

Compléments sur les valeurs propres

Miguel Rodrigues

La partie la plus classique discutée ici peut être trouvée dans les livres classiques d'algèbre linéaire [Ser01], d'analyse numérique matricielle [AK02, Cia07, CMT01] ou d'analyse numérique en général [Fil13, Sch04]. La partie consacrée au calcul fonctionnel se trouve essentiellement dans [LB03, Section 4.3], ou, en anglais, dans les deux premiers chapitres de [Kat95] également parus séparément comme [Kat82]. Pour des aspects plus algébriques, on pourra consulter [Mne06].

Ces notes de cours sont consacrées au spectre des matrices carrées sur le corps des complexes \mathbf{C} . Les résultats les plus élémentaires sont essentiellement supposés acquis. L'accent est mis sur l'utilisation du spectre d'une part, et la stabilité du spectre par rapport aux perturbations sur la matrice d'autre part.

Pour les matrices à coefficients réels, on pourrait certes appliquer les définitions générales associés au corps des réels. Mais ce n'est pas la notion utile d'un point de vue pratique. Quant au spectre, nous les verrons donc toujours comme des matrices à coefficients complexes.

1 Généralités

Définition 1 Soit A une matrice carrée de taille $d \in \mathbf{N}^*$.

1. On appelle spectre de A l'ensemble, noté $\sigma(A)$ ou $sp(A)$, des scalaires $\lambda \in \mathbf{C}$ tels que $\lambda I_d - A$ ne soit pas inversible. Son complémentaire $\mathbf{C} \setminus \sigma(A)$ est appelé ensemble résolvant¹ et pour tout $\lambda \notin \sigma(A)$, $(\lambda I_d - A)^{-1}$ est appelée matrice résolvante.
2. On appelle valeur propre de A tout scalaire $\lambda \in \mathbf{C}$ tel que $\lambda I_d - A$ ne soit pas injectif. Si λ est une valeur propre, alors il existe $x \in \mathbf{C}^d \setminus \{0_d\}$ tel que $Ax = \lambda x$. Un tel vecteur est appelé vecteur propre associé à la valeur propre λ . Un tel couple (λ, x) est appelé mode propre. On appelle espace propre associé à la valeur propre λ le sous-espace

$$\ker(\lambda I_d - A).$$

On appelle espace caractéristique associé à la valeur propre λ le sous-espace

$$E_\lambda = \bigcup_{k \in \mathbf{N}^*} \ker(\lambda I_d - A)^k.$$

La réunion définissant E_λ est une réunion croissante d'espaces vectoriels. Puisque la dimensions de \mathbf{C}^d est finie, la suite croissante est stationnaire.

Comme une conséquence du théorème fondamental sur l'inversibilité en dimension finie, on a le résultat suivant.

1. Souvent noté $\rho(A)$, mais nous réserverons cette notation pour le rayon spectral.

Lemme 2 Le spectre coïncide avec l'ensemble des valeurs propres. De plus $\lambda \in \mathbf{C}$ est une valeur propre si et seulement si λ est une racine du polynôme caractéristique $\chi_A(X) := \det(XI_d - A)$.

Remarque 3 Tout polynôme de degré d unitaire $X^d + \sum_{j=0}^{d-1} \alpha_j X^j$ peut être réalisé comme un polynôme caractéristique, par exemple par sa matrice compagne

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -\alpha_0 & \cdots & \cdots & \cdots & -\alpha_{d-1} \end{pmatrix}.$$

Définition 4 1. On appelle multiplicité géométrique d'une valeur propre λ la dimension de l'espace propre $\ker(\lambda I_d - A)$ et multiplicité algébrique sa multiplicité en tant que racine du polynôme caractéristique. On verra plus tard que celle-ci coïncide avec la dimension de E_λ .

2. Une valeur propre est dite simple si sa multiplicité algébrique vaut un. Elle est dite double si celle-ci vaut deux, triple si elle vaut trois, etc. et multiple si elle est supérieure ou égale à deux.

3. Une valeur propre est dite semi-simple si ses multiplicités algébrique et géométrique coïncident.

Théorème 5 (Factorisation de Schur) Soit A une matrice carrée de taille $d \in \mathbf{N}^*$. Alors il existe P unitaire de taille d et T triangulaire supérieure de taille d tels que $A = PTP^{-1}$. Si de plus A est réelle et les valeurs propres de A sont toutes réelles, alors P et T peuvent être choisis réels.

Démonstration. La trigonalisation avec un P non nécessairement inversible provient d'une récurrence basée sur le fait que toute matrice possède un vecteur propre. Il suffit ensuite d'appliquer une décomposition QR à $P = QR$, avec Q unitaire et R triangulaire, c'est-à-dire d'orthonormaliser par le processus de Gram-Schmidt les colonnes de P . ■

Corollaire 6 Si A une matrice carrée normale, c'est-à-dire telle que $AA^* = A^*A$.

Alors il existe P unitaire de taille d et D diagonale de taille d tels que $A = PDP^{-1}$.

Si de plus A est réelle et auto-adjointe, alors P et D peuvent être choisis à coefficients réels.

Démonstration. Pour le premier point, il suffit de montrer qu'une matrice triangulaire et normale est diagonale. Cela se montre par récurrence à partir de : si T est triangulaire supérieure carrée de taille d , alors $(TT^*)_{d,d} = |T_{d,d}|^2$ et $(T^*T)_{d,d} = \sum_{j=1}^d |T_{j,n}|^2$.

Pour le second point, cela vient du fait que si A est réelle et auto-adjointe alors son spectre est réel. En effet si x est un vecteur propre associé à λ , alors

$$\lambda \|x\|^2 = \langle x, Ax \rangle = \langle A^*x, x \rangle \stackrel{A=A^*}{=} \langle Ax, x \rangle = \bar{\lambda} \|x\|^2$$

où $\langle \cdot; \cdot \rangle$ désigne le produit scalaire hermitien canonique sur \mathbf{C}^n (linéaire à droite) et $\|\cdot\|$ la norme associée. ■

Lemme 7 Si A une matrice carrée de taille d et $\|\cdot\|$ est une norme subordonnée telle que $\|A\| < 1$, alors $I_d - A$ est inversible et

$$(I_d - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Démonstration. Pour tout $n \in \mathbf{N}$, on a

$$I_d - A^n = I_d^n - A^n = (I_d - A) \sum_{k=0}^{n-1} A^k$$

avec

$$\|A^n\| \leq \|A\|^n \xrightarrow{n \rightarrow \infty} 0 \quad \text{et} \quad \sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{\infty} \|A\|^k = \frac{1}{1 - \|A\|} < \infty.$$

■

Théorème 8 (Théorème de Gershgorin-Hadamard) 1. Toute matrice A à diagonale strictement dominante par ligne (resp. par colonne) c'est-à-dire vérifiant

$$\forall i \in \llbracket 1, d \rrbracket, \quad |A_{i,i}| > \sum_{j \neq i} |A_{i,j}|, \quad (\text{resp. } \forall j \in \llbracket 1, d \rrbracket, \quad |A_{j,j}| > \sum_{i \neq j} |A_{i,j}|),$$

est inversible.

2. Pour toute matrice carrée A , on a

$$\sigma(A) \subset \bigcup_{1 \leq i \leq d} B(A_{i,i}, \sum_{j \neq i} |A_{i,j}|) \quad \text{et} \quad \sigma(A) \subset \bigcup_{1 \leq j \leq d} B(A_{j,j}, \sum_{i \neq j} |A_{i,j}|).$$

De plus dans chaque composante connexe d'une de ses réunions, le nombre de valeurs propres de A comptées avec multiplicité algébrique est égal au nombre d'éléments de la diagonale comptés avec multiplicité.

Démonstration. La première partie est un corollaire du lemme puisque si D désigne la partie diagonale de A alors D est inversible et $A = D(I_d + D^{-1}(A - D)) = (I_d + (A - D)D^{-1})D$ avec

$$\|D^{-1}(A - D)\|_{\infty} = \max_i \frac{\sum_{j \neq i} |A_{i,j}|}{|A_{i,i}|}, \quad \|(A - D)D^{-1}\|_1 = \max_j \frac{\sum_{i \neq j} |A_{i,j}|}{|A_{j,j}|},$$

où, pour $1 \leq p \leq \infty$, $\|\cdot\|_p$ désigne la norme subordonnée à la norme ℓ^p .

L'inclusion du spectre dans les réunions résulte de la première partie, puisque si λ n'appartient pas à la première (resp. seconde) union, alors $\lambda I_d - A$ est à diagonale strictement dominante par ligne (resp. par colonne).

Le point manquant se montre par un argument d'homotopie. Considérons pour $t \in [0, 1]$ la matrice $A^{[t]} := (t A_{i,j} + (1 - t) A_{i,i} \delta_{i,j})_{i,j}$. Soit \mathcal{C} une composante connexe d'une des réunions et γ un lacet² simple n'intersectant pas la réunion en question et entourant \mathcal{C} positivement. Alors d'après la première partie l'application continue suivante est bien définie

$$[0, 1] \rightarrow \mathbf{N}, \quad t \mapsto \frac{1}{2i\pi} \int_{\gamma} \frac{\chi'_{A^{[t]}}(z)}{\chi_{A^{[t]}}(z)} dz.$$

2. Ici les lacets sont toujours supposés rectifiables.

Comme \mathbf{N} est discret et $[0, 1]$ est connexe, cette application, qui compte le nombre de zéros de $\chi_{A[t]}$ à l'intérieur de γ , est constante. Or $A^{[1]} = A$ et $A^{[0]}$ est la partie diagonale de A . D'où le résultat en faisant tendre γ vers le bord³ de \mathcal{C} . ■

L'argument de comptage employé ci-dessus est plus général et donne la continuité du spectre en le sens suivant.

Théorème 9 1. Soit A une matrice carrée possédant une valeur propre λ_0 de multiplicité algébrique m_0 et $\|\cdot\|$ une norme subordonnée. Alors il existe $\varepsilon_0 > 0$ et $C_0 > 0$ telle que, pour tout $0 < \varepsilon \leq \varepsilon_0$, toute matrice B vérifiant

$$\|B - A\| \leq C_0 \varepsilon^{m_0}$$

possède m_0 valeurs propres comptées avec multiplicité algébrique dans $B(\lambda_0, \varepsilon)$.

2. Notons $\widehat{\mathbf{C}}^d$ le quotient de \mathbf{C}^d par l'action des permutations sur les coordonnées et \widehat{d} une distance⁴ quotient associée. Soit A une matrice carrée de plus grande multiplicité algébrique m_0 et $\|\cdot\|$ une norme subordonnée. Alors il existe $R_0 > 0$ et $C_0 > 0$ telle que toute matrice B vérifiant $\|B - A\| \leq R_0$ satisfait à

$$\widehat{d}(\Sigma(A), \Sigma(B)) \leq C_0 \|B - A\|^{\frac{1}{m_0}}$$

où $\Sigma(\cdot)$ est l'application qui à une matrice associe la classe d'équivalence du d -uplet de ses valeurs propres comptées avec multiplicité algébrique.

Démonstration. Il suffit de montrer que l'on peut définir l'application qui à B associe son nombre de valeurs propres dans $B(\lambda_0, \varepsilon)$ quand B appartient au connexe $B(A, C_0 \varepsilon^{m_0})$. Or il existe $c' > 0$ et $C' > 0$ tels que si $\|B - A\| \leq 1$ et $|z - \lambda_0|$ est assez petit

$$|\chi_A(z)| \geq c' |z - \lambda_0|^{m_0}, \quad |\chi_B(z) - \chi_A(z)| \leq C' \|B - A\|.$$

D'où le résultat avec $C_0 < c'/C'$ et ε_0 assez petit. ■

L'exemple de la famille de matrices $(A_\varepsilon)_{\varepsilon \in \mathbf{R}}$ obtenue comme matrices compagnes de la famille de polynômes $(X^d - \varepsilon)_{\varepsilon \in \mathbf{R}}$ dont le spectre est

$$\sigma(A_\varepsilon) = \begin{cases} \left\{ \left\{ \varepsilon^{\frac{1}{d}} z \mid z^d = 1 \right\} \right\} & \text{if } \varepsilon \geq 0 \\ \left\{ \left\{ |\varepsilon|^{\frac{1}{d}} e^{\frac{i\pi}{d}} z \mid z^d = 1 \right\} \right\} & \text{if } \varepsilon < 0 \end{cases}$$

montre que la borne peut être saturée du point de vue de la puissance. Cependant, comme on le verra plus tard l'exposant optimal est en fait lié à la multiplicité comme racine du polynôme minimal plutôt qu'à la multiplicité algébrique.

C'est facile à démontrer dans le cas diagonalisable.

Lemme 10 Soit $\|\cdot\|$ une norme subordonnée telle que pour toute matrice diagonale D , $\|D\| = \max_i |D_{i,i}|$. Pour toute matrice A diagonalisable, il existe $R_0 > 0$ telle que toute matrice B vérifiant $\|B - A\| \leq R_0$ l'on ait

$$\widehat{d}(\Sigma(A), \Sigma(B)) \leq C_0 \|B - A\|$$

3. Ce bord n'est pas toujours l'image d'un lacet simple.

4. Notons qu'elle contrôle la distance de Hausdorff entre les ensembles associés.

(avec les notations du Théorème 9) où

$$C_0 = \inf_{\substack{P \text{ inversible} \\ P^{-1}AP \text{ diagonale}}} \|P\| \|P^{-1}\|.$$

La constante C_0 est appelée conditionnement spectral de A (par rapport à $\|\cdot\|$).

Démonstration. Il suffit de montrer que si $d(\lambda, \sigma(A)) > C_0 \|B - A\|$ alors $\lambda \notin \sigma(B)$. Considérons un tel λ . Il existe P inversible tel que $D := P^{-1}AP$ soit diagonale et $d(\lambda, \sigma(A)) > \|P\| \|P^{-1}\| \|B - A\|$.

On a $\lambda I_d - B = P(\lambda I_d - D - (P^{-1}(B - A)P))P^{-1}$ avec

$$\|(\lambda I_d - D)^{-1}(P^{-1}(B - A)P)\| \leq \|(\lambda I_d - D)^{-1}\| \|P^{-1}\| \|B - A\| \|P\| = \frac{\|P\| \|P^{-1}\| \|B - A\|}{d(\lambda, \sigma(A))} < 1.$$

D'où le résultat. ■

2 Calcul fonctionnel

Nous allons à quoi cela peut servir de s'intéresser au spectre d'une matrice et à ses résolvantes. Il est clair que connaître une diagonalisation de A permet de calculer facilement $f(A)$ pour n'importe quelle fonction définie dans un voisinage de $\sigma(A)$. Nous allons voir comment généraliser cela via des formules de Cauchy.

Proposition 11 *Soit A une matrice carrée de taille d .*

1. L'application $z \mapsto (zI_d - A)^{-1}$ est méromorphe⁵. L'ensemble de ses pôles coïncide avec le spectre de A , et leur ordre est inférieur⁶ ou égal à leur multiplicité algébrique.
2. Si $\|\cdot\|$ est une norme subordonnée et $|\lambda| > \|A\|$ alors $\lambda \notin \sigma(A)$ et

$$(\lambda I_d - A)^{-1} = \sum_{k=0}^{\infty} \lambda^{-(k+1)} A^k.$$

Démonstration. L'analyticité sur $\mathbf{C} \setminus \sigma(A)$ et le deuxième point sont laissés en exercice. La formule de Cramer (aussi appelée formule de Laplace) montre l'inclusion des pôles dans le spectre et la borne sur l'ordre. Reste à montrer qu'il y a bien un pôle en chaque valeur propre. Or⁷ si $\|\cdot\|$ désigne une norme subordonnée on sait que pour tout $\lambda \notin \sigma(A)$,

$$B(\lambda, \|(\lambda I_d - A)^{-1}\|^{-1}) \subset \mathbf{C} \setminus \sigma(A)$$

d'où

$$\|(\lambda I_d - A)^{-1}\| \geq \frac{1}{d(\lambda, \sigma(A))} \xrightarrow{d(\lambda, \sigma(A)) \rightarrow 0} +\infty$$

et le résultat. ■

5. Au sens où chacune de ses coordonnées est méromorphe.

6. En fait égal, comme on le verra par la suite.

7. Alternativement : si $z \mapsto (zI_d - A)^{-1}$ est borné au voisinage de λ_0 alors elle possède une valeur d'adhérence R_{λ_0} en λ_0 et par passage à la limite $(\lambda_0 I_d - A)R_{\lambda_0} = I_d$ donc $\lambda_0 \notin \sigma(A)$.

Proposition 12 (Formule de Cauchy) Si A est une matrice carrée et f est une fonction holomorphe sur tout \mathbf{C} . Notons

$$f(A) := \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!} A^k.$$

Si γ est un lacet simple entourant positivement $\sigma(A)$ alors

$$f(A) = \frac{1}{2i\pi} \int_{\gamma} f(z) (zI_d - A)^{-1} dz.$$

Démonstration. Par holomorphie l'intégrale ne dépend pas du contour. On peut donc le déformer vers l'infini pour qu'il entoure positivement 0 et pouvoir obtenir le développement

$$\begin{aligned} \frac{1}{2i\pi} \int_{\gamma} f(z) (zI_d - A)^{-1} dz &= \sum_{k=0}^{\infty} \frac{1}{2i\pi} \int_{\gamma} f(z) z^{-(k+1)} A^k dz \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{1}{2i\pi} \int_{\gamma} \frac{f^{(\ell)}(0)}{\ell!} z^{\ell-(k+1)} A^k dz \\ &= \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!} A^k. \end{aligned}$$

■

Corollaire 13 Soit A une matrice carrée de taille d .

1. Notons

$$\rho(A) := \max_{\lambda \in \sigma(A)} |\lambda|$$

le rayon spectral. Alors pour tout $\varepsilon > 0$ et toute norme subordonnée $\|\cdot\|$, il existe $C_{\varepsilon, \|\cdot\|}$ tel que pour tout $n \in \mathbf{N}$

$$\|A^n\| \leq C_{\varepsilon, \|\cdot\|} (\rho(A) + \varepsilon)^n.$$

En particulier si $\rho(A) < 1$ alors, $A^n \xrightarrow{n \rightarrow \infty} 0$ géométriquement.

2. Pour tout $\varepsilon > 0$ et toute norme subordonnée $\|\cdot\|$, il existe $C_{\varepsilon, \|\cdot\|}$ tel que pour tout $t \geq 0$

$$\|e^{tA}\| \leq C_{\varepsilon, \|\cdot\|} e^{t(\max_{\lambda \in \sigma(A)} \operatorname{Re}(\lambda) + \varepsilon)}.$$

En particulier si $\max_{\lambda \in \sigma(A)} \operatorname{Re}(\lambda) < 0$ alors, $e^{tA} \xrightarrow{t \rightarrow \infty} 0$ exponentiellement.

Démonstration. Dans le premier cas il suffit de choisir un lacet simple entourant positivement le spectre et inclus dans $\bar{B}(0, \rho(A) + \varepsilon)$. On obtient le résultat avec

$$C_{\varepsilon, \|\cdot\|} = \frac{1}{2\pi} \int_{\gamma} \|(zI_d - A)^{-1}\| |dz|.$$

Dans le deuxième cas il suffit de choisir un lacet simple entourant positivement le spectre et inclus dans $\{z \mid \operatorname{Re}(z) \leq \max_{\lambda \in \sigma(A)} \operatorname{Re}(\lambda) + \varepsilon\}$. On obtient le résultat également avec la même formule pour $C_{\varepsilon, \|\cdot\|}$.

■

L'étude des puissances est reliée à celles des suites récurrentes linéaires ou affines, $X_{n+1} = AX_n + b$. Celle des exponentielles aux équations différentielles ordinaires linéaires ou affines, $X'(t) = AX(t) + b$. La définition de $f(A)$ par l'intégrale de contour est en fait plus souple. Elle permet de localiser proche de $\sigma(A)$.

Théorème 14 (Calcul fonctionnel) *Soit A une matrice carrée de taille d et Ω un voisinage ouvert de $\sigma(A)$. Alors il existe une unique application $\Phi_A : \text{Hol}(\Omega) \rightarrow \mathcal{M}_d(\mathbf{C})$ telle que*

1. Φ_A soit un morphisme d'algèbre avec unité;
2. Φ_A soit continue pour la topologie de la convergence uniforme sur tout compact;
3. $\Phi_A(\text{Id}_\Omega) = A$ et $\Phi_A(\mathbf{1}_\Omega) = \text{I}_d$.

De plus on a :

1. si γ est une union finie de lacets simples décrivant positivement les bords de disques disjoints centrés en un élément de $\sigma(A)$ alors pour tout $f \in \text{Hol}(\Omega)$

$$\Phi_A(f) = \frac{1}{2i\pi} \int_{\gamma} f(z) (z\text{I}_d - A)^{-1} dz;$$

2. pour tout $f \in \text{Hol}(\Omega)$, $\sigma(\Phi_A(f)) = f(\sigma(A))$;
3. $\Phi_A(f)$ ne dépend pas d'un choix de Ω tel que $f \in \text{Hol}(\Omega)$.

On notera désormais $f(A) := \Phi_A(f)$ quand f est holomorphe dans un voisinage de $\sigma(A)$.

Notons que la propriété de morphisme d'algèbre avec unité assure que si f est holomorphe dans un voisinage de A et que Q est une fraction rationnelle sans pôle en $f(\sigma(A))$ alors $Q(f(A)) = (Q(f))(A)$.

Démonstration. L'unicité vient de la densité des fonctions polynomiales. Quant à l'existence, il suffit de vérifier que la formule fournie convient. Seule l'égalité $\Phi_A(fg) = \Phi_A(f)\Phi_A(g)$ demande un peu de travail. On peut réduire la taille des disques pour déformer γ en γ' intérieur à γ . Alors

$$\Phi_A(f)\Phi_A(g) = \frac{1}{(2i\pi)^2} \int_{\gamma} \int_{\gamma'} f(z)g(z') (z\text{I}_d - A)^{-1}(z'\text{I}_d - A)^{-1} dz dz'.$$

Or

$$\begin{aligned} (z\text{I}_d - A)^{-1} &= (z\text{I}_d - A)^{-1}(z'\text{I}_d - A)(z'\text{I}_d - A)^{-1} \\ (z'\text{I}_d - A)^{-1} &= (z\text{I}_d - A)^{-1}(z\text{I}_d - A)(z'\text{I}_d - A)^{-1} \end{aligned}$$

d'où la décomposition en éléments simples (appelée identité de la résolvante)

$$\frac{1}{z' - z} ((z\text{I}_d - A)^{-1} - (z'\text{I}_d - A)^{-1}) = (z\text{I}_d - A)^{-1}(z'\text{I}_d - A)^{-1}.$$

Par intégration (via la formule de Cauchy) on récupère

$$\Phi_A(f)\Phi_A(g) = \frac{1}{2i\pi} \int_{\gamma} f(z)g(z) (z\text{I}_d - A)^{-1} dz = \Phi_A(fg).$$

Seule la propriété sur le spectre reste à montrer. Si $\mu \notin f(\sigma(A))$ alors $g : z \mapsto (\mu - f(z))^{-1}$ est holomorphe dans un voisinage de $\sigma(A)$ et $g(A)$ fournit un inverse pour $\mu\text{I}_d - f(A)$. D'où $\sigma(f(A)) \subset f(\sigma(A))$. Réciproquement si $\lambda \in \sigma(A)$, alors $g : z \mapsto (f(\lambda) - f(z))/(\lambda - z)$ est holomorphe dans un voisinage de $\sigma(A)$ et $f(\lambda)\text{I}_d - A = g(A)(\lambda\text{I}_d - A)$ n'est pas inversible. ■

Corollaire 15 Pour toute matrice A et toute norme subordonnée $\|\cdot\|$, on a

$$\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n} = \inf_{n \geq 1} \|A^n\|^{1/n}.$$

Démonstration. On a déjà que pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$ tel que

$$\forall n \geq 0, \quad \|A^n\|^{1/n} \leq C_\varepsilon^{1/n} (\rho(A) + \varepsilon)$$

donc que pour tout $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} \|A^n\|^{1/n} \leq \rho(A) + \varepsilon.$$

On a maintenant pour tout $n \geq 1$

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

D'où

$$\rho(A) \leq \inf_{n \geq 1} \|A^n\|^{1/n} \leq \limsup_{n \rightarrow \infty} \|A^n\|^{1/n} \leq \rho(A)$$

et le résultat. ■

Donnons une première application non triviale du calcul fonctionnel.

Proposition 16 Si A est inversible, alors il existe B tel que $e^B = A$.

Démonstration. Supposons A inversible et donnons-nous F une détermination analytique du logarithme dans un voisinage de $\sigma(A)$. Montrons que $B = F(A)$ convient. Pour cela donnons-nous γ et γ' des unions de cercles paramétrés dans le sens direct centrés sur les valeurs propres de A , avec les cercles définissant γ' à l'intérieur de ceux définissant γ . Notons que l'image de γ par F entoure le spectre de B convenablement. Alors

$$\begin{aligned} e^B &= \frac{1}{2i\pi} \int_{F(\gamma)} e^z (zI_d - B)^{-1} dz \\ &= \frac{1}{(2i\pi)^2} \int_{F(\gamma)} \int_{\gamma'} \frac{e^z}{z - F(z')} (z'I_d - A)^{-1} dz' dz \\ &\stackrel{u=e^z}{=} \frac{1}{(2i\pi)^2} \int_{\gamma} \int_{\gamma'} \frac{1}{F(u) - F(z')} (z'I_d - A)^{-1} dz' du \\ &= \frac{1}{(2i\pi)^2} \int_{\gamma} \int_{\gamma'} \frac{u - z'}{F(u) - F(z')} \frac{1}{u - z'} (z'I_d - A)^{-1} dz' du \\ &\stackrel{F'(z')=1/z'}{=} \frac{1}{2i\pi} \int_{\gamma'} z' (z'I_d - A)^{-1} dz' = A. \end{aligned}$$

■

Théorème 17 (Décomposition spectrale) Soit A une matrice carré de taille d , de valeurs propres $\lambda_1, \dots, \lambda_r$, $r \in \llbracket 1, d \rrbracket$.

Alors il existe $\Pi_1, \dots, \Pi_r, N_1, \dots, N_r$ donnés comme des fonctions⁸ de A tels que

$$A = \sum_{i=1}^r (\lambda_i I_d + N_i) \Pi_i$$

et

⁸. Donc commutant entre eux et avec A .

1. Π_1, \dots, Π_r , étant des projecteurs supplémentaires, l'image de chaque Π_i étant égale à l'espace caractéristique E_{λ_i} ;
2. chaque N_i est nilpotent de degré de nilpotence inférieur ou égal au rang de Π_i .

De plus si f est holomorphe dans un voisinage de $\sigma(A)$ alors

$$f(A) = \sum_{j=1}^r \sum_{k=0}^{m_j-1} \frac{f^{(k)}(\lambda_j)}{k!} N_j^k \Pi_j.$$

où, pour $1 \leq j \leq r$, m_j est le degré de nilpotence de N_j . Les Π_j sont appelés projecteurs spectraux ou projecteurs de Riesz.

Il s'agit d'un raffinement de la décomposition de Dunford. En choisissant une base adaptée aux matrices nilpotentes on déduit la décomposition de Jordan. On peut démontrer le Théorème 17 en utilisant le lemme des noyaux (qui est un corollaire du théorème de Bézout) et le théorème de Cayley-Hamilton. Les N_i, Π_i sont alors obtenus comme des polynômes en A assez peu explicite. L'avantage de l'approche suivie ici est le caractère assez explicite des formules pour les Π_j .

L'expression de $f(A)$ montre que l'on peut étendre par continuité le calcul fonctionnel à des espaces de fonctions avec une quantité limitée de régularité.

Démonstration. Il suffit de poser $\Pi_i = g_i(A)$ avec g_i constant égal à 1 dans un voisinage de λ_i et constant égal à 0 dans un voisinage de $\sigma(A) \setminus \{\lambda_i\}$ et $N_i = (A - \lambda_i I_d) \Pi_i = h_i(A)$ avec h_i coïncidant avec $z \mapsto (z - \lambda_i)$ dans un voisinage de λ_i et constant égal à 0 dans un voisinage de $\sigma(A) \setminus \{\lambda_i\}$.

Les propriétés découlent alors d'une part du fait que dans un voisinage de $\sigma(A)$,

1. $\sum_{i=1}^r g_i$ est constant égal à 1 ;
2. pour tout (i, j) , $g_i g_j = g_i \delta_{i,j}$;
3. $\sum_{i=1}^r (\lambda_i \text{Id} + h_i) g_i$ coïncide avec Id ;

et d'autre part au fait que pour tout i , $\sigma(N_i) = h_i(\sigma(A)) = \{0\}$ et $N_i = (A - \lambda_i I_d) \Pi_i$ a un rang inférieur à celui de Π_i . De la nilpotence de N_i on déduit que l'image de Π_i est incluse dans E_{λ_i} . Réciproquement, introduisons H_i coïncidant avec $z \mapsto (z - \lambda_i)^{-1}$ dans un voisinage de $\sigma(A) \setminus \{\lambda_i\}$ et constant égal à 0 dans un voisinage de λ_i de sorte que pour tout $m \in \mathbf{N}^*$, $H_i(A)^m (\lambda_i I_d - A)^m = \sum_{j \neq i} \Pi_j = I_d - \Pi_i$ ce qui implique que le noyau de $(\lambda_i I_d - A)^m$ est inclus dans l'image de Π_i .

La formule pour $f(A)$ est obtenue en insérant

$$\begin{aligned} (z I_d - A)^{-1} &= \sum_{j=1}^r (z I_d - (\lambda_j I + N_j))^{-1} \Pi_j \\ &= \sum_{j=1}^r \sum_{k=0}^{m_j-1} (z - \lambda_j)^{-k-1} N_j^k \Pi_j \end{aligned}$$

dans les intégrales de contour. ■

On a utilisé le lemme suivant qui découle du théorème de Cayley-Hamilton mais que l'on peut montrer avec de l'analyse complexe.

Lemme 18 *Si N est une matrice carrée de spectre $\sigma(N) = \{0\}$, alors N est nilpotente.*

Démonstration. L'application $z \mapsto (zI_d - N)^{-1}$ est méromorphe sur $\mathbf{C} \setminus \{0\}$. Or son développement en série de Laurent peut être calculé (avec des intégrales de contour) à partir du développement obtenu pour z assez grand

$$(zI_d - N)^{-1} = \sum_{k=0}^{\infty} z^{-(k+1)} N^k.$$

Par méromorphie le développement est fini, N est nilpotente. ■

Corollaire 19 1. *Toute matrice à valeurs propres distinctes est diagonalisable.*

2. *Avec les notations du théorème 17, $\prod_{j=1}^r (X - \lambda_j)^{m_j}$ est le polynôme minimal de A .*

3. *La multiplicité algébrique coïncide avec la dimension de l'espace caractéristique. De plus on a $\chi_A(A) = 0$ (théorème de Cayley-Hamilton).*

4. *Toute fonction de A coïncide avec un polynôme en A .*

Démonstration. Pour le premier point, notons que si $r = d$ alors tous les projecteurs sont de rang 1 et donc tous les N_i ont degré de nilpotence 1 donc sont nuls.

On laisse le deuxième point en exercice.

Pour le troisième point il suffit de calculer le polynôme caractéristique de A dans une base adaptée à la décomposition (de type Jordan) et d'utiliser $\prod_{j=1}^r (\lambda_j I_d - A)^{m_j} = 0$.

Pour le troisième, notons que pour avoir $f(A) = P(A)$, il suffit⁹ d'avoir pour tout $1 \leq j \leq r$, pour tout $0 \leq k \leq m_j - 1$,

$$f^{(k)}(\lambda_j) = P^{(k)}(\lambda_j).$$

L'interpolation¹⁰ fournit un tel polynôme de degré inférieur ou égal à¹¹ $(\sum_{j=1}^r m_j) - 1$. ■

On peut déduire bien d'autres théorèmes.

Corollaire 20 *Soit λ_0 est une valeur propre simple de A de vecteur propre V_0 . Il existe Ω un voisinage de A et $\Lambda : \Omega \rightarrow \mathbf{C}$ et $V : \Omega \rightarrow \mathbf{C}^d$ analytiques tels que*

- *pour tout $B \in \Omega$, $\Lambda(B)$ est une valeur propre simple de B de vecteur propre $V(B)$;*
- *$\Lambda(A) = \lambda_0$ et $V(A) = V_0$.*

Démonstration. Soit γ un lacet simple entourant λ_0 positivement mais n'entourant pas $\sigma(A) \setminus \{\lambda_0\}$. Introduisons

$$\Pi(B) := \frac{1}{2i\pi} \int_{\gamma} (zI_d - B)^{-1} dz$$

qui est une fonction analytique en B à valeur dans les projecteurs de rang 1 définie sur un voisinage connexe de A . Il suffit d'introduire

$$V(B) = \Pi(B) V_0, \quad \Lambda(B) = \frac{\langle V(B), B V(B) \rangle}{\|V(B)\|^2}$$

9. C'est également nécessaire.

10. Il suffit de montrer l'unicité et d'invoquer le théorème du rang.

11. C'est cohérent avec le fait que le polynôme d'annulation soit de degré $\sum_{j=1}^r m_j$.

(quitte à réduire le voisinage pour assurer que $V(B)$ est non nul) où $\langle \cdot; \cdot \rangle$ désigne le produit scalaire hermitien canonique sur \mathbf{C}^n (linéaire à droite) et $\| \cdot \|$ la norme associée. ■

Corollaire 21 1. Soit A une matrice carrée possédant une valeur propre λ_0 de multiplicité algébrique m_0 et de multiplicité m_1 comme racine du polynôme minimal et $\| \cdot \|$ une norme subordonnée. Alors il existe $\varepsilon_0 > 0$ et $C_0 > 0$ telle que, pour tout $0 < \varepsilon \leq \varepsilon_0$, toute matrice B vérifiant

$$\|B - A\| \leq C_0 \varepsilon^{m_1}$$

possède m_0 valeurs propres comptées avec multiplicité algébrique dans $B(\lambda_0, \varepsilon)$.

2. Notons $\widehat{\mathbf{C}}^d$ le quotient de \mathbf{C}^d par l'action des permutations sur les coordonnées et \widehat{d} une distance quotient associée. Soit A une matrice carrée de plus grande multiplicité m_0 pour le polynôme minimal et $\| \cdot \|$ une norme subordonnée. Alors il existe $R_0 > 0$ et $C_0 > 0$ telle que toute matrice B vérifiant $\|B - A\| \leq R_0$ satisfait à

$$\widehat{d}(\Sigma(A), \Sigma(B)) \leq C_0 \|B - A\|^{\frac{1}{m_0}}$$

où $\Sigma(\cdot)$ est l'application qui à une matrice associe la classe d'équivalence du d -uplet de ses valeurs propres comptées avec multiplicité algébrique.

Démonstration. Il suffit de montrer le premier point.

Notons que lorsqu'il est bien défini

$$\Pi_\varepsilon(B) := \frac{1}{2i\pi} \int_{\gamma_\varepsilon} (zI_d - B)^{-1} dz$$

(où γ_ε est un lacet simple paramétrant dans le sens direct le bord de $B(\lambda_0, \varepsilon)$) est la somme des projecteurs spectraux associés à des valeurs propres de B dans $B(\lambda_0, \varepsilon)$. Son rang est égal à la somme des rangs de ces projecteurs donc à la somme des multiplicités algébriques des valeurs propres de B dans $B(\lambda_0, \varepsilon)$.

Il suffit donc de montrer

1. que $\Pi_\varepsilon(B)$ est bien défini quand B et ε sont comme dans le théorème ;
2. que le rang des projecteurs est localement constant.

Le deuxième point est l'objet du lemme qui suit¹². Quant au premier, cela vient du fait qu'il existe c' tel que pour $|z - \lambda_0|$ assez petit

$$\|(zI_d - B)^{-1}\| \leq \frac{c'}{|z - \lambda_0|^{m_0}}.$$

D'où le résultat pour n'importe quel C_0 tel que $C_0 c' < 1$. ■

Lemme 22 Soit $\| \cdot \|$ une norme subordonnée et P et Q deux projecteurs tels que $\|P - Q\| < 1$. Alors P et Q ont même rang.

12. Une version moins quantitative suffit et peut être obtenue en utilisant la continuité du spectre et le fait que pour un projecteur le rang coïncide avec le nombre de valeurs propres de module strictement plus grand que $1/2$.

Démonstration. L'application $\text{Im}P \rightarrow \text{Im}Q$, $x \mapsto Qx$ est injective puisque, pour $x \in \text{Im}P$,

$$Qx = (Q - P)x + x, \quad \|Qx\| \geq (1 - \|P - Q\|)\|x\|,$$

où $\|\cdot\|$ est la norme dont $\|\cdot\|$ est la norme subordonnée. Cela implique que le rang de Q est inférieur ou égal à celui de P . D'où le résultat par symétrie. \blacksquare

Corollaire 23 1. $(A^n)_{n \in \mathbf{N}}$ est bornée si et seulement si $\rho(A) \leq 1$ et toute valeur propre de A de module 1 est semi-simple.

2. $(A^n)_{n \in \mathbf{N}}$ converge si et seulement si $\rho(A) \leq 1$, 1 est la seule valeur propre de A de module 1 et 1 est une valeur propre semi-simple. Dans ce cas la suite converge vers le projecteur spectral associé à la valeur propre 1.

3. On a que, pour tout choix de vecteur b , une suite $(X_n)_{n \in \mathbf{N}}$ vérifiant

$$\forall n \in \mathbf{N}, \quad X_{n+1} = AX_n + b$$

converge si et seulement si $\rho(A) < 1$.

4. Supposons que, dans les notations du théorème 17, λ_1 soit la seule valeur propre de module $\rho(A)$. Alors

(a) si λ_1 n'est pas semi-simple et V_0 est tel que $N_1^{m_1-1}(V_0) \neq 0$, on a, pour tout $n \geq 1$, $A^n V_0$ ne s'annule pas et

$$W_n := \frac{A^n V_0}{\|A^n V_0\|} \stackrel{n \rightarrow \infty}{\cong} \left(\frac{\lambda_1}{|\lambda_1|} \right)^{n-(m_1-1)} \frac{N_1^{m_1-1}(V_0)}{\|N_1^{m_1-1}(V_0)\|} + \mathcal{O}\left(\frac{1}{n}\right),$$

$$\langle W_n, A W_n \rangle \stackrel{n \rightarrow \infty}{\cong} \lambda_1 + \mathcal{O}\left(\frac{1}{n}\right);$$

(b) si λ_1 est semi-simple et V_0 est tel que $\Pi_1(V_0) \neq 0$, on a, pour tout $n \geq 1$, $A^n V_0$ ne s'annule pas et

$$W_n := \frac{A^n V_0}{\|A^n V_0\|} \stackrel{n \rightarrow \infty}{\cong} \left(\frac{\lambda_1}{|\lambda_1|} \right)^n \frac{\Pi_1(V_0)}{\|\Pi_1(V_0)\|} + \mathcal{O}\left(n^{m_0-1} \left(\frac{\rho_0}{\rho(A)} \right)^n\right),$$

$$\langle W_n, A W_n \rangle \stackrel{n \rightarrow \infty}{\cong} \lambda_1 + \mathcal{O}\left(n^{m_0-1} \left(\frac{\rho_0}{\rho(A)} \right)^n\right),$$

avec $\rho_0 := \rho(A(I_d - \Pi_1)) = \max_{j \geq 2} |\lambda_j|$ le deuxième plus grand module des valeurs propres et $m_0 := \max_{j, |\lambda_j| = \rho_0} m_j$;

où $\langle \cdot; \cdot \rangle$ désigne le produit scalaire hermitien canonique sur \mathbf{C}^n (linéaire à droite) et $\|\cdot\|$ la norme associée.

Le dernier point fournit une méthode de calcul approché de la plus grande valeur propre et d'un vecteur propre associé. La suite $(W_n)_{n \geq 0}$ est en fait calculée par récurrence

$$\forall n \geq 0, \quad W_{n+1} = \frac{A W_n}{\|A W_n\|}.$$

Il s'agit de la *méthode de la puissance*.

- Corollaire 24**
1. $(e^{tA})_{t \geq 0}$ est bornée si et seulement si $\max_{\lambda \in \sigma(A)} \operatorname{Re}(\lambda) \leq 0$ et toute valeur propre de A imaginaire pure est semi-simple.
 2. $(e^{tA})_{t \geq 0}$ converge en $t = +\infty$ si et seulement si $\max_{\lambda \in \sigma(A)} \operatorname{Re}(\lambda) \leq 0$, 0 est la seule valeur propre de A de partie réelle nulle et 0 est une valeur propre semi-simple. Dans ce cas la suite converge vers le projecteur spectral associé à la valeur propre 0.
 3. $(e^{tA})_{t \in \mathbf{R}}$ est bornée si et seulement si A est diagonalisable et $\sigma(A) \subset i\mathbf{R}$.

Proposition 25 Si A est inversible ou 0 est une valeur propre semi-simple de A , alors pour tout $m \in \mathbf{N}^*$ il existe B tel que $B^m = A$.

Démonstration. Donnons-nous f une détermination analytique de la racine m -ième dans un voisinage de $\sigma(A) \setminus \{0\}$. Définissons g comme coïncidant avec f dans un voisinage de $\sigma(A) \setminus \{0\}$ et nulle dans un voisinage de 0. Alors $B = g(A)$ convient. ■

Références

- [AK02] G. Allaire and S. M. Kaber. *Algèbre linéaire numérique*. Ellipses, 2002.
- [Cia07] P. G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Dunod, 2007.
- [CMT01] P. G. Ciarlet, B. Miara, and J.-M. Thomas. *Exercices d'analyse numérique matricielle et d'optimisation*. Dunod, 2001.
- [Fil13] F. Filbet. *Analyse numérique - Algorithmes et étude mathématique*. Dunod, 2013.
- [Kat82] T. Kato. *A short introduction to perturbation theory for linear operators*. Springer-Verlag, New York-Berlin, 1982. En anglais.
- [Kat95] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. En anglais.
- [LB03] P. Lévy-Bruhl. *Introduction à la théorie spectrale*. Dunod, 2003.
- [Mne06] R. Mneimné. *Réduction des endomorphismes*. Calvage et Mounet, 2006.
- [Sch04] M. Schatzman. *Analyse numérique*. Dunod, 2004.
- [Ser01] D. Serre. *Les Matrices - Théorie et pratique*. Dunod, 2001.