Valeurs propres

Agrégation de mathématiques
Option modélisation

Septembre 2019

Benjamin Boutin

Université de Rennes 1

Ces notes de cours sont consacrées aux propriétés et à l'approximation des valeurs propres de matrices de $\mathsf{M}_n(\mathbb{C})$. La théorie de la réduction des endomorphismes en dimension finie est supposée acquise, aussi on mettra plutôt l'accent sur les résultats de localisation du spectre, les propriétés de continuité du spectre et sur les techniques d'approximation numérique des valeurs et vecteurs propres.

Plan du cours

1 Rappels et résultats élémentaires

1.1 Réductions remarquables

Sur le corps algébriquement clos $\mathbb{K} = \mathbb{C}$, tout polynôme de $\mathbb{K}[X]$ est scindé et donc tout endomorphisme sur $\mathbb{K}^n$ est trigonalisable. On note $\sigma(A)$ le spectre d'une matrice A de $\mathsf{M}_n(\mathbb{C})$:

$$\sigma(A) := \big\{\lambda \in \mathbb{C},\ \ker(A - \lambda I_n) \neq \{0\}\big\}.$$

Pour débuter, on rappelle les deux résultats de trigonalisation remarquables suivants.

Théorème 1 (*Réduction de Jordan*)

Soit A une matrice de $\mathsf{M}_n(\mathbb{C})$. Il existe $P \in \mathsf{GL}_n(\mathbb{C})$ et $T \in \mathsf{M}_n(\mathbb{C})$ triangulaire supérieure vérifiant $A = PTP^{-1}$ et telles que

$$T = \begin{pmatrix} J_{k_1}(\mu_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\mu_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{k_m}(\mu_m) \end{pmatrix},$$

avec les blocs $J_k(\lambda) \in \mathsf{M}_k(\mathbb{C})$ de la forme :

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

Démonstration (voir [2, 6]). La preuve repose sur la réduction des endomorphismes nilpotents.

- Factorisons le polynôme minimal de A : $\mu_A(X) = \prod_{i=1}^d (X - \lambda_i)^{\alpha_i}$.
  Par le lemme de décomposition des noyaux on fait intervenir les sous-espaces caractéristiques

$$\mathbb{C}^n = \bigoplus_{i=1}^d \underbrace{\ker(A - \lambda_i I_n)^{\alpha_i}}_{=:F_i}.$$

- Fixons désormais $i \in 1, \dots, d$ et notons f_i la restriction à F_i de l'endomorphisme canoniquement associé à A. C'est un endomorphisme sur F_i qui s'écrit sous la forme $f_i = \lambda_i \operatorname{id}_{F_i} + n_i$ avec $n_i \in \mathsf{L}(F_i)$ nilpotent d'indice α_i exactement.
- On démontre par récurrence que l'on peut décomposer F_i comme somme directe de sous-espaces cycliques $F_i = \bigoplus_{\ell=1}^{b_i} F_{i,\ell}$ où $F_{i,\ell} := \operatorname{Vect}(\underbrace{n_i^{k_{i,\ell}-1}(x_{i,\ell})}_{\neq 0}, \dots, n_i(x_{i,\ell}), x_{i,\ell})$ avec $n_i^{k_{i,\ell}}(x_{i,\ell}) = 0$. Les détails sont laissés au soin du lecteur.
- Dans la base considérée pour $F_{i,\ell}$, la matrice de n_i est précisément $J_{k_{i,\ell}}(0)$.
- Dans la base concaténée obtenue pour F_i, la matrice de f_i est diagonale par blocs, blocs de la forme $J_{k_{i,\ell}}(\lambda_i)$.
- Dans la base concaténée obtenue pour $\mathbb{C}^n$, la matrice de l'endomorphisme associé à A est de la forme T.

$\square$

Remarque.

- *Chaque bloc dans T correspond à un et un seul espace cyclique.*

- *Les valeurs propres* $(\mu_i)_{1\leq i\leq m}$ *dans le théorème ne sont pas nécessairement distinctes deux à deux. Les* $(\lambda_i)_{1\leq i\leq d}$ *dans la preuve sont distincts deux à deux.*
- *Pour tout* $i=1,\dots,d$, *l'entier* $\dim(F_i)=\sum_{\ell=1}^{b_i} k_{i,\ell}=\sum_{\mu_j=\lambda_i} k_j$ *est la multiplicité de la valeur propre* λ_i *dans le polynôme caractéristique de* A *(multiplicité algébrique).*
- *Pour tout* $i=1,\dots,d$, *l'entier* $b_i=\sum_{\mu_j=\lambda_i} 1$ *est la multiplicité géométrique de* λ_i.
- *Pour tout* $i=1,\dots,d$, *l'entier* $\alpha_i=\max_{\ell=1}^{b_i} k_{i,\ell}=\max_{\mu_j=\lambda_i} k_j$ *est la multiplicité de* λ_i *dans le polynôme* μ_A.
- *La matrice* A *est diagonalisable si et seulement si pour tout* $i=1,\dots,d$, *on a* $\dim(F_i)=b_i$, *i.e.* $k_{i,\ell}=1$ *(ce qui ne requiert ni* $b_i=1$*), ou de manière équivalente si et seulement si* $\alpha_i=1$.

Un résultat intéressant dès que l'on souhaite tirer profit de la structure hermitienne (euclidienne) et du produit scalaire canonique sur $\mathbb{C}^n$ est le théorème de réduction en base orthonormée suivant.

Théorème 2 (***Théorème de Schur***)

Soit A *une matrice de* $\mathsf{M}_n(\mathbb{C})$. *Il existe* $P\in\mathsf{U}_n(\mathbb{C})$ *unitaire et* $T\in\mathsf{M}_n(\mathbb{C})$ *triangulaire supérieure vérifiant* $A=PTP^\star$ *et* $PP^\star=I$.

Démonstration. On peut démontrer ce résultat de trigonalisation en base orthonormée (pour le produit scalaire hermitien) par exemple à partir d'une trigonalisation quelconque de A ($A=PTP^{-1}$) et d'un procédé d'orthonormalisation (Gram-Schmidt) de la base de trigonalisation (colonnes de P), qui se traduit matriciellement par l'existence d'une factorisation QR : $P=QR$ avec $Q\in\mathsf{U}_n(\mathbb{C})$ et $R\in\mathsf{M}_n(\mathbb{C})$ triangulaire supérieure. Dès lors $A=Q(RTR^{-1})Q^\star$ où le produit RTR^{-1} demeure triangulaire supérieure. □

Exemple.

- *Si* A *est hermitienne, i.e.* $A=A^\star$, *alors* $T=T^\star$ *et donc* T *est diagonale à coefficients réels (mais* $P\in\mathsf{U}_n(\mathbb{C})$*).*
- *Pour le cas* A *réelle symétrique, il y a un théorème de Schur réel, avec* P *orthogonale réelle.*
- *Si* A *est normale, i.e.* $AA^\star=A^\star A$, *alors* T *est également normale en plus d'être triangulaire, donc nécessairement diagonale. Ainsi* A *diagonalise en base orthonormée (et réciproquement).*

1.2 Rayon spectral

Définition 3

Soit $A\in\mathsf{M}_n(\mathbb{C})$. *On définit le rayon spectral de* A :

$$\rho(A):=\max\big\{|\lambda|,\ \lambda\in\sigma(A)\big\}.$$

Afin de localiser les valeurs propres d'une matrice, on se donne $\|\cdot\|$ une norme subordonnée quelconque sur $\mathsf{M}_n(\mathbb{C})$, alors l'inégalité suivante permet de borner l'ensemble des valeurs propres dans le plan complexe :

$$\rho(A)\leq\|A\|.$$

Un exercice consiste à prouver qu'on a plus précisément

$$\begin{aligned}\rho(A)&=\inf\{\|A\|,\ \|\cdot\|\text{ norme subordonnée}\}\\ \Rightarrow\quad&=\lim_{k\to\infty}\|A^k\|^{1/k}.\end{aligned}$$

Attention cependant à ne pas assimiler $\rho(A)$ à une norme sur $\mathsf{M}_n(\mathbb{C})$!

Exemple (Localisation des racines d'un polynôme)**.**
On peut démontrer les formules explicites pour les normes matricielles explicites suivantes,

$$\|A\|_1 = \max_j \sum_i |a_{ij}|, \quad \|A\|_\infty = \max_i \sum_j |a_{ij}|, \quad \|A\|_2 = \sqrt{\rho(A^\star A)}.$$

Soit $p \in \mathbb{C}_n[X]$ un polynôme de coefficient dominant égal à 1 : $p(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_0$. En utilisant la matrice compagnon *associée à p, dont les racines sont les valeurs propres de A*

$$C_p = \begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-2} \\ 0 & & & 1 & -a_{n-1} \end{pmatrix}$$

on obtient les estimations suivantes sur les racines $z \in \mathbb{C}$, respectivement dues initialement à Cauchy, à Montel, et à Carmichael-Mason :

$$\frac{|a_0|}{|a_0| + \max(1, |a_1|, \ldots, |a_{n-1}|)} \leq |z| \leq 1 + \max(|a_0|, \ldots, |a_{n-1}|),$$

$$\frac{|a_0|}{1 + |a_0| + \ldots + |a_{n-1}|} \leq |z| \leq \max(1, |a_0| + \ldots + |a_{n-1}|),$$

$$\frac{|a_0|}{(1 + |a_0|^2 + |a_1|^2 + \ldots + |a_{n-1}|^2)^{1/2}} \leq |z| \leq (1 + |a_0|^2 + |a_1|^2 + \ldots + |a_{n-1}|^2)^{1/2}.$$

Un résultat élémentaire mais essentiel concerne le comportement des suites de puissances $(A^k)_{k\in\mathbb{N}}$ et de la série $(\Sigma_{k\in\mathbb{N}}A^k)$ qui sont intrinséquement liés aux propriétés spectrales de A.

Proposition 4

Soit $A \in \mathsf{M}_n(\mathbb{C})$.

- *La suite $(A^k)_{k\in\mathbb{N}}$ converge vers 0 dans $\mathsf{M}_n(\mathbb{C})$, et de manière équivalente la série $\Sigma_{k\in\mathbb{N}}A^k$ est absolument convergente, si et seulement si $\rho(A) < 1$.*
- *La suite $(A^k)_{k\in\mathbb{N}}$ est bornée si et seulement si*

$$\rho(A) < 1 \quad \text{ou} \quad \rho(A) = 1 \text{ et } \forall\lambda \in \sigma(A),\ |\lambda| = 1 \Rightarrow \lambda \text{ est semi-simple.}$$

Démonstration. Le premier point est un exercice classique. Pour le deuxième énoncé dans le cas $\rho(A) = 1$, on peut considérer le cas de blocs de Jordan associés à des valeurs propres de module 1 (binôme de Newton sur $J_k(\lambda) = \lambda I_k + J_k(0)$). Le détail est laissé en exercice. □

Remarque. *Une petite question de vocabulaire. Les propriétés suivantes sont équivalentes :*

- *La valeur propre λ est semi-simple (ou non-défective)*
- *La multiplicité de λ dans le polynôme minimal est 1*
- *Les multiplicités algébriques et géométriques de λ coïncident*
- *Les sous-espaces propres et caractéristiques associés à λ coïncident*
- *La restriction de l'endomorphisme canonique associé à A au sous-espace caractéristique associé à λ est diagonalisable*
- *Les blocs de Jordan associés à λ sont tous de taille 1.*

1.3 Applications

Convergence d'une méthode linéaire récurrente Les méthodes itératives pour la résolution d'un système inversible $Ax = b$ reposent sur la définition d'une suite récurrente d'éléments de $\mathbb{C}^n$, notés $(x_k)_{k\geq 0}$ et définis à partir de la reformulation de l'équation sous forme d'équation de point fixe : $x = M^{-1}(Nx + b)$ avec $A = M - N$ et $M \in \mathsf{GL}_n(\mathbb{C})$:

$$\begin{cases} x_0 \in \mathbb{C}^n \\ x_{k+1} = \underbrace{M^{-1}N}_{=:G} x_k + M^{-1}b, \qquad k \geq 0. \end{cases}$$

En pratique, on choisit M diagonale (méthode de Jacobi) ou triangulaire (méthode de Gauss-Seidel) par exemple, de sorte à ce que l'inversion de $Mx_{k+1} = NX_k + b$ soit effectuée par un algorithme simple et rapide.
La convergence de la suite ainsi définie est garantie (et équivalente à) sous la condition $\rho(G) < 1$. On peut le vérifier en notant $e_k = x_k - x$ l'erreur à l'étape $k \geq 0$ qui vérifie : $e_k = G^k(x_0 - x)$. Si on souhaite garantir la convergence de $(e_k)_{k\geq 0}$ vers 0 sans avoir à se soucier du choix de la donnée initiale $x_0 \in \mathbb{C}^n$, il faut requérir que la suite matricielle (G^k) tende vers 0 lorsque k tend vers l'infini, ce qui est équivalent à la propriété mentionnée.

Plus loin : stabilité d'une méthode linéaire récurrente Considérons que la suite définie dans l'exemple précédent est imparfaitement calculée et entâchée d'erreurs (d'arrondi par exemple). Notons (y_k) les éléments réellement calculés, et $\eta_0, \epsilon_0, \epsilon_1, \ldots \in \mathbb{C}^n$ ces différentes erreurs avec :

$$\begin{cases} y_0 = x_0 + \eta_0 \in \mathbb{C}^n \\ y_{k+1} = Gy_k + M^{-1}b + \epsilon_k, \qquad k \geq 0. \end{cases}$$

On peut examiner la convergence de la suite (y_k) en s'intéressant à l'écart $\eta_k = y_k - x_k$. On observe facilement que $\eta_{k+1} = (M^{-1}N)\eta_k + \epsilon_k$ et on obtient par récurrence la formule de Duhamel discrète suivante :

$$\eta_k = G^k \eta_0 + \sum_{j=0}^{k-1} G^{k-1-j} \epsilon_j.$$

L'hypothèse $\rho(G) < 1$ garantit qu'on peut se donner une norme $\|\cdot\|$ sur $\mathsf{M}_n(\mathbb{C})$ subordonnée à une norme vectorielle sur $\mathbb{C}^n$ telle que $\rho(G) \leq C = \|G\| < 1$. Alors on a

$$\|\eta_k\| \leq \|\eta_0\| \sup_{k\in\mathbb{N}} \|G^k\| + \sup_{j\in\mathbb{N}} \|\epsilon_j\| \sum_{j\geq 0} \|G^j\|.$$

En conséquence, si les erreurs sont uniformément bornées, alors l'écart le sera également :

$$\exists C > 0,\ \forall k \geq 0,\ \|y_k - x_k\| \leq C\|\eta_0\| + \frac{1}{1-C} \sup_j \|\epsilon_j\|.$$

À la limite $y_k = x + (x_k - x) + (y_k - x_k)$ où le deuxième terme tend géométriquement vers 0, et le troisième terme est borné par les erreurs d'arrondi. Observons ici que la constante $(1 - C)^{-1}$ en facteur des erreurs d'arrondi est d'autant plus grande que C est voisin de 1, i.e. lorsque la convergence de l'algorithme est lente.

Stabilité des EDO linéaires à coefficients constants Considérons le problème de Cauchy d'inconnue à valeurs dans $\mathbb{C}^n$:

$$\begin{cases} Y(0) = Y_0 \in \mathbb{C}^n, \\ Y'(t) = AY(t). \end{cases}$$

L'unique solution $t \mapsto \exp(tA)Y_0$ tend-elle vers 0 lorsque t tend vers l'infini ? Est-elle bornée pour $t \in [0, +\infty)$? Peut-on préciser un résultat de continuité de la solution du problème suivant par rapport aux perturbations η_0 et $\epsilon(t)$, et si oui dans quelles topologies ?

$$\begin{cases} Y(0) = Y_0 + \eta_0 \in \mathbb{C}^n, \\ Y'(t) = AY(t) + \epsilon(t). \end{cases}$$

1.4 Méthode de la puissance

Algorithme Soit $x_0 \in \mathbb{C}^n$, on définit les suites $\nu \in \mathbb{C}^{\mathbb{N}}$ et $x \in (\mathbb{C}^n)^{\mathbb{N}}$ par les récurrences :

$$\nu_k = \langle x_k, Ax_k \rangle, \quad x_{k+1} = \frac{Ax_k}{\|Ax_k\|_2}, \quad k \geq 0.$$

De manière équivalente on a pour tout $k \geq 0$, $x_k = \dfrac{A^k x_0}{\|A^k x_0\|_2}$, l'algorithme étant bien défini si et seulement si la condition suivante est réalisée :

$$x_0 \notin \bigcup_{k \in \mathbb{N}} \ker(A^k) =: K.$$

Théorème 5 (*Convergence de la méthode de la puissance*)

Soit $A \in \mathsf{M}_n(\mathbb{C})$ dont on suppose que son spectre $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_p\}$ admet une valeur propre dominante : $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_p|$. Alors pour tout $x_0 \in \mathbb{C}^n \setminus \bigoplus_{i=2}^p \mathrm{Ker}(A - \lambda_i I_n)^{m_i}$, il existe $e \in \ker(A - \lambda_1 I_n)$ tel que

$$\lim_{k \to \infty} \nu_k = \lambda_1, \qquad \lim_{k \to \infty} \underbrace{\left(\frac{|\lambda_1|}{\lambda_1}\right)^k x_k}_{=:q_k} = e.$$

À noter que les hypothèses couvrent le cas de valeurs propres multiples et d'éventuels défauts de diagonalisabilité.

Démonstration. La condition de convergence sur la donnée initiale x_0 correspond à la conjonction du caractère bien défini de l'algorithme (domaine K exclus) et de la présence d'une composante pour x_0 dans $\ker(A - \lambda_1 I_n)$, au sens des projecteurs spectraux associés à A.
Par exemple si A est diagonalisable, on peut décomposer x_0 dans la base des vecteurs propres (e_i) :

$$x_0 = \sum_{i=1}^n \alpha_i e_i, \qquad A^k x_0 = \sum_{i=1}^n \lambda_i^k \alpha_i e_i, \quad k \in \mathbb{N}.$$

L'hypothèse sur x_0 garantit que $\alpha_1 \neq 0$ de sorte que par le caractère dominant de λ_1 on a :

$$x_k = \left(\frac{\lambda_1}{|\lambda_1|}\right)^k \frac{\alpha_1}{|\alpha_1|} e_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|\right)^k$$

□

Remarque (Vitesse de convergence). *Sous les hypothèses du théorème précédent, on peut préciser les comportements asymptotiques suivants :*

— *Si λ_1 et toute valeur propre de module égal à $|\lambda_2|$ sont non-défectives, alors*

$$\|q_k - e\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

— *Si λ_1 n'est pas défective mais qu'au moins une valeur propre de module égal à $|\lambda_2|$ est défective alors en notant $r \in [\![2, n]\!]$ la taille du plus grand bloc de Jordan associé concerné on a*

$$\|q_k - e\| = O\left(k^{r-1}\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

— *Si λ_1 est défective, alors*

$$\|q_k - e\| = O\left(\frac{1}{k}\right)$$

Le détail de la preuve du théorème et le calcul de la vitesse de convergence est laissée en exercice. Encore une fois, il faut s'appliquer à quantifier le comportement asymptotique des puissances de blocs de Jordan et décomposer A à travers ses sous-espaces caractéristiques (sous-espaces propres pour les valeurs propres semi-simples).

Remarque (Cas d'une matrice hermitienne)**.** *Si la matrice A est hermitienne (ou symétrique dans le cas réel) alors la vitesse de convergence est "doublée" grâce à l'orthogonalité des sous-espaces propres :*

$$|\nu_k - \lambda_1| \leq O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right).$$

Méthode de la puissance inverse avec translation Étant donnée $A \in \mathsf{M}_n(\mathbb{C})$, la méthode précédente se généralise de sorte à calculer d'autres valeurs propres de A, par exemple la valeur propre de plus petit module de A. Pour ce faire, il suffit d'appliquer l'algorithme à la matrice A^{-1} si $A \in \mathcal{GL}_n(\mathbb{C})$, ce qui revient en pratique à résoudre à chaque itération un système de la forme $Ax = b$.
En considérant des translations dans le plan complexe, on peut encore déterminer la valeur propre de A la plus proche d'un complexe $\mu \in \mathbb{C}$ donné, pourvu que les hypothèses du théorème s'appliquent à la matrice $(A - \mu I_n)^{-1}$. En particulier cela suppose que $\mu \notin \sigma(A)$ (ainsi que le caractère dominant de la valeur propre considérée).

2 Localisation du spectre

2.1 Théorème de Gershgörin-Hadamard

Théorème 6 (*de Hadamard*)

Soit $A \in \mathsf{M}_n(\mathbb{C})$ une matrice à diagonale strictement dominante, i.e. telle que

$$\forall i \in \{1, \ldots, n\},\ |a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Alors A est une matrice inversible.

Démonstration. Par l'absurde, supposons qu'il existe $x \in \ker A$ avec $x \neq 0$, on travaille à partir de $Ax = 0$ en isolant les termes diagonaux. Le détail est laissé en exercice.
Autre preuve : considérons la matrice inversible $D = \operatorname{diag}(A)$, et posons $B = I_n - D^{-1}A$ qui est à diagonale nulle avec de plus $\|B\|_\infty = \max_i \sum_{j \neq i} \frac{|a_{i,j}|}{|a_{i,i}|} < 1$ par l'hypothèse. Donc $I_n - B$ et donc A est inversible. □

Théorème 7 (*de Gershgörin*)

Soit $A \in \mathsf{M}_n(\mathbb{C})$. On note, pour $i \in \{1, \ldots, n\}$, $D_i = \left\{z \in \mathbb{C} \ /\ |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|\right\}$. Alors

$$\sigma(A) \subset \bigcup_{i=1}^{n} D_i.$$

Démonstration. En appliquant la contraposée du théorème de Hadamard à $A - \lambda I$, si $\lambda \in \sigma(A)$ alors $A - \lambda I$ n'est pas inversible et donc n'est pas à diagonale strictement dominante. Ainsi il existe un indice i_0 tel que $\lambda \in D_{i_0}$ d'où $\lambda \in \cup_i D_i$. □

Théorème 8 (*Deuxième de Gershgörin*)

Dans chaque composante connexe de $\cup_i D_i$, on compte exactement autant de valeurs propres de A

(comptées avec leur multiplicité algébrique) que de disques de Gershgörin.

Démonstration. C'est un résultat déduit d'un principe de continuité des racines d'un polynôme en fonction de ses coefficients (ou de continuité des valeurs propres en fonction des coefficients de la matrice). On abordera plus loin ce résultat utilisant le théorème de Rouché d'analyse complexe.
Il suffit alors d'appliquer ce principe au polynôme caractéristique de la matrice $B(t) \in \mathsf{M}_n(\mathbb{C})$ pour $t \in [0,1]$ par

$$B(t) = D + t(A - D), \qquad D = \mathrm{diag}(A).$$

vérifiant $B(1) = A$ et $\sigma(B(0)) = \{a_{ii},\ 1 \le i \le n\}$. Le relèvement continu des racines garantit qu'il existe n fonctions continues $\lambda_i : [0,1] \to \mathbb{C}$ telles que $\lambda_i(0) = a_{ii}$ et $\cup_i\{\lambda_i(1)\} = \sigma(A)$.
Soit $t \in [0,1]$. Par le premier théorème de Gershgörin appliqué à $B(t)$, après avoir noté $D_i(t) := \{z \in \mathbb{C}/\ |a_{ii} - z| \le t\sum_{j\ne i}|a_{ij}|\}$ les disques correspondant, on obtient :

$$\sigma(B(t)) = \bigcup_i\{\lambda_i(t)\} \subset K(t) := \bigcup_i D_i(t).$$

Notons à présent $I \subset \{1,\dots,n\}$ un sous-ensemble d'indices tel que $M := \cup_{i\in I}D_i(1)$ est une composante connexe de $K := K(1) = M \sqcup N$ avec, M et N étant fermés disjoints : $d(M,N) > 0$.
Chacune des applications $t \mapsto D_i(t)$ est croissante au sens de l'inclusion. Donc pour tout $t \in [0,1]$, $M(t) := \cup_{i\in I}D_i(t)$ est aussi un sous-ensemble de M, totalement disjoint de N.
Fixons $i \in I$, l'application continue λ_i est à valeurs dans $K = M \sqcup N$, initialement dans M et donc à valeurs dans M pour tout $t \in [0,1]$. En particulier $\{\lambda_i(1)\}_{i\in I} \subset M$ qui contient donc $\mathsf{card}(I)$ valeurs propres de A (avec multiplicité). □

Cas des matrices irréductibles Les résultats précédents admettent une version affaiblie dans le cas de matrices à diagonale fortement dominante et irréductibles. On ne traitera pas ici ces propriétés mais le lecteur pourra se reporter par exemple à [4, Sec. 4.5].

2.2 Continuité des valeurs propres par le théorème des fonctions implicites

Certains résultats de cette partie sont une adaptation de la présentation du livre de D. Serre [4, Sec. 3.1.1]. Ils portent sur la continuité des valeurs propres.

Proposition 9

Soit A une matrice de $\mathsf{M}_n(\mathbb{C})$ supposée diagonalisable et soit $\lambda \in \mathbb{C}$ une valeur propre simple de A. Alors il existe un voisinage ouvert $\mathcal{V}$ de A dans $\mathsf{M}_n(\mathbb{C})$ et une fonction $\mu : \mathcal{V} \to \mathbb{C}$ de classe $\mathcal{C}^\infty$ telle que

$$\begin{aligned}&\mu(A) = \lambda,\\ &\forall B \in \mathcal{V}, \quad \mu(B) \in \sigma(B).\end{aligned}$$

Démonstration. La preuve s'appuie sur le théorème des fonctions implicites. Définissons l'application suivante :

$$\begin{aligned}\psi : \mathsf{M}_n(\mathbb{C}) \times \mathbb{C} \times \mathbb{C}^n &\to \mathbb{C} \times \mathbb{C}^n\\ (B,\mu,Y) &\mapsto (\|Y\|_2^2 - 1, BY - \mu Y)\end{aligned}$$

Les zéros de cette application sont ni plus ni moins que les triplets (B,μ,Y) pour lesquels μ est valeur propre de B et Y un vecteur propre associé unitaire.
On se place au voisinage d'un triplet (A,λ,X) satisfaisant aux hypothèses requises par la proposition $\psi(A,\lambda,Y) = (0,0)$. Afin d mettre en œuvre le théorème des fonctions implicites pour exprimer les zeros de ψ sous la forme $(\mu,Y) = \phi(B)$, calculons la différentielle partielle de ψ par rapport à (μ,Y) :

$$\begin{aligned}\psi(A,\lambda+\mu,X+Y) &= (\|X+Y\|_2^2 - 1, A(X+Y) - (\lambda+\mu)(X+Y))\\ &= (\|X\|_2^2 - 1, AX - \lambda X) + (2X^\star Y, AY - \mu X - \lambda Y) + o(\mu,Y)\end{aligned}$$

Ainsi $d_{(\mu,Y)}\psi(A,\lambda,X) : (\mu, Y) \mapsto (2X^\star Y, AY - \mu X - \lambda Y)$ est un endomorphisme de $\mathbb{C} \times \mathbb{C}^n$ dont on doit à présent simplement vérifier l'inversibilité.
Supposons $(\mu, Y) \in \ker(d_{(\mu,Y)}\psi(A, \lambda, X))$. Cela se traduit par $2X^\star Y = 0$ et $AY - \mu X - \lambda Y = 0$. En appliquant $A - \lambda I$ à la deuxième égalité on obtient : $(A - \lambda I)^2 Y = \mu(A - \lambda I)X = 0$ puisque X est vecteur propre. Ainsi $Y \in \ker(A - \lambda I)^2 = \ker(A - \lambda I)$ puisque A est diagonalisable, et $\ker(A - \lambda I) = \mathsf{Vect}(X)$ puisque λ est valeur propre simple. Donc il existe $\alpha \in \mathbb{C}$ tel que $Y = \alpha X$. La première équation $X^\star Y = 0$ indique alors que $\alpha\|X\|_2^2 = 0$ et donc $\alpha = 0$. Par suite $Y = 0$ et $\mu = 0$.
Le théorème des fonctions implicites s'applique et garantit l'existence d'ouverts $\mathcal{V}$ dans $\mathsf{M}_n(\mathbb{C})$, J dans $\mathbb{C}$ et $\mathcal{O}$ dans $\mathsf{M}_n(\mathbb{C})$, d'une fonction ayant la même régularité que ψ, i.e. de classe $\mathcal{C}^\infty$, définie sur $\mathcal{V}$ à valeurs dans $J \times \mathcal{O}$ telle que

$$\begin{aligned} &A \in \mathcal{V},\ (\lambda, X) \in J \times \mathcal{O}, \\ &\forall (B, \mu, Y) \in \mathcal{V} \times J \times \mathcal{O},\ \psi(B, \mu, Y) = 0 \Leftrightarrow (\mu, Y) = \phi(B). \end{aligned}$$

□

Le résultat précédent est seulement local, et tombe en défaut dès qu'on approche une matrice comportant par exemple un bloc de Jordan non-trivial, le contre-exemple suivant illustre ce fait explicitement.

Exemple (Contre-exemple). *On considère la fonction de classe $\mathcal{C}^\infty(\mathbb{R}, \mathcal{M}_2(\mathbb{R}))$, définie par*

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

On a facilement : $\sigma(A(x)) = \{-\sqrt{x}, +\sqrt{x}\}$ si $x \geq 0$ et $\sigma(A(x)) = \{-i\sqrt{-x}, i\sqrt{-x}\}$ si $x \leq 0$. On observe que le spectre est continuement paramétrable au voisinage de $x = 0$ mais pas de manière $\mathcal{C}^1$.

Un exemple plus général, dans $\mathcal{C}^\infty(\mathbb{R}, \mathcal{M}_n(\mathbb{R}))$, on définit la matrice

$$A_n(x) = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ x & & & 0 \end{pmatrix}.$$

dont le spectre est l'ensemble des racines n-ièmes de x.

$$\sigma(A_n(x)) = \{z \in \mathbb{C} \ t.q.\ z^n = x\}.$$

2.3 Continuité des valeurs propres par l'analyse complexe

Dans la suite de cette partie, nous proposons une approche plus globale qui permet de démontrer le résultat de continuité que l'on peut déjà observer sur le contre-exemple précédent. Bien sûr, en contrepartie on n'obtiendra pas la régularité plus précise des valeurs propres comme précédemment.

Lemme 10 (*Une conséquence du théorème de Rouché*)

Soit un entier $n \in \mathbb{N}$. On munit $\mathbb{C}_n[X]$ de la norme notée $\|\cdot\|$ définie par $\|P\| = \max_{0\leq j\leq n} |p_j|$ où les (p_j) sont les coefficients de P dans la base canonique.
Fixons $P \in \mathbb{C}_n[X]$ non-constant et $x \in \mathbb{C}$ une de ses racines, de multiplicité notée $\mu \in \mathbb{N}^\star$.
Notons $\eta > 0$ la distance de x à l'ensemble des autres racines de P et $D = \{z \in \mathbb{C} : |z - x| \leq \rho\}$ où $\rho \in]0, \eta[$ est fixé.
Alors il existe $\delta > 0$ tel que tout polynôme $Q \in \mathbb{C}_n[X]$ à distance au plus δ de P compte exactement μ racines dans D (avec leur multiplicité).

Démonstration. Avec les hypothèses annoncées, notons γ le lacet orienté correspondant à la frontière du disque ∂D. Puisque P ne s'annule pas sur γ et que γ est un domaine borné, il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que

$$\forall z \in \gamma \quad |P(z)| \geq \alpha, \quad \sum_{j=0}^{n} |z|^j \leq \beta.$$

Par des majorations rudimentaires, on obtient alors pour tout $z \in \gamma$, la majoration $|P(z) - Q(z)| \leq \|P - Q\|\beta$ de sorte que pourvu que l'on choisisse $\delta > 0$ vérifiant $\delta < \alpha/\beta$ alors pour tout $z \in \gamma$ et tout $Q \in \mathbb{C}_n[X]$ tel que $\|P - Q\| \leq \delta$:

$$|P(z) - Q(z)| < |P(z)|.$$

Le théorème de Rouché assure alors que le nombre de zéros de P et Q dans D sont égaux. □

Une difficulté subsiste pour en arriver à la continuité des valeurs propres, qui concerne la topologie retenue pour les ensemble spectraux, afin de tenir compte des éventuelles multiplicités. Pour ce faire, on définit l'application $\tilde{\sigma}$ qui à une matrice $A \in \mathsf{M}_n(\mathbb{C})$ associe le n-uplet $\tilde{\sigma}(A) \in \mathbb{C}^n$ de ses valeurs propres, chacune étant répétée avec sa multiplicité algébrique. Afin de supprimer l'arbitraire, on se place plutôt au niveau du quotient $\mathbb{C}^n/\sim$ de $\mathbb{C}^n$ par la relation d'équivalence induite par les permutations d'indices. Autrement dit : étant donnés $a, b \in \mathbb{C}^n$, $a \sim b$ si et seulement si il existe une permutation s de $\{1, \ldots, n\}$ telle que pour tout i, $a_{s(i)} = b_i$, on note alors $\mathrm{cl}\,(a)$ la classe d'équivalence correspondante.
Le résultat de continuité ne concerne pas à proprement parler l'application de spectre σ mais plutôt celle qui à A associe la classe d'équivalence $\mathrm{cl}\,(\tilde{\sigma}(A))$ qui tient compte des multiplicités algébriques éventuelles. Il reste à préciser la topologie pour laquelle on est en mesure de quantifier le résultat.

On munit $\mathbb{C}^n/\sim$ de la distance suivante, qui est la restriction de la distance de Hausdorff sur les parties de $\mathbb{C}$ aux ensembles à n éléments avec répétition. Étant donnés $a, b \in \mathbb{C}^n/\sim$,

$$d_H(a, b) = \min_{s,t \in \Sigma_n} \max_{1 \leq i \leq n} |a_{s(i)} - b_{t(i)}|.$$

Un corrolaire de la proposition précédente est le suivant :

Théorème 11 (*Continuité des valeurs propres*)

L'application $\mathrm{cl}\,(\tilde{\sigma}(\cdot))$ est continue de $\mathsf{M}_n(\mathbb{C})$ muni de la topologie induite par une norme (subordonnée par exemple), dans $\mathbb{C}^n/\sim$ muni de la topologie pour la distance d_H.

Démonstration. L'application $\chi : \mathsf{M}_n(\mathbb{C}) \to \mathbb{C}_n[X]$ qui à une matrice associe son polynôme caractéristique est clairement continue, en particulier en munissant $\mathbb{C}_n[X]$ de la topologie induite par la distance $d(P, Q) = \|P - Q\|$ introduite précédemment. Fixons désormais un polynôme $P \in \mathbb{C}_n[X]$ et notons $\eta = \min\{|x - y|,\ P(x) = P(y) = 0,\ x \neq y\}$ la plus petite distance entre deux de ses racines.
Soit $\epsilon > 0$ qu'on peut supposer inférieur à $\eta/3$. D'après le lemme précédent, pour toute racine x de P, il existe alors $\delta_x > 0$ tel que pour tout polynôme $Q \in \mathbb{C}_n[X]$ vérifiant $\|P - Q\| < \delta_x$ l'ensemble $D(x, \epsilon)$ contient autant de zéros de P avec multiplicité (i.e. seulement x avec sa multiplicité algébrique μ_x) que de zéros de Q comptés avec multiplicité.
Pour s'abstraire de la dépendance apparente en x, on peut retenir $\delta = \min_x \delta_x$ de sorte que pour tout polynôme $Q \in \mathbb{C}_n[X]$ vérifiant $\|P - Q\| < \delta$, chacun des ensemble $D(x, \epsilon)$ contienne autant de zéros de Q que de zéros de P (i.e. μ_x). Par ailleurs, en comptant les racines ainsi constituées, Q n'admet pas de racines en dehors des $D(x, \epsilon)$, où x décrit les racines de P.
Par conséquent en notant Z_P le n-uplet des zéros de P avec multiplicité et Z_Q celui de Q, on peut majorer le $\min_{s,t\in\Sigma_n}$ par le choix de permutation obtenue via la propriété précédente, à savoir :

$$d_H(\mathrm{cl}\,(Z_P)\,, \mathrm{cl}\,(Z_Q)) \leq \epsilon.$$

Un dessin éclairera sans doute le lecteur ! □

3 Conditionnement du problème aux valeurs propres

Le théorème suivant permet de préciser, au moins dans le cas d'une matrice diagonalisable, la manière donc le spectre dépend de cette matrice. On a vu dans les contre-exemples précédents que sans cette hypothèse, on perd généralement le caractère lipschitizien.

Théorème 12 (*Bauer-Fike*)

On munit $\mathbb{C}^n$ d'une norme d'espace vectoriel et $\mathsf{M}_n(\mathbb{C})$ de la norme d'algèbre induite. Soit $A \in \mathsf{M}_n(\mathbb{C})$ une matrice supposée diagonalisable, $A = PDP^{-1}$. Soit $E \in \mathsf{M}_n(\mathbb{C})$ une perturbation donnée. Alors

$$\forall \mu \in \sigma(A+E), \quad d(\mu, \sigma(A)) \leq \mathrm{cond}(P)\|E\|.$$

Démonstration. Supposons $\mu \notin \sigma(A)$, sans quoi le résultat est immédiat. On peut alors écrire $A + E - \mu I = P(D-\mu I)[I + (D-\mu I)^{-1}P^{-1}EP]P^{-1}$ de sorte que si $\mu \in \sigma(A+E)$ alors -1 est valeur propre de $C = (D-\mu I)^{-1}P^{-1}EP$ et donc

$$1 \leq \rho(C) \leq \|(D-\mu I)^{-1}\|\|E\|\|P\|\|P^{-1}\|.$$

Le résultat suit facilement, en raison du fait que $D - \mu I$ est diagonale et donc $\|(D-\mu I)^{-1}\| = (\min_{\lambda \in \sigma(A)} |\lambda - \mu|)^{-1}$. □

Remarque. *Avec les notations précédemment introduites, on a donc*

$$d_H(\mathrm{cl}\,(\tilde{\sigma}(A+E)), \mathrm{cl}\,(\tilde{\sigma}(A))) \leq \mathrm{cond}(P)\|E\|.$$

Lorsque A n'est pas diagonalisable, le dernier point de la preuve fait défaut.
Ce théorème révèle par ailleurs une propriété essentielle des matrices normales pour lesquelles la constante multiplicative est optimale (en retenant la norme hermitienne $\|\cdot\|_2$ sur $\mathbb{C}^n$). On a en effet $\mathrm{cond}(P) = \|P\|_2\|P^{-1}\|_2 = 1$ lorsque P est unitaire.

Se pose également la question de la précision du spectre d'une matrice, ayant obtenu au préalable une approximation de $Ax = \lambda x$.

Proposition 13 (*Estimation a posteriori*)

Soit $A \in \mathsf{M}_n(\mathbb{C})$ hermitienne et soit $(\tilde{\lambda}, \tilde{x}) \in \mathbb{C} \times \mathbb{C}^n$ l'approximation d'un couple propre (λ, x) de A. On définit le résidu par $\tilde{r} = A\tilde{x} - \tilde{\lambda}\tilde{x}$. Alors on a

$$d(\tilde{\lambda}, \sigma(A)) \leq \frac{\|\tilde{r}\|_2}{\|\tilde{x}\|_2}.$$

Démonstration. Notons (u_i) une base orthonormée de vecteurs propres de A et (λ_i) les valeurs propres correspondantes. Décomposons $\tilde{x}$ sous la forme $\tilde{x} = \sum \alpha_i u_i$ de sorte qu'on peut obtenir

$$\tilde{r} = A\tilde{x} - \tilde{\lambda}\tilde{x} = \sum \alpha_i(\lambda_i - \tilde{\lambda})u_i,$$

et donc

$$\|\tilde{r}\|_2^2 = \sum |\alpha_i|^2|\lambda_i - \tilde{\lambda}|^2, \quad \|\tilde{x}\|_2^2 = \sum |\alpha_i|^2.$$

et par suite

$$\frac{\|\tilde{r}\|_2^2}{\|\tilde{x}\|_2^2} \geq \min_i |\lambda_i - \tilde{\lambda}|^2.$$

□

Remarque (Admis [3])**.** *Dans le cas plus général d'une matrice seulement supposée diagonalisable, avec $A = PDP^{-1}$. Avec les notations précédentes, notons $\epsilon > 0$ tel que $\|\tilde{r}\|_2 \leq \epsilon\|\tilde{x}\|_2$, alors $d(\tilde{\lambda}, \sigma(A)) \leq \epsilon\ \mathrm{cond}(P)$.*

Références

[1] P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod, 2007.

[2] Xavier Gourdon, *Algèbre*, Ellipses, 1994.

[3] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri, *Méthodes numériques. Algorithmes, analyse et applications*, Springer, Milan, 2007.

[4] Denis Serre, *Les matrices : théorie et pratique*, Masson, 2000.

[5] Rached Mneimné, *Réduction des endomorphismes*, Calvage & Mounet, 2006.

[6] Henri Roudier, *Algèbre linéaire*, Vuibert, 2003.

[7] Félix Gantmacher, *Théorie des matrices*, Ed. Jacques Gabay, 2000.