Valeurs propres

Agrégation de mathématiques
Option modélisation

Septembre 2017

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

Rappels

On se placera ici systématiquement sur le corps algébriquement clos $\mathbb{K} = \mathbb{C}$ de sorte que tout polynôme de $\mathbb{K}[X]$ soit scindé et donc tout endomorphisme sur $\mathbb{K}^n$ trigonalisable. Pour débuter, on rappelle les 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

$$\forall i \in \{1, \dots, n-1\},\ \forall j \in \{i+2, \dots, n\}, \quad T_{i,i+1} \in \{0,1\},\ T_{i,j} = 0.$$

Démonstration (principe) : L'idée est d'obtenir le résultat de réduction de Jordan dans le cas d'endomorphismes nilpotents, cela se fait par les propriétés d'inclusion des noyaux itérés et la construction d'une base adaptée à la forme de Jordan. Cette réduction particulière s'applique ici à la restriction de A à chacun de ses sous-espaces caractéristiques (stables par A) : $\ker(A - \lambda_i I)^{\alpha_i}$.
Pour une preuve détaillée, on pourra se reporter par exemple au livre de Gourdon [2]. ■

En conséquence, toute matrice $A \in \mathsf{M}_n(\mathbb{C})$ est semblable à une matrice diagonale par blocs dont chaque bloc prend la forme d'un bloc de Jordan :

$$T = P^{-1}AP = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{k_m}(\lambda_m) \end{pmatrix}.$$

On a ici noté les valeurs propres $(\lambda_i)_{1 \le i \le m}$ de A, qui ne sont pas nécessairement distinctes deux à deux, et pour tout $i \in [\![1, m]\!]$, $J_{k_i}(\lambda_i)$ désigne un bloc de Jordan de taille $k_i \geq 1$ prenant la forme

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

Dans le cas diagonalisable, tous les blocs sont de taille $k = 1$. À noter que dans l'écriture par bloc ci-dessus, les différents blocs de Jordan révèlent immédiatement une décomposition de Dunford de A (comme somme d'une matrice diagonalisable et d'une matrice nilpotente qui commutent) : $J_k(\lambda) = \lambda I_k + J_k(0)$.

Un résultat intéressant pour pouvoir 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 (principe) : 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 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. ■

1 Rayon spectral, résultats élémentaires

Définition 3

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

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

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

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

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

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

Application à la localisation des racines d'un polynôme :

En s'appuyant sur les *matrices compagnons* de polynômes et les normes matricielles explicites suivantes,

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

on obtient les estimations suivantes sur les racines $z \in \mathbb{C}$ d'un polynôme $p \in \mathbb{C}_n[X]$ de coefficient dominant égal à 1, 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 (principe) : 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. ■

Application à la 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 $M \in \mathsf{GL}_n(\mathbb{C})$:

$$x_0 \in \mathbb{C}^n$$
$$\forall k \geq 0,\ x_{k+1} = M^{-1}Nx_k + M^{-1}b.$$

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 $\rho(M^{-1}N) < 1$. On peut le vérifier en notant $e_k = x_k - x$ l'erreur à l'étape $k \geq 0$ qui vérifie : $e_k = (M^{-1}N)^k(x_0 - x)$. 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 $(M^{-1}N)^k$ tende vers 0 lorsque k tend vers l'infini, ce qui est équivalent à la propriété mentionnée.

Application à la stabilité d'une méthode linéaire récurrente : Considérons que la suite définie dans l'exemple précédent soit 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 :

$$y_0 = x_0 + \eta_0 \in \mathbb{C}^n$$
$$\forall k \geq 0,\ y_{k+1} = M^{-1}Ny_k + M^{-1}b + \epsilon_k.$$

On peut examiner la convergence de (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 :

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

L'hypothèse $\rho(M^{-1}N) < 1$ garantit qu'on peut se donner une norme subordonnée $\|\cdot\|$ sur $\mathsf{M}_n(\mathbb{C})$ et la norme vectorielle attenante sur $\mathbb{C}^n$ telle que $\rho(M^{-1}N) \leq C = \|M^{-1}N\| < 1$ et alors

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

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.

2 Théorème de Gershgörin-Hadamard

Théorème 5 (*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.

Théorème 6 (*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 du théorème de Gershgörin : 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 $\lambda \in \cup_i D_i$. ■

Démonstration du théorème de Hadamard : 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. ■

Théorème 7 (*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 une preuve de ce résultat utilisant le théorème de Rouché d'analyse complexe. Il suffit alors d'appliquer ce principe au polynôme caractéristique de la matrice $B(x) \in \mathsf{M}_n(\mathbb{C})$ définie à travers le paramètre $x \in [0, 1]$ par

$$B(x)_{ij} = \begin{cases} A_{ij}, & i = j \\ xA_{ij}, & i \neq j \end{cases}.$$

On a bien sûr $B(1) = A$ et $\sigma(B(0)) = \{a_{ii},\ 1 \leq i \leq n\}$. 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 $\cup_i\{\lambda_i(1)\} = \sigma(A)$.
Appliquons le premier théorème de Gershgörin à $B(x)$, après avoir noté $D_i(x) = \{z \in \mathbb{C}/\ |a_{ii} - z| \leq x \sum_{j \neq i} |a_{ij}|\}$ les disques correspondant :

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

Chacune des applications $x \mapsto D_i(x)$ et de même l'application $x \mapsto K(x)$ sont croissantes au sens de l'inclusion. En conséquence, si on note $I \subset \{1, \dots, n\}$ tel que $M = \cup_{i \in I} D_i(1)$ est une composante connexe de $K(1)$, alors pour tout $x \in [0, 1]$, $\cup_{i \in I} D_i(x)$ est aussi un sous-ensemble de M, totalement disjoint des autres composantes connexes de $K(1)$. En conséquence, les applications continues $(\lambda_i)_{i \in I}$, qui sont à valeurs dans $K(1)$ et initialement dans M, sont également à valeurs dans M pour tout $x \in [0, 1]$. 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 [5, Sec. 4.5].

3 Continuité des valeurs propres

Certains résultats de cette partie sont une adaptation de la présentation du livre de D. Serre [5, Sec. 3.1.1]. Ils portent sur la continuité des valeurs propres.

Proposition 8

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

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

Démonstration : La preuve repose sur l'emploi du théorème des fonctions implicites. Définissons l'application suivante :

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

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. De sorte à mettre en œuvre le théorème des fonctions implicites pour exprimer ces zeros sous la forme $(\mu, Y) = \phi(B)$, calculons la différentielle partielle de ψ en le couple (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 $(\mu, Y) \in \ker(d_{(\mu,Y)}\psi(A, \lambda, X))$. Alors $2X^\star Y = 0$ 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 cette observation.

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 $\mathcal{C}^1$.

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 de résultat de régularité plus forte sur les valeurs propres comme par la méthode précédente.

Lemme 9 (*Une conséquence du théorème de Rouché*)

Soit un entier $n \in \mathbb{N}$. On munit $\mathbb{C}_n[X]$ de la norme notée $\|\cdot\|$, obtenue par $\|P\| = \max_{0 \leq j \leq n} |p_j|$ où les (p_j) sont les coefficients de P dans la base canonique.
Fixons $P \in \mathbb{C}_n[X]$ 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 ∂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 ailleurs par des majorations élémentaires on obtient $|P(z) - Q(z)| \leq \|P - Q\|\beta$ 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 avec laquelle on travaille sur les spectres, 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 $\{1, \dots, n\}$ 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 10 (***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 ! ■

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 :

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

De manière équivalente on a pour tout $k \geq 0$, $x_k = \dfrac{A^k x_0}{\|A^k x_0\|}$, ceci pourvu que

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

ce qui constitue une condition nécessaire et suffisante pour que l'algorithme soit bien défini.

Théorème 11 (***Convergence de la méthode de la puissance***)

Soit $A \in \mathsf{M}_n(\mathbb{C})$ dont on suppose qu'elle admet une valeur propre dominante :
$\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_p\}$ avec $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_p|$, chaque valeur propre pouvant être multiple et la matrice A non-diagonalisable.
Alors pour "presque tout" $x_0 \in \mathbb{C}^n$, il existe $e \in \ker(A - \lambda_1 I_n)$ tel que

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

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. ■

Remarque 1 (***Vitesse de convergence***)

Sous les hypothèses du théorème précédent, on a 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 (propres dans le cas diagonalisable).

Remarque 2 (***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).

5 Conditionnement du problème aux valeurs propres

Commençons par quelques rappels sur les matrices normales qui joueront dans la suite un rôle particulier. On munit $\mathbb{C}^n$ de la structure d'espace hilbertien complexe pour le produit scalaire hermitien canonique, dont on rappelle le lien avec l'opération d'adjonction complexe par la propriété

$$\forall A \in \mathsf{M}_n(\mathbb{C}), \forall x, y \in \mathbb{C}^n, \quad \langle Ax, y\rangle = \langle x, A^\star y\rangle.$$

Définition 12

Une matrice $A \in \mathsf{M}_n(\mathbb{C})$ est dite **normale** *si elle commute avec son adjoint : $A^\star A = AA^\star$.*

Parmi les matrices normales, on trouve en particulier
- les matrices symétriques réelles ou hermitiennes, dont les valeurs propres sont réelles
- les matrices antisymétriques réelles ou antihermitiennes, dont les valeurs propres sont imaginaires pures
- les matrices orthogonales réelles ou unitaires, dont les valeurs propres sont de module 1.

Une excellente caractérisation des matrices normales est la suivante.

Proposition 13

Soit $A \in \mathsf{M}_n(\mathbb{C})$. Il y a équivalence entre :
- *A diagonalise en base orthonormée,*
- *A est normale.*

Démonstration : Le sens direct est immédiat. La réciproque peut s'obtenir par exemple facilement via le théorème de Schur. On a donc $A = PTP^\star$, avec T triangulaire et P unitaire. Par suite, les égalités $A^\star A = AA^\star$ permettent d'obtenir que $T^\star T = TT^\star$. Mais toute matrice triangulaire et normale est nécessairement diagonale (exercice !), d'où le résultat. ■

Le théorème suivant permet de préciser, au moins dans le cas d'une matrice diagonalisable, la manière donc le spectre dépend de cette matrice.

Théorème 14 (*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 3

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 du spectre d'une matrice, ayant obtenu au préalable une approximation de $Ax=\lambda x$.

Proposition 15 (*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

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

et par suite

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

Remarque 4 (*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] Michelle Schatzman, *Analyse numérique : cours et exercices pour la licence*, Dunod, 1991.

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