, K dénote toujours un corps et V un espace vectoriel sur K. On se rappelle que une matrice $N \in K^{n\times n}$ est *nilpotente* s'il existe un $k \in \mathbb{N}$ tel que $N^k = 0$.

Définition 8.1. Un *bloc Jordan* est une matrice de la forme

$$J(\lambda, k) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathbb{C}^{k\times k}$$

où les éléments non décrits sont zéro. Le nombre k est appelé la *longueur* our la *dimension* du bloc Jordan.

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

$$A = \begin{pmatrix} J(\lambda_1, n_1) & & & \\ & J(\lambda_2, n_2) & & \\ & & \ddots & \\ & & & J(\lambda_k, n_k) \end{pmatrix}$$

où les matrices $J(\lambda_j, n_j) \in \mathbb{C}^{n_j\times n_j}$ sont des blocs de Jordan.

Notre but de ce chapitre est de montrer le théorème suivant, qui est une version spécifique du Théorème 7.5.

Théorème 8.1. *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 8.2. Soit $V = \mathbb{C}^n$. Le *décalage* est l'application linéaire

$$U \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ 0 \end{pmatrix}.$$

Le décalage plus une constante est aussi une application linéaire

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

Il est facile de voir que la matrice représentant le décalage plus λ par rapport à la base canonique de $\mathbb{C}$ est un seul bloc de Jordan

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

8.1 Déterminer la forme normale der Jordan

Avant de démontrer le Théorème 8.1, on se pose la question : *Comment calculer une forme normale de Jordan d'une matrice $A \in \mathbb{C}^{n\times n}$?*

On ce rappelle du fait suivant.

> Soit K un corps. Si $A, B \in K^{n\times n}$ sont similaires, alors leurs polynômes caractéristiques sont les mêmes.

Soit alors le polynôme caractéristique de A,

$$p_A(x) = (-1)^n \prod_{i=1}^{n} (x - \lambda_i). \tag{8.1}$$

Alors $\lambda_1, \ldots, \lambda_n$ sont les éléments sur la diagonale de J.

Exemple 8.1. Soit $A \in \mathbb{C}^{5\times 5}$ la matrice

$$A = \begin{pmatrix} -2 & 5 & -1 & 1 & 22 \\ 8 & -7 & 0 & -4 & -32 \\ 2 & -4 & 4 & 1 & -23 \\ -10 & 13 & -2 & 5 & 55 \\ -2 & 3 & 0 & 1 & 13 \end{pmatrix}$$

Le polynôme caractéristique de A est

$$p_A(x) = -(x-2)^4(x-5) \in \mathbb{C}[x].$$

D'ici on sait alors que J a la forme

$$J = \begin{pmatrix} 2 & * & & & \\ & 2 & * & & \\ & & 2 & * & \\ & & & 2 & \\ & & & & 5 \end{pmatrix},$$

ou les $*$ peuvent être soit 0 ou 1. Si, par exemple, le deuxième $*$ est 0 et les autres sont 1, alors, J est composé de 3 blocs Jordan et totale. Si tous $*$ sont 1, alors, J est composé de 2 blocs Jordan.

Maintenant $A = P^{-1}JP$, où $P \in \mathbb{C}^{n\times n}$ est une matrice inversible. Pour $\lambda \in \mathbb{C}$ on voit

$$(A-\lambda I)^k = \left(P^{-1}(J-\lambda I)P\right)^k = P^{-1}(J-\lambda I)^k P.$$

Un élément $x \in \mathbb{C}^n$ est an élément du $\ker\left((A-\lambda I)^k\right)$, si et seulement si $Px \in \ker\left((J-\lambda I)^k\right)$. Alors on peut déduire le lemme suivant.

Lemme 8.2. *Soit $k \in \mathbb{N}$, alors*

$$\dim(\ker\left((A-\lambda I)^k\right)) = \dim(\ker\left((J-\lambda I)^k\right)).$$

Exemple 8.2 (Exemple 8.1 continué). La dimension du $\ker(A-2I)$ est 2 et $J-2I$ est

$$J-2I = \begin{pmatrix} 0 & * & & & \\ & 0 & * & & \\ & & 0 & * & \\ & & & 0 & \\ & & & & 3 \end{pmatrix}.$$

Alors, un des $*$ est zéro, les autres deux sont égaux à 1.

On voit directement le lemme suivant.

Lemme 8.3. *Soit $\lambda \in \mathbb{C}$ une valeur propre de A, la dimension de*

$$\ker(A-\lambda I)$$

est le nombre de bloc Jordan de J dont λ est sur la diagonale.

Exemple 8.3 (Exemple 8.1 continué). La dimension du $\ker(A-2I)$ est 2 et $J-2I$ est

$$J-2I = \begin{pmatrix} 0 & * & & & \\ & 0 & * & & \\ & & 0 & * & \\ & & & 0 & \\ & & & & 3 \end{pmatrix}.$$

Alors, un des $*$ est zéro, les autres deux sont égaux à 1.

On a deux blocs avec diagonale 2. Les possibilités (à l'ordre pré) sont

$$J = J_1 = \begin{pmatrix} 2 & 1 & & & \\ & 2 & 0 & & \\ & & 2 & 1 & \\ & & & 2 & \\ & & & & 5 \end{pmatrix}$$

et

$$J = J_2 = \begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & 0 & \\ & & & 2 & \\ & & & & 5 \end{pmatrix}$$

Maintenant

$$(J_1 - 2I)^2 = \begin{pmatrix} 0 & 0 & & & \\ & 0 & 0 & & \\ & & 0 & 0 & \\ & & & 0 & \\ & & & & 9 \end{pmatrix}$$

et

$$(J_2 - 2I)^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & & \\ & & 0 & 0 & \\ & & & 0 & \\ & & & & 9 \end{pmatrix}$$

La dimension du noyaux de $(A - 2I)^2$ est

$$\dim(\ker((A - 2I)^2)) = 3,$$

alors

$$J = J_2 = \begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & 0 & \\ & & & 2 & \\ & & & & 5 \end{pmatrix}.$$

Lemme 8.4. *Soit $k \in \mathbb{N}$, $\lambda \in \mathbb{C}$ une valeur propre de $A \in \mathbb{C}^{n\times n}$ et J une forme normale de Jordan de A. Le nombre de bloc Jordan dont λ est l'élément diagonale et de longueur aux moins $k+1$ est donné par*

$$\dim(\ker\left((A - \lambda I)^{k+1}\right)) - \dim(\ker\left((A - \lambda I)^{k}\right)).$$

Démonstration. Supposons que

$$J = \begin{pmatrix} B_1 & & & & & \\ & \ddots & & & & \\ & & B_s & & & \\ & & & B_{s+1} & & \\ & & & & \ddots & \\ & & & & & B_k \end{pmatrix}$$

Où les $B_1, \ldots, B_k$ sont des blocs Jordan et $B_1, \ldots, B_s$ ont λ sur la diagonale et les éléments sur la diagonale des $B_{s+1}, \ldots, B_k$ sont différents de λ. Au plus, on suppose $B_i \in \mathbb{C}^{n_i \times n_i}$ pour $i \in \{1, \ldots, k\}$. Pour $j \in \mathbb{N}$, matrice $(J - \lambda I)^j$ est de la forme

$$(J - \lambda I)^j = \begin{pmatrix} C_1^j & & & & & \\ & \ddots & & & & \\ & & C_s^j & & & \\ & & & C_{s+1}^j & & \\ & & & & \ddots & \\ & & & & & C_k^j \end{pmatrix}$$

où $C_i = J(0, n_i)$ pour $i = 1, \ldots, s$ et $C_i = J(\tilde{\lambda}_i, n_i)$ pour $i = s+1, \ldots, k$, où $\tilde{\lambda}_i \neq 0$.

Pour $j \in \mathbb{N}$ on a

$$\begin{aligned} \dim \ker(A - \lambda I)^j &= \dim \ker(J - \lambda I)^j && (8.2) \\ &= \sum_{i=1}^{s} \dim \ker(J(0, n_i)^j) && (8.3) \\ &= \sum_{i=1}^{s} \min\{n_i, j\}. && (8.4) \end{aligned}$$

Si on compare alors $\dim \ker(A - \lambda I)^{k+1}$ avec $\dim \ker(A - \lambda I)^k$, la somme (8.4) augmente par 1 pour chaque bloc Jordan de longueur aux moins $k+1$. Pour les blocs plus courts, alors $n_i \leq k$, on a $\min\{n_i, k\} = n_i = \min\{n_i, k+1\}$. □

Corollaire 8.5. *Soit $A \in \mathbb{C}^{n \times n}$, $J \in \mathbb{C}^{n \times n}$ une forme normale Jordan de A et $\lambda \in \mathbb{C}$ une valeur propre de A. Soit $k \in \{1, \ldots, n\}$ et n_k le nombre de blocs Jordan avec λ sur la diagonale de longueur k. Alors*

$$n_k = 2 \cdot \dim \ker(A - \lambda I)^k - \dim \ker(A - \lambda I)^{k+1} - \dim \ker(A - \lambda I)^{k-1}.$$

Corollaire 8.6. *Soit $A \in \mathbb{C}^{n \times n}$. La forme normale de Jordan est unique a l'ordre des blocs pré.*

8.2 Décomposition selon le polynôme caractéristique

Définition 8.3. Soient V un espace vectoriel sur un corps K, $T : V \to V$ un endomorphisme et $f(x) = a_0 + \cdots + a_n x^n \in K[x]$. L'*évaluation de f sur T* est l'endomorphisme $f(T) : V \to V$

$$f(T) = a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 \mathrm{id},$$

où $T^n = \underbrace{T \circ T \circ \cdots \circ T}_{n \text{ fois}}$ et $\circ$ dénote la composition de fonctions.

Remarque 8.7. Comme la démonstration du Théorème 2.4, où T joue le rôle de x et id le rôle de $1 = x^0$, on démontre :

Pour $f, g \in K[x]$ et $T : V \to V$ un endomorphisme, on a

i) $(f \cdot g)(T) = f(T) \circ g(T)$ et

ii) $(f + g)(T) = f(T) + g(T)$.

Définition 8.4. Soit K un corps, $f(x) = a_0 + \cdots + a_n x^n \in K[x]$ et $A \in K^{n \times n}$. L'*évaluation de f sur A* est la matrice $f(A) \in K^{n \times n}$

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

Remarque 8.8. Pour $f, g \in K[x]$ et $A \in K^{n \times n}$ une matrice, on a

i) $(f \cdot g)(A) = f(A) \cdot g(A)$ et

ii) $(f + g)(A) = f(A) + g(A)$.

Ici, les opérations $\cdot$ et $+$ sont la multiplication et l'addition de matrices respectivement.

Lemme 8.9. *Soient $p(x) \in K[x]$ et $A, B \in K^{n \times n}$ deux matrices similaires. Alors $p(A)$ et $p(B)$ sont similaires.*

Démonstration. Soit $P \in K^{n \times n}$ inversible et $A = P^{-1} B P$. Alors pour $i \in \mathbb{N}_0$ on a

$$A^i = P^{-1} B^i P$$

et alors

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

□

Exemple 8.4. On considère la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -6 & 0 & 5 & 0 \end{pmatrix}.$$

Le polynôme caractéristique de A est

$$f_A(A) = x^4 - 5x^2 + 6 = (x^2 - 2)(x^2 - 3).$$

On vérifie le théorème de Cayleigh-Hamilton (Théorème 3.11) :

$$\begin{aligned} f_A(A) &= (A^2 - 2I) \cdot (A^2 - 3I) \\ &= \begin{pmatrix} -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ -6 & 0 & 3 & 0 \\ 0 & -6 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} -3 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \\ -6 & 0 & 2 & 0 \\ 0 & -6 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Définition 8.5. Soient $T : V \to V$ un endomorphisme et $W \subseteq V$ un sous-espace de V. On dit que W est *invariant sous* T si $T(x) \in W$ pour tout $x \in W$.

Soit $A \in K^{n\times n}$ une matrice et $W \subseteq K^n$ un sous-espace de K^n. On dit que W est *invariant sous* A si $Ax \in W$ pour tout $x \in W$.

Remarque 8.10. Supposons que $\dim(V) = n$, $T\colon V \to V$ un endomorphisme et $W \subseteq V$ un sous-espace invariant sous T. Si

$$B = \{w_1, \ldots, w_k, v_1 \ldots, v_{n-k}\}$$

est une base de V où $\{w_1, \ldots, w_k\}$ est une base de W, alors la matrice de T associée à la base B est de la forme

$$A_B^T = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

où $B \in K^{k\times k}$ et $D \in K^{(n-k)\times(n-k)}$. Si, au plus, $\operatorname{span}(v_1 \ldots, v_{n-k})$ est invariant sous T, alors

$$A_B^T = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$$

Lemme 8.11. *Soient $f(x) \in K[x]$ et $T : V \longrightarrow V$ un endomorphisme, alors* $\ker(f(T))$ *est invariant sous T.*

Démonstration. Si $v \in \ker(f(T))$ on trouve que

$$\begin{aligned} f(T)(T(v)) &= [f(T) \circ T]\,(v) \\ &= [(f(x) \cdot x)(T)]\,(v) \\ &= [(x \cdot f(x))(T)]\,(v) \\ &= T(f(T)(v)) \\ &= 0. \end{aligned}$$

Alors, $T(v) \in \ker(f(T))$. □

Théorème 8.12. *Soit $T : V \to V$ un endomorphisme et soit $f(x) = f_1(x) \cdot f_2(x) \in K[x]$ tel que*

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

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

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

Démonstration. Parce que $\gcd(f_1, f_2) = 1$ il existe $g_1(x), g_2(x) \in K[x]$ tels que

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

et alors

$$(g_1 f_1 + g_2 f_2)(T) = \mathrm{id}\,. \tag{8.5}$$

Pour $v \in \ker(f(T))$, alors

$$[(g_1 f_1 + g_2 f_2)(T)](v) = v.$$

Mais $[(g_1 \cdot f_1)(T)](v) \in \ker(f_2(T))$, parce que

$$\begin{aligned} f_2(T) \circ g_1(T) \circ f_1(T)\,(v) &= g_1(T) \circ f_1(T) \circ f_2(T)\,(v) \\ &= g_1(T) \circ f(T)\,(v) = 0 \end{aligned}$$

et d'une manière similaire on voit que $g_2(T) \circ f_2(T)\,(v) \in \ker(f_1(T))$. Il reste à démontrer que la somme est directe.

Soit alors $v \in \ker(f_1(T)) \cap \ker(f_2(T))$. L'équation (8.5) montre

$$v = g_1(T) \circ f_1(T)\,(v) + g_2(T) \circ f_2(T)\,(v) = 0,$$

qui démontre que la somme est directe. □

Exemple 8.5 (Exemple 8.4 contunué). Le noyaux $\ker(f_A(A))$ égale à $\mathbb{C}^4$. On a

$$f_A(x) = (x^2 - 2)(x^2 - 3) \text{ et } \gcd(x^2 - 2, x^2 - 3) = 1.$$

On trouve les matrices

$$(x^2 - 2)(A) = A^2 - 2 \cdot I_4 = \begin{pmatrix} -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ -6 & 0 & 3 & 0 \\ 0 & -6 & 0 & 3 \end{pmatrix}$$

et

$$(x^2 - 3)(A) = A^2 - 3 \cdot I_4 = \begin{pmatrix} -3 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \\ -6 & 0 & 2 & 0 \\ 0 & -6 & 0 & 2 \end{pmatrix}$$

Leurs noyaux sont générés par

$$\left\{ \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\} \text{ et } \left\{ \begin{pmatrix} \frac{1}{3} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{3} \\ 0 \\ 1 \end{pmatrix} \right\}$$

respectivement. Pour la matrice

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

on trouve

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

une matrice bloc diagonale composé de deux blocs 2×2 respectivement.

Le Théorème suivant est maintenant facilement montré par récurrence.

Théorème 8.13. *Soit $T : V \to V$ un endomorphisme et soit $f(x) = f_1(x) \cdots f_k(x) \in K[x]$ tel que*

i) $\deg(f_i) > 0$ *pour tout* i

ii) $\gcd(f_i, f_j) = 1$ *pour tout* $i \neq j$

alors $\ker(f(T)) = \ker(f_1(T)) \oplus \cdots \oplus \ker(f_k(T))$.

Lemme 8.14. *Soit V un espace vectoriel de dimension finie sur $\mathbb{C}$ et soit $T\colon V \longrightarrow V$ une application linéaire. Alors V est la somme directe de sous-espaces $V = V_1 \oplus \cdots \oplus V_K$ tels que*

i) $T(V_i) \subseteq V_i$ c.à.d. V_i invariant sous T pour tout i et

ii) $T_{|V_i}\colon V_i \longrightarrow V_i$ est de la forme $N_i + \lambda \operatorname{id}$ où N_i est nilpotente.

Démonstration. Soit $\dim(V) = n$ et $p(x) =$ le polynôme caractéristique de T, alors $p(T) = 0$. Le coefficient dominant de $p(x)$ est $(-1)^n$. Le théorème fondamental de l'algèbre implique que

$$p(x) = (-1)^n (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$$

avec des λ_i différents. Le diviseur le plus grand de $(x - \lambda_i)^{m_i}$ et $p(x)/(x - \lambda_i)^{m_i}$ est 1. En utilisant théorème 8.13, alors

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

et avec $V_i = \ker(T - \lambda_i \operatorname{id})^{m_i}$ on a $V = V_1 \oplus \cdots \oplus V_K$ et i) avec lemme 8.11.

De plus,

$$T_{|V_i} = (T - \lambda_i \operatorname{id})_{|V_i} + \lambda_i \operatorname{id}_{|V_i} =: N_i + \lambda_i \operatorname{id}$$

et $N_i = (T - \lambda_i \operatorname{id})_{|V_i}$ est bien nilpotente, car $V_i = \ker(T - \lambda_i \operatorname{id})^{m_i}$ et donc $N_i^{m_i} = (T - \lambda_i \operatorname{id})^{m_i}_{|V_i} = 0$. □

Remarque 8.15. Lemme 8.14 démontre que pour tout espace vectoriel V sur $\mathbb{C}$ il existe une base

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

où $b_1^i, \ldots, b_{\ell_i}^i$ est une base de V_i telle que la matrice $A^T_{\mathscr{B}}$ de T par rapport à la base $\mathscr{B}$ est une matrice bloc diagonale

$$A^T_{\mathscr{B}} = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}$$

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

Des que les N_i et $\lambda_i I$ commutent, ça démontre le Théorème 7.5.

Exemple 8.6. Soit

$$A = \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$$

Le polynôme caractéristique de A est $p_A(x) = -(x-1)^2(x-3)$. Alors

$$\ker(p_A(A)) = \mathbb{R}^3 = \ker\left((A-I)^2\right) \oplus \ker(A-3I).$$

On a

$$(A-I)^2 = \begin{pmatrix} 28 & 28 & 56 \\ -16 & -16 & -32 \\ -4 & -4 & -8 \end{pmatrix}$$

et une base du $\ker\left((A-I)^2\right)$ est

$$\left\{ \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \right\}$$

Et

$$A - 3I = \begin{pmatrix} 22 & 34 & 18 \\ -14 & -22 & -10 \\ -4 & -6 & -4 \end{pmatrix}$$

avec base de noyaux

$$\left\{ \begin{pmatrix} -7 \\ 4 \\ 1 \end{pmatrix} \right\}$$

Avec matrice

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

on a

$$P^{-1}AP = \begin{pmatrix} 16 & 25 & 0 \\ -9 & -14 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

et

$$\begin{pmatrix} 15 & 25 \\ -9 & -15 \end{pmatrix}^2 = 0$$

Alors

$$P^{-1}AP = \begin{pmatrix} N_1 + 1 \cdot I_2 & 0 \\ 0 & N_2 + 3 \cdot I_1 \end{pmatrix}$$

est en forme blocque diagonale, où N_1 et N_2 sont nilpotentes. En fait, $N_2 = 0$.

8.2.1 Le polynôme minimal et la diagonalisation

Pendent cette section, V est toujours un espace vectoriel sur le corps K, $\dim(V) = n$ et $T\colon V \to V$ un endomorphisme. Soit $f_T(x) \in K[x]$ le polynôme caractéristique de T. Le théorème de Cayleigh-Hamilton (Théorème 3.11) implique $f_T(T) = 0$.

Théorème 8.16. *Soit $p(x) \in K[x] \setminus \{0\}$ un polynôme unitaire de degré minimal tel que*

$$p(T) = 0.$$

Alors $p(x)$ est unique.

Démonstration. Supposons que $g(x) \in K[x] \setminus \{0\}$ est aussi unitaire, $\deg(g) = \deg(p)$. On effectue division avec reste

$$g = q \cdot p + r \text{ où } \deg(r) < \deg(p).$$

Parce que $r(T) = 0$ on a forcement $r = 0$. (p est de degré minimal parmi les polynômes unitaires évaluant à zéro). Parce que g et p sont unitaires et de même degré, alors $q = 1$ et $p = g$. □

Définition 8.6. Le polynôme $p(x) \in K[x] \setminus \{0\}$ du Théorème 8.16 est appelé le *polynôme minimal* de T.

Le polynôme minimale d'une matrice $A \in K^{n\times n}$ est le polynôme minimal de l'endomorphisme $T\colon K^n \to K^n$, $T(x) = Ax$.

Remarque 8.17. Le polynôme minimal $p(x)$ de T divise chaque polynôme $f(x) \in K[x]$ tel que $f(T) = 0$. Particulièrement $p(x)$ divise le polynôme caractéristique $f_T(x)$ de T.

Théorème 8.18. *Une matrice $A \in K^{n\times n}$ est diagonalisable, si et seulement si le polynôme minimal $p(x) \in K[x]$ de A factorise en termes linéaires distinctes, c.à.d.*

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

où $\lambda_1, \ldots, \lambda_k \in K$ sont les valeurs propres distinctes de A.

Démonstration. Supposons que A est diagonalisable. Alors il existe une matrice diagonale $D \in K^{n\times n}$ et une matrice inversible $P \in K^{n\times n}$ tel que

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

Le polynôme minimale de A et celui de D est le même, voir Lemme 8.9. Soient $\lambda_1, \ldots, \lambda_K$ les éléments distinctes sur la diagonale de D. Il est facile de vérifier

$$0 = (D - \lambda_1 I) \cdots (D - \lambda_k I).$$

Et si un des facteurs $D - \lambda_i I$ manque, on obtient pas 0. Alors le polynôme minimale de A est $p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$.

Supposons maintenant que le polynôme minimal de A est

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

où $\lambda_1, \ldots, \lambda_k \in K$ sont distinctes. Pour $i \neq j$ on a

$$\gcd(x - \lambda_i, x - \lambda_j) = 1,$$

alors avec Théorème 8.13

$$K^n = \ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_k I)$$

est la somme directe d'espaces propres de A. Alors A est diagonalisable.

□

Exercices

1. Cet exercice concerne Remarque 8.7. Soient V un espace vectoriel, $T : V \to V$ un endomorphisme et $f(x) = x(x + 1)$, $g(x) = (x + 2)(x - 1)$.
   i) Calculez $f(x) - g(x)$.
   ii) Vérifiez (à pied) le fait que
   $$\mathrm{id} = \frac{1}{2} \left(T \circ (T + \mathrm{id}) - (T + 2\mathrm{id}) \circ (T - \mathrm{id}) \right).$$
   iii) Dans le Remarque 8.7. Est-ce qu'il faut exiger que T est un endomorphisme ou est-ce que le remarque est juste pour toute fonction $T : V \to V$?
2. Soit V un espace vectoriel de dimension fini sur $\mathbb{C}$, $T : V \longrightarrow V$ un endomorphisme et $f(x) = (x - \lambda)^m \in \mathbb{C}[x]$. Montrer que $\ker(f(T)) \neq \{0\}$ si et seulement si λ est un valeur propre de T.

8.3 Les endomorphismes nilpotent

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 8.1, qui, ensemble avec Lemme 8.14 nous donne Théorème 8.1 du début de ce chapitre.

Pendent ce sous-chapitre, V est un espace vectoriel sur corps K de dimension $n < \infty$ et $N : V \to V$ est un endomorphisme nilpotent.

Théorème 8.19. *Soit V un espace vectoriel sur K de dimension finie et $N\colon V \longrightarrow V$ une application linéaire nilpotente. Alors V possède une base $\mathscr{B}$ de la forme*

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

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

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

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \downarrow{\scriptstyle \phi_{\mathscr{B}}} & & \downarrow{\scriptstyle \phi_{\mathscr{B}}} \\ \mathbb{C}^n & \xrightarrow{A^T_{\mathscr{B}}\cdot x} & \mathbb{C}^n \end{array}$$

Remarque 8.20. Si on inverse l'ordre de chaque orbite $x_i, Nx_i, \ldots, N^{m_i-1}x_i$, on obtient une base $\mathscr{B}'$ et la matrice $A^N_{\mathscr{B}'}$ de l'application N a la forme

$$A^N_{\mathscr{B}'} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

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 conséquent, $N + \lambda\,\mathrm{id}$ est représentée par

$$A^{N+\lambda I}_{\mathscr{B}'} = A^N_{\mathscr{B}'} + \lambda I_n$$

en forme normale de Jordan.

Définition 8.7. Pour $x \in V \setminus \{0\}$ on appelle

$$m_x = \min\{i\colon N^i x = 0\}$$

la *durée de vie* de x. La séquence

$$x, Nx, \ldots, N^{m_x-1}x$$

est l'*orbite* de x sous N.

Lemme 8.21. *Soient $x_1, \dots, x_\ell \in V \setminus \{0\}$ et soient les orbites*

$$x_1, Nx_1, \dots, N^{m_1-1}x_1, \dots, x_\ell, Nx_\ell, \dots, N^{m_\ell-1}x_\ell \tag{8.6}$$

linéairement dépendant. Alors

$$N^{m_1-1}x_1, \dots, N^{m_\ell-1}x_\ell$$

sont déjà linéairement dépendant, où les $m_1, \dots, m_k$ sont les durées de vie des $x_1, \dots, x_k$.

Exemple 8.7. Supposons que $\ell = 3$, $m_1 = 2, m_2 = 2, m_3 = 3$ et soit

$$0 = 0 \cdot x_1 + \alpha \cdot N(x_1) + \beta \cdot x_2 + 0 \cdot N(x_2) + 0 \cdot x_3 + \gamma \cdot N(x_3) + \delta N^2(x_3)$$

une combinaison non-triviale de zéro, où $\alpha, \beta, \gamma, \delta \in K \setminus \{0\}$. On obtient

$$\begin{aligned} 0 &= N(0) \\ &= 0 \cdot N(x_1) + \alpha \cdot N^2(x_1) + \beta \cdot N(x_2) + 0 \cdot N^2(x_2) + 0 \cdot N(x_3) + \\ &\qquad \gamma \cdot N^2(x_3) + \delta N^3(x_3) \\ &= \beta \cdot N(x_2) + \gamma \cdot N^2(x_3). \end{aligned}$$

Alors $N(x_2), N^2(x_3)$ sont linéairement dépendant. L'idée c'est d'appliquer N jusqu'au dernier moment, ou des éléments $N^i(x_j)$, $0 \leq i < m_j$ ont un scalaire non-zéro dans la combinaison non-triviale de zéro.

Démonstration du Lemme 8.21. Soit

$$0 = \beta_0^1 x_1 + \dots + \beta_{m_1-1}^1 N^{m_1-1}x_1 + \dots + \beta_0^\ell x_\ell + \dots + \beta_{m_\ell-1}^\ell N^{m_\ell-1}x_\ell$$

une combinaison non-triviale de zéro. Maintenant, nous allons appliquer l'application N k-fois, où $k \geq 0$ est le plus grand entier tel que les termes

$$\beta_i^j N^{k+i} x_j$$

ne sont pas tous égaux à zéro. Ainsi, nous avons trouvé un sous-ensemble $J \subseteq \{1, \dots, \ell\}$ et des $\gamma_j \neq 0$ tels que

$$\sum_{j \in J} \gamma_j N^{m_j-1} x_j = 0.$$

□

Démonstration du Théorème 8.19. En concaténant les orbites des éléments d'une base de V nous obtiendrons un ensemble de vecteurs qui engendrent V. Supposons alors qu'au début de l'étape q, nous avons un ensemble $x_1, \dots, x_\ell$ avec $x_1, \dots, x_\ell \neq 0$ dont les orbites

$$x_1, Nx_1, \dots, N^{m_1-1}x_1, \dots, x_\ell, Nx_\ell, \dots, N^{m_\ell-1}x_\ell \tag{8.7}$$

engendrent V (pour la première étape, on prend $\ell = n$ avec des x_i formant une base de V). Ici m_i est la durée de vie de x_i. Si (8.7) est linéairement dépendant, nous allons soit supprimer un x_i et son orbite (car superflus), soit remplacer un x_i par un vecteur y tel que

i) Les orbites de $x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_\ell$ engendrent aussi l'ensemble V,

ii) la somme des durées de vie de $x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_\ell$ est strictement plus petite que la somme des durées de vie de $x_1, \ldots, x_\ell$.

Cela prouvera le théorème parce qu'un tel procédé doit se terminer.

Dès que l'ensemble (8.7) est linéairement dépendant, il existe une combinaison linéaire non triviale de (8.7) qui est égale a 0 :

$$0 = \beta_0^1 x_1 + \beta_1^1 N x_1 + \cdots + \beta_{m_1-1}^1 N^{m_1-1} x_1 + \cdots + \beta_0^\ell x_\ell + \beta_1^\ell N x_\ell + \cdots + \beta_{m_\ell-1}^\ell N^{m_\ell-1} x_\ell$$

Lemme 8.21 implique qu'il existe un sous-ensemble $J \subseteq \{1, \ldots, \ell\}$ et des $\gamma_j \neq 0$ tels que

$$\sum_{j \in J} \gamma_j N^{m_j-1} x_j = 0.$$

Cas 1 : Supposons qu'il existe un $j \in J$ tel que $m_j = 1$. Dans ce cas, un peut supprimer x_j et les orbites des autres $x_i, i \neq j$ engendrent ensemble V aussi.

Cas 2 : Autrement soit $m = \min_{j \in J} m_j - 1 \geq 1$ et soit $i \in J$ un index où le minimum est atteint. Alors

$$0 = N^m \sum_{j \in J} \gamma_j N^{m_j-1-m} x_j = N^m \left(\gamma_i x_i + \sum_{j \in J, j \neq i} \gamma_j N^{m_j-1-m} x_j \right)$$

Maintenant, en posant

$$y = \sum_{j \in J} \gamma_j N^{m_j-1-m} x_j = \gamma_i x_i + \sum_{j \in J, j \neq i} \gamma_j N^{m_j-1-m} x_j$$

Si $y \neq 0$, on remplace x_i par y. Il est alors facile de voir que les orbites de

$$x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_\ell$$

engendrent encore V. Et la durée de vie de y est au plus $m < m_i$.
Sinon, les orbites de

$$x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_\ell$$

suffisent alors à engendrer V.

On a alors démontré le théorème. □

Exemple 8.8. On considère la matrice

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

On a

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

et

$$N^4 = 0.$$

La durée de vie de la deuxième colonne de N est 4.

Exercices

1. Montrer que les **orbites** de $x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_\ell$ engendrent encore V. (Voir démonstration du théorème 8.19).
2. Le but de cet exercice est de faire la preuve du Théorème 8.1 "à l'envers".
   Soit $T\colon V \to V$ un endomorphisme. Soit $\phi : V \to \mathbb{C}^n$ l'isomophisme associé à une base B de V et à la base canonique E de $\mathbb{C}^n$. Supposons que $A_B = ([T(b_1)]_E, ..., [T(b_n)]_E)$, la matrice de T relativement à la base B, admette une forme normale de Jordan J avec matrice de passage $P = (p_1, ..., p_n)$.
   Montrer qu'il existe des sous-espaces $V_1, ..., V_k$ de V tels que pour tout i :
   a) $V = V_1 \oplus \cdots \oplus V_k$;
   b) $V_i = \phi(\text{span}(p_{k_i}, ..., p_{k_i+l_i}))$;
   c) $T(V_i) \subset V_i$;
   d) $T_{|V_i} = N_i + \lambda_i I$, où $N_i\colon V_i \to V_i$ est nilpotente ;
   e) $\{\lambda_1, ..., \lambda_k\} = \{J_{11}, ..., J_{nn}\}$.
3. Le but de cet exercice est de montrer les propriétés des décompositions comme dans le Lemme 8.14.
   Soit $T\colon V \to V$ un endomorphisme et soit $V_1, ..., V_k$ une décomposition de V tel que $V = V_1 \oplus \cdots \oplus V_k$, $T(V_i) \subset V_i$ et $T_{|V_i} = N_i + \lambda_i I$, où $N_i : V_i \to V_i$ est nilpotente. Montrer que :
   a) $V_i \subset \ker(T - \lambda_i I)^{a_i}$ pour un entier a_i tel que $N_i^{a_i} = 0$.
   b) Les $\lambda_1, ..., \lambda_k$ sont des valeurs propres (pas forcément distinctes) de T. (*Indice* : Utiliser par exemple le premier point).
   c) Le polynôme $f(x) = \prod_{i=1}^{k}(x - \lambda_i)^{a_i}$ annule T. (*Indice* : Montrer que $f(T)v = 0$ pour tout $v \in V$ en utilisant la décomposition de V et le premier point).
   d) En déduire que l'ensemble $\{\lambda_1, ..., \lambda_q\}$ contient toutes les valeurs propres de T (*Indice* : Si $v \neq 0$ est un vecteur propre de T de valeur propres λ, exprimer $f(T)v$ en fonction de f, λ, et v).
   e) En déduire que les valeurs sur la diagonale de n'importe quelle forme normale de Jordan de T constituent l'ensemble des valeurs propres de T. (*Indice* : Utiliser l'exercice 2.)

4. Comparer les polynômes caractéristiques de J et A. En déduire que les éléments diagonaux de J contiennent exactement l'ensemble des valeurs propres de A et le nombre d'apparitions de chaque valeur propre sur la diagonale de J est égale à la multiplicité algébrique de ladite valeur propre.
5. Soit $A \in \mathbb{C}^{n\times n}$ et soient J une forme normale de Jordan de A, P la matrice de passage associée ($A = PJP^{-1}$).
   Le but de cet exercice est de montrer que le nombre de blocs de Jordan sur J associé à une valeur propre λ est exactement $\dim\ker(A-\lambda I)$.
   a) Soit $S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}$ une matrice blocs diagonale. Montrer que
   $$\operatorname{rang}(S) = \operatorname{rang}(S_1) + \operatorname{rang}(S_2)$$
   Généraliser pour p blocs sur la diagonale. (*Indice* : Considérer les lignes linéairement indépendantes de S_1, S_2).
   b) Soit $B = U + \lambda I \in \mathbb{C}^{q\times q}$ un bloc de Jordan, où U est l'application de décalage. Montrez que la seule valeur propre de B est λ et que l'espace propre associé est engendré par e_1. Déduisez $\dim\ker(B-\lambda I) = 1$ et $\dim\operatorname{Im}(B-\lambda I) = q-1$.
   c) Soient $B_1, ..., B_k$ l'ensemble des blocs de Jordan sur J associé à une valeur propre λ. Déduire de a) et b) que $\dim\operatorname{Im}(J-\lambda I) = n-k$.
   d) En déduire que $\dim\ker(A-\lambda I) = k$ et que $\ker(A-\lambda I) = \operatorname{span}(Pe_{i_1}, ..., Pe_{i_k})$, où les i_j sont les indices des premières lignes/colonnes des $B_1, ..., B_k$ dans J.
6. Déduire des exercices 4 et 5 que si A est diagonalisable, la forme normale de Jordan J de A est diagonale.
7. Soit $A \in \mathbb{C}^{n\times n}$ une matrice diagonalisable. Montrer que $\ker(A-\lambda I) = \ker(A-\lambda I)^k$ pour toute valeur propre λ de A et pour tout $k > 0$. (*Indice* : Diagonaliser d'abord $(A-\lambda I)$ et $(A-\lambda I)^k$ de manière simultanée).
8. Trouver deux matrices $A \in \mathbb{C}^{n\times n}$ et $B \in \mathbb{C}^{n\times n}$ qui ont le même polynôme caractéristique, mais qui ne sont pas similaires (*Rappel* : A et B sont dites similaires s'il existe une matrice P inversible telle que $B = P^{-1}AP$).
9. Soit $J \in \mathbb{C}^{n\times n}$ une matrice en forme normale de Jordan. Montrer que J et J^T sont similaires. En déduire que pour tout $A \in \mathbb{C}^{n\times n}$, les matrices A et A^T sont similaires.

9 Algèbre linéaire sur les entiers

Quand est-ce qu'un système d'équations linéaires possède une solution en nombre entiers ? Étant donnés $A \in \mathbb{Z}^{m\times n}$ et $b \in \mathbb{Z}^m$, on aimerait décider si

$$Ax = b,\ x \in \mathbb{Z}^n \tag{9.1}$$

est résoluble et trouver toutes les solutions.

Une condition nécessaire pour que le système (9.1) soit soluble est qu'il existe une solution $x \in \mathbb{Q}^n$, et donc que $\operatorname{rang}(A) = \operatorname{rang}(A|b)$. Aussi, on peut supprimer une ligne de $(A|b)$ qui est dans le span des autres lignes. On peut alors supposer que A est de rang ligne plein, c'est-à-dire que $\operatorname{rang}(A) = m$.

Définition 9.1. Un nombre entier $d \in \mathbb{Z}$ *divise* un nombre entier $a \in \mathbb{Z}$ s'il existe un nombre entier $x \in \mathbb{Z}$ tel que $d \cdot x = b$. On note alors $d \mid b$, et si d ne divise pas b on écrit $d \nmid b$.

Définition 9.2. Un nombre $d \in \mathbb{Z}$ est un diviseur commun de $a \in \mathbb{Z}$ et $b \in \mathbb{Z}$, si $d \mid a$ et $d \mid b$. Si $\max\{|a|, |b|\} \geq 1$, l'ensemble des diviseurs commun de a et b est un ensemble fini. Dans ce cas, on dénote le *plus grand diviseur commun* de a et b par $\gcd(a, b)$.

Théorème 9.1. *Soient $a, b \in \mathbb{Z}$ et $\max\{|a|, |b|\} \geq 1$. On a*

$$\gcd(a, b) = \min\{x \cdot a + y \cdot b \colon x, y \in \mathbb{Z}, x \cdot a + y \cdot b \geq 1\}.$$

Démonstration. Soit d un diviseur commun de a et b. Alors il existe $x^*, y^* \in \mathbb{Z}$ tel que $a = d \cdot x^*$ et $b = d \cdot y^*$. Si $x \cdot a + y \cdot b \geq 1$, où $x, y \in \mathbb{Z}$, alors

$$x \cdot a + y \cdot b = (x \cdot x^* + y \cdot y^*)d \geq |d|.$$

Par conséquent, on a $d \leq \min\{x \cdot a + y \cdot b \colon x, y \in \mathbb{Z}, x \cdot a + y \cdot b \geq 1\}$. Montrons que $\min\{x \cdot a + y \cdot b \colon x, y \in \mathbb{Z}, x \cdot a + y \cdot b \geq 1\}$ est un diviseur commun de a et b. Supposons que $\min \nmid a$. Alors la division avec reste implique l'existence de $q, r \in \mathbb{Z}$ tels que

$$a = q \cdot \min + r \quad \text{et} \quad 1 \leq r < \min.$$

Soient x, y les entiers qui vérifient $\min = x \cdot a + y \cdot b$, alors

$$1 \leq r = a - q \cdot (x \cdot a + y \cdot b) = (1 - q \cdot x)a - qy \cdot b.$$

Or r est strictement plus petit que $\min$, ce qui est absurde. Il suit donc que $\min \mid a$ et, de façon analogue, $\min \mid b$. Aussi, on a montré que pour tout diviseur commun d de a et b, $\min\{x \cdot a + y \cdot b \colon x, y \in \mathbb{Z}, x \cdot a + y \cdot b \geq 1\} \geq d$ et donc que $\min\{x \cdot a + y \cdot b \colon x, y \in \mathbb{Z}, x \cdot a + y \cdot b \geq 1\} = \gcd(a, b)$

□

Corollaire 9.2. *Soient $a, b \in \mathbb{Z}$ et $\max\{|a|, |b|\} \geq 1$. Le $\gcd(a, b)$ est le diviseur commun positif qui est divisé par chaque diviseur commun de a et b.*

Pour calculer le plus grand diviseur commun de a et b on peut utiliser l'algorithme d'Euclide. Soient $a_0 \geq a_1 \in \mathbb{Z}$ pas tout les deux nuls. Si $a_1 = 0$, alors

$$\gcd(a_0, a_1) = a_0.$$

Autrement, on applique la division avec reste

$$a_0 = q_1 a_1 + a_2,$$

où $q_1, a_2 \in \mathbb{Z}$ et $0 \leq a_2 < a_1$. Un nombre entier $d \in \mathbb{Z}$ est un diviseur commun de a_0 et a_1 si et seulement si d est un diviseur commun de a_1 et a_2. L'algorithme d'Euclide est le procédé de calculer la suite $a_0, a_1, a_2, \ldots, a_{k-1}, a_k \in \mathbb{Z}$ où $a_{k-1} > 0$, $a_k = 0$ et

$$a_{i-1} = q_i a_i + a_{i+1}$$

est le résultat de la division avec reste de a_{i-1} par a_i.

Exemple 9.1. On calcule le plus grand diviseur commun de $a_0 = 52$ et $a_1 = 22$:

$$52 = 2 \cdot 22 + 8, \quad 22 = 2 \cdot 8 + 6, \quad 8 = 1 \cdot 6 + 2, \quad 6 = 3 \cdot 2 + 0.$$

La suite est alors $a_0 = 52, a_1 = 22, a_3 = 8, a_4 = 6, a_5 = 2, a_6 = 0$. Ainsi $\gcd(52, 22) = 2$.

Le calcul des suites a_i et q_i donne aussi une représentation $\gcd(a_0, a_1) = x \cdot a_0 + y \cdot a_1$, $x, y \in \mathbb{Z}$. En effet

$$\begin{pmatrix} a_i \\ a_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \begin{pmatrix} a_{i-1} \\ a_i \end{pmatrix}$$

et alors

$$\begin{pmatrix} a_{k-1} \\ a_k \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_{k-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}.$$

Exemple 9.2. On continue l'exemple 9.1.

$$\begin{aligned} \begin{pmatrix} 2 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 5 \\ 22 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -7 \\ -11 & 26 \end{pmatrix} \begin{pmatrix} 52 \\ 22 \end{pmatrix} \end{aligned}$$

Alors $2 = \gcd(52, 22) = 3 \cdot 52 - 7 \cdot 22$.

Nous pouvons donc résoudre le problème (9.1) dans le cas où $m = 1$ et $n = 2$.

Théorème 9.3. *Soient $a, b \in \mathbb{Z}$ pas tous les deux nuls et $c \in \mathbb{Z}$. L'équation*

$$x \cdot a + y \cdot b = c,\ x, y \in \mathbb{Z} \tag{9.2}$$

possède une solution si et seulement si $\gcd(a, b) \mid c$.

Démonstration. Soient $x', y' \in \mathbb{Z}$ tel que $x'a + y'b = \gcd(a, b)$. Si $\gcd(a, b) \mid c$ alors il existe un $z \in \mathbb{Z}$ tel que $z \cdot \gcd(a, b) = c$ et $zx'a + zy'b = c$ est une solution en nombre entiers de (9.2).

S'il existe une solution de (9.2), alors chaque diviseur commun de a et b est aussi un diviseur de c. □

Exercices

1. Démontrer le Corollaire 9.2.
2. Soient $n \geq 2$ et $a_1, \ldots, a_n \in \mathbb{Z}$ pas tous égaux à zéro. On définit $\gcd(a_1, \ldots, a_n)$ comme étant le plus grand diviseur commun de $a_1, \ldots, a_n$. Montrer :
   i) $\gcd(a_1, \ldots, a_n) = \min\{x_1a_1 + \cdots + x_na_n \colon x_1a_1 + \cdots + x_na_n \geq 1,\ x_i \in \mathbb{Z},\ i = 1, \ldots, n\}$.
   ii) $\gcd(a_1, \ldots, a_n) = \gcd(\gcd(a_1, a_2), a_3, \ldots, a_n)$ pour $n \geq 3$.
3. Soit $a_0 \geq a_1 \geq \ldots \geq a_k = 0$ la suite calculée par l'algorithme d'Euclide. Montrer $a_{i-1} \geq 2 \cdot a_{i+1}$ et conclure que $k \leq 2 \cdot \log_2(a_0) + 1$.

9.1 La forme normale d'Hermite

Maintenant on s'occupe du problème (9.1) où $A \in \mathbb{Z}^{m \times n}$ et $b \in \mathbb{Z}^m$ et $\operatorname{rang}(A) = m$.

Lemme 9.4. *Soit $A \in \mathbb{Z}^{n \times n}$ une matrice inversible (sur $\mathbb{Q}$), alors $A^{-1} \in \mathbb{Z}^{n \times n}$ si et seulement si $\det(A) = \pm 1$.*

Démonstration. Supposons que $A^{-1} \in \mathbb{Z}^{n \times n}$. Alors $1 = \det(I_n) = \det(A^{-1}) \det(A)$. Les deux facteurs $\det(A^{-1})$ et $\det(A)$ sont des nombres entiers. Les seul diviseurs de 1 en nombre entiers sont 1 et -1.

Réciproquement, si $\det(A) = \pm 1$, on a

$$A^{-1} = \operatorname{ad}(A)/\det(A) \in \mathbb{Z}^{n \times n}.$$

où $\operatorname{ad}(A) \in \mathbb{Z}^{n \times n}$ est la matrice complémentaire de A. On se rappelle que $(\operatorname{ad}(A))_{ij} = (-1)^{i+j} \det(A_{ji})$ où $A_{ji} \in \mathbb{Z}^{(n-1) \times (n-1)}$ est la matrice qu'on obtient de A en supprimant la j-ème ligne et i-ème colonne. □

Définition 9.3. Une matrice $U \in \mathbb{Z}^{n \times n}$ telle que $\det(U) = \pm 1$ est appelée *unimodulaire*.

Remarque 9.5. Soit $U \in \mathbb{Z}^{n \times n}$ une matrice unimodulaire. Un $x^* \in \mathbb{Z}^n$ est une solution du problème (9.1) si et seulement $U^{-1}x^* \in \mathbb{Z}^n$ est une solution du problème

$$AUx = b,\ x \in \mathbb{Z}^n. \tag{9.3}$$

L'idée est maintenant de trouver une matrice unimodulaire $U \in \mathbb{Z}^{n\times n}$ telle que

$$A \cdot U = [H|0] \tag{9.4}$$

où

$$H = \begin{pmatrix} h_{11} & & & \\ h_{21} & h_{22} & & \\ & & \ddots & \\ h_{m1} & \dots & \dots & h_{mm} \end{pmatrix}$$

est une matrice triangulaire. Le problème (9.1) est alors soluble si et seulement si $H^{-1}b \in \mathbb{Z}^n$.

Définition 9.4. Soit $A \in \mathbb{Z}^{m\times n}$ une matrice. Une *opération élémentaire unimodulaire* est l'une des trois opérations suivantes

i) Multiplier une colonne par -1.

ii) Échanger deux colonnes de A.

iii) Additionner un multiple entier d'une colonne de A à une autre colonne de A.

Exemple 9.3. Une suite d'opérations élémentaire unimodulaire sur A correspond à la multiplication $A \cdot U$ où $U \in \mathbb{Z}^{n\times n}$ est une matrice unimodulaire. Soit

$$A = \begin{pmatrix} 3 & 6 & 2 \\ 11 & 5 & 7 \end{pmatrix}.$$

Échanger les colonnes 1 et 2 correspond à la multiplication à droite de A avec la matrice unimodulaire

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

$$\begin{pmatrix} 6 & 3 & 2 \\ 5 & 11 & 7 \end{pmatrix} = A \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Additionner -3 fois la colonne 3 sur la colonne 1 est la multiplication à droite de A avec la matrice unimodulaire

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

$$\begin{pmatrix} 0 & 3 & 2 \\ -16 & 11 & 7 \end{pmatrix} = A \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$

Exemple 9.4. Est-ce que le système

$$\begin{pmatrix} 3 & 6 & 2 \\ 11 & 5 & 10 \end{pmatrix} x = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, x \in \mathbb{Z}^3 \tag{9.5}$$

possède une solution ? Soustrayons la colonne 3 à la colonne 1 :

$$\begin{pmatrix} 1 & 6 & 2 \\ 1 & 5 & 10 \end{pmatrix}$$

Ensuite, on soustrait six fois la colonne 1 à la colonne 2 et deux fois la colonne 1 à la colonne 3 :

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

Puis, on additionne huit fois la colonne 2 à la colonne 3 :

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

Le système (9.5) se réduit finalement à résoudre

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{pmatrix} y = b, y \in \mathbb{Z}^3$$

où $y = U^{-1}x$ pour une matrice unimodulaire U adéquate. On conclut donc qu'il possède une solution entière. En fait

$$\begin{pmatrix} 3 & 6 & 2 \\ 11 & 5 & 10 \end{pmatrix} x = b, x \in \mathbb{Z}^3$$

est soluble pour tout $b \in \mathbb{Z}^2$.

Lemme 9.6. *Soit $A \in \mathbb{Z}^{m \times n}$ une matrice à coefficients entiers, alors il existe une matrice unimudulaire $U \in \mathbb{Z}^{n \times n}$ telle que la première ligne de AU est de la forme $(d, 0, \cdots, 0)$, où $d \in \mathbb{Z}$.*

Démonstration. Si la première ligne n'est pas de cette forme, et si elle possède seulement une composante qui n'est pas égale à zéro, on échange les colonnes de sorte que la matrice résultante soit de la forme souhaitée.

Autrement, il existe deux indices de colonne $j_1 \neq j_2$, tels que $a_{1j_1} \neq 0$ et $a_{1j_2} \neq 0$. On peut supposer, quitte à permutter les colonnes j_1 et j_2, que $|a_{1j_1}| \geq |a_{1j_2}|$. La division avec reste nous donne des entiers $q \in \mathbb{Z}$ et $0 \leq r < |a_{1j_2}|$ tels que

$$a_{1j_1} = q \cdot a_{1j_2} + r.$$

On applique l'opération unimodulaire : *Soustraire q fois la colonne j_2 à la colonne j_1*, ce qui a pour effet de remplacer a_{1j_1} par r et de laisser les autres composantes de la première ligne intactes. Comme

$$0 < |r| + |a_{1j_2}| < |a_{1j_1}| + |a_{1j_2}|$$

ce procédé ne peut être répété infiniment. Il existe alors une matrice unimodulaire qui transforme A en une matrice dont la première ligne possède une seule composante non nulle. Un échange de colonnes adéquat donne la forme désirée. □

Corollaire 9.7. *Soit $A \in \mathbb{Z}^{m\times n}$ une matrice en nombre entiers. Alors il existe une matrice unimudulaire $U \in \mathbb{Z}^{n\times n}$ telle que $A \cdot U$ est de la forme* (9.4).

Démonstration. On raisonne par récurrence sur m. Le cas $m = 1$ suit directement du Lemme 9.6. Soit $m > 1$. Le Lemme 9.6 implique qu'il existe une matrice unimodulaire $U_1 \in \mathbb{Z}^{n\times n}$ telle que

$$A \cdot U_1 = \begin{pmatrix} d & 0\cdots 0 \\ a & A' \end{pmatrix}$$

où $d \in \mathbb{Z}$, $a \in \mathbb{Z}^{m-1}$ et $A' \in \mathbb{Z}^{(m-1)\times(n-1)}$. Par l'hypothèse de récurrence, il existe une matrice unimudulaire $U_2 \in \mathbb{Z}^{(n-1)\times(n-1)}$ telle que $A'U_2$ est de la forme désirée. Clairement

$$\begin{pmatrix} 1 & 0^T \\ 0 & U_2 \end{pmatrix} \in \mathbb{Z}^{n\times n}$$

est une matrice unimodulaire et

$$AU_1 \begin{pmatrix} 1 & 0^T \\ 0 & U_2 \end{pmatrix}$$

est de la forme (9.4). □

Définition 9.5. Une matrice $A \in \mathbb{Z}^{m\times n}$ est en *forme normale d'Hermite*, si elle est de la forme (9.4), où $h_{ii} > 0$ pour tout $1 \le i \le m$ et $0 \le h_{ij} < h_{ii}$ pour tout $1 \le j < i \le m$.

Théorème 9.8. *Soit $A \in \mathbb{Z}^{m\times n}$ une matrice en nombre entiers. Alors il existe une matrice unimudulaire $U \in \mathbb{Z}^{n\times n}$ telle que $A \cdot U$ est en* forme normale d'Hermite.

Définition 9.6. Soit $A \in \mathbb{Z}^{m\times n}$ et $rang(A) = m$. L'ensemble $\Lambda(A) := \{Ax : x \in \mathbb{Z}^n\}$ est un *réseau entier généré* de A. Une matrice $B \in \mathbb{Z}^{m\times m}$ telle que $\Lambda(A) = \Lambda(B)$ est appelée base de $\Lambda(A)$.

Remarque 9.9. Une *base* $B \in \mathbb{Z}^{m\times m}$ de $\Lambda(A)$ est inversible comme matrice réelle. C'est à dire $\det(B) \in \mathbb{Z} \setminus \{0\}$.

Corollaire 9.10. *Chaque réseau entier possède une base.*

Théorème 9.11. *Soient $A, B \in \mathbb{Z}^{m\times m}$ en forme normale d'Hermite. Alors $\Lambda(A) = \Lambda(B)$ si et seulement si $A = B$.*

Démonstration. Si $A = B$, alors $\Lambda(A) = \Lambda(B)$.

Supposons alors que $\Lambda(A) = \Lambda(B)$. Maintenant on montre qui si $A \neq B$ alors $\Lambda(A) \neq \Lambda(B)$. On note $A = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ a_{m1} & & a_{mm} \end{pmatrix}$ et $B = \begin{pmatrix} b_{11} & & \\ & \ddots & \\ b_{m1} & & b_{mm} \end{pmatrix}$. Soit i l'indice minimal tel que la i-ème ligne de A et celle de B soient différentes. Alors $\exists j \in \{1, \ldots, i\}$ tel que $a_{ij} \neq b_{ij}$ et sans perte de généralité on a $a_{ij} > b_{ij}$. Clairement en notant A_j (resp. B_j) la j-ème colonne de A (resp. B), on a $A_j - B_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{ij} - b_{ij} \\ \vdots \end{pmatrix} \in \Lambda(A)$ avec $a_{ii} > a_{ij} - b_{ij} > 0$.

Il existe donc un vecteur entier $z \in \mathbb{Z}^n$ tel que $A_j - B_j = Az$. On montre aisément que les $i - 1$ premières coordonées de z doivent être nulles car aucun élément diagonal de A ne l'est (ou, de manière équivalente, car A et de rang plein). Or, en comparant la i-ème coordonnée, on trouve que $a_{ij} - b_{ij} = a_{ii} z_i$. Dès lors, $a_{ii} | a_{ij} - b_{ij}$ ce qui est une contradiction. □

Remarque 9.12. Le Théorème 9.8 nous permet de vérifier, pour $A \in \mathbb{Z}^{m \times n_1}$ et $B \in \mathbb{Z}^{m \times n_2}$ de rang ligne pleins, si $\Lambda(A) = \Lambda(B)$. On calcule $(H_A | 0)$ et $(H_B | 0)$ les formes normales d'Hermite de A et B. Comme $\Lambda(A) = \Lambda(H_A)$ et $\Lambda(B) = \Lambda(H_B)$, on a que $\Lambda(A) = \Lambda(B)$ si et seulement si $H_A = H_B$.

Exemple 9.5. On va transformer $A = \begin{pmatrix} 4 & 6 & 10 \\ 6 & 12 & 9 \end{pmatrix}$ en forme normale d'Hermite afin de trouver toutes les solutions entières de

$$\begin{pmatrix} 4 & 6 & 10 \\ 6 & 12 & 9 \end{pmatrix} x = \begin{pmatrix} 6 \\ 3 \end{pmatrix}, \; x \in \mathbb{Z}^3. \tag{9.6}$$

$$\begin{array}{ll}
\begin{pmatrix} 4 & 6 & 10 \\ 6 & 12 & 9 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 4 & 2 & 2 \\ 6 & 6 & -3 \end{pmatrix} & \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 2 & 4 & 2 \\ 6 & 6 & -3 \end{pmatrix} & \begin{pmatrix} -1 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 2 & 0 & 0 \\ 6 & -6 & -9 \end{pmatrix} & \begin{pmatrix} -1 & 3 & -1 \\ 1 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 2 & 0 & 0 \\ 6 & 3 & -9 \end{pmatrix} & \begin{pmatrix} -1 & 4 & -1 \\ 1 & -1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \\
\begin{pmatrix} 2 & 0 & 0 \\ 6 & 3 & 0 \end{pmatrix} & \begin{pmatrix} -1 & 4 & 11 \\ 1 & -1 & -4 \\ 0 & -1 & -2 \end{pmatrix} \\
\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} & \begin{pmatrix} -9 & 4 & 11 \\ 3 & -1 & -4 \\ 2 & -1 & -2 \end{pmatrix}
\end{array} \tag{9.7}$$

L'ensemble des solutions entières de (9.6) est

$$\left\{ \begin{pmatrix} -23 \\ 8 \\ 5 \end{pmatrix} + \begin{pmatrix} 11 \\ -4 \\ -2 \end{pmatrix} \cdot z \colon z \in \mathbb{Z} \right\}$$

Définition 9.7. Soit $A \in \mathbb{Z}^{m \times n}$. Le *noyau* de A sur $\mathbb{Z}$ est défini par $ker_{\mathbb{Z}}(A) := \{y \in \mathbb{Z}^n, Ay = 0\}$.

Théorème 9.13. *Soient $A, B \in \mathbb{Z}^{m \times m}$ en forme normale de Hermite avec $rang(A) = rang(B) = m$. Alors $\Lambda(A) = \Lambda(B) \Leftrightarrow \exists U \in \mathbb{Z}^{m \times m}$ unimodulaire telle que $B = AU$.*

Démonstration. Si $\Lambda(A) = \Lambda(B)$ alors $A = B \cdot P, P \in \mathbb{Z}^{m \times m}$ et $B = A \cdot Q, Q \in \mathbb{Z}^{m \times m}$. Alors on obtient $A = AQP$ ce qui implique $QP = I_m$ et donc on a bien P, Q unimodulaires.

Si $B = AU, U \in \mathbb{Z}^{m \times m}$ unimodulaire alors comme $U\mathbb{Z}^m = \mathbb{Z}^m$ on obtient immédiatement $\Lambda(A) = \Lambda(B)$. □

Définition 9.8. Soient $A \in \mathbb{Z}^{m \times n}$ avec $rang(A) = m$ et $\Lambda(A)$ le réseau entier de A. Le *déterminant* du réseau $\Lambda(A)$ est donné par $\det(\Lambda(A)) := |\det(B)|$ où $B \in \mathbb{Z}^{m \times m}$ est une base de $\Lambda(A)$

Remarque 9.14. Le théorème 9.13 assure que $|\det(B)|$ ne dépend pas du choix de B et donc que $\det(\Lambda(A))$ a du sens.

Exercices

1. Soit $A \in \mathbb{Z}^{m\times n}$ avec rang ligne plein. Le noyaux entier de A est l'ensemble

$$\ker_{\mathbb{Z}}(A) = \{x \in \mathbb{Z}^n \colon Ax = 0\}.$$

Soit $U \in \mathbb{Z}^{n\times n}$ une matrice unimodulaire telle que

$$A \cdot U = (H|0)$$

est la forme normale d'Hermite. Montrer que $\ker_{\mathbb{Z}}(A) = \{y_1u_1 + \cdots + y_{n-m}u_{n-m} \colon y_i \in \mathbb{Z}\}$ où $u_1, \ldots, u_{n-m}$ sont les dernières $n-m$ colonnes de U.

La forme normale de Smith

Théorème 9.15. *Soit $A \in \mathbb{Z}^{m\times n}$, $A \neq 0$ alors il existent $U \in \mathbb{Z}^{m\times m}$ et $V \in \mathbb{Z}^{n\times n}$ unimodulaires telle que*

$$U \cdot A \cdot V = \begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_k \\ & & \end{pmatrix}, \qquad \text{avec } \delta_i \in \mathbb{N}_{\geq 1} \text{ et } \delta_1 | \cdots | \delta_k \tag{9.8}$$

et les coefficients non spécifiés sont 0.

La matrice (9.8) est appelée la *forme normale de Smith* de A.

Démonstration. On raisonne par récurrence sur $\min\{m, n\}$. Si $\min\{m, n\} = 1$, la forme normale de Hermite de A ou son transposé est en forme normale Smith.

Maintenant, soit $\min\{m, n\} > 1$. Soient $U \in \mathbb{Z}^{m\times m}$ et $V \in \mathbb{Z}^{n\times n}$ unimodulaires telle que la valeur absolue minimale des éléments non-zéro de la matrice $U \cdot A \cdot V$ est minimale, alors tel que

$$\min\{|(U \cdot A \cdot V)_{i,j}| : (U \cdot A \cdot V)_{i,j} \neq 0,\ 1 \leq i \leq m,\ 1 \leq j \leq n\}$$

est minimal. Soit cette valeur minimale d. En échangeant lignes et colonnes et en multipliant la par -1 si nécessaire, on peut supposer que $d = (U \cdot A \cdot V)_{1,1}$. Tous composantes de la première ligne de $U \cdot A \cdot V$ sont des multiples de d. Autrement, si $(U \cdot A \cdot V)_{1,j} = z$ où z n'est pas un multiple de d, la division avec reste de z par d donne

$$z = q \cdot d + r,\ 0 < r < d.$$

Une opération élémentaire de colonne (soustraire q fois colonne 1 de colonne j) donne composante i, j égal à r. Ceci est une contradiction à d minimal. Également, d divise toute composante de la première colonne de $U \cdot A \cdot V$. Alors on peut éliminer toute autre composante de la première ligne et colonne. Alors on peut supposer que

$$U \cdot A \cdot V = \begin{pmatrix} d & \\ & B \end{pmatrix}, \qquad \text{où } B \in \mathbb{Z}^{(m-1)\times(n-1)}. \tag{9.9}$$

Maintenant, si $b_{i,j}$ n'est pas un multiple de d, on peut additionner la ligne $i+1$ de $U \cdot A \cdot V$ sur la première ligne. Comme avant on obtient le reste de la division comme composante après une opération élémentaire de colonnes. Ceci démontre que $d \in \mathbb{N}_+$ divise toute composante de B.

Par récurrence, B possède une forme normale Smith et ses composantes sont tous des multiples de d.

□

Sous groupes de $\mathbb{Z}^n$

À partir de maintenant, on se donne une matrice $A \in \mathbb{Z}^{m\times n}$ avec $rang(A) = k$ et k n'est pas forcément m.

Définition 9.9. Soit $A \in \mathbb{Z}^{m\times n}$ alors $\Lambda(A) = \{Ax : x \in \mathbb{Z}^n\}$ est le *réseau entier général* de A.

Théorème 9.16. *Soit $A \in \mathbb{Z}^{m\times n}$ avec $rang(A) = k$ alors $\exists B \in \mathbb{Z}^{m\times k}$ tq. $\Lambda(A) = \Lambda(B)$. B est alors appelée une base générale de $\Lambda(A)$.*

Remarque 9.17. $rang(A) = rang(B) = k$.

Démonstration. Supposons que les k premières lignes de A sont linéairement indépendantes. Dès lors on a $A = \begin{pmatrix} A' \\ A'' \end{pmatrix}$ avec $A' \in \mathbb{Z}^{k\times n}$ et $rang(A') = k$ soit alors $U \in \mathbb{Z}^{n\times n}$ unimodulaire tq.

$$A' \cdot U = [H \mid 0]$$

ou $H \in \mathbb{Z}^{k\times k}$ est en forme normale de Hermite. Alors on peut se convaincre en raisonnant sur les rangs que $AU = \begin{pmatrix} H & 0 \\ C & 0 \end{pmatrix}$ où $C \in \mathbb{Z}^{(m-k)\times k}$. Dés lors on a

$$\begin{aligned} \Lambda(A) &= A\mathbb{Z}^n \\ &= A \cdot U\, \mathbb{Z}^n \\ &= \begin{bmatrix} H \\ C \end{bmatrix} \mathbb{Z}^k \\ &= \Lambda\left(\begin{bmatrix} H \\ C \end{bmatrix}\right). \end{aligned}$$

□

Théorème 9.18. *Soit G un sous groupe de $\mathbb{Z}^n$ alors il existe $B \in \mathbb{Z}^{n\times k}$ avec $rang(B) = k$ et tq. $\Lambda(B) = G$*

Démonstration. Soit $(v_1, \ldots, v_k) \in G^k$ tq. $(v_1, \ldots, v_k)$ est une base du sous espace $\text{span}(G) \subseteq \mathbb{R}^n$. Alors posons $B = (v_1 \ldots v_k) \in \mathbb{Z}^{n\times k}$. Si $\Lambda(B) = G$ et c'est terminé. Si $\Lambda(B) \subsetneq G$ alors il existe $v^* \in G - \Lambda(B)$. Soit alors $B^* \in \mathbb{Z}^{n\times k}$ une base générale du

réseau général $G \supseteq \Lambda(v_1 \dots v_k v^*) \supsetneq \Lambda(B)$. Alors il existe $U \in \mathbb{Z}^{k \times k}$ tq. $B = B^* U$ et on a nécessairement $|\det(U)| \in \mathbb{N}_{\geq 2}$. Dés lors on peut remarquer que $|\det(B^T B)| = \det(U)^2 \times |\det({B^*}^T B^*)|$ et donc $|\det(B^T B)| \leq \frac{1}{4}|\det(B^T B)|$ mais comme $|\det({B^*}^T B^*)| \geq 1$ et on peut répéter ces opérations sur B^* mais ce procédé ne peut pas continuer indéfiniment. □

10 Groupes

Dans ce chapitre on s'intéresse à l'étude des groupes.

10.1 Groupes abéliens engendrés finis

Définition 10.1. Soit $(G,+)$ un groupe alors $H \subset G$ est un sous groupe si $(H,+|_H)$ est un groupe.

Lemme 10.1. *$(H,\times|_H)$ est un sous groupe de $(G,\times) \Leftrightarrow \forall a,b \in H \quad a \times b^{-1} \in H$*

Définition 10.2. $H \leq G$ est un sous groupe normal de G ssi $\forall g \in G \quad Hg = gH$ On note $H \trianglelefteq G$.

Définition 10.3. Soient $g, g' \in G$ alors on pose $Hg \circ Hg' := H(gg')$.

Théorème 10.2. *$\circ$ munit $G/_H := \{gH, \quad g \in G\}$ d'une structure de groupe ssi $H \trianglelefteq G$.*

Définition 10.4. Soient $(G,+),(G',\times)$ deux groupes et $\phi : G \to G'$ une application alors ϕ est un homomorphisme ssi $\forall a,b \in G, \phi(a+b) = \phi(a) \times \phi(b)$.

Remarque 10.3. *$ker(\phi) \trianglelefteq G$.*

Théorème 10.4. *Si ϕ est un morphisme alors $G/_{ker(\phi)} \cong Im(\phi)$.*

Définition 10.5. Soit $(G,+)$ un groupe abélien. On dit que $(G,+)$ est engendré fini si $\exists g_1,\ldots,g_n \in G$ tqs. $G = \{x_1g_1 + \cdots + x_ng_n, x_i \in \mathbb{Z}\}$.

Définition 10.6. Soient $(G_1,+),(G_2,\times)$ deux groupes alors $G_1 \otimes G_2 := (G_1 \times G_2, \bullet)$ avec $\forall (g_1,g_2),(g_1',g_2') \in G_1 \times G_2, (g_1,g_2) \bullet (g_1',g_2') = (g_1+g_2, g_1' \times g_2')$.

Remarque 10.5. *$G_1 \otimes G_2$ est un groupe.*

Lemme 10.6. *Soit $\phi : G \to G$ un automorphisme et $H \trianglelefteq G$ alors $G/_H \cong G/_{\phi(H)}$.*

Théorème 10.7. *Soit G un groupe abélien engendré fini alors $\exists d_1,\ldots,d_k \in \mathbb{N}_{\geq 1}$ et $l \in \mathbb{N}$ tq $G \cong \mathbb{Z}_{d_1} \otimes \cdots \otimes \mathbb{Z}_{d_k} \otimes \mathbb{Z}^l$ avec $\mathbb{Z}^l := \mathbb{Z} \otimes \cdots \otimes \mathbb{Z}$ l fois. De plus on a $d_1 | \ldots | d_k$.*

Démonstration. Soient $g_1,\ldots,g_n \in G$ tqs. $G = \{x_1g_1 + \cdots + x_ng_n, x_i \in \mathbb{Z}\}$ alors on peut vérifier que $\phi : \mathbb{Z}^n \to G, (x_1,\ldots,x_n) \mapsto x_1g_1 + \cdots + x_ng_n$ est un homomorphisme surjectif. Mais alors $ker(\phi) \leq \mathbb{Z}^n$ et donc $\exists B \in \mathbb{Z}^{n \times k}$ avec $rang(B) = k$ et $\Lambda(B) = ker(\phi)$. En calculant la forme normale de Smith de B on obtient que $\exists U \in \mathbb{Z}^{n \times n}, V \in$

$\mathbb{Z}^{k\times k}$ unimodulaires tqs. $B = U\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_k \\ 0 & & \end{pmatrix}V$ avec $1 \leq d_1|\dots|d_k$ et donc on a $ker(\phi) = \left\{ U\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_k \\ 0 & & \end{pmatrix} y, y \in \mathbb{Z}^k \right\}$. Alors par les théorèmes explicités plus haut on obtient $G \cong \mathbb{Z}^n/\left\{ U\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_k \\ 0 & & \end{pmatrix} y, y \in \mathbb{Z}^k \right\}$. Puis comme $U : \mathbb{Z}^n \to \mathbb{Z}^n$ est un automorphisme on arrive finalement à
$G \cong \mathbb{Z}^n/\{(d_1y_1, \dots, d_ky_k, 0, \dots, 0), y_i \in \mathbb{Z}\}$. Il suffit maintenant de constater que
$\mathbb{Z}^n/\{(d_1y_1, \dots, d_ky_k, 0, \dots, 0), y_i \in \mathbb{Z}\} \cong \mathbb{Z}_{d_1} \otimes \dots \otimes \mathbb{Z}_{d_k} \otimes \mathbb{Z}^{n-k}$. □

Lemme 10.8. *Soient L, K deux groupes abéliens, M un sous groupe normal de K $f : K \to L$ un isomorphisme. Alors $K/_M \cong L/_{\phi(M)}$.*

Lemme 10.9. *Soient G, H_1, H_2 trois groupes abéliens avec $|H_i| < \infty$ pour $i = 1, 2$ alors si $G \cong H_1 \otimes \mathbb{Z}^{n_1}$ et $G \cong H_2 \otimes \mathbb{Z}^{n_2}$ alors $H_1 \cong H_2$ et $n_1 = n_2$.*

Démonstration. Soit $\phi : H_1 \otimes \mathbb{Z}^{n_1} \to H_2 \otimes \mathbb{Z}^{n_2}$ un isomorphisme. Soit $h \in H_1$ alors $\phi(h, 0) = (h', x) \in H_2 \times \mathbb{Z}^{n_2}$ mais comme ϕ préserve l'ordre on a nécessairement $x = 0$ on a ainsi $\phi(H_1, 0) \subset (H_2, 0)$ et de même $\phi^{-1}(H_2, 0) \subset (H_1, 0)$ et donc finalement $\phi(H_1, 0) = (H_2, 0)$ ce qui permet de conclure que $H_1 \cong H_2$.
Maintenant on a d'après le lemme ci dessus : $\mathbb{Z}^{n_2} \cong$
$H_2 \otimes \mathbb{Z}^{n_2}/_{H_2 \otimes \{0_{\mathbb{Z}^{n_2}}\}} \cong H_1 \otimes \mathbb{Z}^{n_1}/_{H_1 \otimes \{0_{\mathbb{Z}^{n_1}}\}} \cong \mathbb{Z}^{n_1}$. Mais alors on a nécessairement $n_1 = n_2$. En effet et sans perte de généralité supposons $n_1 > n_2$ alors soient $x_1, \dots, x_{n_1} \in \mathbb{Z}^{n_1}$ des vecteurs $\mathbb{Q}$ linéairement indépendants alors on a forcement $\phi(x_1), \dots, \phi(x_{n1}) \in \mathbb{Z}^{n_2}\mathbb{Q}$ linéairement dépendants. Ainsi $\exists \alpha_1, \dots, \alpha_{n_1} \in \mathbb{Z}$ non tous nuls avec $\alpha_1\phi(x_1) + \dots + \alpha_{n_1}\phi(x_{n1}) = 0$ mais comme ϕ est un isomorphisme on a $\phi(\alpha_1x_1 + \dots + \alpha_{n_1}x_{n_1}) = 0$. Cela implique par injectivité que $\alpha_1x_1 + \dots + \alpha_{n_1}x_{n_1} = 0$. Ceci est absurde. □

Théorème 10.10. *Soient $m_1, m_2 \in \mathbb{N}_{\geq}$ alors $\mathbb{Z}_m \cong \mathbb{Z}_{m_1} \otimes \mathbb{Z}_{m_2}$ ssi $m = m_1 \times m_2$ et $pgcd(m_1, m_2) = 1$.*

Démonstration. $\boxed{\Leftarrow}$ Tout d'abord il s'agit de remarquer que $\phi : \mathbb{Z}_m \to \mathbb{Z}_{m_1} \otimes \mathbb{Z}_{m_2}, a \mapsto (a, a)$ est bien définie et est un morphisme de groupe. Ensuite on a $|\mathbb{Z}_m| = |\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}|$ donc il suffit de montrer que ϕ est surjective. En effet soit $(a, b) \in \mathbb{Z}_{m_1} \otimes \mathbb{Z}_{m_2}$ alors on a $1 = pgcd(m_1, m_2) = rm_1 + sm_2, r, s \in \mathbb{Z}$. Dès lors on vérifie que $\phi(rm_1b + sm_2a) = (a, b)$.

$\boxed{\Rightarrow}$ En égalisant les cardinaux on a $m = m_1m_2$ de plus en posant $d = pgcd(m_1, m_2)$ on a $\frac{m}{d} = ordre(d) = ordre(\phi(d)) \leq \frac{m}{d^2}$ et donc $1 \leq d \leq 1$ et donc $d = 1$. □

Remarque 10.11. Soit $n = p_1^{a_1} \dots p_k^{a_k}, p_i \in \mathbb{P} a_i \in \mathbb{N}_{\geq 1}$ alors $Z_n \cong \mathbb{Z}_{p_1^{a_1}} \otimes \dots \otimes \mathbb{Z}_{p_k^{a_k}}$.

Lemme 10.12. *Soient $p \in \mathbb{P}$ et $\alpha_1 \leq \dots \leq \alpha_k, \beta_1 \leq \dots \leq \beta_l \in \mathbb{N}_{\geq 1}$ alors $\mathbb{Z}_{p^{\alpha_1}} \otimes \dots \otimes \mathbb{Z}_{p^{\alpha_k}} \cong \mathbb{Z}_{p^{\beta_1}} \otimes \dots \otimes \mathbb{Z}_{p^{\beta_l}}$ ssi $k = l$ et $\alpha_i = \beta_i \forall i \in \{1, \dots, k\}$.*

Démonstration. $\boxed{\Leftarrow}$ Trivial.
$\boxed{\Rightarrow}$ Soit ϕ l'isomorphisme entre ces deux groupes. On a $p^{\alpha_k} = ordre((0, \dots, 1)) = ordre(\phi((0, \dots, 1))) \leq p^{\beta_l}$. De même on obtient l'inégalité inverse puis finalement on trouve donc $\alpha_k = \beta_l$. Ainsi on a $\mathbb{Z}_{p^{\alpha_1}} \otimes \dots \otimes \mathbb{Z}_{p^{\alpha_{k-1}}} \cong \mathbb{Z}_{p^{\beta_1}} \otimes \dots \otimes \mathbb{Z}_{p^{\beta_{l-1}}}$ puis on prouve le théorème par induction. □

Remarque 10.13. Soit G abélien engendré fini. Alors $G \cong \mathbb{Z}_{d_1} \otimes \dots \otimes \mathbb{Z}_{d_k} \otimes \mathbb{Z}^l$ avec $d_1, \dots, d_k \in \mathbb{N}_{\geq 1}$ tqs $d_1 | \dots | d_k$ et $l \in \mathbb{N}$. Alors $d_k = p_1^{a_1} \dots p_n^{a_n}, p_i \in \mathbb{P}, a_i \in \mathbb{N}_{\geq 1}$. On obtient ainsi $d_i = p_1^{e_1^i} \dots p_n^{e_n^i}$ avec $0 \leq e_j^i \leq a_j$. Dès lors on a que $G \cong \mathbb{Z}_{p_1^{e_1^1}} \otimes \dots \otimes \mathbb{Z}_{p_n^{e_n^1}} \otimes \dots \otimes \mathbb{Z}_{p_1^{a_1}} \otimes \dots \otimes \mathbb{Z}_{p_n^{a_n}} \otimes \mathbb{Z}^l$.

Bibliographie

[Mic26] Philippe Michel. Algèbre linéaire avancée, 2026. Notes du cours 1-er Semestre.