Algèbre linéaire numérique

Pierron Théo

ENS Ker Lann

Table des matières

Chapitre 1

Réduction des matrices carrées

1.1 Définitions

Définition 1.1 On appelle produit hermitien une application $\langle \cdot, \cdot \rangle : \mathbb{C}^2 \to \mathbb{C}$ qui est :

- linéaire à droite : $\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle x, z \rangle$
- hermitienne : $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- définie positive : $\langle x, x \rangle > 0$

Remarque 1.1 $\langle \cdot, \cdot \rangle$ *est semi-linéaire à gauche :* $\langle \lambda x + y, z \rangle = \overline{\lambda} \langle x, z \rangle + \langle y, z \rangle$. *Elle est donc sesquilinéaire.*

Exemple : $\langle x, y \rangle = \sum_{i=1}^{n} \overline{x_i} y_i$ est le produit scalaire hermitien canonique sur $\mathbb{C}^n$.

Remarque 1.2 $x \mapsto \sqrt{\langle x, x \rangle}$ *définit une norme.*

Définition 1.2 Pour $A \in \mathfrak{M}_n(\mathbb{C})$, on note $A^* = \overline{{}^t A}$.

Remarque 1.3 $A^{**} = A$ *et* $\langle Ax, y \rangle = \langle x, A^* y \rangle$.

Notations :

- On note $\mathscr{H}_n(\mathbb{C})$ l'ensemble des matrices $n \times n$ hermitiennes (telles que $A^* = A$).
- On note $\mathscr{H}_n^+(\mathbb{C})$ l'ensemble des matrices $n \times n$ hermitiennes positives.
- On note $\mathscr{H}_n^{++}(\mathbb{C})$ l'ensemble des matrices $n \times n$ hermitiennes définies positives.
- On note $\mathcal{U}_n(\mathbb{C})$ les matrices $n \times n$ unitaires (telles que $A^* A = I_n$).
- On note $SU_n(\mathbb{C})$ l'ensemble des matrices $n \times n$ unitaires de déterminant 1.

Remarque 1.4 *Si* $A \in \mathscr{H}_n(\mathbb{C})$, $\langle Ax, x \rangle \in \mathbb{R}$.

Exemples :
- Si $A \in \mathscr{H}_n^{++}(\mathbb{C})$, $(x,y) \mapsto x^*Ay = \langle Ax, y\rangle$ définit un produit scalaire hermitien.
- Réciproquement, tout produit scalaire hermitien s'écrit sous la forme $A = (\langle e_i, e_j\rangle)_{i,j}$.

1.2 Réduction de matrices carrées

1.2.1 Définitions

Définition 1.3 Pour $A \in \mathfrak{M}_n(\mathbb{C})$, on définit :
- $\chi_A = \det(A - XI_n)$ le polynôme caractéristique de A.
- $\mathrm{Sp}(A) = \{\lambda \in \mathbb{C}, \chi_A(\lambda) = 0\}$ le spectre de A.
- Les élements de $\mathrm{Sp}(A)$ sont les valeurs propres de A.
- Un vecteur propre de A associé à λ est un $x \neq 0$ tel que $Ax = \lambda x$.

Définition 1.4 Pour $\lambda \in \mathrm{Sp}(A)$, on définit :
- la multiplicité algébrique de λ est le plus grand k tel que $(X-\lambda)^k | \chi_A$.
- la multiplicité géométrique de λ est $l = \dim(E_\lambda) = \dim(\mathrm{Ker}(A - \lambda I_n))$.
- λ est dite défective si $l < k$.

Exemple :
- $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: $\chi_A = (1-X)^2$, $k = 2$, $l = 1$.
- $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$: $\chi_A = (1-X)(2-X)$, les valeurs propres sont non-défectives et valent 1 et 2.

Définition 1.5 $A \in \mathfrak{M}_n(\mathbb{K})$ est dite diagonalisable sur $\mathbb{K}$ ssi elle admet une base de vecteurs propres ssi il existe $P \in GL_n(\mathbb{K})$ et $D \in D_n(\mathbb{K})$ telles que $A = P^{-1}DP$.

Proposition 1.1 $A \in \mathfrak{M}_n(\mathbb{K})$ est diagonalisable sur $\mathbb{K}$ ssi χ_A est scindé sur $\mathbb{K}$ et A n'a aucune valeur propre défective.

1.2.2 Polynôme de matrices

Définition 1.6 Soit $A \in \mathfrak{M}_n(\mathbb{K})$. On note $\mathbb{K}[A]$ la sous-algèbre engendrée par A. $\mathbb{K}[A] = \mathrm{Vect}\,\{A^i, i \in \mathbb{N}\} = \{P(A), P \in \mathbb{K}[X]\}$.

Remarque 1.5 $I = \{P \in \mathbb{K}[X], P(A) = 0\}$ *est un idéal de* $\mathbb{K}[X]$. $\mathbb{K}[X]$ *est principal donc il existe* $\mu_A \in \mathbb{K}[X]$ *tel que* $I = \langle \mu_A \rangle$. *Le* μ_A *unitaire est appelé polynôme minimal de* A.

THÉORÈME 1.1 *Les racines de μ_A sont les valeurs propres de A.*

Démonstration.

$\subset$ $0 = \mu_A(A) = (A - \lambda I_n)Q(A)$. Or $Q(A) \neq 0$ (contredirait la minimalité de μ_A) donc $A - \lambda I_n \notin GL_n(\mathbb{K})$ donc $\lambda \in \mathrm{Sp}(A)$.

$\supset$ $A \sim \begin{pmatrix} \lambda & 0 \\ 0 & A' \end{pmatrix}$ donc $P(A) \sim \begin{pmatrix} P(\lambda) & 0 \\ 0 & P(A') \end{pmatrix}$. On a $\mu_A(A) = 0$ donc $\mu_A(\lambda) = 0$. ■

THÉORÈME 1.2 DE CAYLEY-HAMILTON *Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$, $\mu_A | \chi_A$ ie $\chi_A(A) = 0$.*

Démonstration.

- Si $A = 0$, $0|X$.
- Si $A \neq 0$, il existe $x \in \mathbb{K}^n \setminus \{0\}$ tel que $Ax \neq 0$.
  Posons $E_x = \mathrm{Vect}\left\{A^k x, k \in \mathbb{N}\right\} \subset \mathbb{K}^n$ et p_x l'entier maximal tel que $\mathcal{F} = \{A^k x, k \in [\![0, p_x - 1]\!]\}$ soit libre.
  $\mathcal{F}$ est une base de E_x car elle est libre et génératrice ($\{A^k x, k \in [\![0, p_x]\!]\}$ est liée + récurrence).
  On note $A^{p_x} x = \sum_{k=0}^{p_x - 1} a_k A^k x$ et $\Pi_x = X^{p_x} - \sum_{k=0}^{p_x - 1} a_k X^k$.
  E_x est stable par A donc A induit un endomorphisme A_x sur E_x.
  Sur E_x, $\Pi_x(A) = 0$ et Π_x est le polynôme minimal de A_x. En effet, la matrice de A_x dans $\mathcal{F}$ est $\begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & 1 & a_{p_x - 1} \end{pmatrix}$. On a alors $\chi_{A_x} = (-1)^{p_x} \Pi_x$.
  On complète $\mathcal{F}$ en une base de $\mathbb{C}^n$ de sorte que la matrice de A dans cette base soit :
  $$\begin{pmatrix} A_x & * \\ 0 & B_x \end{pmatrix}$$
  On a alors $\chi_A = \chi_{A_x} \chi_{B_x}$ donc $\chi_A(A) = 0$. ■

Remarque 1.6 Si A et B sont semblables, $\mu_A = \mu_B$. La réciproque est fausse.

Démonstration.

- Si $A = PBP^{-1}$, $\mu_B(A) = P\mu_B(B)P^{-1} = 0$ donc $\mu_A | \mu_B$. Par symétrie, $\mu_A = \mu_B$.

- $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ ne sont pas semblables ($\chi_A \neq \chi_B$) mais ont même polynôme minimal (à savoir $(X-1)(X+1)$). ■

THÉORÈME 1.3 DE DÉCOMPOSITION DES NOYAUX

- *Soit $A \in \mathfrak{M}_n(\mathbb{K})$ tel que χ_A soit scindé dans $\mathbb{K}$. On note alors $\chi_A = (-1)^n \prod_{i=1}^{s} (X-\lambda_i)^{k_i}$.*
  On a alors :
  $$\mathbb{K}^n = \bigoplus_{i=1}^{s} \operatorname{Ker}(A - \lambda_i \operatorname{Id})^{k_i} = \bigoplus_{i=1}^{s} E_{\lambda_i}$$
- *Plus généralement, si $P \in \mathbb{K}[X]$ vérifie $P = \prod_{i=1}^{p} P_i$ où les P_i sont premiers deux à deux, on a :*
  $$\operatorname{Ker}(P(A)) = \bigoplus_{i=1}^{p} \operatorname{Ker}(P_i(A))$$

Remarque 1.7 Le premier point est une conséquence du second à cause de Cayley-Hamilton.

Démonstration.

- Résultat préliminaire :
  On pose $Q_i = \dfrac{P}{P_i}$ pour tout i. Les Q_i sont premiers entre eux donc (Bezout) il existe $(V_1, \ldots, V_p)$ tels que $\sum_{i=1}^{p} V_i Q_i = 1$.
  On a donc $\sum_{i=1}^{p} V_i(A) Q_i(A) = I_n$ (1).
- Notons $F_i = \operatorname{Ker}(P_i(A))$ et $F = \operatorname{Ker}(P(A))$. Montrons que $\sum_{i=1}^{p} F_i = \bigoplus_{i=1}^{p} F_i$.
  Soit $(x_1, \ldots, x_p) \in \prod_{i=1}^{p} F_i$ tel que $\sum_{i=1}^{p} x_i = 0$.
  Pour tout $j \neq i$, $P_j | Q_i$ donc $Q_i(A)(x_j) = 0$ donc $\sum_{j=1}^{p} Q_i(A)(x_j) = Q_i(A)(x_i)$.
  Or, par hypothèse, $Q_i(A)\left(\sum_{j=1}^{p} x_j\right) = 0$ donc $Q_i(A)(x_i) = 0$.

On a de plus, d'après (1), $x_i = \sum_{j=1}^{p} V_j(A) \underbrace{Q_j(A)(x_i)}_{=0} = 0$.

Donc les F_i sont en somme directe.

- Montrons $F \subset \bigoplus_{i=1}^{p} F_i$. Soit $x \in F$.

  On pose, pour tout j, $x_j = (V_j(A)Q_j(A))(x)$.

  On a :

$$\begin{aligned} P_i(A)(x_i) &= (P_i(A)V_i(A)Q_i(A))(x) \\ &= (P(A)V_i(A))(x) \\ &= V_i(A)(P(A)(x)) \\ &= 0 \end{aligned}$$

  Donc pour tout i, $x_i \in F_i$ et $x = \sum_{i=1}^{p} x_i$. Donc $F = \bigoplus_{i=1}^{p} F_i$. ■

COROLLAIRE 1.1 *Soit $M \in \mathfrak{M}_n(\mathbb{K})$.*

M est diagonalisable ssi χ_M est scindé sans valeurs propres défectives ssi μ_M est scindé à racines simples.

Démonstration.

$1 \Rightarrow 2$ Si $M = PDP^{-1}$, $\chi_M = \chi_D$ et le calcul de χ_D assure qu'il est scindé sans valeurs propres défectives.

$2 \Rightarrow 3$ On a $\chi_M = \prod_{i=1}^{p} (\lambda_i - X)^{k_i}$ avec $k_i = \dim E_{\lambda_i}$.

Posons $Q = \prod_{i=1}^{p} (X - \lambda_i)$. $Q | \mu_M$ car les racines de μ_M sont les valeurs propres de M.

De plus, $\mathrm{Ker}(Q(M)) = \bigoplus_{i=1}^{p} E_{\lambda_i} = \mathbb{K}^n$. Donc $\mu_M = Q$ qui est scindé à racines simples.

$3 \Rightarrow 1$ Si $\mu_M = \prod_{i=1}^{p} (X - \lambda_i)$ avec λ_i distinctes deux à deux, le théorème de décomposition des noyaux assure que $\mathbb{K}^n = \bigoplus_{i=1}^{p} E_{\lambda_i}$.

M est donc diagonalisable. ■

1.2.3 Trigonalisation de matrices

THÉORÈME 1.4 *Soit $A \in \mathfrak{M}_n(\mathbb{K})$.*

A est trigonalisable ssi χ_A est scindé sur $\mathbb{K}$.

Démonstration.

$\Rightarrow$ On a $A = P \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_p \end{pmatrix} P^{-1}$ donc $\chi_A = (-1)^n \prod_{i=1}^{p} (X - \lambda_i)$.

$\Leftarrow$ Par récurrence sur n. Le théorème est clair pour $n = 1$.

Supposons qu'il soit vrai au rang $n-1$. Soit $\lambda \in \mathrm{Sp}(A)$ et x un vecteur propre associé.

Il existe $(e_2, \ldots, e_n)$ tels que $(x, e_2, \ldots, e_n)$ soit une base de $\mathbb{K}^n$.

Dans cette base, A s'écrit $\begin{pmatrix} \lambda & * \\ 0 & A' \end{pmatrix}$

L'hypothèse de récurrence assure que A' est trigonalisable ($\chi_{A'} = \frac{\chi_A}{X-\lambda}$) donc A l'est.

Le principe de récurrence démontre le théorème. ■

Remarque 1.8

- *Les éléments diagonaux de la matrice triangulaire sont les valeurs propres de la matrice de départ.*
- *Si les coefficients de A sont connus à une erreur près, on ne peut pas toujours maîtriser l'erreur sur T.*

THÉORÈME 1.5 SCHUR *Soit $A \in \mathfrak{M}_n(\mathbb{C})$.*

A est trigonalisable en base orthonormale (ie il existe $U \in \mathcal{U}_n$ et T triangulaire supérieure telles que $A = UTU^*$).

Démonstration. χ_A est scindé sur $\mathbb{C}$ donc A est trigonalisable. La preuve est identique à la précédente (on doit choisir $\|x\| = 1$ et $(x, e_2, \ldots, e_n)$ orthonormale). ■

THÉORÈME 1.6

- *Si $A \in S_n(\mathbb{R})$, il existe $P \in O_n$ et D diagonale telles que $A = PDP^{-1} = PD^tP$.*
- *Si $A \in \mathscr{H}_n(\mathbb{C})$, il existe $P \in \mathcal{U}_n$ et D diagonale telles que $A = UDU^*$.*

Démonstration. Par récurrence sur n. Clair pour $n = 1$.

Si le théorème est vrai au rang $n-1$, soit $A \in \mathscr{H}_n(\mathbb{C})$ et $x \in E_\lambda$ tel que $\|x\| = 1$.

On complète x en $(x, e_2, \ldots, e_n)$ orthonormale.

Dans cette base, A s'écrit $\begin{pmatrix} \lambda & * \\ 0 & A' \end{pmatrix}$. Or $A^* = A$ donc $* = 0$ et $A'^* = A'$.

L'hypothèse de récurrence conclut. ■

THÉORÈME 1.7 *Pour tout $A \in S_n^+(\mathbb{R})$, il existe une unique $\sqrt{A} \in S_n^+(\mathbb{R})$ tel que $A = \sqrt{A}^2$.*

Démonstration.

$\exists$ On a $A = P\,\mathrm{diag}(\lambda_1, \dots, \lambda_p)^t P$. $P\,\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})^t P$ convient.
$!$ A et $\sqrt{A}$ commutent donc sont co-diagonalisables.
On a $A = P\,\mathrm{diag}(\lambda_1, \dots, \lambda_p)^t P$ et $\sqrt{A} = P\,\mathrm{diag}(\lambda'_1, \dots, \lambda'_{p'})^t P$.
On a donc obligatoirement $p = p'$ et $\lambda_i = \lambda_i'^2$ pour tout i. ■

1.3 Réduction de Jordan

Le but est, étant donnée une matrice A, de trouver J semblable à A de la forme $\begin{pmatrix} \lambda_1 & \varepsilon & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \varepsilon \\ 0 & \cdots & \cdots & 0 & \lambda_p \end{pmatrix}$ où $\varepsilon \in \{0, 1\}$.

1.3.1 Matrices nilpotentes

THÉORÈME 1.8 *Soit N une matrice nilpotente.*

Il existe $s \geqslant 1$ et $(d_1, \dots, d_s)$ supérieurs ou égaux à 1 tels que $\sum_{i=1}^{s} d_i = n$ et N est semblable à J où J est diagonale par blocs dont les blocs $(J_i)_{i \in [\![1,s]\!]}$ sont de taille $d_i \times d_i$ et valent $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$.

Définition 1.7 Chaque J_i est appelé bloc de Jordan nilpotent de taille d_i.

Démonstration. Par récurrence sur n. Pour $n = 1$, $N = (0)$ donc le résultat est clair.

Supposons que celui-ci est vrai en dimension inférieure à $n - 1$. N est nilpotente d'ordre k dans $\mathfrak{M}_n(\mathbb{K})$.

Il existe donc x tel que $N^{k-1}x \neq 0$. Notons $\mathcal{B} = (x, Nx, \dots, N^{k-1}x)$ et $E_x = \mathrm{Vect}\,\{\mathcal{B}\}$.

- Montrons que $\dim(E_x) = k$ ie que $\mathcal{B}$ est libre.
  Soient $(\lambda_0, \dots, \lambda_{k-1})$ tels que $\sum_{i=0}^{k-1} \lambda_i N^i x = 0$.
  Supposons qu'il existe i tel que $\lambda_i \neq 0$. Il existe alors i_0 minimal tel que $\lambda_{i_0} \neq 0$.

On a $0 = N^{k-i_0-1}\sum_{j=0}^{k-1}\lambda_i N^i x = \lambda_{i_0} N^{k-1}x$. Donc $\lambda_{i_0} = 0$.
$\mathcal{B}$ est donc libre et $\dim(E_x) = k$.

- Montrons que E est stable par N.
  On a $N\mathcal{B} = (Nx, N^2x, \dots, N^{k-1}x, 0)$ dont chacun des termes appartient à E_x. Donc E_x est stable par N.
  Dans $\overline{B} = (N^{k-1}x, \dots, x)$, N s'écrit comme un bloc de Jordan de taille $n \times n$.
- Si $k = n$, le problème est résolu. Sinon, $k < n$. On cherche alors un supplémentaire de E_x stable par N.
  On a $N^{k-1}x \neq 0$ donc il existe y tel que $\langle N^{k-1}x | y\rangle \neq 0$.
  On pose $\mathcal{B}^* = (y, N^*y, \dots, (N^*)^{k-1}y)$ et $G = (\text{Vect}\,\{\mathcal{B}^*\})^\perp$.
  - On montre $\mathcal{B}^*$ libre. C'est le même principe que pour le liberté de $\mathcal{B}$ puisque $(N^*)^{k-1}y \neq 0$.
  - On montre que G est stable par N.
    Soit $j \in [\![0, k-1]\!]$ et $v \in G$.
    $\langle (N^*)^j y | Nv\rangle = \langle (N^*)^{j+1}y | v\rangle = 0$ car $(N^*)^k = 0$ et car $v \in G$.
  - On montre $E_x \oplus G = E$. Les dimensions permettent de se limiter à $E_x \cap G = \{0\}$.
    Si $v = \sum_{i=0}^{k-1} a_i N^i x \in E_x \cap G$, pour tout j, $\langle (N^*)^j y | v\rangle = 0$.
    Donc $0 = \sum_{i=0}^{k-1} \overline{a_i}\langle (N^*)^j y, N^i x\rangle = \sum_{i=0}^{k-1} \overline{a_i}\langle (N^*)^{i+j} y, x\rangle$.
    S'il existe i_0 minimal tel que $a_i \neq 0$, $j = k - i_0 - 1$ apporte une contradiction. Donc $v = 0$ et $E_x \oplus G = E$.
- On complète $\overline{\mathcal{B}}$ en une base adaptée à $E = E_x \oplus G$. On note Q la matrice de passage associée et on a :

$$N = Q \begin{pmatrix} J_1 & 0 \\ 0 & N' \end{pmatrix} Q^{-1}$$

  $\mu_{N'} | X^k$ donc N' est nilpotente. Le principe de récurrence permet de mettre N' sous la forme voulue : $N' = P'J'P'^{-1}$.
  On a donc N semblable à une matrice J de la forme recherchée via $P = \begin{pmatrix} Q & 0 \\ 0 & P' \end{pmatrix}$. ■

Exemple : Quelle est la réduite de Jordan de $M = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$?

$M^3 = 0$. Les formes possibles de la réduites sont : 0, $E_{1,2}$ (ou $E_{2,3}$) ou $E_{1,2} + E_{2,3}$.

- Si $a = b = c = 0$, la réduite est la matrice nulle. On suppose maintenant $(a, b, c) \neq (0, 0, 0)$.
- Si $M^2 = 0$, ie $ac = 0$, la réduite est $E_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.
- Sinon, la réduite est $E_{1,2} + E_{2,3} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Proposition 1.2 Soit M une matrice nilpotente et J sa réduite de Jordan. Le nombre de blocs de Jordan de J est égal à $\dim(\mathrm{Ker}(M))$.

Démonstration. On a :

$$\dim(\mathrm{Ker}(M)) = \dim(\mathrm{Ker}(J)) = \sum_{i=1}^{s} \dim(\mathrm{Ker}(J_i)) = \sum_{i=1}^{s} 1 = s$$

■

Proposition 1.3 Le nombre de blocs de taille supérieure à k est la différence $\dim(\mathrm{Ker}(M^k)) - \dim(\mathrm{Ker}(M^{k-1}))$.

Exemple : $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ne sont pas semblables.

1.3.2 Cas général : matrices complexes

THÉORÈME 1.9 *Soit $A \in \mathfrak{M}_n(\mathbb{C})$. A est semblable sur $\mathbb{C}$ à J matrice composée de blocs $J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix}$ où les λ_i ne sont pas nécéssairement distinctes.*

Démonstration. $\chi_A = \prod_{i=1}^{p} (X - \lambda_i)^{\alpha_i}$ donc, d'après le lemme des noyaux, $\mathbb{C}^n = \bigoplus_{i=1}^{p} \mathrm{Ker}(E_{\lambda_i})$.

Les E_{λ_i} sont stables par A, donc, dans une base de E_{λ_i}, $A|_{E_{\lambda_i}}$ s'écrit $\lambda_i I_i + N_i$ avec N_i nilpotente.

Le théorème précédent s'applique à N_i et fournit le résultat en juxtaposant les blocs. ■

Remarque 1.9

- *On retrouve la décomposition de Dunford $A = D + N$ avec $DN = ND$.*
- *Le nombre de blocs associés à λ_i est* $\dim(E_{\lambda_i})$*. Donc, si λ est non défective, le bloc est diagonal.*

Exemple : $M = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -1 \end{pmatrix}$.

Deux réduites sont possibles : $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Si $a \neq 0$, c'est la première, sinon, c'est la seconde.

1.3.3 Réduction de Jordan sur $\mathbb{R}$

Si χ_A est scindé, on fait comme dans $\mathbb{C}$.

Sinon, il existe $\lambda_k = a_k + ib_k$ avec $b_k \neq 0$. On pose $\Lambda_k = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$.

THÉORÈME 1.10 *Soit $A \in \mathfrak{M}_n(\mathbb{R})$.*

$$A \text{ est semblable à } J = \begin{pmatrix} J_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & J_s & \ddots & & \vdots \\ \vdots & & \ddots & K_1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & K_r \end{pmatrix}$$

$$\text{avec } J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix} \text{ et } K_i = \begin{pmatrix} \Lambda_i & I_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & I_2 \\ 0 & \cdots & \cdots & 0 & \Lambda_i \end{pmatrix}.$$

Démonstration. Sur $\mathbb{C}$, A et J sont semblables car elles ont même décomposition de Jordan.

En effet, les K_i sont semblables à des

$$J_{s+k} = \begin{pmatrix} (a_k + ib_k)\,\mathrm{Id} & 0 \\ 0 & (a_k - ib_k)\,\mathrm{Id} \end{pmatrix}$$

Donc $A = PJP^{-1}$ avec $P = P_1 + iP_2$. L'application $\det(P_1 + tP_2)$ est polynômiale non nulle (en i) donc elle n'est pas nulle sur $\mathbb{R}$.

Donc il existe $\lambda \in \mathbb{R}$, $Q = P_1 + \lambda P_2 \in GL_n(\mathbb{R})$. On a alors $A = QJQ^{-1}$. ■

1.3.4 Applications : suites récurrentes linéaires

Problème : On donne $(u_0, \dots, u_{k-1}) \in \mathbb{C}^k$ et $(a_0, \dots, a_{k-1}) \in \mathbb{C}^k$.

Le but est de d'expliciter la suite u telle que, pour tout $n \in \mathbb{N}$, $u_{n+k} = \sum_{i=0}^{k-1} a_i u_{n-i}$.

On pose $U_n = \begin{pmatrix} u_n \\ \vdots \\ u_{n+k-1} \end{pmatrix}$. U_0 est donné. La relation de récurrence se transforme en $U_{n+1} = AU_n$ avec A la transposée de la matrice compagnon de $\sum_{i=0}^{k-1} a_i X^i$.

On a alors $U_n = A^n U_0$. Il faut donc calculer A^n.

Si $A = PJP^{-1}$ avec J composée de blocs de $J_i = \lambda_i I_{\alpha_i} + N_i$, $J_i^k = \sum_{j=0}^{d_i-1} \binom{k}{j} \lambda_i^{k-j} N_i^j$.

Donc $J_i^k = \begin{pmatrix} \lambda_i^k & k\lambda_i^{k-1} & \cdots & \cdots & \binom{k}{d_i-1}\lambda_i^{k-d_i+1} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & k\lambda_i^{k-1} \\ 0 & \cdots & \cdots & 0 & \lambda_i^k \end{pmatrix}$.

Il faut alors recomposer J^k (matrice des J_i^k) puis $A^k = PJ^kP^{-1}$.

Chapitre 2

Topologie matricielle

$\mathfrak{M}_n(\mathbb{K})$ est de dimension finie donc les normes sont équivalentes donc on étudie une seule topologie d'espace vectoriel normé.

2.1 Norme subordonnée induite

2.1.1 Définition

Définition 2.1 Si $\|\cdot\|$ est une norme sur $\mathbb{K}^n$, on appelle norme induite par $\|\cdot\|$ l'application :

$$|||\cdot||| : \begin{cases} \mathfrak{M}_n(\mathbb{K}) & \to & \mathbb{R}^+ \\ A & \mapsto & \sup\limits_{x\neq 0} \dfrac{\|Ax\|}{\|x\|} \end{cases}$$

Exemple : $|||A|||_\infty = \max_i \sum_{j=1}^{n} |a_{i,j}|$ et $|||A|||_1 = \max_j \sum_{i=1}^{n} |a_{i,j}|$.

Soit $x \in \mathbb{K}^n$ de norme 1.

$\|Ax\|_\infty = \max_i \left| \sum_{j=1}^{n} a_{i,j} x_j \right| \leqslant \max_i \sum_{j=1}^{n} |a_{i,j}|$.

La borne est atteinte pour $x_j = \dfrac{\overline{a_{i_0,j}}}{|a_{i_0,j}|}$ si $a_{i_0,j} \neq 0$ et 1 sinon. (i_0 est tel que $\sum_{j=1}^{n} |a_{i_0,j}| = \max\limits_i \sum_{j=1}^{n} |a_{i,j}|$).

2.1.2 Propriétés

Proposition 2.1 $|||\cdot|||$ est une norme d'algèbre.

Démonstration. Si $\|x\| = 1$, $\|ABx\| \leqslant |||A||| \cdot \|Bx\| \leqslant |||A||| \cdot |||B||| \cdot \|x\| = |||A||| \cdot |||B|||$.

Donc $|||AB||| \leqslant |||A||| \cdot |||B|||$. ∎

Remarque 2.1

- *Il existe des normes sur $\mathfrak{M}_n(\mathbb{K})$ qui ne sont pas d'algèbre (par exemple la norme infinie des coefficients avec $A = B = J_n$)*
- *La réciproque de la propriété est fausse : $\|A\| = \sqrt{\mathrm{tr}(A^*A)}$ n'est pas une mesure induite ($\|I_n\| = \sqrt{n} \neq 1$) mais elle est d'algèbre.*

2.1.3 Cas de la norme euclidienne

Définition 2.2 Soit $A \in \mathfrak{M}_n(\mathbb{C})$. Le rayon spectral de A, noté $\rho(A)$ est le maximum des modules des valeurs propres de A.

Proposition 2.2 Si A est diagonalisable en base orthonormée, $|||A|||_2 = \rho(A)$.

Démonstration. $A = UDU^*$. Soit $x \in \mathbb{K}^n$ tel que $\|x\|_2 = 1$.

$\|Ax\|_2 = \|UDU^*x\|_2 = \|DU^*x\|_2$.

Donc $\max\limits_{\|x\|_2=1} \|Ax\|_2 = \max\limits_{\|x\|_2=1} \|Dx\|_2 = \rho(D) = \rho(A)$ car U et U^* sont des isométries. ∎

Théorème 2.1 *Pour tout $A \in \mathfrak{M}_n(\mathbb{C})$, $|||A|||_2 = \sqrt{\rho(A^*A)}$.*

Démonstration. Montrons que $|||A||| = |||A^*|||$.

$\|Ax\|_2^2 = \langle Ax|Ax\rangle \leqslant \|A^*Ax\|_2 \|x\|_2 \leqslant |||A^*A|||_2$.

Donc $|||A|||_2^2 \leqslant |||A^*A|||_2 \leqslant |||A^*|||_2 |||A|||_2$. (1)

Donc, si $A \neq 0$, $|||A|||_2 \leqslant |||A^*|||_2$.

De même, $|||A|||_2 = |||A^*|||_2$. D'après (1), $|||A|||_2^2 = |||A^*A|||_2$.

La propriété précédente conclut car A^*A est hermitienne. ∎

Définition 2.3 Les racines carrées des valeurs propres de A^*A sont appelées valeurs singulières de A.

Remarque 2.2 $\mathrm{Sp}(A^*A) \subset \mathbb{R}^+$ *car* $A^*Ax = \lambda x \Rightarrow \|Ax\|^2 = \lambda \|x\|^2 \Rightarrow x \in \mathbb{R}^+$.

Théorème 2.2

- *Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$ et pour tout norme subordonnée $|||\cdot|||$, $|||A||| \geqslant \rho(A)$.*
- *Soit $A \in \mathfrak{M}_n(\mathbb{C})$ et $\varepsilon > 0$.*
  Il existe une norme subordonnée $|||\cdot|||_{A,\varepsilon}$ telle que $|||A|||_{A,\varepsilon} \leqslant \rho(A) + \varepsilon$.
- *$\rho(A) = \inf\{|||A|||, |||\cdot||| \text{ subordonnée}\}$.*

Démonstration.

- Soit x un vecteur propre de norme 1 associé à λ telle que $|\lambda| = \rho(A)$.
  $|||A||| = |||A||| \cdot \|x\| \geqslant \|Ax\| = |\lambda| \, \|x\| = \rho(A)$.
- D'après le théorème de Schur, $A = UTU^*$.
  On pose $D_\eta = \mathrm{diag}(\eta, \ldots, \eta^n)$ et $T_\eta = D_\eta T D_\eta^{-1}$ $(\eta > 0)$.
  $|||T_\eta|||_1 \underset{\eta \to 0}{\to} \rho(A)$ donc pour η suffisament petit, $|||T_\eta|||_1 \leqslant \rho(A) + \varepsilon$.
  On définit donc $\|x\|_{A,\varepsilon} = \left\| D_\eta^{-1} U^* x \right\|_1$.

$$\begin{aligned} \|Ax\|_{A,\varepsilon} &= \left\| D_\eta^{-1} U^* U T U^* x \right\|_1 = \left\| D_\eta^{-1} T U^* x \right\|_1 \\ &= \left\| T_\eta D_\eta^{-1} U^* x \right\| \leqslant |||T_n|||_1 \, \|x\|_{A,\varepsilon} \end{aligned}$$

  Donc $|||A|||_{A,\varepsilon} \leqslant |||T_\eta|||_1 \leqslant \rho(A) + \varepsilon$.
- Pour tout $\varepsilon > 0$, $\rho(A) \leqslant |||A|||_{A,\varepsilon} \leqslant \rho(A) + \varepsilon$ donc $\rho(A)$ est bien la borne inférieure des normes subordonnées de A. ∎

COROLLAIRE 2.1 *Pour tout $A \in \mathfrak{M}_n(\mathbb{C})$, $\lim\limits_{k \to +\infty} A^k = 0 \Rightarrow \rho(A) < 1$.*

Démonstration.

$\Rightarrow$ Soit $\lambda \in \mathrm{Sp}(A)$ tel que $|\lambda| = \rho(A)$ et x un vecteur propre associé à λ.
$A^k x = \lambda^k x$ donc $\lim\limits_{k \to +\infty} \lambda^k = 0$ donc $|\lambda| < 1$.

$\Leftarrow$ Soit A telle que $\rho(A) < 1$. Il existe une norme subordonnée $|||\cdot|||$ telle que $|||A||| < 1$.
$|||A^k||| \leqslant |||A|||^k$ donc converge vers 0. ∎

COROLLAIRE 2.2 *Pour toute norme subordonnée $|||\cdot|||$,*

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

Démonstration. Si $\lambda \in \mathrm{Sp}(A)$ vérifie $|\lambda| = \rho(A)$, $\lambda^k \in \mathrm{Sp}(A^k)$ donc $\rho(A^k) \geqslant |\lambda|^k = \rho(A)^k$.

Soit $|||\cdot|||$ une norme subordonnée. On a $\rho(A^k) < |||A^k|||$.

Donc $\rho(A) < |||A^k|||^{\frac{1}{k}}$

De plus, pour tout $\varepsilon > 0$, $A_\varepsilon = \dfrac{1}{\rho(A) + \varepsilon} A$ vérifie $\rho(A_\varepsilon) = \dfrac{\rho(A)}{\rho(A) + \varepsilon} < 1$.

Donc $\lim\limits_{k \to +\infty} A_\varepsilon^k = 0$.

Il existe donc k_ε tel que pour tout $k \geqslant k_\varepsilon$, $|||A_\varepsilon^k||| < 1$.

On a donc $|||A^k||| < (\rho(A) + \varepsilon)^k$ donc $\rho(A) \leqslant |||A^k|||^{\frac{1}{k}} < \rho(A) + \varepsilon$.

Donc $\lim\limits_{k \to +\infty} |||A^k|||^{\frac{1}{k}} = \rho(A)$. ∎

2.1.4 Application aux systèmes différentiels linéaires

Cas $n = 1 : y' = ay$, $y(0) = y_0$ a pour solution $y : t \mapsto y_0 \mathrm{e}^{at}$. $|y|$ est croissante si $\Re(a) > 0$, décroissante si $\Re(a) > 0$ et constante sinon.

Cas $n \geqslant 2 : Y' = AY, Y(0) = Y_0$ a pour solution $Y : t \mapsto \mathrm{e}^{tA} Y_0$. Comment se comporte Y en l'infini ?

Proposition 2.3

- $\lim\limits_{t \to +\infty} \mathrm{e}^{tA} = 0 \quad \text{ssi} \quad \forall \lambda \in \mathrm{Sp}(A), \Re(\lambda) < 0.$
- $\{\mathrm{e}^{tA}, t \geqslant 0\}$ est borné ssi $\forall \lambda \in \mathrm{Sp}(A), \Re(\lambda) \leqslant 0$ et $\Re(\lambda) = 0 \Rightarrow \lambda$ non défective.

Démonstration. On va montrer le résultat pour les blocs de Jordan $J_i = \lambda_i I_{d_i} + N_i$.

$$\mathrm{e}^{tJ_i} = \mathrm{e}^{\lambda_i I_{d_i}} \mathrm{e}^{N_i} = \mathrm{e}^{t\lambda_i} \sum_{k=0}^{d_i - 1} t^k \frac{N_i^k}{k!}$$

Si $\Re(\lambda_i) < 0$, $\left\| \mathrm{e}^{tJ_i} \right\|$ tend vers 0 car l'exponentielle l'emporte sur le polynôme.

Si $\Re(\lambda_i) > 0$, $\left\| \mathrm{e}^{tJ_i} \right\|$ tend vers $+\infty$.

Si $\Re(\lambda_i) = 0$, $t \mapsto \mathrm{e}^{tJ_i}$ est bornée ssi $d_i = 1$.

En recollant les blocs de Jordan, on a :

- S'il existe $\lambda \in \mathrm{Sp}(A)$ tel que $\Re(\lambda) > 0$, $\left\| \mathrm{e}^{tA} \right\|$ tend vers $+\infty$.
- Sinon, $t \mapsto \mathrm{e}^{tA}$ est bornée si tous les blocs de Jordan associés à une valeur propre imaginaire pure sont de taille 1, ie ces valeurs propres sont non défectives.
- De plus $\left\| \mathrm{e}^{tA} \right\|$ tend vers 0 ssi tous les blocs convergent vers 0 ssi $\forall \lambda \in \mathrm{Sp}(A)$, $\Re(\lambda) < 0$. ∎

2.2 Conditionnement d'une matrice

2.2.1 Exemple classique

On considère le système $AX = B$ avec :

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}, \quad B = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}, \quad X = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

A est inversible car $\det(A) = 1$ et X est bien solution du système.

Effet d'une perturbation du second membre

On perturbe maintenant légèrement (mais pas n'importe comment) le vecteur B, et on calcule la nouvelle solution. On a par exemple $A\widetilde{X} = \widetilde{B}$ avec :

$$\widetilde{B} = \begin{pmatrix} 32,1 \\ 22,9 \\ 33,1 \\ 30,9 \end{pmatrix}, \quad \widetilde{X} = \begin{pmatrix} 9,2 \\ -12,6 \\ 4,5 \\ -1,1 \end{pmatrix}$$

Le résultat surprend : le vecteur $\widetilde{B}$ étant proche de B on s'attend à trouver une solution $\widetilde{X}$ proche de X, ce qui ne semble pas être le cas ! De manière plus précise, si on calcule les erreurs relatives :

$$\frac{\left\|\widetilde{B} - B\right\|_\infty}{\|B\|_\infty} = \frac{0,1}{33} \approx 3.10^{-3}$$

et

$$\frac{\left\|\widetilde{X} - X\right\|_\infty}{\|X\|_\infty} = \frac{13,6}{1} = 13,6$$

on remarque que, par la résoltuion du système linéaire, l'erreur relative sur B est multipliée par 4488, ce qui est assez effrayant !

Cette situation n'est pas seulement liée au choix de la norme, on obtient avec la norme euclidienne des taux d'amplification moindres, mais très élevés aussi :

$$\frac{\left\|\widetilde{B} - B\right\|_2}{\|B\|_2} \approx \frac{0,2}{60} \approx 3,3.10^{-3}$$

et

$$\frac{\left\|\widetilde{X} - X\right\|_2}{\|X\|_2} \approx \frac{16,4}{2} = 8.2$$

soit un facteur d'amplification de l'erreur égal à 2460 environ.

Effet d'une perturbation de la matrice

On perturbe maintenant la matrice A, on considère donc le système $\widehat{A}\widehat{X} = B$ avec :

$$\widehat{A} = \begin{pmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,98 & 9,89 & 9 \\ 6,99 & 4,99 & 9 & 9,98 \end{pmatrix}, \quad \widehat{X} = \begin{pmatrix} -8,1 \\ 137 \\ -34 \\ 22 \end{pmatrix}$$

Encore une fois, le calcul des erreurs relatives :

$$\frac{\left\|\widehat{A}-A\right\|_\infty}{\|A\|_\infty}=\frac{0,3}{33}\approx 9.10^{-3}$$

et

$$\frac{\left\|\widehat{X}-X\right\|_\infty}{\|X\|_\infty}=136$$

révèle une forte amplification de l'ordre de 15000 !

2.2.2 Explication

Définition 2.4 Pour tout $A\in GL_n(\mathbb{K})$, et pour toute $|||\cdot|||$ norme induite, on appelle conditionnement de A et on note $c(A)$ le scalaire $|||A|||A^{-1}|||$.

Remarque 2.3 *$c(A)$ dépend de la norme et $c(A)\geqslant 1$ (car $AA^{-1}=I_n$).*

THÉORÈME 2.3 *Pour une norme vectorielle $\|\cdot\|$ donnée, et la norme matricielle $|||\cdot|||$ induite, on a, avec $A\in GL_n(\mathbb{K})$ et $X,B\in\mathbb{K}^n$ tels que $AX=B$,*

- *Si $\delta B\in\mathbb{K}^n$ et $\delta X\in\mathbb{K}^n$ tel que $A(X+\delta X)=B+\delta B$,*

$$\frac{\|\delta X\|}{\|X\|}\leqslant c(A)\frac{\|\delta B\|}{\|B\|}$$

- *Si $\delta A\in\mathfrak{M}_n(\mathbb{K})$ et $\delta X\in\mathbb{K}^n$ vérifient $(A+\delta A)(X+\delta X)=b$,*

$$\frac{\|\delta X\|}{\|X+\delta X\|}\leqslant c(A)\frac{|||\delta A|||}{|||A|||}$$

Démonstration.

- $\delta X=A^{-1}\delta B$ donc $\|\delta X\|\leqslant |||A^{-1}|||\,\|\delta B\|$.
  De plus, $AX=B$ donc $\|B\|\leqslant |||A|||\,\|X\|$.
  Donc $\frac{\|\delta X\|}{\|X\|}\leqslant c(A)\frac{\|\delta B\|}{\|B\|}$.
- $(A+\delta A)(X+\delta X)=B$ donc $A\delta X+\delta A(X+\delta X)=0$.
  Donc $\delta X=-A^{-1}\delta A(X+\delta X)$.
  On a donc $\|\delta X\|\leqslant |||A^{-1}|||\,|||\delta A|||\,\|X+\delta X\|$.
  Donc $\frac{\|\delta X\|}{\|X+\delta X\|}\leqslant c(A)\frac{|||\delta A|||}{|||A|||}$. ∎

Remarque 2.4 *$c(A)$ est donc le facteur minimal d'amplification des erreurs relatives. Dans notre exemple, $c_\infty(A)=4455$ et $c_2(A)=3030$.*

Proposition 2.4

- Pour tout $A\in GL_n(\mathbb{K})$, $c_2(A)=\sqrt{\dfrac{\max\{|\lambda|,\lambda\in\mathrm{Sp}(A^*A)\}}{\min\{|\lambda|,\lambda\in\mathrm{Sp}(A^*A)\}}}$

- Si A est ortho-diagonalisable, $c_2(A) = \dfrac{\rho(A)}{\min \mathrm{Sp}(A)}$.

Démonstration. C'est une conséquence des formules sur $\|\cdot\|_2$. ∎

Remarque 2.5 *Pour $A \in \mathcal{U}_n(\mathbb{C})$, $c_2(A) = 1$.*

2.3 Topologie dans $\mathfrak{M}_n(\mathbb{K})$

2.3.1 Groupe linéaire

Proposition 2.5 $GL_n(\mathbb{K})$ est un ouvert dense.

Démonstration. $GL_n(\mathbb{K})$ est l'image réciproque d'un ouvert ($\mathbb{K}^*$) par une fonction continue (det).

Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$, on pose $A_t = A + tI_n$. Pour $t \in]0, r[\setminus \mathrm{Sp}(A)$, A_t est inversible et A_t tend vers A. ∎

Remarque 2.6 *Si $A \in GL_n(\mathbb{K})$, $B(A, \frac{1}{\|A^{-1}\|}) \subset GL_n(\mathbb{K})$.*

En effet, si $\|H\| < \frac{1}{\|A^{-1}\|}$, $A + H = A(I_n + A^{-1}H)$ et $\|A^{-1}H\| \leqslant \|H\|\|A^{-1}\| < 1$ donc $I_n + A^{-1}H \in GL_n(\mathbb{K})$.

On a de plus $(A + H)^{-1} = \sum_{n=0}^{\infty} (-A^{-1}H)^n A^{-1}$ (qui converge dès que $\|A^{-1}H\| < 1$).

Proposition 2.6 $GL_n(\mathbb{R})$ n'est pas connexe mais $GL_n(\mathbb{C})$ l'est (par arcs).

Démonstration.

- $GL_n(\mathbb{R}) = \{A, \det(A) > 0\} \cup \{A, \det(A) < 0\}$ qui sont deux ouverts disjoints non vides.
- $\mathrm{Sp}(A)$ est fini donc il existe $z_0 \in \mathbb{C}^*$ tel que $\mathbb{R}^- z_0 \cap \mathrm{Sp}(A) = \varnothing$.
  On pose :

$$\psi : \begin{cases}]0,1] & \to & \mathbb{R}^- z_0 \\ t & \mapsto & z_0 \dfrac{t-1}{t} \end{cases} \text{ et } \varphi : \begin{cases}]0,1] & \to & \mathbb{C} \\ t & \mapsto & \dfrac{1}{1-\psi(t)} \end{cases}$$

  On prolonge φ par continuité avec $\varphi(0) = 0$.
  φ est continue, $\varphi(0) = 0$ et $\varphi(1) = 1$. $A_t = \varphi(t)A + (1 - \varphi(t))I_n \in GL_n(\mathbb{C})$ car $\frac{\varphi(t)-1}{\varphi(t)} = \psi(t) \notin \mathrm{Sp}(A)$.
  Donc $GL_n(\mathbb{C})$ est connexe par arcs donc connexe. ∎

Application : Il n'existe aucune norme stable par changement de base. En effet, si $\|\cdot\|$ en était une, on aurait $\|PAP^{-1}\| = \|A\|$ pour tout A, P.

Donc $\|AP\| = \|P(AP)P^{-1}\| = \|PA\|$. Or la norme est continue donc ce résultat est vrai pour tout $P \in \mathfrak{M}_n(\mathbb{R})$.

Cependant, les matrices $A = E_{1,2}$ et $B = E_{2,3}$ sont telles que $\|AB\| \neq 0$ et $\|BA\| = 0$.

2.3.2 Groupe orthogonal, groupe unitaire

THÉORÈME 2.4 *$O_n(\mathbb{R})$ et $\mathcal{U}_n(\mathbb{C})$ sont compacts. Le premier est non connexe et possède deux composantes connexes. Le second est connexe par arcs.*

Démonstration.

- Ils sont fermés car images réciproques de $\{I_n\}$ par $A \mapsto A^*A$ et $A \mapsto A^tA$ qui sont continues.
  Ils sont bornés car inclus dans $S_{\|\cdot\|_2}(0,1)$.
- $O_n = SO_n \cup O_n^-$ qui sont images réciproques par det des ouverts $\{1\}$ et $\{-1\}$ donc ils sont ouverts, non vides et disjoints.
- Si $A = PDP^*$ avec $D = \operatorname{diag}(\mathrm{e}^{i\theta_1}, \dots, \mathrm{e}^{i\theta_n})$, on a un arc continu inclus dans $\mathcal{U}_n$ donné par $t \mapsto P \operatorname{diag}(\mathrm{e}^{it\theta_1}, \dots, \mathrm{e}^{it\theta_n})P^*$ qui joint A à I_n.
- Montrons que SO_n est connexe par arcs en reliant tout le monde à I_n.
  Soit $A \in SO_n$. $A = P^{-1}BP$ avec $B = \begin{pmatrix} I_p & & & & \\ & -I_{2q} & & 0 & \\ & & R_{\theta_1} & & \\ & 0 & & \ddots & \\ & & & & R_{\theta_r} \end{pmatrix}$.
  On pose $B_t = \begin{pmatrix} I_p & & & & & & \\ & R_{t\pi} & & & 0 & & \\ & & \ddots & & & & \\ & & & R_{t\pi} & & & \\ & & & & R_{\theta_1} & & \\ & 0 & & & & \ddots & \\ & & & & & & R_{\theta_r} \end{pmatrix}$ et $A_t = P^{-1}B_tP$.
  A_t est un arc continu joignant A et I_n.
- De même O_n^- est connexe par arcs. ∎

THÉORÈME 2.5 *Pour tout $A \in GL_n(\mathbb{R})$, il existe un unique $(\Omega, S) \in O_n \times S_n^{++}$ tel que $A = \Omega S$.*

De plus $A \in GL_n^+(\mathbb{R})$ ssi $\Omega \in SO_n(\mathbb{R})$.

Démonstration.

! Si $A = \Omega S$; ${}^tAA = S^2$ donc $S = ({}^tAA)^{\frac{1}{2}}$. Par unicité de la racine carrée (${}^tAA \in S_n^{++}$), S est unique donc Ω aussi.

$\exists$ $S = ({}^tAA)^{\frac{1}{2}} \in S_n^{++}$ par construction. Il faut vérifier que $\Omega = AS^{-1}$ est orthogonale.
${}^t\Omega\Omega = {}^t S^{-1t}AAS^{-1} = S^{-1}S^2S^{-1} = I_n$.

– $\det(S) > 0$ donc $\det(A) > 0$ ssi $\det(\Omega) > 0$ ssi $\det(\Omega) = 1$. ■

COROLLAIRE 2.3 *$GL_n(\mathbb{R})$ a deux composantes connexes.*

2.3.3 Matrices diagonalisables, trigonalisables

On pose Δ_n l'ensemble des matrices diagonalisables, Θ_n l'ensemble des matrices trigonalisables et $\Delta'_n = \{M \in \Delta_n, \text{Card}(\text{Sp}(M)) = n\}$.

Remarque 2.7 *Δ_n n'est ni ouvert ni fermé :*

$$\Delta_n \ni \begin{pmatrix} 0 & 1 \\ 0 & \frac{1}{k} \end{pmatrix} \to \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \notin \Delta_n$$

$$\Delta_n \not\ni \begin{pmatrix} 1 & \frac{1}{k} \\ 0 & 1 \end{pmatrix} \to I_2 \in \Delta_n$$

THÉORÈME 2.6 $\overline{\Delta_n(\mathbb{C})} = \mathfrak{M}_n(\mathbb{C}) = \overline{\Delta'_n(\mathbb{C})}$.

Démonstration. Si $A \in \mathfrak{M}_n(\mathbb{C})$, $A = PTP^{-1}$ avec $T = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_p \end{pmatrix}$.

Si les λ_i sont distincts, $A \in \Delta'_n(\mathbb{C})$.

Sinon, on pose $\alpha = \min\{|\lambda_i - \lambda_j|, \lambda_k \neq \lambda_j\}$ si cet ensemble est différent de $\{0\}$ et 1 sinon.

On pose $\Delta_k = \text{diag}(\frac{\alpha}{k}, \ldots, \frac{\alpha}{nk})$, $T_k = T + \Delta_k$ et $A_k = PT_kP^{-1}$.

A_k tend vers A et $\text{Sp}(A_k) = (\lambda_i + \frac{\alpha}{ik})_{1 \leqslant i \leqslant n}$

Or :

$$\lambda_i + \frac{\alpha}{ik} = \lambda_j + \frac{\alpha}{jk} \Rightarrow |\lambda_i - \lambda_j| = \frac{\alpha}{k}\left|\frac{1}{j} - \frac{1}{i}\right| < \alpha \Rightarrow \lambda_i = \lambda_j \Rightarrow i = j$$

Donc $A_k \in \Delta'_n(\mathbb{C})$ donc $\overline{\Delta'_n(\mathbb{C})} = \mathfrak{M}_n(\mathbb{C})$.

Donc $\overline{\Delta_n(\mathbb{C})} = \mathfrak{M}_n(\mathbb{C})$. ■

Remarque 2.8 *Ceci est faux sur $\mathbb{R}$: avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\chi_A = X^2 + 1$ de discriminant -4.*

Si $A = \lim_{n \to +\infty} A_n$, avec $A_n \in \Delta_n(\mathbb{R})$, -4 serait la limite d'une suite strictement positive.

THÉORÈME 2.7 $\overline{\Delta_n(\mathbb{R})} = \Theta_n(\mathbb{R})$.

Démonstration.

$\supset$ Il suffit de réécrire la preuve précédente.

$\subset$ Montrons $\Theta_n(\mathbb{R})$ fermé.

Soit $(A_k)_k \in \Theta_n(\mathbb{R})^{\mathbb{N}}$ qui converge vers A.

χ_{A_k} est scindé donc $\chi_{A_k} = \prod_{i=1}^{n}(X - \lambda_i^{(k)})$ avec $\lambda_1^{(k)} < \cdots < \lambda_n^{(k)}$.

$\rho(A_k) \leqslant |||A_k|||$ qui est bornée car $(|||A_k|||)_k$ converge donc il existe φ extractrice telle que $(\lambda_i^{(\varphi(k))})_i$ tende vers $(\lambda_i)_i$.

Donc $\chi_{A_{\varphi(k)}}$ tend vers χ_A donc χ_A est scindé et $A \in \Theta_n(\mathbb{R})$. ∎

Application : On peut prouver le théorème de Cayley-Hamilton dans $\mathbb{C}$ via ce résutat : ce théorème se vérifie facilement sur les matrices diagonalisables et $\chi_\cdot(\cdot)$ est continue car polynômiale.

THÉORÈME 2.8 *Pour tout $A \in \mathfrak{M}_n(\mathbb{C})$, il existe $(U, H) \in \mathcal{U}_n \times \mathscr{H}_n$ avec H positive tel que $A = UH$.*

Si de plus A est inversible, H est définie positive et la décomposition est unique.

Démonstration. Si A est inversible, on fait la même preuve que dans $\mathbb{R}$.

Sinon, on prend une suite de matrices inversibles $(A_k)_k$ tendant vers A. $A_k = U_k H_k$.

Or $\mathcal{U}_n$ est compact donc il existe φ extractrice telle que $U_{\varphi(k)} \to U \in \mathcal{U}_n$.

$\mathscr{H}_n^+$ est fermé donc $H_{\varphi(k)}$ converge vers H donc $A = UH$. ∎

Chapitre 3

Décompositions usuelles

Motivations

- Stockage de données : les images numériques sont des matrices à cœfficients dans $[0,1]$. Besoin d'optimiser la taille mémoire et le temps d'exécution.
- Discrétisation d'équations différentielles.
  Exemple : le problème $u-u''=f$, $u(0)=u(1)=0$ peut être discrétisé : $[0,1]=\bigcup\limits_{0\leqslant i\leqslant n-1}[\frac{i}{n},\frac{i+1}{n}]$.

$$u'(t_i)=\frac{u(t_{i+1})-u(t_i)}{h},\qquad u''(t_i)=\frac{u(t_(i+1))-2u(t_i)+u(t_{i-1})}{h^2}$$

  On peut donc obtenir un problème approché : $AU=F$ avec $U=(u(t_i))_i$, $F=(f(t_i))_i$ et :

$$A=\frac{1}{h^2}\begin{pmatrix}2+h^2 & -1 & & 0\\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1\\ 0 & & -1 & 2+h^2\end{pmatrix}$$

  avec $n\approx 10^5$.
- Problèmes aux valeurs propres : $u-u''=\lambda u$ d'inconnues u et λ se ramène à chercher un spectre.
- Les matrices favorables à la résoltuion des ces problèmes sont : triangulaires, diagonales, unitaires, hermitiennes,...

3.1 Décomposition polaire

THÉORÈME 3.1
- $\forall A \in GL_n, \exists! U, H \in \mathcal{U}_n \times \mathscr{H}_n^{++}$, $A = UH$.
- $\forall A \in \mathfrak{M}_n, \exists U, H \in \mathcal{U}_n \times \mathscr{H}_n$, $A = UH$ *et* H *positive.*
- $GL_n(\mathbb{C})$ *et* $\mathcal{U}_n \times \mathscr{H}_n^{++}$ *sont homéomorphes.*

Démonstration. Voir chapitre II ∎

3.2 Décomposition LU

3.2.1 Mineurs fondamentaux

Définition 3.1 Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$, on appelle mineurs fondamentaux de A les $(\det((a_{i,j})_{i,j\leqslant k}))_k$.

THÉORÈME 3.2 *Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$ dont aucun mineur fondamental n'est nul, il existe un unique couple $(L, U) \in T_n^- \times T_n^+$ tel que L n'a que des 1 sur sa diagonale, U n'a que des coefficients non nuls sur la diagonale et $A = LU$.*

Démonstration.

∃ Par récurrence sur n.
Le cas $n = 1$ est clair.
Si le résultat est vrai pour $n-1$, et $A \in \mathfrak{M}_n$ comme il faut, A s'écrit $\begin{pmatrix} A' & X \\ {}^tY & \lambda \end{pmatrix}$.
A' vérifie les hypothèses du théorème donc $A' = L'U'$.
On pose $L = \begin{pmatrix} L' & 0 \\ {}^tl & 1 \end{pmatrix}$ et $U = \begin{pmatrix} U' & Z \\ 0 & u \end{pmatrix}$.
On veut $LU = A$ donc $L'Z = X$, ${}^tlU' = {}^t Y$ et ${}^tlZ + u = \lambda$ donc on choisit $Z = L'^{-1}X$, $l = {}^t U'^{-1}Y$ et $u = \lambda - {}^t YU'^{-1}L'^{-1}X$.
Le principe de récurrence assure l'existence.

! Si $A = L_1U_1 = L_2U_2$, $L_2^{-1}L_1 = U_2U_1^{-1}$ donc $L_2^{-1}L_1 \in D_n$ avec des 1 sur la diagonale.
Donc $L_2^{-1}L_1 = I_n$ donc $L_1 = L_2$ et $U_1 = U_2$. ∎

Remarque 3.1 Réciproquement, si A admet une décomposition LU, A a ses mineurs fondamentaux non nuls. En effet, pour tout $k \in [\![1, n]\!]$, on a :

$$A = \left(\begin{array}{c|c} A_k & B_k \\ \hline C_k & D_k \end{array}\right), \quad L = \left(\begin{array}{c|c} L_k & 0 \\ \hline L'_k & L''_k \end{array}\right) \text{ et } U = \left(\begin{array}{c|c} U_k & U'_k \\ \hline 0 & U''_k \end{array}\right)$$

$A_k = L_k U_k$ *donc* $\det(A_k) = \det(U_k) \neq 0$ *car* U_k *est inversible.*

3.2.2 Cas général

THÉORÈME 3.3 *Pour tout* $A \in GL_n(\mathbb{K})$*, il existe* P *matrice de permutation et* L, U *triangulaires inférieure et supérieure tel que* $A = PLU$ *et que* L *n'ait que des* 1 *sur sa diagonale et* $U \in GL_n(\mathbb{K})$*.*

Démonstration. Pivot de Gauss : cf TP ■

Remarque 3.2 *L'algorithme est en* $\frac{2n^3}{3} + O(n^2)$*.*

3.2.3 Décomposition LDU

THÉORÈME 3.4 *Si* A *a ses mineurs fondamentaux non nuls, il existe un unique triplet* (L, D, U) *avec* D *diagonale inversible,* L *triangulaire inférieure à diagonale composée de* 1 *et* U *triangulaire supérieure à diagonale composée de* 1 *tel que* $A = LDU$*.*

Démonstration.

∃ Si A a ses mineurs fondamentaux non nuls, A se décompose sous la forme LU.
Posons $D = \text{diag}(u_{1,1}, \ldots, u_{n,n})$ et $U' = D^{-1}U$. On a $A = LDU$.

! Si $L_1 D_1 U_1 = L_2 D_2 U_2$, par unicité de la décomposition LU, $L_1 = L_2$ et $D_1 U_1 = D_2 U_2$.
En regardant les cœfficients diagonaux, on obtient, pour tout $i \in [\![1, n]\!]$, $d_{i,i}^{(2)} = d_{i,i}^{(1)}$. Donc $D_1 = D_2$ et $U_1 = U_2$. ■

Remarque 3.3 $\text{Sp}(D) \neq \text{Sp}(A)$*. Par exemple :*

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

Or $\text{tr}(A) \neq 2$*.*

3.2.4 Décomposition LD^{t}L

THÉORÈME 3.5 *Pour tout* $A \in S_n(\mathbb{R})$ *ayant ses mineurs fondamentaux non nuls, il existe un unique* L, D *triangulaire inférieure et diagonale tel que* $A = LD^tL$*.*

Démonstration. $A = LDU$ et $A = {}^tA = {}^tUD^tL$ donc par unicité de la décomposition LDU, $U = {}^tL$.
Donc $A = LD^tL$. ■

Remarque 3.4

- *Ceci est valable dans $\mathscr{H}_n(\mathbb{C})$.*
- *On dit que A et D sont congruentes. En fait, elles représentent la même forme quadratique : pour tout $x \in \mathfrak{M}_{n,1}(\mathbb{R})$, on a*

$$\langle Ax|x\rangle = \langle LD^tLx|x\rangle = \langle D^tLx|^tL\rangle = \langle Dy|y\rangle$$

- *A est définie positive ssi D l'est.*
- *A et D ont même signature (ie $\mathrm{Card}(\mathrm{Sp}(A)\cap\mathbb{R}_+^*) = \mathrm{Card}(\mathrm{Sp}(D)\cap\mathbb{R}_+^*)$).*

3.2.5 Décomposition de Choleski

Théorème 3.6 *Pour tout $A \in S_n^{++}(\mathbb{R})$, il existe une unique B triangulaire inférieure à diagonale positive telle que $A = B^tB$.*

Démonstration.

∃ Montrons que A a ses mineurs fondamentaux non nuls.

$$A = \left(\begin{array}{c|c} A_k & B_k \\ \hline {}^tB_k & C_k \end{array}\right)$$

Si $A_k \notin GL_n(\mathbb{R})$, il existe $x_k \neq 0$ tel que $A_kx_k = 0$. Posons $x = \begin{pmatrix} x_k \\ 0 \end{pmatrix}$.
On a $\langle Ax|x\rangle = \left\langle \begin{pmatrix} A_kx_k \\ {}^tB_kx_k \end{pmatrix}, \begin{pmatrix} x_k \\ 0 \end{pmatrix} \right\rangle = 0$.
Donc A n'est pas définie, ce qui est contradictoire.
Donc A peut se décomposer sous la forme LD^tL. On pose $B = L\sqrt{D}$ et on a $A = B^tB$.

! Si $A = B^tB$, on pose $\widetilde{D} = \mathrm{diag}(b_{i,i}^2)$, $\widetilde{L} = B\sqrt{\widetilde{D}^{-1}}$ (dont les cœfficients diagonaux valent 1).
$\widetilde{L}\widetilde{D}^t\widetilde{L} = B^tB$ donc, par unicité de la décomposition LD^tL, $L = \widetilde{L}$ et $D = \widetilde{D}$ donc $B = L\sqrt{D}$. ∎

Remarque 3.5

- *La réciproque est vraie : B^tB est symétrique réelle et pour tout x,*

$$\langle Ax|x\rangle = \left\|{}^tBx\right\|^2 \geqslant 0$$

De plus, si $\langle Ax|x\rangle = 0$, $x \in \mathrm{Ker}({}^tB) = \{0\}$ car B inversible.

- *Cet algorithme est plus avantageux que le pivot de Gauss : avec les formules*

$$\forall i < j, b_{i,j} = \frac{1}{b_{j,j}}\left(a_{i,j} - \sum_{k=1}^{j-1} b_{i,k}b_{j,k}\right)$$

$$\forall j, b_{j,j}^2 = a_{j,j} - \sum_{k=1}^{j-1} b_{j,k}^2$$

on effectue n racines, $\frac{n(n-1)}{2}$ divisions, $\frac{n(n+1)(n-1)}{6}$ multiplications et $\frac{n(n+1)(n-1)}{6}$ additions. On a donc $\frac{n^3}{3} + O(n^2)$ opérations.

- *Pour savoir si $M \in S_n^{++}$, on lui applique cet algorithme et s'il ne marche pas, $M \notin S_n^{++}$.*

3.3 Décomposition QR

3.3.1 Cas des matrices inversibles

THÉORÈME 3.7 *Pour tout $A \in GL_n(\mathbb{R})$, il existe un unique $(Q, R) \in O_n \times T_n^{++}$ tel que $A = QR$.*

Démonstration.

∃ $A = (C_1|\cdots|C_n)$ est inversible donc $(C_1, \ldots, C_n) = C$ est une base de $\mathfrak{M}_{n,1}(\mathbb{R})$.
D'après Gram-Schmidt, il existe ε base orthonormée de $\mathfrak{M}_{n,1}(\mathbb{R})$ telle que pour tout $i \in [\![1, n]\!]$, $\text{Vect}\{C_1, \ldots, C_i\} = \text{Vect}\{\varepsilon_1, \ldots, \varepsilon_i\}$ et de plus $\langle C_i, \varepsilon_i \rangle > 0$.
Si E est la base canonique de $\mathfrak{M}_{n,1}(\mathbb{R})$, on a :

$$A = P_{E\to C} = P_{E\to\varepsilon} P_{\varepsilon\to C}$$

Or $P_{E\to\varepsilon} \in O_n$ car ε est orthonormée et $P_{\varepsilon\to C}$ est triangulaire par respet des drapeaux et à diagonale strictement positive car les cœfficients diagonaux sont, par construction, les $\|\varepsilon_i\| > 0$.

! Si $A = Q_1R_1 = Q_2R_2$, $Q_2^{-1}Q_1 = R_2R_1^{-1}$ est une matrice orthogonale et triangulaire supérieure à coefficients diagonaux positifs donc vaut I_n.
Donc $Q_1 = Q_2$ et $R_1 = R_2$. ■

Remarque 3.6 On a un algorithme via Gram-Schmidt mais pas très performant (car les erreurs d'arrondis s'accumulent).

3.3.2 Matrices rectangulaires de $\mathfrak{M}_{n,p}(\mathbb{R})$ avec $n > p$

THÉORÈME 3.8 *Pour tout $A \in \mathfrak{M}_{n,p}(\mathbb{R})$ avec $\text{rg}(A) = n$, il existe $(Q, R) \in O_n \times T_{n,p}^{+}$ tel que $A = QR$.*

Démonstration. Notons $A = (C_1|\cdots|C_p)$. $(C_1, \ldots, C_p)$ est libre donc on peut la compléter en une base $(C_1, \ldots, C_n)$.

$\widetilde{A} = (C_1|\cdots|C_n)$ est inversible sonc il existe $Q, \widetilde{R} \in O_n \times T_n^+$ tel que $\widetilde{A} = Q\widetilde{R}$.

On pose $\widetilde{R} = (R|R')$ avec R de taille $n \times p$ et on a $A = QR$. ∎

3.3.3 Applications

- Résolution de $AX = B$.

$$AX = B \quad \text{ssi} \quad QRX = B \quad \text{ssi} \quad RX = {}^tQB$$

  et ce système est alos triangulaire.
- Résolution de systèmes sur-déterminés (méthode des moindres carrés)
- Calcul approché des valeurs propres.

3.3.4 Cas non inversible

THÉORÈME 3.9 *Pour tout $A \in \mathfrak{M}_n(\mathbb{R})$, il existe $(Q, R) \in O_n \times T_n^+$ tel que $A = QR$ et R à diagonale positive.*

Démonstration. Il existe $(A_n)_n$ inversibles tel que $\lim\limits_{n\to+\infty} A_k = A$.

Pour tout n, $A_n = Q_nR_n$ donc, comme O_n est compact, il existe φ, Q tel que $\lim\limits_{n\to+\infty} Q_{\varphi(n)} = Q$.

On a de plus $R_{\varphi(n)} = {}^tQ_{\varphi(n)}A_{\varphi(n)}$ donc converge vers ${}^tQA = R$.

$R \in T_n^+$ car T_n^+ est fermé. De plus, pour tout i, $r_{i,i} \geqslant 0$ car $r_{i,i} = \lim\limits_{n\to+\infty} r_{i,i}^{\varphi(n)}$. ∎

3.4 Décomposition en valeurs singulières

On l'appelle aussi SVD (singular values decomposition).

THÉORÈME 3.10 *Pour tout $A \in \mathfrak{M}_{n,p}(\mathbb{C})$ de rang k, on note $\sigma_1 \geqslant \cdots \geqslant \sigma_k > 0$ les valeurs singulières de A (racines des valeurs propres de A^*A).*

On pose $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$. Il existe $U \in \mathcal{U}_n$ et $V \in U_p$ tel que $A = USV^*$ avec :

$$S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \in \mathfrak{M}_{n,p}(\mathbb{C})$$

Démonstration.

- $\mathrm{Ker}(A^*A) = \mathrm{Ker}(A)$ donc $\mathrm{Sp}(A^*A) = \{0, \sigma_1^2, \dots, \sigma_k^2\}$ avec 0 de multiplicité $p-k$.
  Soit $(v_1, \dots, v_p)$ une base orthonormée de vecteurs propres colonne de A^*A tel que v_i soit associé à σ_i^2.
  Donc $V = (v_1|\cdots|v_p) \in U_p$.
- On pose $u_i = \frac{1}{\sigma_i} A v_i$. Montrons que $(u_1, \dots, u_k)$ est orthonormée dans $\mathfrak{M}_{n,1}(\mathbb{C})$.
  On a $AA^*u_i = \frac{1}{\sigma_i} AA^*Av_i = \sigma_i A v_i = \sigma_i^2 u_i$.
  De plus,

$$\begin{aligned}\langle u_i | u_j \rangle &= \frac{\langle Av_i | Av_j \rangle}{\sigma_1 \sigma_2} \\ &= \frac{\langle v_i | A^*Av_j \rangle}{\sigma_i \sigma_j} \\ &= \frac{\sigma_i \langle v_i | v_j \rangle}{\sigma_j} \\ &= \delta_{i,j}\end{aligned}$$

  Donc $(u_1, \dots, u_k)$ est orthonormée donc libre donc on peut la compléter en une base orthonormée $(u_1, \dots, u_n)$ de $\mathfrak{M}_{n,1}(\mathbb{C})$. On pose $U = (u_1|\cdots|u_n) \in \mathcal{U}_n$.

$$\begin{aligned}U^*AV &= U^*(Av_1|\cdots|Av_p|0) \\ &= U^*(\sigma_1 u_1|\cdots|\sigma_k u_k|0) \\ &= \begin{pmatrix} \mathrm{diag}(\sigma_1, \dots, \sigma_k) & 0 \\ 0 & 0 \end{pmatrix} \\ &= S\end{aligned}$$

Proposition 3.1 Si $A = USV^* \in \mathfrak{M}_{n,p}(\mathbb{C})$ et $S = \mathrm{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ et $\sigma_1 \geqslant \cdots \geqslant \sigma_r > 0$.

Pour $k \in [\![1, r-1]\!]$, la meilleure approximation de A par une matrice de rang k est $A_k = US_kV^*$ où $S_k = \mathrm{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0)$.

De plus, $\|A - A_k\|_2 = \sigma_{k+1}$.

Démonstration.

- Montrer que $\|A - A_k\|_2 = \sigma_{k+1}$.
  $A - A_k = U(S - S_k)V$ et $S - S_k = \mathrm{diag}(0, \dots, 0, \sigma_{k+1}, \dots, \sigma_r, 0, \dots, 0)$.
  Donc $\|A - A_k\|_2 = \|S - S_k\|_2 = \sigma_{k+1}$.
- Soit B de rang k. On veut trouver $x \neq 0$ tel que $\|(A-B)x\|_2 \geqslant \sigma_{k+1} \|x\|_2$.

On pose $E = \text{Vect}\{v_1, \dots, v_{k+1}\}$ avec $V = (v_1 | \cdots | v_p)$.
$\dim(E) = k+1$ donc, comme $\text{rg}(B) = k$, $\dim(\text{Ker}(B)) = n-k$ et $\text{Ker}(B) \cap E \neq \{0\}$.
Donc il existe $x \in \text{Ker}(B) \cap E$ non nul. On note $x = \sum_{i=1}^{k+1} \lambda_i v_i$.

$$\|(A-B)x\|_2 = \|Ax\|_2 = \|SV^*x\|_2$$

Or $V^*x = \sum_{i=1}^{k+1} \lambda_i V^* v_i = \sum_{i=1}^{k+1} \lambda_i e_i$.
Donc $SV^*x = {}^t(\sigma_1\lambda_1, \dots, \sigma_{k+1}\lambda_{k+1}, 0, \dots, 0)$.
On a alors $\|(A-B)x\|_2 \geqslant \sigma_{k+1} \|x\|_2$. ■

Application : Compression d'images.

Un image en noir et blanc est représentée par une matrice de $\mathfrak{M}_{558,768}$ (généralement) qui est de rang 555 environ.

Si on approxime avec une matrice de rang environ 50, on obtient une image très ressemblante et on gagne de la place.

Chapitre 4

Analyse spectrale en dimension finie

4.1 Localisation des valeurs propres

4.1.1 Disques de Gerschgörin

THÉORÈME 4.1 *Soit* $A \in \mathfrak{M}_n(\mathbb{C})$, $\lambda \in \mathrm{Sp}(A)$.

$$\lambda \in \bigcup_{i=1}^{n} D_i \text{ avec } D_i = \left\{ z \in \mathbb{C}, |z - a_{i,i}| \leqslant \sum_{j \neq i} |a_{i,j}| \right\}.$$

Démonstration. Soit x un vecteur propre de A associé à λ. Notons i_0 l'indice tel que $|x_{i_0}| = \|x\|_\infty$.

$Ax = \lambda x$ donc $\sum_{j=1}^{n} a_{i_0,j} x_j = \lambda x_{i_0}$ donc $(\lambda - a_{i_0,i_0}) = \sum_{j \neq i_0} \frac{a_{i_0,j} x_j}{x_{i_0}}$.

L'inégalité triangulaire assure que $\lambda \in D_{i_0}$. ∎

Remarque 4.1 $\mathrm{Sp}(A) = \mathrm{Sp}({}^tA)$ *donc* $\mathrm{Sp}(A) \subset \bigcup_{i=1}^{n} D'_i$ *avec*

$$D'_i = \left\{ z \in \mathbb{C}, |z - a_{i,i}| \leqslant \sum_{j \neq i} |a_{j,i}| \right\}$$

Définition 4.1 $A \in \mathfrak{M}_n(\mathbb{C})$ est dite à diagonale strictement dominante ssi pour tout $i \in [\![1, n]\!]$, $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$.

THÉORÈME 4.2 *Toute matrice à diagonale strictement dominante est inversible.*

Démonstration. Pour tout i, $0 \notin D_i$ donc $0 \notin \mathrm{Sp}(A)$. ∎

4.1.2 Continuité des valeurs propres

Lemme 4.2.1
Si $\lim\limits_{k\to+\infty} A_k = A$ alors pour tout $\lambda \in \mathrm{Sp}(A)$, il existe $(\lambda_k)_k \in \mathbb{C}^{\mathbb{N}}$ tel que $\lambda_k \in \mathrm{Sp}(A_k)$ et $\lim\limits_{k\to+\infty} \lambda_k = \lambda$.

Démonstration. Soit $\lambda \in \mathrm{Sp}(A)$.

$\chi_\cdot$ est continue donc $\lim\limits_{n\to+\infty} \chi_{A_n}(\lambda) = \chi_A(\lambda) = 0$.

Or $|\chi_{A_k}(\lambda)| = \prod\limits_{j=1}^{n} |\lambda - \lambda_j^{(k)}| \geqslant (\min\limits_{1\leqslant j\leqslant n} |\lambda - \lambda_j^{(k)}|)^n$.

Donc $\lim\limits_{k\to+\infty} \min\limits_{1\leqslant j\leqslant n} |\lambda - \lambda_j^{(k)}| = 0$.

En prenant λ_k la valeur propre qui réalise la minimum, on obtient le résultat. ∎

Remarque 4.2 *Si $(A_t)_{t\in[0,1]}$ est une famille continue de matrices, il existe $(\lambda_1^{(t)})_t, \dots, (\lambda_n^{(t)})_t$ tel que $\mathrm{Sp}(A_t) = \{\lambda_1^{(t)}, \dots, \lambda_n^{(t)}\}$ et $t \mapsto \lambda_k^{(t)}$ soit continue pour tout k.*

Corollaire 4.1 *Soit $A \in \mathfrak{M}_n(\mathbb{C})$.*

S'il existe k tel que $\Omega_k = \bigcup\limits_{i=1}^{k} D_i$ et $\Gamma_k = \bigcup\limits_{i=k+1}^{n} D_i$ soient disjoints, alors Ω_k contient exactement k valeurs propres (comptées avec ordre de multiplicité) et Γ_k en contient $n-k$.

Démonstration. On pose $D = \mathrm{diag}(a_{i,i})_i$, $B = A - D$, $A_t = D + tB$, $\Omega_k^t = \bigcup\limits_{i=1}^{k} D_i^t$ et $\Gamma_k^t = \bigcup\limits_{i=k+1}^{n} D_i^t$ avec $D_i^t = \left\{ z \in \mathbb{C}, |z - a_{i,i}| \leqslant t\sum\limits_{j\neq i} |a_{i,j}| \right\}$.

On définit les λ_i^t de la remarque précédente pour $i \leqslant k$.

S'il existe $t > 0$ tel que $\lambda_i^t \notin \Omega_k$, on pose $t_0 = \inf\{t, \lambda_i^t \notin \Omega_k\}$.

Il existe donc $(t_n) \in [0,1]^{\mathbb{N}}$ tel que $\lim\limits_{n\to+\infty} t_n = t_0$.

On a $\lambda_i^{t_n} \in \bigcup\limits_{i=1}^{n} D_i^{t_n} \subset \bigcup\limits_{i=1}^{n} D_i = \Omega_k \cup \Gamma_k$.

Donc $\lambda_i^{t_n} \in \Gamma_k$ qui est fermé donc $\lambda_i^{t_0} \in \Gamma_k$.

De même, on a $\lambda_i^{t_0} \in \Omega_k$ donc contradiction.

Donc, pour tout $t \in [0,1]$, $\lambda_i^t \in \Omega_k$ donc $\lambda_i \in \Omega_k$. ∎

4.1.3 Perturbation des valeurs propres

Cas diagonal

Proposition 4.1 Si A est diagonale et $B \in \mathfrak{M}_n(\mathbb{C})$, pour tout $\lambda \in \mathrm{Sp}(A+B)$, il existe $\lambda' \in \mathrm{Sp}(A)$ tel qie $|\lambda - \lambda'| \leqslant |||B|||_\infty$.

Démonstration.

$$\begin{aligned}\mathrm{Sp}(A+B) &\subset \bigcup_{i=1}^{n} \left\{ z \in \mathbb{C}, |z - \lambda_i - b_{i,i}| \leqslant \sum_{j \neq i} |b_{i,j}| \right\} \\ &\subset \bigcup_{i=1}^{n} \left\{ z \in \mathbb{C}, |z - \lambda_i| \leqslant \sum_{j=1}^{n} |b_{i,j}| \right\} \\ &\subset \bigcup_{i=1}^{n} \{ z \in \mathbb{C}, |z - \lambda_i| \leqslant |||B|||_\infty \}\end{aligned}$$

Cas diagonalisable

Proposition 4.2 Si A est diagonalisable et $B \in \mathfrak{M}_n(\mathbb{C})$, pour tout $\lambda \in \mathrm{Sp}(A+B)$, il existe $\lambda' \in \mathrm{Sp}(A)$ tel qie $|\lambda - \lambda'| \leqslant c_\infty(P)|||B|||_\infty$.

Démonstration. $\mathrm{Sp}(A+B) = \mathrm{Sp}(D + P^{-1}BP)$ conclut. ∎

Cas général

On n'a pas de tel résultat et on ne peut en avoir : $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ \varepsilon & 0 \end{pmatrix}$ avec $\varepsilon > 0$ sont telles que $\mathrm{Sp}(A) = \{0\}$, $\mathrm{Sp}(A+B) = \{\pm\sqrt{\varepsilon}\}$ et $\sqrt{(}\varepsilon) \leqslant c\varepsilon$ n'est possible pour aucun c.

Remarque 4.3 *$A+B$ est diagonalisable via $P = \begin{pmatrix} 1 & 1 \\ \sqrt{\varepsilon} & -\sqrt{\varepsilon} \end{pmatrix}$ et $c_\infty(P) = \frac{1+\sqrt{\varepsilon}}{\sqrt{\varepsilon}}$.*

4.2 Cas hermitien

4.2.1 Définitions

Définition 4.2 Soit $A \in \mathscr{H}_n$.

- On appelle quotient de Rayleigh associé à A la fonction :

$$r_A : \begin{cases} \mathbb{C}^n \setminus \{0\} & \to & \mathbb{C} \\ x & \mapsto & \frac{\langle Ax, x\rangle}{\|x\|_2^2} \end{cases}$$

- On appelle hausdorffien de A et on note $\mathscr{H}(A)$ l'ensemble $r_A(\mathbb{C}^n \setminus \{0\})$.

Remarque 4.4

- *Pour tout* λ, x, $r_A(\lambda x) = \lambda r_A(x)$ *donc* $\mathscr{H}(A) = r_A(S(O,1))$.
- $\mathscr{H}(A) \subset \mathbb{R}$ *car* $A \in \mathscr{H}_n$.

THÉORÈME 4.3 RAYLEY-RITZ *Pour tout* $A \in \mathbb{H}_n$ *dont on note* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *les valeurs propres,*

$$\lambda_1 = \min_{x \in \mathbb{C}^n \setminus \{0\}} r_A(x) = \min_{x \in S(O,1)} r_A(x)$$

$$\lambda_n = \max_{x \in \mathbb{C}^n \setminus \{0\}} r_A(x) = \max_{x \in S(O,1)} r_A(x)$$

Démonstration. A est orthogonalement diagonalisable donc $A = UDU^*$ avec $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Pour $x \neq 0$, on a :

$$r_A(x) = \frac{\langle Ax, x\rangle}{\|x\|_2^2} = \frac{\langle DU^*x, U^*x\rangle}{\|x\|_2^2}$$

Donc $\min_{x \in S(O,1)} r_A(x)) = \min_{x \in S(O,1)} r_D(U^*x) = \min_{y \in S(O,1)} r_D(y) = \lambda_1$.

On a de même le résultat avec le max. ∎

COROLLAIRE 4.2 $\mathscr{H}(A) = [\lambda_1, \lambda_n]$.

Démonstration. $\mathscr{A} \subset [\lambda_1, \lambda_n]$ et r_A est continue donc le TVI conclut. ∎

4.2.2 Caractérisation min-max de COURANT-FISHER

THÉORÈME 4.4 *Soit* $A \in \mathscr{H}_n$ *et* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ *ses valeurs propres.*

Pour tout $k \in [\![1, n]\!]$,

$$\lambda_k = \min_{\substack{F \in \mathscr{G}(\mathbb{C}^n)^{\,3} \\ \dim(F) = k}} \max_{x \in F \setminus \{0\}} r_A(x)$$

Remarque 4.5 *Les cas* $k = 1$ *et* $k = n$ *proviennent du théorème précédent.*

Démonstration. Soit $k \in [\![1, n]\!]$, et $F \in \mathscr{G}(\mathbb{C}^n)$ de dimension k.

$A = UDU^*$ donc pour tout $x \in F \setminus \{0\}$, $r_A(x) = r_D(U^*x)$.

On a $\max\limits_{x \in F\setminus\{0\}} r_A(x) = \max\limits_{x \in U^*(F)\setminus\{0\}} r_D(y)$.

Donc $\min\limits_{\substack{F \in \mathscr{G}(\mathbb{C}^n) \\ \dim(F)=k}} \max\limits_{x \in F\setminus\{0\}} r_A(x) = \min\limits_{\substack{F \in \mathscr{G}(\mathbb{C}^n) \\ \dim(F)=k}} \max\limits_{x \in U^*(F)\setminus\{0\}} r_D(y)$.

Le sous-espace pour lequel le min est atteint est $G = \text{Vect}\{e_1, \ldots, e_k\}$ et on a $\max\limits_{y \in G} r_D(y) = \lambda_k$, ce qui assure le résultat. ■

COROLLAIRE 4.3 *(WEYL) Soient $A, B \in \mathscr{H}_n$.*

Pour tout M, on note $\lambda_k(M)$ la k-ième plus petite valeur propre de M.

Pour tout $k \in [\![1, n]\!]$, $\lambda_k(A) + \lambda_1(B) \leqslant \lambda_k(A + B) \leqslant \lambda_k(A) + \lambda_n(B)$.

Démonstration. Soit $x \neq 0$.

$r_\cdot(x)$ est linéaire donc :

$$\lambda_k(A+B) = \min_{\substack{F \in \mathscr{G}(\mathbb{C}^n) \\ \dim(F)=k}} \max_{x \in F\setminus\{0\}} r_{A+B}(x) = \min_{\substack{F \in \mathscr{G}(\mathbb{C}^n) \\ \dim(F)=k}} \max_{x \in F\setminus\{0\}} r_A(x) + r_B(x)$$

Or $\lambda_1(B) \leqslant \lambda_k(B) \leqslant \lambda_n(B)$, ce qui conclut. ■

Remarque 4.6 Si A est diagonalisable en base orthonormée, la partie 1 donne l'encardrement $|\lambda_k(A + B) - \lambda| \leqslant c_\infty(U)|||B|||_\infty$ et la partie 2 donne $|\lambda_k(A + B) - \lambda| \leqslant \rho(B) \leqslant |||B|||_\infty$.

4.3 Spectre des matrices positives

4.3.1 Définitions

Définition 4.3 On définit la relation d'ordre partiel $\preceq$ sur $\mathfrak{M}_{n,p}(\mathbb{R})$ par :

$$A \preceq B \quad \text{ssi} \quad \forall (i,j), a_{i,j} \leqslant b_{i,j}$$

On dit que :

$$A \prec B \quad \text{ssi} \quad \forall (i,j), a_{i,j} < b_{i,j}$$

Définition 4.4 On dit que M est positive (resp. strictement positive) ssi $A \succeq 0$ (resp. $A \succ 0$).

On note de plus $|A| = (|a_{i,j}|)_{i,j}$.

Proposition 4.3 Soit $A, B \in \mathfrak{M}_n(\mathbb{R})$ et $x \in \mathfrak{M}_{n,1}(\mathbb{R})$.

- $|A| \succeq 0$ et $|A| = 0 \quad$ ssi $\quad A = 0$.
- Si $A \succeq 0$ et $x \succeq 0$, $Ax \succeq 0$.
- Si $A \succeq 0$ et $B \succeq 0$, $AB \succeq 0$.

- $|Ax| \leqslant |A||x|$ et $|AB| \leqslant |A||B|$.
- Si $0 \preceq A \preceq B$, $\rho(A) \leqslant \rho(B)$.
- Si $A \succ 0$, $\rho(A) > 0$.

Démonstration.

- Les quatre premiers points sont faciles.
- $\rho(A) = \|\!|A^k|\!\|^{\frac{1}{k}}$.
  Pour tout k, on a $0 \preceq A^k \preceq B^k$ donc $0 \leqslant \|\!|A^k|\!\| \leqslant \|\!|B^k|\!\|$.
  En passant à la limite, $\rho(A) \leqslant \rho(B)$.
- Si $\rho(A) = 0$, $A^n = 0$ donc on a une contradiction ($A \succ 0 \Rightarrow A^n \succ 0$). ■

Proposition 4.4 Soit $A \in \mathfrak{M}_n(\mathbb{R})$ et $x \in \mathfrak{M}_{n,1}(\mathbb{R})$ tel que $A \succeq 0$ et $x \succ 0$.

$$\underbrace{\min_{1\leqslant i\leqslant n} \sum_{j=1}^{n} a_{i,j}}_{=\alpha} \leqslant \rho(A) \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} a_{i,j} = \|\!|A|\!\|_\infty$$

$$\min_{1\leqslant i\leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{i,j}x_j \leqslant \rho(A) \leqslant \max_{1\leqslant i\leqslant n} \frac{1}{x_i} \sum_{j=1}^{n} a_{i,j}x_j$$

Démonstration.

- On a déjà vu $\rho(A) \leqslant \|\!|A|\!\|_\infty$.
  On va construire B tel que $0 \preceq B \preceq A$ et $\rho(B) = \alpha$. Posons $B = \left(\frac{\alpha a_{i,j}}{\sum_{k=1}^{n} a_{i,k}} \right)_{i,j}$.
  On a clairement $B \succeq 0$ et $\frac{\alpha}{\sum_{k=1}^{n} a_{i,k}} < 1$ donc $B \preceq A$.
  On a de plus $\rho(B) \leqslant \|\!|B|\!\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} b_{i,j} = \alpha$.
  Or $v = (1, \dots, 1)^t$ est un vecteur propre de B associé à α donc $\rho(B) \geqslant \alpha$ et $\rho(A) \geqslant \rho(B) = \alpha$.
- Si $D = \mathrm{diag}(x_1, \dots, x_n)$, on conclut en appliquant le point précédent à $A' = D^{-1}AD$. ■

4.3.2 Matrices strictement positives

Lemme 4.4.1
Soit $A \succ 0$ et $\lambda \in \mathrm{Sp}(A)$ tel que $|\lambda| = \rho(A)$.
Soit $x \in \mathfrak{M}_{n,1}(\mathbb{C})$ associé à λ.

On a $A|x| = |\lambda||x|$ et $|x| \succ 0$.

Remarque 4.7 *On a ainsi* $\rho(A) \in \mathrm{Sp}(A)$.

Démonstration. $Ax = \lambda x$ donc $A|x| = |A||x| \succeq |Ax| = |\lambda x| = |\lambda||x| = \rho(A)|x|$.

Supposons $y = A|x| - \rho(A)|x|$ non nul.

On a $Ay = AA|x| - \rho(A)\underbrace{A|x|}_{=z} = Az - \rho(A)z$.

On a $z \succ 0$ car $z_j = \sum_{k=1}^{n} \underbrace{a_{i,k}}_{>0 \forall j,k} \underbrace{|x_k|}_{>0 \text{ pour un } k} > 0$.

De même, $y \succeq 0$, $y \neq 0$ et $A \succ 0$ donc $Ay \succ 0$ (de même que précédemment).

On a donc $Az \succ \rho(A)z$ donc $\forall i, \sum_{j=1}^{n} a_{i,j}z_j > \rho(A)z_i$ donc

$$\min_{1 \leqslant i \leqslant n} \frac{1}{z_i} \sum_{j=1}^{n} a_{i,j}z_j > \rho(A)$$

d'où une contradiction.

Donc $y = 0$ et $A|x| = \rho(A)|x|$.

On a de plus $|x| = \frac{1}{\rho(A)}z \succ 0$. ∎

Lemme 4.4.2

Soit $A \succ 0$ et $\lambda \in \mathrm{Sp}(A)$ tel que $|\lambda| = \rho(A)$.

Pour tout $x \in \mathfrak{M}_{n,1}(\mathbb{C})$ associé à λ, il existe $\theta \in \mathbb{R}$ tel que $x = |x|\mathrm{e}^{i\theta}$.

Démonstration. $|x| \succ 0$ et $A|x| = \rho(A)|x|$ donc pour tout i, $\sum_{j=1}^{n} a_{i,j}|x_j| = \rho(A)|x_i|$.

De plus, $\forall i$, $\rho(A)|x_i| = |\lambda x_i| = |(Ax)_i| = \left|\sum_{i=1}^{n} a_{i,j}x_j\right|$ qui est le cas d'égalité dans l'inégalité triangulaire.

Donc pour tout i, j, $\arg(x_i) \equiv \arg(x_j) \mod 2\pi$. ∎

Corollaire 4.4 *Soit* $A \succ 0$ *et* $\lambda \in \mathrm{Sp}(A)$ *tel que* $|\lambda| = \rho(A)$.

On a $\lambda = \rho(A)$.

Démonstration. Soit x un vecteur propre associé à λ. On a $x = |x|\mathrm{e}^{i\theta}$.

$$A|x| = \mathrm{e}^{-i\theta}Ax = \mathrm{e}^{-i\theta}\lambda x = \lambda|x|$$

Donc $\lambda > 0$ et $\lambda = \rho(A)$. ∎

Lemme 4.4.3
Soit $A \succ 0$.

$$\dim(E_{\rho(A)}) = 1$$

Démonstration. Soient w, z deux vecteurs propres associés à $\rho(A)$.

Quitte à multiplier par $e^{i\theta}$, on peut supposer $w \succ 0$ et $z \succ 0$.

On pose $\beta = \min\limits_{1 \leqslant i \leqslant n} \frac{z_i}{w_i} > 0$ et $r = z - \beta w$.

On a $Ar = Az - \beta Aw = \rho(A)r$. Si $r \neq 0$, $r \in E_{\rho(A)}$ donc $|r| \succ 0$ donc, comme $r \succeq 0$, $r \succ 0$.

Or il existe i tel que $z_i = \beta r_i$ donc $r_i = 0$ d'où la contradiction.

Donc $r = 0$ et $z = \beta w$. ■

Théorème 4.5 de Perron-Frobénius *Soit $A \in \mathfrak{M}_n(\mathbb{R})$ telle que $A \succ 0$.*

- $\rho(A) > 0$.
- $\exists x \succ 0$, $Ax = \rho(A)x$ *et* $E_{\rho(A)} = \mathbb{R}x$.
- $\forall \lambda \in \mathrm{Sp}(A) \setminus \{\rho(A)\}$, $|\lambda| < \rho(A)$.

Remarque 4.8 Dans le cas $A \succeq 0$, quelles assertions restent vraies ?

- $\rho(A) > 0$ *devient fausse* ($A = 0$).
- $\dim(E_{\rho(A)}) = 1$ *devient fausse* ($A = I_n$).
- $\exists x \succ 0$, $Ax = \rho(A)x$ *reste vraie (considérer des limites de matrices positives).*
- $\forall \lambda \in \mathrm{Sp}(A) \setminus \{\rho(A)\}$, $|\lambda| < \rho(A)$ *devient fausse* ($A = J_2 - I_2$).

Chapitre 5

Systèmes linéaires

Introduction

Le but est de résoudre $Ax = b$ avec $A \in \mathfrak{M}_n(\mathbb{K})$ (n très grand), ou de résoudre au sens des moindres carrés $Ax = b$ avec $A \in \mathfrak{M}_{n,p}(\mathbb{K})$.

Il y a deux types de méthodes de résolution : les méthodes directes où on calcule directement les solutions exactes, ou bien les méthodes itératives où on construit uns suite qui converge vers la solution.

5.1 Méthodes directes

5.1.1 Cramer

Les formules de Cramer sont inutilisables en pratique car, si A est de taille $n \times n$, on doit faire $O(n \cdot n!)$ opérations.

Par exemple, pour n assez petit ($n = 50$) avec un processeur 3GHz, cette opération nécessite 10^{47} années.

5.1.2 Gauss

THÉORÈME 5.1 *Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$, il existe $M \in GL_n(\mathbb{K})$ telle que $T = AM \in T_n^+$.*

5.1.3 Décomposition LU

THÉORÈME 5.2 *Pour tout $A \in \mathfrak{M}_n(\mathbb{K})$ dont aucun mineur fondamental n'est nul, il existe un unique couple $(L, U) \in T_n^- \times T_n^+$ tel que L n'a que des*

1 *sur sa diagonale, U n'a que des coefficients non nuls sur la diagonale et $A = LU$.*

5.1.4 Matrices creuses

Définition 5.1 On dit qu'une matrice est creuse ssi elle a beaucoup[1] de 0

Définition 5.2 Soit $A \in \mathfrak{M}_n(\mathbb{K})$.

Pour tout i, on note $j_i = \min\{j, a_{i,j} \neq 0\}$ et pour tout j, on note $i_j = \min\{i, a_{i,j} \neq 0\}$.

On appelle profil inférieur (resp. supérieur) et on note ProfInf (resp. ProfSup) l'ensemble $\{(i,j), j_i \leqslant j \leqslant i\}$ (resp. $\{(i,j), i_j \leqslant i \leqslant j\}$).

On appelle profil l'union de ces ensembles. On la note $\text{Prof}(A)$.

Remarque 5.1 $a_{i,j} = 0 \not\Rightarrow (i,j) \notin \text{Prof}(A)$.

Pour des matrices très creuses, on stocke en mémoire seulement les cœfficients du profil et les indices des cœfficients diagonaux.

THÉORÈME 5.3 *Si A admet une décomposition LU,*

$$\text{ProfInf}(L) \subset \text{ProfInf}(A) \text{ et } \text{ProfSup}(U) \subset \text{ProfSup}(A)$$

Démonstration. On a, pour tout (i,j), $a_{i,j} = \sum_{k=1}^{\min(i,j)} l_{i,k}u_{k,j}$.

Donc $l_{i,j} = \frac{1}{u_{j,j}}\left(a_{i,j} - \sum_{k=1}^{j-1} l_{i,k}u_{k,j}\right)$.

Soit i tel que $j_i \geqslant 2$.

Pour $j = 1$, $l_{i,j} = \frac{a_{1,1}}{u_{1,1}} = 0$.

Si $l_{i,1}, \ldots, l_{i,j-1}$ sont nuls et $j \leqslant j_i - 1$, $l_{i,j} = \frac{a_{i,j}}{u_{j,j}} = 0$. Donc par récurrence finie sur j, on obtient le résultat :

$$(i,j) \notin \text{ProfInf}(A) \Rightarrow (i,j) \notin \text{ProfInf}(L)$$

Par contraposition, on obtient le résultat. ∎

Définition 5.3 A est dite matrice bande de demi-largeur $p \in \mathbb{N}$ ssi $a_{i,j} = 0$ pour tout (i,j) tel que $|i-j| > p$.

La largeur de bande de A est $2p+1$.

COROLLAIRE 5.1 *La décomposition LU préserve la structure bande.*

1. Le « beaucoup » peut signifier aucun !

5.1.5 Choleski

La matrice B se calcule en $\frac{n^3}{6} + O(n^2)$ opérations.
De plus, la décomposition préserve le profil.

5.1.6 QR

Si $A = QR$, $Ax = B$ ssi $Rx = Q^*B$ avec R triangulaire.

Algorithme naïf

Le calcul de Q et R se fait en $n^3 + O(n^2)$ opérations.
Le calcul de Q^*B se fait en $O(n^2)$ opérations, de même que la résolution du système triangulaire.
On a donc un algorithme en $n^3 + O(n^2)$.

Algorithme de Householder

On veut améliorer la stabilité de Gram-Schmidt. On va multiplier A à gauche par des matrices unitaires pour obtenir des zéros sous la diagonale.

Définition 5.4 Pour tout $v \in \mathbb{C}^n \setminus \{0\}$, on appelle matrice de Householder associée à v et on note $H[v]$ la matrice $I_n - \frac{2vv^*}{\|v\|^2}$.

Théorème 5.4
- *$H[v]$ est la matrice de la réflexion par rapport à $v^\perp$.*
- *Pour tout $e \in \mathbb{C}^n$ de norme 1, et pour tout v non colinéaire à e, $H[v \pm \|v\| e] = \mp \|v\| e$.*
- *Pour tout $v \in \mathbb{C}^n$, $(vv^*)^2 = \|v\| vv^*$.*

On a donc l'algorithme :

Algorithme 1: Algorithme de Householder

Entrées : $A = (C_1|\cdots|C_p) \in \mathfrak{M}_{n,p}(\mathbb{C})$ $(n \geqslant p)$
Sorties : Q unitaire et R triangulaire telle que $A = QR$

$A_1 := A$
pour $k = 1$ **à** p **faire**
  Notons $A_k = (a^k_{i,j})_{i,j}$ et $a^k = (a^k_{i,k})_{k \leqslant i \leqslant n}$.
  Prenons e'_1 le premier vecteur de base de $\mathbb{R}^{n-k}$.
  $H_k := \begin{pmatrix} I_{k-1} & 0 \\ 0 & H[a^k - \left\|a^k\right\| - e'_1] \end{pmatrix}$
  $A_{k+1} := H_k A_k$
retourner $(Q, R) = ((H_p \cdots H_1)^*, A_{p+1})$

5.2 Systèmes surdéterminés

Le problème est le suivant :

Étant donnés $A \in \mathfrak{M}_{n,p}(\mathbb{C})$ et $b \in \mathbb{R}^n$ avec $n \geqslant p$, on veut trouver $x \in \mathbb{R}^p$ tel que $\|Ax - b\|_2 = \min_{y \in \mathbb{R}^p} \|Ay - b\|_2$.

5.2.1 Conditions d'existence et d'unicité de la solution

THÉORÈME 5.5 *Il existe une solution.*

Démonstration. La projection orthogonale z de b sur $\mathrm{Im}(A)$ vérifie $z = Ax$ et x convient. ∎

THÉORÈME 5.6 *Il y a unicité ssi A est injective.*

Démonstration. Notons $z = p_{\mathrm{Im}(A)}(b)$. $z \in \mathrm{Im}(A)$ donc $z = Ax$.

Si $\mathrm{Ker}(A) = \{0\}$, il existe un unique x tel que $z = Ax$ et x convient.

Si $\mathrm{Ker}(A) \neq \{0\}$, on prend $y \neq 0$ tel que $Ay = 0$.

On a alors $A(x + y) = Ax = z$ donc on a deux solutions. ∎

5.2.2 Équation normale

Lemme 5.6.1
x est solution du problème ssi $A^*Ax = A^*b$. (c'est une équation normale)

Démonstration.

$$
\begin{aligned}
&x \text{ est solution du problème} \quad \text{ssi} \quad \forall y \in \mathbb{R}^p, \|Ay - b\|^2 \geqslant \|Ax - b\|^2 \\
&\quad \text{ssi} \quad \forall t \in \mathbb{R} \forall z \in \mathbb{R}^p, \|Ax + tAz - b\|^2 \geqslant \|Ax - b\|^2 \\
&\quad \text{ssi} \quad \forall t \in \mathbb{R} \forall z \in \mathbb{R}^p, t^2 \|Az\|^2 + 2t\langle Az, Ax - b\rangle \geqslant 0 \\
&\quad \text{ssi} \quad \forall z \in \mathbb{R}^p, \begin{cases} \forall t > 0, t \|Az\|^2 + 2\langle Az, Ax - b\rangle \geqslant 0 \\ \forall t < 0, t \|Az\|^2 + 2\langle Az, Ax - b\rangle \leqslant 0 \end{cases} \\
&\quad \text{ssi} \quad \forall z \in \mathbb{R}^p, \langle Az, Ax - b\rangle = 0 \\
&\quad \text{ssi} \quad \forall z \in \mathbb{R}^p, \langle z, A^*Ax - A^*b\rangle = 0 \\
&\quad \text{ssi} \quad A^*Ax = A^*b
\end{aligned}
$$

Remarque 5.2 $\mathrm{Ker}(A) = \mathrm{Ker}(A^*A)$ *donc si* $\mathrm{Ker}(A) = \{0\}$, $A^*A \in GL_p$ *et on peut appliquer les méthodes précédentes. Mais on évite cette méthode car A^*A est moins bien conditionnée que A.*

5.2.3 QR

Par Householder, on décompose A sous la forme QR.
Pour $y \in \mathbb{R}^p$, $\|Ay - b\| = \|QRy - b\| = \|Ry - Q^*b\|$.
Si $\text{Ker}(A) = \{0\}$, on pose $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$ et $Q = (Q_1|Q_2)$ avec $R_1 \in GL_p$.

$$\|Ry - Q^*b\|^2 = \|R_1y - (Q^*b)_1\|^2 + \|(Q^*b)_2\|^2$$

donc minimiser $\|Ry - Q^*b\|$ revient à minimiser $\|R_1y - (Q^*b)_1\|$ donc à choisir $y = R_1^{-1}(Q^*b)_1$.

Si $\text{Ker}(A) \neq \{0\}$, on permute des lignes, des colonnes et on extrait une sous-matrice inversible.

5.3 Méthodes itératives

5.3.1 Méthodes basées sur des décompositions

On pose $A = M - N$ avec M facile à inverser.

Pour résoudre $Ax = B$, on construit une suite $(x^k)_k$ tel que x^0 est quelconque et $\forall k$, $x^{k+1} = M^{-1}(Nx^k + B)$.

Si x converge vers $\overline{x}$, on a $M\overline{x} = N\overline{x} + B$ et $A\overline{x} = B$.

THÉORÈME 5.7 *$\forall x^0$, x converge ssi $\rho(M^{-1}N) < 1$.*

Démonstration. Voir TD ■

Exemples :

- Jacobi, Gauss-Siedel, relaxation : voir TP
- Méthode de Richardson : $M = \frac{1}{\alpha}I_n$ et $N = M - A$.
  Si $\text{Sp}(A) \subset \mathbb{R}_+^*$, il y a convergence ssi $\alpha \in]0, \frac{2}{\rho(A)}[$.
  Si $\text{Sp}(A) \subset \mathbb{R}_-^*$, il y a convergence ssi $-\alpha \in]0, \frac{2}{\rho(A)}[$.
  Sinon, il n'y a pas convergence.

5.3.2 Méthodes variationnelles

On va remplacer le système linéaire par la minimisation d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ tel que $\overrightarrow{\text{grad}}(f) = y \mapsto Ay - b$.

Dans le cas de $A \in S_n^{++}$, $f = y \mapsto \frac{\langle Ay|y\rangle}{2} - \langle b|y\rangle$.

Méthode de gradient à pas fixe

Pour minimiser la fonction, on va définir une suite de vecteurs qui va tendre vers le vecteur qui atteint le minimum :

$$\begin{cases} x^0 \in \mathbb{R}^n \\ \forall k, x^{k+1} = x^k - \mu \overrightarrow{\text{grad}}(f)(x^k) \end{cases}$$

Si f est la fonction précédente, on retrouve la méthode de Richardson avec $M = \frac{1}{\mu} I_n$ et on a convergence ssi $\mu \in]0, \frac{2}{\rho(A)}[$.

Dans ce cas, cherchons quel est le meilleur pas μ.

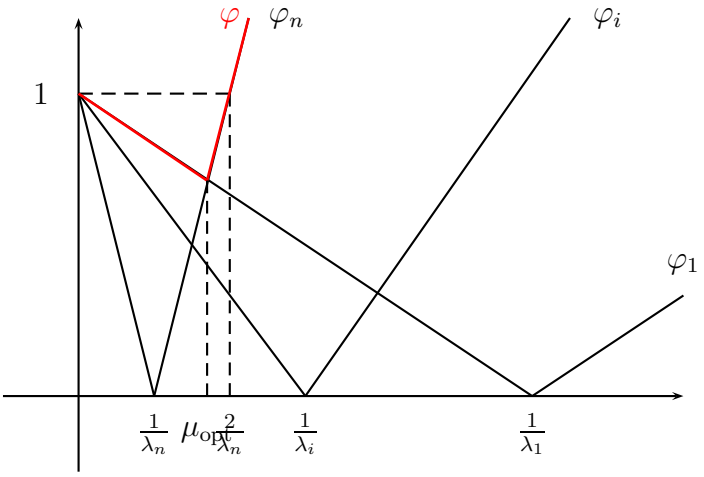
On a $x = A^{-1}b$, $A = P^t D P$. Posons $y^k = Px^k$.

On a $\left\| \underbrace{x^k - x}_{e^k} \right\| = \left\| \underbrace{y^k - Px}_{\varepsilon^k} \right\|$.

De plus, $e^{k+1} = x^{k+1} - x = x^k - \mu(Ax^k - b) - x = e^k - \mu A e^k = (I_n - \mu A)e^k$ donc $\varepsilon^k = Pe^{k+1} = P(I_n - \mu A)e^k = (I_n - \mu D)\varepsilon^k$.

Si on note $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ les valeurs propres de A, le meilleur μ est celui qui minimise $\varphi = \mu \mapsto \max\limits_{1 \leqslant i \leqslant n} |1 - \mu\lambda_i|$.

Notons $\varphi_i = \mu \mapsto |1 - \mu\lambda_i|$.



On a $|1 - \mu_{optimal}\lambda_1| = |1 - \mu_{optimal}\lambda_n|$ donc $\mu_{optimal} = \frac{2}{\lambda_1 + \lambda_n}$.

Méthode de gradient à pas optimal

$$\begin{cases} x^0 \in \mathbb{R}^n \\ \forall k, x^{k+1} = x^k - \mu^k \overrightarrow{\text{grad}}(f)(x^k) \text{ avec} \\ \qquad \mu^k \text{ tel que } f(x^{k+1}) = \min_{\mu \in \mathbb{R}} \underbrace{f(x^k - \mu \overrightarrow{\text{grad}}(f)(x^k))}_{g(\mu)} \end{cases}$$

g est convexe dérivable donc μ^k vérifie $g'(\mu^k) = 0$. Si $\psi = \mu \mapsto x^k - \mu \overrightarrow{\text{grad}}(f)(x^k)$, on a :

$$g'(\mu) = \langle \overrightarrow{\text{grad}}(f)(\psi(\mu)) | \psi'(\mu) \rangle = \langle \overrightarrow{\text{grad}}(f)(\psi(\mu)) | \overrightarrow{\text{grad}}(f)(x^k) \rangle$$

On peut remarquer que, comme $g'(\mu^k) = 0$, $\langle \overrightarrow{\text{grad}}(f)(x^k) | \overrightarrow{\text{grad}}(f)(x^{k+1}) \rangle$, deux directions successives sont équivalentes.

De plus, si $r^k = Ax^k - b$, on a $g'(\mu) = \langle r^k, (I_n - \mu A) r^k \rangle$.

Donc $\mu^k = \frac{\left\| r^k \right\|^2}{\langle r^k | Ar^k \rangle}$.

Ce qui conduit à l'algorithme :

Algorithme 2: Calcul de la suite $(x^k)_k$

Entrées : A, b, $x^0 \in \mathbb{R}^n$ et $r > 0$

Sorties : La suite $(x^k)_{1 \leqslant k \leqslant r}$

```
pour k = 1 à r faire
    r_k := A x^k − b
    μ_k := ‖r_k‖² / ⟨A r_k, r_k⟩
    x^{k+1} := x^k − μ_k r_k
retourner (x^k)_{1 ⩽ k ⩽ r}
```

THÉORÈME 5.8 *Soit $A \in S_n^{++}(\mathbb{R})$.*

La suite des $(x^k)_k$ converge pour tout $x_0 \in \mathbb{R}^n$ vers $\overline{x} = A^{-1}b$.

En notant $e^k = (x^k - \overline{x})$, on a :

$$\frac{\left\| e^{k+1} \right\|_A^2}{\left\| e^k \right\|_A^2} = \frac{\langle Ae^{k+1}, e^{k+1} \rangle}{\langle Ae^k, e^k \rangle} \leqslant \left(\frac{c-1}{c+1} \right)^2$$

avec $c = c_2(A)$.

Remarque 5.3 *$c < 1$ donc $\frac{c-1}{c+1} < 1$ et on a une convergence géométrique.*

De plus, plus c est grand, plus $\frac{c-1}{c+1}$ est proche de 1 donc plus la convergence est lente.

Démonstration. La preuve repose sur le lemme admis suivant :

Lemme 5.8.1 (Inégalité de KANTOROVITCH)
Si $A \in S_n^{++}(\mathbb{R})$, dont les valeurs propres sont $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n$, alors, pour tout $x \in \mathbb{R}^n$,

$$\|x\|^4 \leqslant \langle Ax, x \rangle \langle A^{-1}x, x \rangle \leqslant \left(\sqrt{\frac{\lambda_n}{\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_n}} \right)^2 \frac{\|x\|^4}{4}$$

Soit $k \geqslant 0$.
On a $Ae^k = Ax^k - A\overline{x} = Ax^k - b = r^k = -\frac{x^{k+1}-x^k}{\mu^k}$.
De plus, $\langle Ae^{k+1}, e^{k+1} - e^k\rangle = -\mu^k\langle Ae^{k+1}, r^k\rangle = -\mu^k\langle r^{k+1}, r^k\rangle = 0$.
Donc $\langle Ae^{k+1}, e^{k+1}\rangle = \langle Ae^{k+1}, e^k\rangle$.
D'où :

$$\begin{aligned}\langle Ae^{k+1}, e^{k+1}\rangle &= \langle e^{k+1}, Ae^k\rangle \\ &= \langle e^k - r^k\mu^k, r^k\rangle \\ &= \langle e^k, Ae^k\rangle - \mu^k \left\|r^k\right\|^2\end{aligned}$$

Donc :

$$\frac{\langle Ae^{k+1}, e^{k+1}\rangle}{\langle Ae^k, e^k\rangle} = 1 - \mu^k\frac{\left\|r^k\right\|^2}{\langle e^k, Ae^k\rangle} = 1 - \frac{\left\|r^k\right\|^4}{\langle e^k, Ae^k\rangle\langle Ar^k, r^k\rangle}$$

Or par le lemme, si $c = \frac{\lambda_n}{\lambda_1}$, on a

$$\langle e^k, Ae^k\rangle\langle Ar^k, r^k\rangle = \langle A^{-1}r^k, r^k\rangle\langle Ar^k, r^k\rangle \leqslant \left(\sqrt{c} + \sqrt{\frac{1}{c}}\right)^2 \frac{\left\|r^k\right\|^4}{4}$$

D'où :

$$\frac{\langle Ae^{k+1}, e^{k+1}\rangle}{\langle Ae^k, e^k\rangle} \leqslant 1 - \frac{1}{\left(\sqrt{c} + \sqrt{\frac{1}{c}}\right)^2 \frac{1}{4}} = \frac{c^2 - 2c + 1}{c^2 + 2c + 1} = \left(\frac{c-1}{c+1}\right)^2$$

■

Remarque 5.4 *Dans le cas de la méthode à pas fixe,* $\left\|e^{k+1}\right\| \leqslant f(\mu)\left\|e^k\right\|$ *avec* $f(\mu) = \max\limits_{1\leqslant i\leqslant n}|1 - \lambda_i\mu|$.

On a $\mu_0 = \frac{2}{\lambda_1+\lambda_n}$ *et* $f(\mu_0) = 1 - \mu_0\lambda_1 = \frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}$.

Et on a une convergence géométrique de raison $\frac{c-1}{c+1}$.

Donc l'algorithme du gradient à pas optimal converge plus vite. On s'en doutait un peut.

Méthode du gradient conjugué

Le gradient à pas optimal repose sur les relations de récurrence :

$$\begin{cases} x^0 \in \mathbb{R}^n \\ x^{k+1} = x^k - \mu^k \underbrace{(Ax^k - b)}_{r_k} \\ \\ \mu^k \text{ qui minimise } \mu \mapsto f(x^k - \mu r^k) \end{cases}$$

Plus généralement, on peut définir les algorithmes de descente par :

- On choisit $x_0 \in \mathbb{R}^n$ fixé.
- Pour $k \geqslant 0$, on choisit une direction de descente $d^k \in \mathbb{R}^n$
- on calcule le pas optimal dans cette direction : celui qui minimise $\mu \mapsto f(x^k - \mu d^k)$.

Remarque 5.5 Le choix $d^k = r^k$ est optimal pour des petits pas mais ici on autorise des grands pas donc ce n'est pas forcément la meilleure direction.

Pour obtenir la meilleure direction, on peut choisir d^k en fonction de $x^0, \dots, x^k$.

L'algorithme est le suivant :

- On prend $x^0 \in \mathbb{R}^n$ fixé.
- À la première étape, on pose $d^0 = \overrightarrow{\text{grad}}\, f(x^0)$, μ^0 la pas optimal et $x^1 = x^0 - \mu^0 d^0$.
- À l'étape k, $x^0, \dots, x^k$ sont construits.
  Si $\overrightarrow{\text{grad}}\, f(x^k) = 0$, c'est fini : $x^k = \overline{x}$.
  Sinon, on prend x^{k+1} la valeur qui minimise f sur $x^k + F^k$ avec $F^k = \text{Vect}\left\{\overrightarrow{\text{grad}}\, f(x^0), \dots, \overrightarrow{\text{grad}}\, f(x^k)\right\}$.

En fait, on cherche $x^{k+1} = x^k + \sum_{i=0}^{k} \lambda^i \overrightarrow{\text{grad}}\, f(x^i)$ qui minimise f. On s'est donc ramené à un problème de minimisation dans un espace de dimension $k+1$.

Soit k fixé. Montrons que si $\overrightarrow{\text{grad}}\, f(x^k) \neq 0$, $\dim(F^k) = \dim(F^{k-1}) + 1$ ie que $\overrightarrow{\text{grad}}\, f(x^k) \notin F^k$.

En fait, $\overrightarrow{\text{grad}}\, f(x^k) \in (F^k)^\perp$.

En effet, $x^k \in x^{k-1} + F^k$ donc $x^k = x^{k-1} + \sum_{i=0}^{k-1} \lambda_i \overrightarrow{\text{grad}}\, f(x^i)$.

On pose $g(\lambda_0, \dots, \lambda_{k-1}) = f\left(x^{k-1} + \sum_{i=0}^{k-1} \lambda_i \overrightarrow{\text{grad}}\, f(x^i)\right)$.

On cherche le minimum de g.

La fonction g est quadratique à cœfficients positifs pour le termes de degré 2 donc il existe un minimum caractérisé par $\overrightarrow{\text{grad}}\, g(\lambda_0, \dots, \lambda_{k-1}) = 0$.

Or $\frac{\partial g}{\partial \mu_i}(\lambda_0, \dots, \lambda_{k-1}) = \langle \overrightarrow{\text{grad}}\, f(x^k), \overrightarrow{\text{grad}}\, f(x^i) \rangle$.

Donc

$$\overrightarrow{\text{grad}}\, g(\lambda_0, \dots, \lambda_{k-1}) = 0 \quad \text{ssi} \quad \forall i \leqslant k-1, \langle \overrightarrow{\text{grad}}\, f(x^k), \overrightarrow{\text{grad}}\, f(x^i) \rangle = 0$$

Donc on a bien le résultat.

Finalement, au bout d'au plus $n-1$ itérations, on converge vers $\overline{x} = A^{-1}b$ car $F_{n-1} = \mathbb{R}^n$ si on a pas convergé avant.

Il reste à savoir si on sait calculer en pratique les x^i. C'est possible : on peut écrire cet algorithme sous la forme de méthode de descente avec un calcul simple des d^k.

L'idée est de poser $w^k = x^{k+1} - x^k$. On peut montrer à partir du résultat précédent que $\langle Aw^k, w^l \rangle = 0$ pour $k \neq l$ (on dit que les w^k sont deux à deux conjugués ie orthogonaux pour le produit scalaire associé à A).

Il suffit alors d'utiliser le procédé de Gram-Schmidt pour le produit scalaire associé à A pour trouver les d^k, c'est-à-dire qu'on orthogonalise la base des $(\overrightarrow{\text{grad}}\, f(x^i))_i$.

Algorithme 3: Gradient conjugué

Entrées : Une matrice A de taille n et un vecteur b
Sorties : La solution de $Ax = b$

Prendre un $x_0 \in \mathbb{R}^n$.
$d_0 := Ax_0 - b$
pour $k = 1$ **à** n **faire**
    $\mu_k := \frac{\|r_k\|^2}{\langle Ad_k, d_k \rangle}$
    $x_{k+1} := x_k - \mu_k d_k$
    $r_{k+1} = r_k - \mu_k A d_k$ **si** $r_{k+1} = 0$ **alors**
        Arrêter la boucle et **retourner** x_{k+1}
    sinon
        $d_{k+1} := r_{k+1} - \frac{\|r_{k+1}\|^2}{\|r_k\|^2} d_k$
retourner x_{n+1}

On pourrait croire que c'est une méthode directe. Mais elle nécéssite $2n^3 + O(n^2)$ opérations alors que Choleski fait $\frac{n^3}{3} + O(n^2)$.

Ce n'est donc pas une bonne méthode directe. En revanche, au bout d'un petit nombre d'itérations, on a une bonne approximation de $\overline{x}$.

On montre l'estimation :

$$\frac{\langle Ae^k, e^k \rangle}{\langle Ae^0, e^0 \rangle} \leqslant 4 \left(\frac{\sqrt{c} - 1}{\sqrt{c} + 1} \right)^{2k}$$

avec $c = c_2(A)$.

Chapitre 6

Approximation spectrale

L'équation d'une corde de longueur l qui vibre est donnée par $\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = \frac{\partial^2 U}{\partial x^2}$ avec les conditions au bord $U(0,t) = U(l,t) = 0$ et $c = \sqrt{\frac{\tau}{p}}$ (p : masse linéique, τ : tension)

On cherche $U(x,t)$ sous la forme $u(x)v(t)$ et on trouve l'équation $\frac{u''(x)}{u(x)} = \frac{1}{c^2}\frac{v''(t)}{v(t)} = \lambda \in \mathbb{R}$ constant. On est donc amenés à résoudre un problème aux valeurs propres pour u. En deux dimensions, on peut faire pareil avec une membrane.

Suivant les cas, on peut être intéressés par les vecteurs et valeurs propres, la valeur maximale/minimale, ou bien la valeur propre la plus proche d'un scalaire donné.

On dispose de plusieurs méthodes qu'on va développer après :

- Calculer le polynôme caractéristique et trouver ses racines. Clairement impossible par du calcul exact. Et les méthodes itératives ont un problème de stabilité numérique. En plus en général, on fait l'inverse : on calcule le spectre de la matrice compagnon pour obtenir les racines du polynôme.
- Méthode de la puissance
- Méthode QR

6.1 Conditionnement d'un problème au valeurs propres

On a vu que si $A = PDP^{-1}$ avec D diagonale et si E est une matrice de norme $\varepsilon > 0$, alors pour tout $\lambda_\varepsilon \in \mathrm{Sp}(A+E)$, il existe $\lambda \in \mathrm{Sp}(A)$ tel que $|\lambda - \lambda_\varepsilon| \leqslant c_\infty(P)\varepsilon$.

C'est faux si A n'est pas diagonalisable :

Prenons $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $E = \varepsilon \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Les valeurs propres de $A + E$ tendent vers 1 quand ε tend vers 0, mais à quelle vitesse ?

Soit $\lambda_\varepsilon \in \mathrm{Sp}(A + E)$.

$$\begin{aligned} 0 = \det(A + E - \lambda_\varepsilon I_2) &= \begin{vmatrix} 1 - \lambda_\varepsilon + \varepsilon a & 1 + b\varepsilon \\ c\varepsilon & d\varepsilon \end{vmatrix} \\ &= (1 - \lambda_\varepsilon)^2 + (1 - \lambda_\varepsilon)\varepsilon(a + d) + \varepsilon^2 ad - c\varepsilon - \varepsilon^2 bc \\ &= (1 - \lambda_\varepsilon)^2 - \varepsilon c + o(\varepsilon) \end{aligned}$$

Donc $\lambda_\varepsilon = 1 \pm \sqrt{c\varepsilon} + o(\sqrt{\varepsilon})$.

Dans les blocs de Jordan de taille n, $\lambda_\varepsilon = 1 + \omega_i(\varepsilon a_{1,n})^{\frac{1}{n}} + o(\varepsilon^{\frac{1}{n}})$ avec ω_i les racines n-èmes de l'unité.

Si $n = 4$ est si les cœfficients de A sont connus à 10^{-8} près, λ est calculée à 10^{-2} près.

Conclusion : Plus A possède de blocs de Jordan de grande taille, plus le calcul des valeurs propres sera mal conditionné.

6.2 Méthode de la puissance

Entrées : Une matrice A

Sorties : La valeurs propre de plus grand module de A et un vecteur propre associé

Prendre q^0 de norme 1. **pour** $k \geqslant 1$ **faire**

- $x^k := Aq^{k-1}$
- $q^k := \frac{x^k}{\|x^k\|}$
- $\lambda^k :=$ la moyenne des $\frac{x_j^k}{q_j^{k-1}}$ pour j tel que $q_j^{k-1} \neq 0$.

retourner λ *et* x

6.2.1 Cas diagonalisable

Remarque 6.1

- *Si q^0 n'appartient pas à l'hyperplan engendré par les vecteurs propres associés aux autres valeurs propres, alors il y a convergence.*

Comme q^0 est pris aléatoirement, et que l'hyperplan est de mesure nulle, on a une convergence presque sûre.

- *En général, $(q^k)_k$ ne converge pas car si $\lambda_{\max} = |\lambda_1|e^{i\theta}$, $z^k = q^k e^{-i\theta k}$ tend vers x, $q^k \sim xe^{ik\theta}$ qui oscille quand $\lambda_1 \notin \mathbb{R}^+$.*

THÉORÈME 6.1 *Soit $A \in \Delta_n(\mathbb{K})$ de valeurs propres $\lambda_1, \dots, \lambda_n$. On suppose que A admet une unique valeur propre (éventuellement multiple) de module maximal ie $\lambda_1 = \cdots = \lambda_p$ et $|\lambda_1| > |\lambda_{p+1}| \geqslant \cdots \geqslant |\lambda_n|$.*

Notons $(u^1, \dots, u^n)$ une base de vecteurs propres associés unitaires.

Soit q^0 de norme 1. La suite (q^k) est bien définie et vérifie :

- $\lim_{k\to+\infty} \underbrace{(\frac{\overline{\lambda_1}}{|\lambda_1|})^k q^k}_{z^k} = x$ *avec x un vecteur propre associé) λ_1 et la vitesse de convergence est donnée par $\left\|z^k - x\right\| = O(|\frac{\lambda_{p+1}}{\lambda_1}|^k)$.*
- $\lim_{k\to+\infty} \left\|Aq^k\right\| = |\lambda_1|$ *et pour tout j tel que q_j^k ne tende pas vers 0 quand k tend vers l'infini, on a $\lim_{k\to+\infty} \frac{x_j^{k+1}}{q_j^k} = \lambda_1$.*

Démonstration.

- Cas $p = 1$

  On écrit $q^0 = \sum_{i=1}^n \alpha_i u^i$ avec $\alpha_1 \neq 0$ par hypothèse.

  On a

$$A^k q^0 = \sum_{i=1}^n \alpha_i \lambda_i^k u^i = \alpha_1 \lambda_1^k (u^1 + v^k) \text{ avec } v_k = \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} (\frac{\lambda_i}{\lambda_1})^k u^k$$

  En particulier, $A^k q^0 \neq 0$ donc $q^k = \frac{A^k q^0}{\left\|A^k q^0\right\|}$ est bien défini pour tout k et on a :

$$q^k = \frac{\alpha_1}{|\alpha_1|} \frac{\lambda_1^k}{|\lambda_1|^k} \frac{u^1 + v^k}{\|u^1 + v^k\|}$$

  Par hypothèse, $|\lambda_1| > |\lambda_i|$ donc v^k tend vers 0 avec $\left\|v^k\right\| = O(|\frac{\lambda_2}{\lambda_1}|^k)$.
  Notons $\varepsilon_k = \left\|u^1 + v^k\right\| - \|u^1\| == \left\|u^1 + v^k\right\| - 1$. On a :

$$|\varepsilon_k| \leqslant \left\|v^k\right\| = O(|\frac{\lambda_2}{\lambda_1}|^k)$$

  On a donc :

$$z^k = \frac{\overline{\lambda_1^k}}{|\lambda_1|^k} q^k = \frac{\alpha_1}{|\alpha_1|} \frac{u^1 + v^k}{\|u^1 + v^k\|} \to \frac{\alpha_1}{|\alpha_1|} u^1$$

Posons $x = \frac{\alpha_1}{|\alpha_1|}u^1$. $\|x\| = 1$ et x est un vecteur propre associé à λ_1.
On a aussi :

$$\begin{aligned}
\left\|z^k - x\right\| &= \left\|\frac{\alpha_1}{|\alpha_1|}\left(\frac{u^1 + v^k}{\|u^1 + v^k\|} - u^1\right)\right\| \\
&= \frac{\left\|u^1 + v^k - (1-\varepsilon_k)u^1\right\|}{\|u^1 + v^k\|} \\
&= \frac{\left\|v^k - \varepsilon_k u^1\right\|}{1+\varepsilon_k} \\
&\leqslant \frac{\left\|v^k\right\| + \varepsilon_k}{1+\varepsilon_k} \\
&\leqslant \frac{2\left\|v^k\right\|}{1+\varepsilon_k} = O(|\frac{\lambda_2}{\lambda_1}|^k)
\end{aligned}$$

D'où le premier point.
On remarque de plus que :

$$\left\|Aq^k\right\| = \frac{\left\|A^{k+1}q^0\right\|}{\|A^k q^0\|} = \frac{|\alpha_1||\lambda_1|^{k+1}(1+\varepsilon_{k+1})}{|\alpha_1||\lambda_1|^k(1+\varepsilon_k)} \to |\lambda_1|$$

De plus, si q_j^k ne tend pas vers 0, $u_j^1 \neq 0$ et on a :

$$\frac{x_j^{k+1}}{q_j^k} = \frac{(A^{k+1}q^0)_j}{(A^k q^0)_j} = \frac{\alpha_1\lambda_1^{k+1}(u_j^1 + v_j^{k+1})}{\alpha_1\lambda_1^k(u_j^1 + v_j^k)} \to \lambda_1$$

- Cas p quelconque.

  On a comme dans le premier cas : $A^k q^0 = \sum_{i=1}^{n}\alpha_i\lambda_i^k u^i = \lambda_1^k(u + v^k)$ avec $u = \sum_{i=1}^{p}\alpha_i u^i \neq 0$ et $v^k = \sum_{i=p+1}^{n}\alpha_i\left(\frac{\lambda_i}{\lambda_1}\right)^k u^i$. Comme précédemment, $\left\|v^k\right\| = O(|\frac{\lambda_{p+1}}{\lambda_1}|^k)$ et $\varepsilon_k = \left\|u + v^k\right\| - \|u\| = O(|\frac{\lambda_{p+1}}{\lambda_1}|^k)$.
  On en déduit que :

$$z^k = \frac{\overline{\lambda_1^k}}{|\lambda_1|^k}q^k = \frac{u + v^k}{\|u + v^k\|} \to \frac{u}{\|u\|} = x$$

  avec x un vecteur propre associé à λ_1.

On a de plus l'estimation :

$$\begin{aligned}\left\|z^k - x\right\| &= \frac{\left\|u + v^k - \left\|u + v^k\right\| x\right\|}{\|u + v^k\|} \\ &= \frac{\left\|v^k - \varepsilon_k x\right\|}{\|u\| + \varepsilon_k} \\ &\leqslant \frac{2\|u\|}{\|u\| + \varepsilon_k} \\ &= O\left(\left|\frac{\lambda_{p+1}}{\lambda_1}\right|^k\right)\end{aligned}$$

Le deuxième point se démontre comme précédemment. ■

6.2.2 Cas non diagonalisable

THÉORÈME 6.2 *Supposons que A ait une unique valeur propre de plus grand module (éventuellement multiple) de module maximal et soit ρ_2 le second plus grand module des valeurs propres. Soit $(u^1, \dots, u^n)$ une base de Jordan pour A avec u^1 un vecteur propre associé à λ_1.*

Si $q^0 \notin \text{Vect}\{u^2, \dots, u^n\}$, alors les suites q et z définies précédemment sont bien définies et $\lim_{k\to+\infty} z^k = x$ avecx un vecteur propre unitaire associé à λ_1 avec une vitesse en :

- *$O(\frac{1}{k})$ si λ_1 est défective.*
- *$O(k^{r-1}(\frac{\rho_2}{|\lambda_1|})^k)$ si λ_1 est non défective avec r la taille du plus grand bloc de Jordan associé aux valeurs propres de module ρ_2.*

On a aussi $\lim_{k\to+\infty} \left\|Aq^k\right\| = |\lambda_1|$ et pour tout j tel que q_j^k ne tende pas vers 0, on a :

$$\lim_{k\to+\infty} \frac{x_j^{k+1}}{q_j^k} = \lambda_1$$

Démonstration.

- Si λ_1 est non défective :

$$J = P^{-1}AP = \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \cdots & 0 \\ 0 & J_{r_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{r_p}(\lambda_p) \end{pmatrix}$$

avec $J_{r_i}(\lambda_i)$ le bloc de Jordan de taille r_i associé à la valeur propre λ_i.

J^k est digonale par blocs et on a :

$$\frac{J^k}{\lambda_1^k} = \operatorname{diag}\left(I_{r_1}, \frac{J_{r_2}^k(\lambda_2)}{\lambda_1^k}, \dots, \frac{J_{r_p}^k(\lambda_p)}{\lambda_1^k}\right)$$

où les blocs diagonaux sont, pour $k \geqslant r_i$:

$$\frac{J_{r_i}^k(\lambda_i)}{\lambda_1^k} = \begin{pmatrix} \left(\frac{\lambda_i}{\lambda_1}\right)^k & k\frac{\lambda_i^{k-1}}{\lambda_1^k} & \cdots & \binom{k}{r_i-1}\frac{\lambda_i^{k-r_i+1}}{\lambda_1^k} \\ 0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \left(\frac{\lambda_i}{\lambda_1}\right)^k \end{pmatrix}$$

Dans chaque bloc, c'est le terme en position $(1, r_i)$ qui est prépondérant quand $k \to +\infty$.
Donc $\frac{J^k}{\lambda_1^k} \to \begin{pmatrix} I_{r_1} & 0 \\ 0 & 0 \end{pmatrix}$ et le terme qui tend le plus lentement vers 0 est de la forme $\binom{k}{r_2-1}\frac{\lambda_2^{k-r_2+1}}{\lambda_1^k}$ avec λ_2 tel que $\rho_2 = |\lambda_2|$ et r_2 la taille du plus grand bloc de Jordan associé aux valeurs propres de module ρ_2.
Ainsi, si $P^{-1}q_0 \notin \operatorname{Vect}\{u^{r_1+1}, \dots, u^n\}$ et si les suites q et z sont définies comme précédemment, on a $z^k \to x = \frac{A_\infty q^0}{\|A_\infty q^0\|}$ avec $A_\infty = P\begin{pmatrix} I_{r_1} & 0 \\ 0 & 0 \end{pmatrix}P^{-1}$ et x un vecteur propre associé à λ_1.
De plus, la vitesse de convergence est donnée par :

$$\left\|z^k - x\right\| = O\left(k^{r_2-1}\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

- Si λ_1 est défective :
  Soit $J_r(\lambda_1)$ un bloc de Jordan associé à λ_1 et $(u^1, \dots, u^r)$ une base associée. On a, pour $i \in [\![1, r]\!]$:

$$\begin{aligned} J^k u^i &= \binom{k}{i-1}\lambda_1^{k-i+1}u^1 + \binom{k}{i-2}\lambda_1^{k-i+1}u^2 + \cdots + \lambda_1^k u^i \\ &= \binom{k}{i-1}\lambda_1^{k-i+1}\left(u^1 + \frac{i-1}{k-i+1}\lambda_1 u^2 + o\left(\frac{1}{k}\right)\right) \end{aligned}$$

  Quand $k \to +\infty$, $\frac{\overline{\lambda_1^k}}{|\lambda_1|^k}\frac{J^k u^1}{\left\|J^k u^1\right\|} \to u^1$ avec une vitesse en $O(\frac{1}{k})$. De même, pour les autres vecteurs de base associés aux autres blocs de Jordan, on a comme précédemment, $J^k u^i \to 0$ avec une vitesse en : $O(k^{r_2-1}|\frac{\lambda_2}{\lambda_1}|^k)$.
  ■

Remarque 6.2 La convergence est fortement dégradée, surtout si λ_1 est défective.

6.2.3 Méthode de la puissance inverse

Si A est inversible, on cherche la valeur propre de plus petit module.
On applique la méthode de la puissance à A^{-1}.
En pratique, on ne calcule pas A^{-1}, mais on applique l'algorithme suivant.

Entrées : Une matrice A
Sorties : Sa valeur propre de plus petit module et un vecteur propre associé

1 Prendre q^0 aléatoire **pour** $k \geqslant 1$ **faire**
2 Résoudre $Ax^k = q^{k-1}$
3 $q^k := \frac{x^k}{\|x^k\|}$
4 Prendre λ^k la moyenne des $\frac{q_j^{k-1}}{x_j^k}$ avec j tel que $x_j^k \neq 0$.
5 **retourner** λ^k, x^k

Pour la résolution des systèmes linéaires, on utilise une décomposition PLU de A.

Variante : On peut calculer la valeur propre la plus proche de $\mu \in \mathbb{C}$ donné. Il suffit d'appliquer la méthode de la puissance inverse à $A - \mu I_n$.

6.3 La méthode QR

Remarque 6.3 C'est une méthode très performante, y compris pour calculer les racines d'un polynôme via la matrice compagnon.

On voudrait un algorithme qui trigonalise A.

Par l'algortihme d'Householder, on peut décomposer A sous la forme $A = QR$. Mais R et A n'ont pas les même valeurs propres.

On a $A = QRQQ^T$. Posons $A_2 = RQ$. On peut réitérer : $A_2 = Q_2R_2 = Q_2R_2Q_2Q_2^T$ donc on pose $A_3 = R_2Q_2$,...

6.3.1 Première stratégie (Jacobi, 1846)

THÉORÈME 6.3 *Soit $A \in \Delta_n(\mathbb{R})$. Notons $\mathrm{Sp}(A) = \{\lambda_1, \ldots, \lambda_n\}$ avec $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$.*

Si $A = PDP^{-1}$ avec P admettant une décomposition PLU, alors :

- *La suite (A_k) vérifie $((A_k)_{i,i})_k$ converge vers λ_i pour tout i.*

- *La suite* $((A_k)_{i,j})_k$ *tend vers* 0 *si* $i > j$.

Remarque 6.4 Il existe des améliorations qui garantissent la convergence dans des cas plus généraux.

Démonstration.

- On a $A_{k+1} = R_k Q_k = Q_k^T A_k Q_k = \Omega_k^T A \Omega_k$ avec $\Omega_k = Q_1 \cdots Q_k$.
  On a $A = QR$ donc $A^k = (Q_1 R_1)k = Q_1 (R_1 Q_1)^{k-1} R_1 = Q_1 A_2^{k-1} R_1$.
  Par récurrence, on a $A^k = (Q_1 \cdots Q_k)(R_k \cdots R_1)$.
- On a, d'autre part $A^k = P D^k P^{-1}$ et $P = \widetilde{Q}\widetilde{R}$ et $P^{-1} = LU$.
  Donc :

$$\begin{aligned} A^k &= \widetilde{Q}\widetilde{R}D^k LU \\ &= \widetilde{Q}\,\underbrace{\widetilde{R}D^k L D^{-k}\widetilde{R}^{-1}}_{=\mathcal{Q}_k\mathcal{R}_k}\,\widetilde{R}D^k U \\ &= \underbrace{\widetilde{Q}\mathcal{Q}_k}_{\in O_n}\ \underbrace{\mathcal{R}_k\widetilde{R}D^k U}_{\in T_n^+ \text{ de diagonale } \Lambda_k} \end{aligned}$$

  On a $A^k = \underbrace{\widetilde{Q}\mathcal{Q}_k\widetilde{\Lambda}_k}_{\in O_n}\underbrace{\widetilde{\Lambda}_k^{-1}\mathcal{R}_k\widetilde{R}D^k U}_{\in T_n^{++}}$ avec $\widetilde{\Lambda}_k = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ où $\alpha_i = \frac{\lambda_i^k}{|\lambda_i^k|}$ (λ_i^k sont les cœfficients diagonaux de Λ_k).
- Par unicité de la décomposition QR, $\Omega_k = \widetilde{Q}\mathcal{Q}_k\widetilde{\Lambda}_k$ et $R_k \cdots R_1 = \widetilde{\Lambda}_k^{-1}\mathcal{R}_k\widetilde{R}D^k U$.
  Montrons que $\mathcal{Q}_k \to I_n$.
  On va montrer que $\widetilde{R}D^k L D^{-k}\widetilde{R}^{-1} \to I_n$ ie $D^k L D^{-k} \to I_n$.
  $D^k L D^{-k}$ est triangulaire inférieure avec des 1 sur la diagonale. Ses cœfficients non diagonaux non nuls $m_{i,j}^k$ valent $(\frac{\lambda_i}{\lambda_j})^k L_{i,j}$ qui tend bien vers 0 car $|\lambda_i| < |\lambda_j|$ pour tout $i > j$.
  De plus, $\mathcal{Q}_k$ est une suite de O_n donc elle admet une sous-suite convergeant vers $\mathcal{Q} \in O_n$.
  $\mathcal{Q}_{\varphi(k)} \to \mathcal{Q}$ et $\mathcal{Q}_{\varphi(k)}\mathcal{R}_{\varphi(k)} \to I_n$ donc $\mathcal{R}_{\varphi(k)} \to \mathcal{Q}^T$.
  $\mathcal{Q}^T$ est orthogonale et triangulaire donc $\mathcal{Q} = I_n$.
- On a donc $\Omega_k\widetilde{\Lambda}_k^{-1} \to \widetilde{Q}$.
  On a $A'_{k+1} = \widetilde{\Lambda}_k A_{k+1}\widetilde{\Lambda}_k^{-1} = \widetilde{\Lambda}_k\Omega_k^T A\Omega_k\widetilde{\Lambda}_k^{-1} \to \widetilde{Q}^T A\widetilde{Q}$.
  Donc $A'_{k+1} \to \widetilde{Q}^T P D P^{-1}\widetilde{Q} = \widetilde{R}D\widetilde{R}^{-1}$ qui est triangulaire supérieure dont les cœfficients diagonaux sonc les λ_i.
- $A'_{k+1} = (a'_{i,j})_{i,j}$ avec $a'_{i,j} = \frac{\alpha_i}{\alpha_j}a_{i,j}$ et $|\alpha_i| = |\alpha_j| = 1$.
  Donc $|a'_{i,j}| = |a_{i,j}|$.
  Donc $(A_{k+1})_{i,i} = a'_{i,i} \to \lambda_i$ et $|(A_{k+1})_{i,j}| \to 0$ si $i > j$. ∎

Remarque 6.5 *En notant* $r = \max_{i>j} |\frac{\lambda_i}{\lambda_j}| = \max_{1\leqslant i\leqslant n-1} |\frac{\lambda_{i+1}}{\lambda_i}|$, *on a une convergence en* $O(r^k)$.

6.3.2 Deuxième stratégie

On modifie l'algorithme de Householder : on prend H_1 matrice de Householder telle que $H_1 A H_1^T = \begin{pmatrix} a_{1,1} & * & \cdots & * \\ \alpha & \vdots & & \vdots \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}$

En itérant, on construit une suite de matrices $H_1, \ldots, H_{n-2}$ telle que $Q = H_{n-2} \cdots H_1$ vérifie :

$$QAQ^T = \begin{pmatrix} * & \cdots & \cdots & \cdots & * \\ * & * & \cdots & \cdots & * \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * & * \end{pmatrix}$$

En particulier, si A est symétrique, QAQ^T aussi sont QAQ^T est tridiagonale symétrique et on peut calculer ses valeurs propres en utilisant les suites de Sturm.

