§1 : Fonctions de matrices

Dans des nombreux applications, on est amené à evaluer $f(A)$ ou $f(A)b$ avec $A \in \mathbb{C}^{n\times n}$ de spectre $\sigma(A)$, et $b \in \mathbb{R}^n$. Commençons d'abord à définir proprement cet objet, avant de voir des applications.

§1.1. Définitions et applications

1.1.1. Rappel : forme de Jordan

Toute $A \in \mathbb{C}^{n\times n}$ est similaire à une matrice sous forme de Jordan : $\exists Z \in \mathbb{C}^{n\times n}$ inversible de sorte que $Z^{-1}\cdot A \cdot Z \overset{J=}{=} \mathrm{diag}(J_1, .., J_p)$ diag. par blocs,

avec $J_k = J_{m_k}(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k}$

dit block de Jordan et $m_1 + .. + m_p = n$. Donc $\sigma(A) = \{\lambda_1, \lambda_2, .., \lambda_p\}$ (pas forcément distincts). On appelera indice $m(\lambda)$ d'une valeur propre la taille du plus grand bloc de Jordan associé à λ.

Le nombre de blocs et leur taille est un invariant de A.

Cas particulier si l'indice de toutes les valeurs propres vaut 1 ($m_1 = \dots = m_n = 1$) on dit que A est diagonalisable, ici les colonnes de Z sont des vecteurs propres associés.

cas particulier Si A est hermitienne ($A = A^*$) ou normale ($AA^* = A^*A$) alors elle est diagonalisable et on peut choisir Z unitaire $Z^*Z = I_n$ (base orthonormée de vecteurs propres).

1.1.2. Définition d'une fonction d'une matrice

Soit f une fonction défini sur le spectre de A, c'est à dire $\forall \lambda \in \sigma(A)$ d'indice $m^{(\lambda)}$ on connaît $f^{(j)}(\lambda)$ pour $j = 0, 1, \dots, m^{(\lambda)}-1$. Alors

$$f(A) = Z \cdot \mathrm{diag}\left(f(J_1), \dots, f(J_p)\right) \cdot Z^{-1}$$

où on défini pour un block de Jordan

$$f(J_m(\lambda)) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(m-1)}(\lambda)}{(m-1)!} \\ & f(\lambda) & f'(\lambda) & \vdots \\ & & \ddots & f'(\lambda) \\ & & & f(\lambda) \end{bmatrix}. \quad \square$$

On remarque vite que $f(A)$ ne dépend pas de la forme particulière de la décomposition de Jordan choisie (sinon, voir 1.1.4 b). En particulier, si A est diagonalisable $Z^{-1}AZ = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, alors $f(A) = Z \cdot \mathrm{diag}(f(\lambda_1), \dots f(\lambda_n)) Z^{-1}$.

1.1.3. Lemme : calcul élémentaire

(a) $(f+g)(A) = f(A) + g(A)$

(b) $(f * g)(A) = f(A) * g(A) = g(A) * f(A)$

(c) Pour $f(z) = \alpha \in \mathbb{C}$: $f(A) = \alpha \cdot I_n$.

(d) Soit $f(z) = \dfrac{1}{\alpha - z}$ pour un $\alpha \notin \sigma(A)$.

Alors $f(A) = (\alpha I_n - A)^{-1}$ l'inverse, dite <u>resolvante</u>.

<u>Preuve</u> : <u>de (a), (c)</u> trivial

<u>de (b)</u> Il suffit de montrer que $(f*g)(J_m(\lambda)) = f(J_m(\lambda)) * g(J_m(\lambda))$, découle de la formule de Leibnitz : $\dfrac{(fg)^{(k)}(\lambda)}{k!} = \sum_{j=0}^{k} \dfrac{f^{(j)}(\lambda)}{j!} \cdot \dfrac{g^{(k-j)}(\lambda)}{(k-j)!}$

<u>de (d)</u> avec $g(z) = \alpha - z$, f et g sont définis sur $\sigma(A)$ et $(f*g)(z) = 1$, et alors d'après (a), (b) et (c)

$f(A) * g(A) = I = f(A) \cdot (\alpha I_n - A)$. ■

Le même argument montre que $(1/f)(A) = f(A)^{-1}$.

Il découle de 1.1.3 que

$A^0 = I$, et pour un entier $p>0$: $A^p = A * A^{p-1} = \underbrace{A * \dots * A}_{p \text{ fois}}$.

Donc pour une fonction rationnelle avec poles $\notin \sigma(A)$

$$r(z) = \frac{a_0 + a_1 z + \dots + a_m z^m}{b_0 + b_1 z + \dots + b_n z^n} = c \, \frac{(z-x_1) \cdot \dots \cdot (z-x_m)}{(z-y_1) \cdot \dots \cdot (z-y_n)}$$

$$r(A) = (a_0 I + a_1 A + \dots + a_m A^m)(b_0 I + b_1 A + \dots + b_n A^n)^{-1}$$
$$= c \, (A - x_1 I) \cdot \dots \cdot (A - x_m I)(A - y_1 I)^{-1} \cdot \dots \cdot (A - y_n I)^{-1}$$

où on remarque que deux termes permutent.

1.1.4. Exo :

ⓐ Pour le polynôme charactéristique $\chi(\lambda) = \det(\lambda I - A)$, montrer que $\chi(A) = 0$.

ⓑ M.q. il existe un unique polynôme unitaire ψ de degré minimum (dit polynôme minimal) de sorte que $\psi(A) = 0$.
Verifier la formule $\psi(z) = \prod_{j=1}^{s} (z - \lambda_{k_j})^{m(\lambda_{k_j})}$
avec $\lambda_{k_1}, \dots, \lambda_{k_s}$ les valeurs propres distincts de A, et m_{k_j} l'indice de λ_{k_j}. □

Il est rassurant de savoir qu'il suffit de savoir evaluer $f(A)$ pour f polynôme.

1.1.5 Corollaire

Si le polynôme p interpole f sur $\sigma(A)$ (au sens d'Hermite)

$$\forall \lambda \in \sigma(A) \quad \forall j = 0, \dots, m^{(\lambda)}-1 : f^{(j)}(\lambda) = p^{(j)}(\lambda),$$ avec $m(\lambda)$ indice de λ

alors $f(A) = p(A)$.

Il existe un unique tel p avec $\deg p < \deg \psi$ avec ψ polynôme minimal de A, dit polynôme d'interpolation de (f, A).

Preuve existence/unicité : M206, matrices de Vandermonde généralisées □
unicité découle aussi de 1.1.4ⓑ
[Stoer, Burlisch]

1.1.6. Exercices (en utilisant 1.1.5)

(a) M.q. $f(A^*) = f(A)^*$ si $f(\bar{z}) = \overline{f(z)}$

(b) M.q. $f(X \cdot A \cdot X^{-1}) = X \cdot f(A) \cdot X^{-1}$

(c) Si X permute avec A alors aussi avec $f(A)$.

(d) $B \cdot f(A \cdot B) = f(B \cdot A) \cdot B$ (A, B carrés)

(d) $f(\operatorname{diag}(A,B)) = \operatorname{diag}(f(A), f(B))$.

1.1.7. Exercice La matrice DFT d'ordre n est définie par $F_n = \frac{1}{\sqrt{n}} \left(\exp\left(-2\pi i \frac{jk}{n}\right)\right)_{j,k=0,1,\dots,n-1}$.
Vérifier que F_n est unitaire, symétrique complexe et que $F_n^4 = \mathbb{I}$. En déduire $\exp(\pi \cdot F_n)$.

1.1.8. Exercice :
Vérifier que la matrice triangulaire par blocs est diagonalisable par blocs sous la forme

$$M = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I & +X \\ 0 & I \end{bmatrix}$$

ssi X est solution de l'équation de Sylvester

$$AX - XB = C.$$

Dans ce cas, montrer que $f(M) = \begin{bmatrix} f(A) & f(A)X - Xf(B) \\ 0 & f(B) \end{bmatrix}$.

1.1.9. Exercice :
Soit M une matrice bidiagonale non dégénérée $M = \begin{bmatrix} \lambda_1 & d_1 & & \\ & \ddots & \ddots & \\ & & \ddots & d_{n-1} \\ & & & \lambda_n \end{bmatrix}$ avec $d_j \neq 0$,
et $D = \operatorname{diag}(1, d_1, d_1 d_2, \dots, d_1 d_2 \cdots d_{n-1})$.

M.q. $f(M) = D^{-1} \cdot \begin{bmatrix} f[\lambda_1] & f[\lambda_1,\lambda_2] & \cdots & f[\lambda_1,\dots,\lambda_n] \\ & f[\lambda_2] & \ddots & \vdots \\ & & \ddots & f[\lambda_{n-1},\lambda_n] \\ & & & f[\lambda_n] \end{bmatrix} D$

avec $f[\lambda_i,\dots,\lambda_j]$ différences divisées.

1.1.6.)

(e) Soient $A, B \in \mathbb{C}^{n \times n}$, A inversible.
M.q. $A \cdot B$ et $B \cdot A$ admettent les mêmes blocs de Jordan. En deduire que
$A \cdot f(BA) = f(AB) \cdot A$ (d'abord pour un polynôme).

(f) Soit g définie sur le spectre de A, et supposons que $f^{(j)}(g(\lambda))$ existe $\forall \lambda \in \sigma(A)$ $\forall j = 0, \dots, m(\lambda)-1$.
Montrer que $(f \circ g)(A) = f(g(A))$.
Etape 1 : f polynôme, Etape 2 comparer blocs de Jordan de A et de $g(A)$, construire polynôme pour f.
facile dans [Etape 2]

1.1.10 Exercice: Pour $X \in \mathbb{C}^{n\times r}$, $Y \in \mathbb{C}^{r,n}$, $\alpha \in \mathbb{C}$, YX de rang r
utiliser l'identité

$$\begin{bmatrix} \alpha I_n & X \\ 0 & \alpha I_r + YX \end{bmatrix} \begin{bmatrix} I_n & 0 \\ Y & I_r \end{bmatrix} =$$

$$\begin{bmatrix} I_n & 0 \\ Y & I_r \end{bmatrix} \cdot \begin{bmatrix} \alpha I_n + XY & X \\ 0 & \alpha I_r \end{bmatrix}$$

$$f(\alpha I_n + XY) = [I_n, 0] \cdot f\left(\begin{bmatrix} \alpha I_n & X \\ 0 & (\alpha I_r + YX) \end{bmatrix} \right) \begin{bmatrix} I_n \\ Y \end{bmatrix}$$

$$Z = -X(YX)^{-1} \Leftarrow \alpha I \cdot Z - Z(\alpha I_r + YX) = X$$

$$= [I_n, 0] \begin{bmatrix} f(\alpha) I & f(\alpha) Z - Z f(\alpha I_r + YX) \\ 0 & f(\alpha I_r + YX) \end{bmatrix} \begin{bmatrix} I_n \\ Y \end{bmatrix}$$

$$= f(\alpha) \cdot I_n + X \cdot (YX)^{-1} \cdot [f(\alpha I_r + YX) - f(\alpha) I_r] Y$$

$$\alpha = 1,\ f(z) = \frac{1}{z} : \quad (I + XY)^{-1} = I + X \underbrace{(YX)^{-1}[(I+YX)^{-1} - I]}_{= -(I+YX)^{-1}} Y$$

∎

Pour terminer, donnons une dernière formule pour $f(A)$ basée sur la formule de Cauchy (également généralisable pour des opérateurs sur les espaces de Hilbert)

1.1.11. Thm (formule de Cauchy)
Soit f analytique dans un ouvert Ω et $\Gamma \subset \Omega$ un ensemble de courbes de Jordan encerclant tout $\lambda \in \sigma(A)$ une fois au sens positif. Alors

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(\xi)\,(\xi I - A)^{-1}\, d\xi.$$

Preuve: Avec $A = Z \cdot \mathrm{diag}(J_k) \cdot Z^{-1}$ nous avons

$$\frac{1}{2\pi i} \int_\Gamma f(\xi)(\xi I - A)^{-1} d\xi = Z\, \mathrm{diag}\left(\frac{1}{2\pi i} \int_\Gamma f(\xi)\, (\xi I - J_k)^{-1} d\xi \right) Z^{-1}$$

il suffit alors de montrer notre Thm. pour A un bloc de Jordan.

Notons que d'après 1.1.3 (d) et def 1.1.2. pour $\zeta \notin \sigma(A)$

$$(\zeta I - J_m(\lambda))^{-1} = \begin{bmatrix} g(\lambda) & g'(\lambda) & \cdots & \frac{g^{(m-1)}(\lambda)}{(m-1)!} \\ & \ddots & & \\ & & & g(\lambda) \end{bmatrix}$$

avec $g_\zeta(z) = \frac{1}{\zeta - z}$, $\frac{g_\zeta^{(j)}(\lambda)}{j!} = \frac{1}{(\zeta - z)^{j+1}}$

en comparant élément par élément on trouve bien

$$\frac{1}{2\pi i}\int_\Gamma f(\zeta)\cdot \frac{g_\zeta^{(j)}(\lambda)}{j!}\, d\zeta = \frac{1}{2\pi i j!}\int_\Gamma \frac{f(\zeta)}{(\zeta-\lambda)^{j+1}}\, d\zeta$$

$$\overset{\text{Cauchy}}{=} \frac{f^{(j)}(\lambda)}{j!}\ ,$$ comme désiré ! ▯

1.1.12. Corollaire :

Si $f(z) = \sum_{j=0}^{\infty} a_j z^j$ admet un rayon de convergence $R > \max\{|\lambda| : \lambda \in \sigma(A)\} = \rho(A)$ rayon spectral de A, alors $f(A) = \sum_{j=0}^{\infty} a_j A^j$ (convergence en norme, $\| f(A) - \sum_{j=0}^{n} a_j A^j \| \xrightarrow[n\to\infty]{} 0$.)

Preuve : choisir dans 1.1.11. $\Omega = \{z \in \mathbb{C} : |z| < R\}$
$\Gamma = \{z \in \mathbb{C} : |z| = \frac{R + \rho(A)}{2}\}$ et échanger somme et intégrale par le fait que la somme converge uniformément sur Γ (et la résolvente est bornée sur Γ). ▯

En remplaçant A par $A - z_0 I$, on peut également considérer des résultats similaires pour des développements autour d'un $z_0 \neq 0$.

L'approche 1.1.12 nous donne une manière "facile" de calculer $\exp(A)$, $\cos(A)$, $\cos(\sqrt{A})$, ... mais voir [Moler, Van Loan, 438 (2003)].

Si on veut définir $\log(A)$, $\sqrt{A}$ etc, il faudra choisir d'abord dans 1.1.11 l'ensemble Ω correct, pour définir correctement la racine.

Généralement $\Omega = \mathbb{C} \setminus (-\infty, 0]$ et donc il faut exclure des matrices avec ~~des~~ valeurs propres < 0. Aussi, choix de Γ ?

Encore 6 :

Pourtant, aucune méthode exposée en §1.1. est conseillée comme boîte noire de calcul de fonctions de matrices

* une forme de Jordan n'est pas stable sous perturbations (la taille des blocs change !)

* evaluer P(A) avec P polynôme d'interpolation de (f, A) pour A diagonalisable ? Revient à la déf de départ :

$$A = Z \cdot \mathrm{diag}(\lambda_1, .., \lambda_n)\, Z^{-1}, \quad Z = (y_1 | \cdots | y_n), \quad Z^{-1} = \begin{pmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_n \end{pmatrix}$$

Z^{-1} : v.p. à droite et à gauche

vecteurs propres $\Rightarrow P(z) = \sum f(\lambda_j) \cdot l_j(z)$

$$l_j(z) = \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{z - \lambda_k}{\lambda_j - \lambda_k} \Rightarrow l_j(\mathrm{diag}(\lambda_1, .., \lambda_n))$$

$$= \mathrm{diag}(0, 0, .., 0, 1, 0, .., 0)$$

$j^{ième}$

$$\Rightarrow f(A) = \sum f(\lambda_j) \cdot y_j \cdot \tilde{y}_j^T$$

* par formule de Cauchy, intégration numérique ? Comment choisir Γ ???