

6 Valeurs et Vecteurs Propres

Les premiers vecteurs et valeurs propres viennent des équations différentielles (Lagrange 1759, *théorie du son*; Lagrange 1781, des matrices 6×6 dans le but de calculer les perturbations séculaires des orbites des 6 planètes connues à l'époque, *Oeuvres V*, p. 125-490). Aujourd'hui, le calcul des valeurs et vecteurs propres est indispensable dans toutes les branches de la science, en particulier pour la solution des systèmes des équations différentielles linéaires, en théorie de stabilité, pour les questions de convergence de processus itératifs, et en physique et chimie (mécanique, circuits, cinétique chimique, équation de Schrödinger). FIG. V.1 : Une application

linéaire comme champ de vecteurs (à gauche); transformée sur la base des vecteurs propres (à droite). Observons en figure V.1 (à gauche) le champ de vecteurs d'une équation différentielle

$y' = Ay$. Deux directions sautent aux yeux : ce sont les directions où le vecteur Av prend la même direction que le vecteur v , c'est-à-dire, où

$$Av = \lambda v \quad \text{ou} \quad (A - \lambda I)v = 0 \quad (6.1)$$

Si cette équation est vérifiée, $\lambda \in \mathbb{C}$ s'appelle *valeur propre* de la matrice A et $v \in \mathbb{C}^n (v \neq 0)$ est le *vecteur propre* correspondant. L'équation (6.1) possède une solution v non nulle si et seulement si

$$P_A(\lambda) = \det(A - \lambda I) = 0$$

Le polynôme $P_A(\lambda)$ est le *polynôme caractéristique* de la matrice A . Les valeurs propres de A sont alors les zéros du polynôme caractéristique.

7 La condition du calcul des valeurs propres

A cause des erreurs d'arrondi, les éléments d'une matrice A , pour laquelle on cherche les valeurs propres, ne sont pas exacts. Ils sont plutôt égaux à

$$\tilde{a}_{ij} = a_{ij}(1 + \varepsilon_{ij}) \quad \text{avec} \quad |\varepsilon_{ij}| \leq \text{eps}$$

(*eps* étant la précision de l'ordinateur, est supposée être très petite). Il est alors très important d'étudier l'influence de ces perturbations sur les valeurs propres et sur les vecteurs propres de la matrice. Pour montrer ceci, considérons la famille de matrices

$$A(\varepsilon) = A + \varepsilon C \quad \text{où} \quad |\varepsilon| \leq \text{eps} \quad \text{et} \quad |c_{ij}| \leq |a_{ij}|$$

(souvent, la dernière hypothèse va être remplacée par $\|C\| \leq \|A\|$).

Théorème 236 (Gershgorin). Soit A une matrice $n \times n$ (avec des éléments dans \mathbb{R} ou dans \mathbb{C}).

a) Si λ est une valeur propre de A , alors il existe un indice i tel que

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

c'est-à-dire, que toutes les valeurs propres de A se trouvent dans l'union des disques $D_i = \left\{ \lambda; |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}$

b) Si une composante connexe de $\bigcup_{i=1}^n D_i$ consiste de k disques, elle contient exactement k valeurs propres de A .

Démonstration. Soit $v \neq 0$ un vecteur propre et choisissons l'indice i tel que $|v_i| \geq |v_j|$ pour tout j . La ligne i de l'équation $Av = \lambda v$ donne

$$\sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} v_j = (\lambda - a_{ii}) v_i.$$

En divisant par v_i et en utilisant l'inégalité du triangle, on obtient U

$$|\lambda - a_{ii}| = \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} \frac{v_j}{v_i} \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

L'affirmation (b) est vraie si A est une matrice diagonale. Le cas général est obtenu par un argument de continuité en faisant tendre les éléments en dehors de la diagonale vers zéro. cqfd

Théorème 237. Soit A une matrice diagonalisable, c'est-à-dire, il existe P avec $P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ et soit $A(\varepsilon) = A + \varepsilon C$. Alors, pour chaque valeur propre $\lambda(\varepsilon)$ de $A(\varepsilon)$, il existe un λ_i avec

$$|\lambda(\varepsilon) - \lambda_i| \leq \varepsilon \cdot \kappa_\infty(P) \cdot \|C\|_\infty$$

Démonstration. Nous transformons la matrice $A(\varepsilon) = A + \varepsilon C$ par la même matrice, qui transforme A sous forme diagonale :

$$P^{-1}A(\varepsilon)P = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) + \varepsilon P^{-1}CP$$

Si l'on dénote par e_{ij} les éléments de $P^{-1}CP$, le théorème de Gershgorin implique l'existence d'un indice i tel que $|\lambda(\varepsilon) - (\lambda_i + \varepsilon e_{ii})| \leq \varepsilon \sum_{j \neq i} |e_{ij}|$. L'inégalité triangulaire donne alors

$$|\lambda(\varepsilon) - \lambda_i| \leq \varepsilon \cdot \max_i \left(\sum_j |e_{ij}| \right) \leq \varepsilon \cdot \|P^{-1}CP\|_\infty \leq \varepsilon \cdot \|P^{-1}\|_\infty \cdot \|C\|_\infty \cdot \|P\|_\infty$$

ce qui démontre l'affirmation du théorème, car $\kappa_\infty(P) = \|P^{-1}\|_\infty \cdot \|P\|_\infty$ (condition de T). cqfd

Remarque 238. La condition du calcul des valeurs propres dépend de la condition de la matrice de transformation P . Si la matrice A est symétrique (P est orthogonale), le problème est bien conditionné. Toutefois, observons qu'on obtient seulement une estimation pour l'erreur absolue et non pour l'erreur relative.

Théorème 239 (différentiabilité des valeurs propres). Soit λ_1 une racine simple de $P_A(\lambda) = 0$. Alors, pour $|\varepsilon|$ suffisamment petit, la matrice $A(\varepsilon) = A + \varepsilon C$ possède une valeur propre unique $\lambda_1(\varepsilon)$ proche de λ_1 . La fonction $\lambda_1(\varepsilon)$ est différentiable (même analytique) et on a

$$\lambda_1(\varepsilon) = \lambda_1 + \varepsilon \frac{u_1^* C v_1}{u_1^* v_1} + O(\varepsilon^2) \quad (7.1)$$

où v_1 est le vecteur propre à droite ($Av_1 = \lambda v_1$) et u_1 est le vecteur propre à gauche ($u_1^* A = \lambda u_1^*$). On peut supposer que $\|v_1\| = \|u_1\| = 1$

Démonstration. Soit $p(\lambda, \varepsilon) = P_{A+\varepsilon C}(\lambda) = \det(A + \varepsilon C - \lambda I)$. Comme

$$p(\lambda_1, 0) = 0 \quad \text{et} \quad \frac{\partial p(\lambda_1, 0)}{\partial \lambda} \neq 0$$

le théorème des fonctions implicites garantit l'existence d'une fonction différentiable $\lambda_1(\varepsilon)$ (même analytique), tel que $\lambda_1(0) = \lambda_1$ et $p(\lambda_1(\varepsilon), \varepsilon) = 0$. Il existe donc un vecteur $v_1(\varepsilon)$ tel que

$$(A(\varepsilon) - \lambda_1(\varepsilon)I)v_1(\varepsilon) = 0. \quad (7.2)$$

La matrice dans (11.2) étant de rang $n - 1$, on peut fixer une composante à 1 et appliquer la règle de Cramer. Ceci montre que les autres composantes sont des fonctions rationnelles des éléments de la matrice $A + \varepsilon C - \lambda_1(\varepsilon)I$ et donc différentiables. Après la normalisation à $v_1(\lambda)^T v_1(\lambda) = 1$, la fonction $v_1(\lambda)$ reste différentiable. Pour calculer $\lambda_1'(0)$, nous pouvons dériver l'équation (11.2) par rapport à ε et poser ensuite $\varepsilon = 0$. Ceci donne

$$(A - \lambda_1 I)v_1'(0) + (C - \lambda_1'(0)I)v_1 = 0 \quad (7.3)$$

En multipliant cette relation par u_1^* , on obtient $u_1^*(C - \lambda_1'(0)I)v_1 = 0$, ce qui permet de calculer $\lambda_1'(0)$ et démontre la formule (11.1). cqfd

Conséquences. La formule (11.1) du théorème précédent montre que plus le vecteur propre de droite est parallèle au vecteur propre de gauche, mieux la valeur propre correspondante est bien conditionnée (par exemple, pour les matrices symétriques les deux vecteurs sont identiques); plus ils se rapprochent de l'orthogonalité, plus la valeur propre est mal conditionnée. Si la matrice n'est pas symétrique (ou normale), le calcul de λ_1 (valeur propre simple) peut être mal conditionné. Considérons par exemple la matrice

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 2 \end{pmatrix} \quad \text{où} \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_1 = \frac{1}{\sqrt{1 + \alpha^2}} \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$$

Dans cette situation, la formule (11.1) nous donne $\lambda_1(\varepsilon) - \lambda_1 = \varepsilon.(c_{11} - \alpha c_{21}) + O(\varepsilon^2)$ et le calcul de $\lambda_1 = 1$ est mal conditionné si α est grand. Exemple 1.4 Considérons la matrice (boîte de Jordan)

$$A = \left(\begin{array}{cccc} \lambda_1 & 1 & & \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{array} \right) \Bigg\} n \quad (7.4)$$

Le polynôme caractéristique de $A + \varepsilon C$ satisfait

$$\det(A + \varepsilon C - \lambda I) = (\lambda_1 - \lambda)^n - (-1)^n . \varepsilon . c_{n1} + O(\varepsilon^2) + O(\varepsilon . |\lambda_1 - \lambda|).$$

Si $c_{n1} \neq 0$, les termes $O(\varepsilon^2)$ et $O(\varepsilon . |\lambda_1 - \lambda|)$ sont négligeables par rapport à ε . Les valeurs propres de $A + \varepsilon C$ sont alors approximativement données par les racines de

$$(\lambda_1 - \lambda)^n - (-1)^n . \varepsilon . c_{n1} = 0 \quad (7.5)$$

c'est-à-dire $\lambda = \lambda_1 + (\varepsilon . c_{n1})^{1/n}$ (observer que $(\varepsilon . c_{n1})^{1/n}$ donne n valeurs complexes distinctes - multiples des racines de l'unité). **Expérience numérique.** Prenons la matrice (11.4) avec $\lambda_1 = 1$ et $n = 5$. Les éléments de la matrice C sont des nombres aléatoires dans l'intervalle $[-1, 1]$. Le dessin 7 ci-contre montre les 5 valeurs propres de $A + \varepsilon C$ pour $\varepsilon = 10^{-4}, 10^{-5}, \dots, 10^{-10}$. L'erreur est $\approx 10^{-1}$ pour $\varepsilon = 10^{-5}$ et $\approx 10^{-2}$ pour $\varepsilon = 10^{-10}$, ce qui correspond à la formule (11.5) pour $n = 5$. **Conséquence.** Si la dimension n d'une boîte de Jordan est plus grande que 1, le calcul de la valeur propre de cette matrice est très mal conditionné.

7.1 Condition du calcul des vecteurs propres

Considérons la situation où toutes les valeurs propres de A sont distinctes. La démonstration du théorème sur la différentiabilité des valeurs propres montre (voir formule (11.2)) que les vecteurs propres normalisés $v_i(\varepsilon)$ de $A + \varepsilon C$ sont des fonctions différentiables de ε . Pour étudier la condition du calcul des vecteurs propres, nous exprimons $v_1'(0)$ dans la base des vecteurs propres (de droite)'

$$v_1'(0) = \sum_{i=1}^n \alpha_i v_i. \quad (7.6)$$

La formule (11.3) donne alors

$$\sum_{j=2}^n (\lambda_j - \lambda_1) \alpha_j v_j + (C - \lambda_1'(0)I) v_1 = 0. \quad (7.7)$$

En multipliant (11.7) par le vecteur propre de gauche u_1^* (observer que $u_1^* v_1 = 0$ pour $i \neq j$), on obtient α_i (pour $i \geq 2$) de la relation $(\lambda_i - \lambda_1) \alpha_i u_i^* v_i + u_i^* C v_1 = 0$. La normalisation $\|v_1(\varepsilon)\|_2^2 = 1$ donne (en la dérivant) $v_1^* v_1'(0) = 0$ et on en déduit que $\alpha_1 = -\sum_{j=2}^n \alpha_j v_1^* v_j$. Si l'on insère les formules pour α_i dans (11.6), on obtient pour $v_1(\varepsilon) = v_1 + \varepsilon v_1'(0) + O(\varepsilon^2)$ la relation

$$v_1(\varepsilon) = v_1 + \varepsilon \sum_{j=2}^n \frac{u_j^* C v_1}{(\lambda_1 - \lambda_j) u_j^* v_j} (v_j - v_1 v_1^* v_j) + O(\varepsilon^2). \quad (7.8)$$

De cette formule, on voit que la condition du calcul du vecteur propre v_1 dépend de la grandeur $u_i^* v_i$ (comme c'est le cas pour la valeur propre; voir la formule (11.1)) et aussi de la distance entre λ_1 & et les autres valeurs propres de A . **Un algorithme dangereux** La première méthode (déjà utilisée par Lagrange) pour calculer les valeurs propres d'une matrice A est la suivante : *calculer d'abord les coefficients du polynôme caractéristique $P_A(\lambda)$ et déterminer ensuite les zéros de ce polynôme*. Si la dimension de A est très petite (disons $n \leq 3$) ou si l'on fait le calcul en arithmétique exacte, cet algorithme peut être très utile. Par contre, si l'on fait le calcul en virgule flottante, cet algorithme peut donner des mauvaises surprises. Considérons, par exemple, le problème de calculer les valeurs propres de la matrice diagonale

$$A = \text{diag}(1, 2, 3, \dots, n)$$

dont le polynôme caractéristique est

$$P_A(\lambda) = (1 - \lambda)(2 - \lambda)(3 - \lambda) \cdots (n - \lambda) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 \quad (7.9)$$

Les coefficients calculés satisfont $\tilde{a} = a_i(1 + \varepsilon_i)$ avec $|\varepsilon_i| \leq \text{eps}$. Cette perturbation dans les coefficients provoque une grande erreur dans les zéros de (11.9). Les résultats numériques pour $n = 9, 11, 13, 15$ (avec $\text{eps} \approx 6.10^{-8}$, simple précision) sont dessinés dans la figure V.2. **Conclusion.** Éviter le calcul des coefficients du polynôme caractéristique. Un tel algorithme est numériquement instable.

8 La méthode de la puissance

Un algorithme simple pour calculer les valeurs propres d'une matrice A est basé sur l'itération

$$y_{k+1} = A y_k \quad (8.1)$$

où y_0 est un vecteur arbitraire. Dans le théorème suivant, on démontre que $y_k = A^k y_0$ (*méthode de la puissance*) tend vers un vecteur propre de A et que le *quotient de Rayleigh* $y_k^* A y_k / y_k^* y_k$ est une approximation d'une valeur propre de A .

Théorème 240. Soit A une matrice diagonalisable de valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ et de vecteurs propres v_1, v_2, \dots, v_n (normalisés par $\|v_i\|_2 = 1$). Si $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, les vecteurs y_k de l'itération (12.1) vérifient

$$y_k = \lambda_1^k (a_1 v_1 + O(|\lambda_2/\lambda_1|^k)) \quad (8.2)$$

(le nombre a_1 est défini par $y_0 = \sum_i a_i v_i$). Le quotient de Rayleigh satisfait (si $a_1 \neq 0$)

$$\frac{y_k^* A y_k}{y_k^* y_k} = \lambda_1 + O(|\lambda_2/\lambda_1|^k) \quad (8.3)$$

Si A est une matrice normale (c'est-à-dire que les vecteurs propres sont orthogonaux), l'erreur dans (12.3) est $O(|\lambda_2/\lambda_1|^{2k})$

Démonstration. Exprimons le vecteur de départ y_0 dans la base des vecteurs propres, c'est-à-dire $y_0 = \sum_{i=1}^n a_i v_i$. Par récurrence, on voit que

$$y_k = A^k y_0 = \sum_{i=1}^n a_i \lambda_i^k v_i = \lambda_1^k (a_1 v_1 + \sum_{i=2}^n a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k v_i) \quad (8.4)$$

ce qui démontre la formule (12.2). De cette relation, on déduit que

$$y_k^* A y_k = y_k^* y_{k+1} = \sum_{i=1}^n |a_i|^2 |\lambda_i|^{2k} \lambda_i + \sum_{i \neq j} \tilde{a}_i a_j \tilde{\lambda}_i^k \lambda_j^{k+1} v_i^* v_j \quad (8.5)$$

$$y_k^* y_k = \sum_{i=1}^n |a_i|^2 |\lambda_i|^{2k} + \sum_{i \neq j} \tilde{a}_i a_j \tilde{\lambda}_i^k \lambda_j^k v_i^* v_j. \quad (8.6)$$

Si $a_1 \neq 0$, la formule (12.3) est une conséquence de

$$\frac{y_k^* A y_k}{y_k^* y_k} = \frac{|a_1|^2 \cdot |\lambda_1|^{2k} \cdot \lambda_1 \cdot (1 + O(|\lambda_2/\lambda_1|^k))}{|a_1|^2 \cdot |\lambda_1|^{2k} \cdot (1 + O(|\lambda_2/\lambda_1|^k))}. \quad (8.7)$$

Pour une matrice normale, le deuxième terme dans les formules (12.5) et (12.6) est absent et l'expression $O(|\lambda_2/\lambda_1|^k)$ peut être remplacée par $O(|\lambda_2/\lambda_1|^{2k})$ dans (12.7) et dans (12.3). cqfd

Exemple 241. Considérons la matrice

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

dont la valeur propre la plus grande est $\lambda_1 = 2(1 + \cos(\pi/4)) \approx 3,41421356$. Quelques itérations de la méthode de la puissance nous donnent

$$y_0 = (1, 1, 1)^T \quad y_1 = (3, 4, 3)^T \quad y_2 = (10, 14, 10)^T$$

et une première approximation de λ_1 est obtenue par

$$\frac{y_1^* A y_1}{y_1^* y_1} = \frac{y_1^* y_2}{y_1^* y_1} = \frac{116}{34} \approx 3,41176$$

Remarques. Les éléments du vecteur y_k croissent exponentiellement avec k . Il est alors recommandé de normaliser y_k après chaque itération, c'est-à-dire de remplacer y_k par $y_k / \|y_k\|$. Sinon, on risque un "overflow". Si $|\lambda_2/\lambda_1|$ est proche de 1, la convergence est très lente. Pour accélérer la convergence, on utilise la modification suivante :

9 Méthode de la puissance inverse de Wielandt

Supposons qu'on connaisse une approximation μ de la valeur propre cherchée λ_1 (il n'est pas nécessaire de supposer que λ_1 soit la plus grande valeur propre de A). L'idée est d'appliquer l'itération (12.1) à la matrice $(A - \mu I)^{-1}$ (Les valeurs propres de cette matrice sont $(\lambda_i - \mu)^{-1}$). Si μ est proche de λ_1 , on a

$$\frac{1}{|\lambda_1 - \mu|} \gg \frac{1}{|\lambda_i - \mu|} \quad \text{pour } i \geq 2$$

et la convergence va être très rapide. L'itération devient alors $y_{k+1} = (A - \mu I)^{-1} y_k$ ou

$$(A - \mu I) y_{k+1} = y_k \quad (9.1)$$

Après avoir calculé la décomposition LU de la matrice $A - \mu I$, une itération de (13.1) ne coûte pas plus cher qu'une de (12.1). Pour la matrice A de l'exemple précédent, choisissons $\mu = 3,41$ et $y_0 = (1; 1, 4; 1)^T$. Deux itérations de (13.1) nous donnent

$$y_1 = \begin{pmatrix} 236,134453781513 \\ 333,949579831933 \\ 236,134453781513 \end{pmatrix}, \quad y_2 = \begin{pmatrix} 56041,9461902408 \\ 79255,2785820210 \\ 56041,9461902408 \end{pmatrix}$$

et on obtient

$$\frac{1}{\lambda_1 - 3,41} \approx \frac{y_1^*(A - \mu I)^{-1} y_1}{y_1^* y_1} = \frac{y_1^* y_2}{y_1^* y_1} \approx 237,328870774159$$

De cette relation, on calcule λ_1 et on obtient l'approximation $3,41421356237333$. Les 13 premiers chiffres sont corrects. La méthode de la puissance (et celle de Wielandt) est importante pour la compréhension d'autres algorithmes. Si l'on veut calculer toutes les valeurs propres d'une matrice, on utilise des méthodes encore plus sophistiquées. En pratique, on procède de la manière suivante :

- on distingue les cas : A symétrique ou A quelconque.
- on cherche P telle que $P^{-1}AP$ devienne une matrice de Hessenberg (ou une matrice tridiagonale, si A est symétrique); voir V.3.
- on applique l'algorithme QR à la matrice H (voir V.6).
- si H est une matrice tridiagonale et symétrique, on peut également appliquer la méthode de bissection (voir V.4).

10 VALEURS PROPRES ET VECTEURS PROPRES

Les premiers vecteurs et valeurs propres viennent des équations différentielles (Lagrange 1759, *théorie du son*; Lagrange 1781, des matrices 6×6 dans le but de calculer les perturbations séculaires des orbites des 6 planètes connues à l'époque, *Oeuvres V*, p. 125-490). Aujourd'hui, le calcul des valeurs et vecteurs propres est indispensable dans toutes les branches de la science, en particulier pour la solution des systèmes des équations différentielles linéaires, en théorie de stabilité, pour les questions de convergence de processus itératifs, et en physique et chimie (mécanique, circuits, cinétique chimique, équation de Schrödinger).

FIG. V.1 : Une application linéaire comme champ de vecteurs (à gauche); transformée sur la base des vecteurs propres (à droite).

Observons en figure V.1 (à gauche) le champ de vecteurs d'une équation différentielle $y' = Ay$.

Deux directions sautent aux yeux : ce sont les directions où le vecteur Av prend la même direction que le vecteur v , c'est-à-dire, où

$$Av = \lambda v \quad \text{ou} \quad (A - \lambda I)v = 0 \tag{10.1}$$

Si cette équation est vérifiée, $\lambda \in \mathbb{C}$ s'appelle *valeur propre* de la matrice A et $v \in \mathbb{C}^n (v \neq 0)$ est le *vecteur propre* correspondant. L'équation (10.1) possède une solution v non nulle si et seulement si

$$P_A(\lambda) = \det(A - \lambda I) = 0$$

Le polynôme $P_A(\lambda)$ est le *polynôme caractéristique* de la matrice A . Les valeurs propres de A sont alors les zéros du polynôme caractéristique.

11 LA CONDITION DU CALCUL DES VALEURS PROPRES

A cause des erreurs d'arrondi, les éléments d'une matrice A , pour laquelle on cherche les valeurs propres, ne sont pas exacts. Ils sont plutôt égaux à

$$\tilde{a}_{ij} = a_{ij}(1 + \varepsilon_{ij}) \quad \text{avec} \quad |\varepsilon_{ij}| \leq eps$$

(eps étant la précision de l'ordinateur, est supposée être très petite). Il est alors très important d'étudier l'influence de ces perturbations sur les valeurs propres et sur les vecteurs propres de la matrice.

Pour montrer ceci, considérons la famille de matrices

$$A(\varepsilon) = A + \varepsilon C \quad \text{où} \quad |\varepsilon| \leq eps \quad \text{et} \quad |c_{ij}| \leq |a_{ij}|$$

(souvent, la dernière hypothèse va être remplacée par $\|C\| \leq \|A\|$).

Théorème 242 (Gershgorin). Soit A une matrice $n \times n$ (avec des éléments dans \mathbb{R} ou dans \mathbb{C}).

a) Si λ est une valeur propre de A , alors il existe un indice i tel que

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

c'est-à-dire, que toutes les valeurs propres de A se trouvent dans l'union des disques $D_i =$

$$\left\{ \lambda; |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}$$

b) Si une composante connexe de $\bigcup_{i=1}^n D_i$ consiste de k disques, elle contient exactement k valeurs propres de A .

Démonstration. Soit $v \neq 0$ un vecteur propre et choisissons l'indice i tel que $|v_i| \geq |v_j|$ pour tout j . La ligne i de l'équation $Av = \lambda v$ donne

$$\sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} v_j = (\lambda - a_{ii}) v_i.$$

En divisant par v_i et en utilisant l'inégalité du triangle, on obtient U

$$|\lambda - a_{ii}| = \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} \frac{v_j}{v_i} \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

L'affirmation (b) est vraie si A est une matrice diagonale. Le cas général est obtenu par un argument de continuité en faisant tendre les éléments en dehors de la diagonale vers zéro. cqfd

Théorème 243. Soit A une matrice diagonalisable, c'est-à-dire, il existe P avec $P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ et soit $A(\varepsilon) = A + \varepsilon C$. Alors, pour chaque valeur propre $\lambda(\varepsilon)$ de $A(\varepsilon)$, il existe un λ_i avec

$$|\lambda(\varepsilon) - \lambda_i| \leq \varepsilon \cdot \kappa_\infty(P) \cdot \|C\|_\infty$$

Démonstration. Nous transformons la matrice $A(\varepsilon) = A + \varepsilon C$ par la même matrice, qui transforme A sous forme diagonale :

$$P^{-1}A(\varepsilon)P = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) + \varepsilon P^{-1}CP$$

Si l'on dénote par e_{ij} les éléments de $P^{-1}CP$, le théorème de Gershgorin implique l'existence d'un indice i tel que $|\lambda(\varepsilon) - (\lambda_i + \varepsilon e_{ii})| \leq \varepsilon \sum_{j \neq i} |e_{ij}|$. L'inégalité triangulaire donne alors

$$|\lambda(\varepsilon) - \lambda_i| \leq \varepsilon \cdot \max_i \left(\sum_j |e_{ij}| \right) \leq \varepsilon \cdot \|P^{-1}CP\|_\infty \leq \varepsilon \cdot \|P^{-1}\|_\infty \cdot \|C\|_\infty \cdot \|P\|_\infty$$

ce qui démontre l'affirmation du théorème, car $\kappa_\infty(P) = \|P^{-1}\|_\infty \cdot \|C\|_\infty$ (condition de T). cqfd

Remarque 244. La condition du calcul des valeurs propres dépend de la condition de la matrice de transformation P . Si la matrice A est symétrique (P est orthogonale), le problème est bien conditionné. Toutefois, observons qu'on obtient seulement une estimation pour l'erreur absolue et non pour l'erreur relative.

Théorème 245 (différentiabilité des valeurs propres). Soit λ_1 une racine simple de $P_A(\lambda) = 0$. Alors, pour $|\varepsilon|$ suffisamment petit, la matrice $A(\varepsilon) = A + \varepsilon C$ possède une valeur propre unique $\lambda_1(\varepsilon)$ proche de λ_1 . La fonction $\lambda_1(\varepsilon)$ est différentiable (même analytique) et on a

$$\lambda_1(\varepsilon) = \lambda_1 + \varepsilon \frac{u_1^* C v_1}{u_1^* v_1} + O(\varepsilon^2) \quad (11.1)$$

où v_1 est le vecteur propre à droite ($Av_1 = \lambda_1 v_1$) et u_1 est le vecteur propre à gauche ($u_1^* A = \lambda_1 u_1^*$). On peut supposer que $\|v_1\| = \|u_1\| = 1$

Démonstration. Soit $p(\lambda, \varepsilon) = P_{A+\varepsilon C}(\lambda) = \det(A + \varepsilon C - \lambda I)$. Comme

$$p(\lambda_1, 0) = 0 \quad \text{et} \quad \frac{\partial p(\lambda_1, 0)}{\partial \lambda} \neq 0$$

le théorème des fonctions implicites garantit l'existence d'une fonction différentiable $\lambda_1(\varepsilon)$ (même analytique), tel que $\lambda_1(0) = \lambda_1$ et $p(\lambda_1(\varepsilon), \varepsilon) = 0$. Il existe donc un vecteur $v_1(\varepsilon)$ tel que

$$(A(\varepsilon) - \lambda_1(\varepsilon)I)v_1(\varepsilon) = 0. \quad (11.2)$$

La matrice dans (11.2) étant de rang $n-1$, on peut fixer une composante à 1 et appliquer la règle de Cramer. Ceci montre que les autres composantes sont des fonctions rationnelles des éléments de la matrice $A + \varepsilon C - \lambda_1(\varepsilon)I$ et donc différentiables. Après la normalisation à $v_1(\lambda)^T v_1(\lambda) = 1$, la fonction $v_1(\lambda)$ reste différentiable.

Pour calculer $\lambda_1'(0)$, nous pouvons dériver l'équation (11.2) par rapport à ε et poser ensuite $\varepsilon = 0$. Ceci donne

$$(A - \lambda_1 I)v_1'(0) + (C - \lambda_1'(0)I)v_1 = 0 \quad (11.3)$$

En multipliant cette relation par u_1^* , on obtient $u_1^*(C - \lambda_1'(0)I)v_1 = 0$, ce qui permet de calculer $\lambda_1'(0)$ et démontre la formule (11.1). cqfd

Conséquences. La formule (11.1) du théorème précédent montre que plus le vecteur propre de droite est parallèle au vecteur propre de gauche, mieux la valeur propre correspondante est bien conditionnée (par exemple, pour les matrices symétriques les deux vecteurs sont identiques); plus ils se rapprochent de l'orthogonalité, plus la valeur propre est mal conditionnée.

Si la matrice n'est pas symétrique (ou normale), le calcul de λ_1 (valeur propre simple) peut être mal conditionné. Considérons par exemple la matrice

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 2 \end{pmatrix} \quad \text{où} \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_1 = \frac{1}{\sqrt{1+\alpha^2}} \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$$

Dans cette situation, la formule (11.1) nous donne $\lambda_1(\varepsilon) - \lambda_1 = \varepsilon \cdot (c_{11} - \alpha c_{21}) + O(\varepsilon^2)$ et le calcul de $\lambda_1 = 1$ est mal conditionné si α est grand.

Exemple 1.4 Considérons la matrice (boîte de Jordan)

$$A = \left(\begin{array}{cccc} \lambda_1 & 1 & & \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{array} \right) \Bigg\}^n \quad (11.4)$$

Le polynôme caractéristique de $A + \varepsilon C$ satisfait

$$\det(A + \varepsilon C - \lambda I) = (\lambda_1 - \lambda)^n - (-1)^n \cdot \varepsilon \cdot c_{n1} + O(\varepsilon^2) + O(\varepsilon \cdot |\lambda_1 - \lambda|).$$

Si $c_{n1} \neq 0$, les termes $O(\varepsilon^2)$ et $O(\varepsilon \cdot |\lambda_1 - \lambda|)$ sont négligeables par rapport à ε . Les valeurs propres de $A + \varepsilon C$ sont alors approximativement données par les racines de

$$(\lambda_1 - \lambda)^n - (-1)^n \cdot \varepsilon \cdot c_{n1} = 0 \quad (11.5)$$

c'est-à-dire $\lambda = \lambda_1 + (\varepsilon \cdot c_{n1})^{1/n}$ (observer que $(\varepsilon \cdot c_{n1})^{1/n}$ donne n valeurs complexes distinctes - multiples des racines de l'unité).

Expérience numérique. Prenons la matrice (11.4) avec $\lambda_1 = 1$ et $n = 5$. Les éléments de la matrice C sont des nombres aléatoires dans l'intervalle $[-1, 1]$. Le dessin 7 ci-contre montre les 5 valeurs propres de $A + \varepsilon C$ pour $\varepsilon = 10^{-4}, 10^{-5}, \dots, 10^{-10}$. L'erreur est $\approx 10^{-1}$ pour $\varepsilon = 10^{-5}$ et $\approx 10^{-2}$ pour $\varepsilon = 10^{-10}$, ce qui correspond à la formule (11.5) pour $n = 5$.

Conséquence. Si la dimension n d'une boîte de Jordan est plus grande que 1, le calcul de la valeur propre de cette matrice est très mal conditionné.

11.1 Condition du calcul des vecteurs propres

Considérons la situation où toutes les valeurs propres de A sont distinctes. La démonstration du théorème sur la différentiabilité des valeurs propres montre (voir formule (11.2)) que les vecteurs propres normalisés $v_i(\varepsilon)$ de $A + \varepsilon C$ sont des fonctions différentiables de ε . Pour étudier la condition du calcul des vecteurs propres, nous exprimons $v_1'(0)$ dans la base des vecteurs propres (de droite)'

$$v_1'(0) = \sum_{i=1}^n \alpha_i v_i. \quad (11.6)$$

La formule (11.3) donne alors

$$\sum_{j=2}^n (\lambda_j - \lambda_1) \alpha_j v_j + (C - \lambda_1'(0)I) v_1 = 0. \quad (11.7)$$

En multipliant (11.7) par le vecteur propre de gauche u_1^* (observer que $u_1^* v_1 = 0$ pour $i \neq j$), on obtient α_i (pour $i \geq 2$) de la relation $(\lambda_i - \lambda_1) \alpha_i u_i^* v_i + u_i^* C v_1 = 0$. La normalisation $\|v_1(\varepsilon)\|_2^2 = 1$ donne (en la dérivant) $v_1^* v_1'(0) = 0$ et on en déduit que $\alpha_1 = -\sum_{j=2}^n \alpha_j v_1^* v_j$. Si l'on insère les formules pour α_i dans (11.6), on obtient pour $v_1(\varepsilon) = v_1 + \varepsilon v_1'(0) + O(\varepsilon^2)$ la relation

$$v_1(\varepsilon) = v_1 + \varepsilon \sum_{j=2}^n \frac{u_j^* C v_1}{(\lambda_1 - \lambda_j) u_j^* v_j} (v_j - v_1 v_1^* v_j) + O(\varepsilon^2). \quad (11.8)$$

De cette formule, on voit que la condition du calcul du vecteur propre v_1 dépend de la grandeur $u_i^* v_i$ (comme c'est le cas pour la valeur propre; voir la formule (11.1)) et aussi de la distance entre λ_1 & et les autres valeurs propres de A .

Un algorithme dangereux

La première méthode (déjà utilisée par Lagrange) pour calculer les valeurs propres d'une matrice A est la suivante : *calculer d'abord les coefficients du polynôme caractéristique $P_A(\lambda)$ et déterminer ensuite les zéros de ce polynôme.* Si la dimension de A est très petite (disons $n \leq 3$) ou si l'on fait le calcul en arithmétique exacte, cet algorithme peut être très utile. Par contre, si l'on fait le calcul en virgule flottante, cet algorithme peut donner des mauvaises surprises.

Considérons, par exemple, le problème de calculer les valeurs propres de la matrice diagonale

$$A = \text{diag}(1, 2, 3, \dots, n)$$

dont le polynôme caractéristique est

$$P_A(\lambda) = (1 - \lambda)(2 - \lambda)(3 - \lambda) \cdots (n - \lambda) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 \quad (11.9)$$

Les coefficients calculés satisfont $\tilde{a} = a_i(1 + \varepsilon_i)$ avec $|\varepsilon_i| \leq \text{eps}$. Cette perturbation dans les coefficients provoque une grande erreur dans les zéros de (11.9). Les résultats numériques pour $n = 9, 11, 13, 15$ (avec $\text{eps} \approx 6.10^{-8}$, simple précision) sont dessinés dans la figure V.2.

Conclusion. Eviter le calcul des coefficients du polynôme caractéristique. Un tel algorithme est numériquement instable.

12 LA METHODE DE LA PUISSANCE

Un algorithme simple pour calculer les valeurs propres d'une matrice A est basé sur l'itération

$$y_{k+1} = Ay_k \quad (12.1)$$

où y_0 est un vecteur arbitraire. Dans le théorème suivant, on démontre que $y_k = A^k y_0$ (*méthode de la puissance*) tend vers un vecteur propre de A et que le *quotient de Rayleigh* $y_k^* A y_k / y_k^* y_k$ est une approximation d'une valeur propre de A .

Théorème 246. Soit A une matrice diagonalisable de valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ et de vecteurs propres v_1, v_2, \dots, v_n (normalisés par $\|v_i\|_2 = 1$).

Si $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, les vecteurs y_k de l'itération (12.1) vérifient

$$y_k = \lambda_1^k (a_1 v_1 + O(|\lambda_2/\lambda_1|^k)) \quad (12.2)$$

(le nombre a_1 est défini par $y_0 = \sum_i a_i v_i$). Le quotient de Rayleigh satisfait (si $a_1 \neq 0$)

$$\frac{y_k^* A y_k}{y_k^* y_k} = \lambda_1 + O(|\lambda_2/\lambda_1|^k) \quad (12.3)$$

Si A est une matrice normale (c'est-à-dire, que les vecteurs propres sont orthogonaux), l'erreur dans (12.3) est $O(|\lambda_2/\lambda_1|^{2k})$

Démonstration. Exprimons le vecteur de départ y_0 dans la base des vecteurs propres, c'est-à-dire $y_0 = \sum_{i=1}^n a_i v_i$. Par récurrence, on voit que

$$y_k = A^k y_0 = \sum_{i=1}^n a_i \lambda_i^k v_i = \lambda_1^k (a_1 v_1 + \sum_{i=2}^n a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k v_i) \quad (12.4)$$

ce qui démontre la formule (12.2). De cette relation, on déduit que

$$y_k^* A y_k = y_k^* y_{k+1} = \sum_{i=1}^n |a_i|^2 |\lambda_i|^{2k} \lambda_i + \sum_{i \neq j} \tilde{a}_i a_j \tilde{\lambda}_i^k \lambda_j^{k+1} v_i^* v_j \quad (12.5)$$

$$y_k^* y_k = \sum_{i=1}^n |a_i|^2 |\lambda_i|^{2k} + \sum_{i \neq j} \tilde{a}_i a_j \tilde{\lambda}_i^k \lambda_j^k v_i^* v_j. \quad (12.6)$$

Si $a_1 \neq 0$, la formule (12.3) est une conséquence de

$$\frac{y_k^* A y_k}{y_k^* y_k} = \frac{|a_1|^2 \cdot |\lambda_1|^{2k} \cdot \lambda_1 \cdot (1 + O(|\lambda_2/\lambda_1|^k))}{|a_1|^2 \cdot |\lambda_1|^{2k} \cdot (1 + O(|\lambda_2/\lambda_1|^k))} \tag{12.7}$$

Pour une matrice normale, le deuxième terme dans les formules (12.5) et (12.6) est absent et l'expression $O(|\lambda_2/\lambda_1|^k)$ peut être remplacée par $O(|\lambda_2/\lambda_1|^{2k})$ dans (12.7) et dans (12.3). cqfd

Exemple 247. Considérons la matrice

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

dont la valeur propre la plus grande est $\lambda_1 = 2(1 + \cos(\pi/4)) \approx 3,41421356$. Quelques itérations de la méthode de la puissance nous donnent

$$y_0 = (1, 1, 1)^T \quad y_1 = (3, 4, 3)^T \quad y_2 = (10, 14, 10)^T$$

et une première approximation de λ_1 est obtenue par

$$\frac{y_1^* A y_1}{y_1^* y_1} = \frac{y_1^* y_2}{y_1^* y_1} = \frac{116}{34} \approx 3,41176$$

Remarques. Les éléments du vecteur y_k croissent exponentiellement avec k . Il est alors recommandé de normaliser y_k après chaque itération, c'est-à-dire. de remplacer y_k par $y_k / \|y_k\|$. Sinon, on risque un "overflow". Si $|\lambda_2/\lambda_1|$ est proche de 1, la convergence est très lente. Pour accélérer la convergence, on utilise la modification suivante :

13 METHODE DE LA PUISSANCE INVERSE DE WIELANDT

Supposons qu'on connaisse une approximation μ de la valeur propre cherchée λ_1 (il n'est pas nécessaire de supposer que λ_1 soit la plus grande valeur propre de A). L'idée est d'appliquer l'itération (12.1) à la matrice $(A - \mu I)^{-1}$ (Les valeurs propres de cette matrice sont $(\lambda_i - \mu)^{-1}$. Si μ est proche de λ_1 , on a :

$$\frac{1}{|\lambda_1 - \mu|} \gg \frac{1}{|\lambda_i - \mu|} \quad \text{pour } i \geq 2$$

et la convergence va être très rapide. L'itération devient alors $y_{k+1} = (A - \mu I)^{-1} y_k$ ou

$$(A - \mu I) y_{k+1} = y_k \tag{13.1}$$

Après avoir calculé la décomposition LU de la matrice $A - \mu I$, une itération de (13.1) ne coûte pas plus cher qu'une de (12.1).

Pour la matrice A de l'exemple précédent, choisissons $\mu = 3,41$ et $y_0 = (1; 1, 4; 1)^T$ Deux itérations de (13.1) nous donnent

$$y_1 = \begin{pmatrix} 236,134453781513 & 56041,9461902408 \\ 333,949579831933, & y_2 = 79255,2785820210 \\ 236,134453781513 & 56041,9461902408 \end{pmatrix}$$

et on obtient

$$\frac{1}{\lambda_1 - 3,41} \approx \frac{y_1^* (A - \mu I)^{-1} y_1}{y_1^* y_1} = \frac{y_1^* y_2}{y_1^* y_1} \approx 237,328870774159$$

De cette relation, on calcule λ_1 et on obtient l'approximation 3,41421356237333. Les 13 premiers chiffres sont corrects.

La méthode de la puissance (et celle de Wielandt) est importante pour la compréhension d'autres algorithmes. Si l'on veut calculer toutes les valeurs propres d'une matrice, on utilise des méthodes encore plus sophistiquées. En pratique, on procède de la manière suivante :

- on distingue les cas : A symétrique ou A quelconque.
- on cherche P telle que $P^{-1}AP$ devienne une matrice de Hessenberg (ou une matrice tridiagonale, si A est symétrique); voir V.3.
- on applique l'algorithme QR à la matrice H (voir V.6).
- si H est une matrice tridiagonale et symétrique, on peut également appliquer la méthode de bisection (voir V.4).