

Chapitre 5

CALCUL DE VALEURS PROPRES ET DE VECTEURS PROPRES

Chapitre 5

CALCUL DE VALEURS PROPRES ET DE VECTEURS PROPRES

5.1 Conditionnement d'un problème de valeurs propres

Considérons la matrice d'ordre 4

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

où ε est un paramètre positif.

En choisissant $\varepsilon = 0$, il est facile de voir que toutes les valeurs propres de $A(0)$ sont nulles. Par contre, pour $\varepsilon = 10^{-4}$, les valeurs propres de $A(10^{-4})$ sont toutes de module égal à 10^{-1} . Et par suite, si on utilise un ordinateur qui effectue les calculs à 10^{-4} près, alors le nombre $\varepsilon = 10^{-4}$ peut être remplacé par zéro, et le calcul des valeurs propres de la matrice $A(10^{-4})$ peut être entaché d'une erreur absolue de 10^{-1} , ou encore d'une erreur relative $\Delta\lambda/\Delta\varepsilon = 10^3$. On est ainsi en présence d'un problème où une "petite perturbation" sur les éléments de la matrice entraîne une "grande perturbation" sur les valeurs propres de la matrice. Il est donc nécessaire de pouvoir mesurer le conditionnement d'un problème de valeurs propres afin de pouvoir identifier les matrices pour lesquelles le risque de commettre une grande erreur dans le calcul des valeurs propres existe.

Plus précisément, si on considère le problème suivant :

$$\begin{cases} \text{trouver } \lambda \in \mathbf{C} \text{ et } x \in \mathbf{C}^n \setminus \{0\} \text{ tels que} \\ Ax = \lambda x \end{cases}$$

alors comment sont situées les valeurs propres de $A + \delta A$ par rapport à celles de A , où δA est une matrice de même ordre que A . La réponse à cette question sera donnée au théorème 5.1 où l'on va considérer des matrices diagonalisables. On donnera au théorème 5.2 un résultat qui "améliore" celui du théorème 5.1 dans le cas des matrices symétriques.

Les normes matricielles que nous utilisons dans cette section sont telles que

$$(5.1) \quad \|D\| = \max_{1 \leq i \leq n} |d_i|,$$

où $D = \text{diag}(d_i)_{1 \leq i \leq n}$ est une matrice diagonale d'ordre n .

Remarque :

Les normes matricielles subordonnées $\|\cdot\|_p$ vérifient la propriété (5.1). ■

Définition 5.1

Soit A une matrice d'ordre n . On appelle spectre de A , l'ensemble de toutes les valeurs propres de A , noté $Sp(A)$. ■

Théorème 5.1 (de Bauer-Fike)

Soit A une matrice d'ordre n et diagonalisable, et soit

$$\mathcal{P} = \{ \text{matrice } P \text{ d'ordre } n \text{ et inversible ; } P^{-1}AP = \text{diag}(\lambda_i)_{1 \leq i \leq n} \},$$

où λ_i , $1 \leq i \leq n$, sont les valeurs propres de A comptées avec leur ordre de multiplicité.

Si δA est une matrice de perturbation quelconque (de même ordre que A), alors les valeurs propres de $A + \delta A$ sont contenues dans $\bigcup_{i=1}^n D_i$:

$$(5.2) \quad Sp(A + \delta A) \subset \bigcup_{i=1}^n D_i,$$

où

$$(5.3) \quad D_i = \{z \in \mathbf{C}; |z - \lambda_i| \leq \tau(A) \|\delta A\|\}$$

avec $\tau(A) = \inf_{P \in \mathcal{P}} \text{cond}(P)$, et $\text{cond}(P) = \|P\| \|P^{-1}\|$.

Démonstration :

Soit λ une valeur propre de $A + \delta A$. S'il existe $j \in \{1, 2, \dots, n\}$ tel que $\lambda = \lambda_j$, alors il est clair que $\lambda \in D_j$ et donc $\lambda \in \bigcup_{i=1}^n D_i$. Supposons dans la suite que pour tout $j \in \{1, 2, \dots, n\}$, $\lambda \neq \lambda_j$.

Soit x un vecteur propre associé à λ : $(A + \delta A)x = \lambda x$. La matrice $(A + \delta A - \lambda I)$ est alors non inversible.

La matrice A étant diagonalisable, alors il existe une matrice inversible P telle que $D = P^{-1}AP$, où $D = \text{diag}(\lambda_i)_{1 \leq i \leq n}$. Utilisant le fait que la matrice $D - \lambda I$ est inversible, puisque $\lambda \neq \lambda_i$ pour $i \in \{1, 2, \dots, n\}$, on obtient

$$\begin{aligned} P^{-1}(A + \delta A - \lambda I)P &= (D - \lambda I) + P^{-1}(\delta A)P \\ &= (D - \lambda I)(I + (D - \lambda I)^{-1}P^{-1}(\delta A)P), \end{aligned}$$

et par suite, puisque $A + \delta A - \lambda I$ n'est pas inversible, on a

$$\det(I + (D - \lambda I)^{-1}P^{-1}(\delta A)P) = 0.$$

Ainsi -1 est une valeur propre de la matrice $(D - \lambda I)^{-1}P^{-1}(\delta A)P$, d'où :

$$\rho((D - \lambda I)^{-1}P^{-1}(\delta A)P) \geq 1.$$

Il s'en suit, grâce au théorème 1.1, que pour toute norme matricielle $\|\cdot\|$, on a

$$\|(D - \lambda I)^{-1}P^{-1}(\delta A)P\| \geq 1$$

et donc

$$(5.4) \quad \|(D - \lambda I)^{-1}\| \|P^{-1}\| \|\delta A\| \|P\| \geq 1.$$

Par ailleurs, on a

$$(D - \lambda I)^{-1} = \begin{pmatrix} \frac{1}{\lambda_1 - \lambda} & & \circ \\ & \ddots & \\ \circ & & \frac{1}{\lambda_n - \lambda} \end{pmatrix}$$

d'où, grâce à (5.1), on obtient

$$\|(D - \lambda I)^{-1}\| = \max_{1 \leq i \leq n} \frac{1}{|\lambda_i - \lambda|} = \frac{1}{\min_{1 \leq i \leq n} |\lambda_i - \lambda|}.$$

Il en découle alors, grâce à (5.4), que

$$\min_{1 \leq i \leq n} |\lambda_i - \lambda| \leq \text{cond}(P) \|\delta A\|,$$

et donc, si on note $\tau(A) = \inf_{P \in \mathcal{P}} \text{cond}(P)$, on obtient

$$(5.5) \quad \min_{1 \leq i \leq n} |\lambda_i - \lambda| \leq \tau(A) \|\delta A\|.$$

Soit $i_0 \in \{1, 2, \dots, n\}$ tel que $\min_{1 \leq i \leq n} |\lambda_i - \lambda| = |\lambda_{i_0} - \lambda|$. Alors, d'après (5.5), $\lambda \in D_{i_0}$, où D_{i_0} est donné par (5.3). Ainsi $\text{Sp}(A + \delta A) \subset \bigcup_{i=1}^n D_i$. Ce qui prouve (5.2) et achève la démonstration. ■

Définition 5.2

Le nombre

$$\tau(A) = \inf_{P \in \mathcal{P}} \text{cond}(P) \geq 1$$

est appelé conditionnement de la matrice A relativement au calcul de ses valeurs propres. ■

Remarques :

1) Si la matrice A est symétrique, alors la matrice de passage P est orthogonale, et donc $\text{cond}_2(P) = 1$. Ainsi

$$\tau_2(A) = \inf_{P \in \mathcal{P}} \text{cond}_2(P) = 1.$$

Les matrices symétriques sont alors les mieux conditionnées vis-à-vis de la recherche de valeurs propres.

2) Étant donnée la k -ème valeur propre β_k de la matrice $A + \delta A$, il résulte de ce qui précède, qu'il existe au moins une valeur propre α_i de A telle que

$$(5.6) \quad |\alpha_i - \beta_k| \leq \tau(A) \|\delta A\|.$$

Mais, on ne peut pas affirmer que β_k est proche de la k -ème valeur propre α_k de A . Le théorème suivant conduit à un résultat "plus fort", dans le cas où les matrices A et δA sont symétriques, à savoir que précisément la k -ème valeur propre α_k de A vérifie l'inégalité $|\alpha_k - \beta_k| \leq \|\delta A\|_2$. ■

Théorème 5.2

Soit A une matrice symétrique (ou hermitienne) de valeurs propres $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$, et soit δA une matrice de perturbation symétrique (ou hermitienne). Soient $\beta_1 \leq \beta_2 \leq \dots \leq \beta_n$ les valeurs propres de la matrice "perturbée" $A + \delta A$. Alors, on a :

$$|\alpha_k - \beta_k| \leq \|\delta A\|_2, \quad 1 \leq k \leq n,$$

où $\|\cdot\|_2$ désigne la norme matricielle subordonnée à la norme vectorielle $\|\cdot\|_2$. ■

Avant de prouver ce résultat, nous énonçons et démontrons le lemme suivant.

Lemme 5.1

Si A est une matrice symétrique de valeurs propres $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$, alors

$$(5.7) \quad \alpha_i = \min_{E_i \in \mathcal{E}_i} \max_{x \in E_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)}, \quad 1 \leq i \leq n$$

où $\mathcal{E}_i$ est l'ensemble des sous-espaces vectoriels de $\mathbb{R}^n$ de dimension égale à i ($\dim E_i = i$).

Démonstration :

Soit u_i un vecteur propre associé à la valeur propre α_i . La famille $(u_i)_{1 \leq i \leq n}$ forme une base orthonormée de $\mathbb{R}^n$.

Soit $\mathcal{U}_i = \text{Vect}\{u_1, u_2, \dots, u_i\}$, $1 \leq i \leq n$, le sous-espace vectoriel engendré par $(u_1, u_2, \dots, u_i)$. Alors, pour tout $x \in \mathcal{U}_i \setminus \{0\}$, on a $x = \sum_{j=1}^i \gamma_j u_j$, et par suite

$$\frac{(Ax, x)}{(x, x)} = \frac{\left(\sum_{j=1}^i \gamma_j A u_j, \sum_{k=1}^i \gamma_k u_k\right)}{\sum_{j=1}^i \gamma_j^2} = \frac{\sum_{j=1}^i \gamma_j^2 \alpha_j}{\sum_{j=1}^i \gamma_j^2}$$

et donc, on a

$$\forall x \in \mathcal{U}_i \setminus \{0\}, \quad \frac{(Ax, x)}{(x, x)} \leq \alpha_i$$

puisque $\alpha_j \leq \alpha_i$ pour tout $j \in \{1, 2, \dots, i\}$. Ainsi, on obtient

$$\max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)} \leq \alpha_i, \quad 1 \leq i \leq n.$$

D'où

$$(5.8) \quad \min_{E_i \in \mathcal{E}_i} \max_{x \in E_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)} \leq \alpha_i, \quad 1 \leq i \leq n.$$

Montrons l'inégalité inverse. Considérons pour cela l'application

$$\begin{aligned} \varphi : E &\longrightarrow \mathbb{R}^{i-1} \\ x &\longmapsto ((x, u_1), (x, u_2), \dots, (x, u_{i-1})) \end{aligned}$$

où E est un sous-espace vectoriel (quelconque) de $\mathbb{R}^n$ de dimension i . L'application φ est linéaire et on a $\dim \text{Im} \varphi \leq i - 1$. Or, $\dim \ker \varphi + \dim \text{Im} \varphi = \dim E$ et $\dim E = i$, alors on a nécessairement $\dim \ker \varphi \geq 1$. Et par suite, il existe au moins un élément $\bar{x} \in E \setminus \{0\}$ tel que $\varphi(\bar{x}) = 0$. Il s'en suit que

$$\bar{x} = \sum_{j=i}^n (\bar{x}, u_j) u_j = \sum_{j=i}^n \gamma_j u_j,$$

où $\gamma_j = (\bar{x}, u_j)$. D'où

$$\frac{(A\bar{x}, \bar{x})}{(\bar{x}, \bar{x})} = \frac{\sum_{j=i}^n \gamma_j^2 \alpha_j}{\sum_{j=i}^n \gamma_j^2} \geq \alpha_i$$

puisque $\alpha_j \geq \alpha_i$ pour tout $j \in \{i, i+1, \dots, n\}$. Ainsi, on a montré que pour tout sous-espace vectoriel E de $\mathbb{R}^n$, de dimension i , il existe $\bar{x} \in E \setminus \{0\}$, tel que

$$\frac{(A\bar{x}, \bar{x})}{(\bar{x}, \bar{x})} \geq \alpha_i.$$

Et par suite, on a

$$\max_{x \in E \setminus \{0\}} \frac{(Ax, x)}{(x, x)} \geq \alpha_i,$$

d'où

$$(5.9) \quad \min_{E_i \in \mathcal{E}_i} \max_{x \in E_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)} \geq \alpha_i, \quad 1 \leq i \leq n.$$

Ce qui prouve, grâce à (5.8), l'égalité (5.7) et achève la démonstration. ■

Définition 5.3

Le nombre

$$R_A(x) = \frac{(Ax, x)}{(x, x)}, \quad x \neq 0$$

est appelé le quotient de Rayleigh de la matrice A au point x . ■

Démonstration du théorème 5.2 :

Soient $(u_1, u_2, \dots, u_n)$ une base de vecteurs propres de A correspondant aux valeurs propres α_i , $1 \leq i \leq n$, et $(v_1, v_2, \dots, v_n)$ une base de vecteurs propres de $A + \delta A$ correspondant aux valeurs propres β_i , $1 \leq i \leq n$.

Considérons le sous-espace vectoriel $\mathcal{U}_i = \text{Vect}\{u_1, u_2, \dots, u_i\}$ de $\mathbb{R}^n$ engendré par $(u_1, u_2, \dots, u_i)$ et le sous-espace vectoriel $\mathcal{V}_i = \text{Vect}\{v_1, v_2, \dots, v_i\}$ de $\mathbb{R}^n$ engendré par $(v_1, v_2, \dots, v_i)$. Utilisant le lemme 5.1, on obtient :

$$\beta_i = \min_{E_i \in \mathcal{E}_i} \max_{x \in E_i \setminus \{0\}} \frac{((A + \delta A)x, x)}{(x, x)} \leq \max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{((A + \delta A)x, x)}{(x, x)}, \quad 1 \leq i \leq n.$$

Et par suite, on obtient

$$(5.10) \quad \beta_i \leq \max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)} + \max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{((\delta A)x, x)}{(x, x)}$$

Par ailleurs, (c.f. démonstration du lemme 5.1), on a

$$\alpha_i = \max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{(Ax, x)}{(x, x)}$$

et

$$\max_{x \in \mathcal{U}_i \setminus \{0\}} \frac{((\delta A)x, x)}{(x, x)} \leq \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{((\delta A)x, x)}{(x, x)} \leq \rho(\delta A) = \|\delta A\|_2.$$

La deuxième inégalité est obtenue en utilisant le lemme 1.1. Il s'en suit, d'après (5.10), que

$$(5.11) \quad \beta_i \leq \alpha_i + \|\delta A\|_2, \quad 1 \leq i \leq n.$$

Échangeant les rôles de A et $A + \delta A$ et de $\mathcal{U}_i$ et $\mathcal{V}_i$, on obtient de la même façon

$$(5.12) \quad \alpha_i \leq \beta_i + \|\delta A\|_2, \quad 1 \leq i \leq n.$$

Le résultat du théorème 5.2 découle alors des inégalités (5.11) et (5.12). ■

5.2 La méthode de la puissance

On va donner dans cette section la méthode de la puissance qui permet de calculer une approximation de la valeur propre de plus grand module ainsi que celle d'un vecteur propre associé à cette valeur propre.

Algorithme de la méthode de la puissance

$$(5.13) \quad \left\{ \begin{array}{l} x^{(0)} \quad \text{donné dans } \mathbf{C}^n \text{ de norme } 1 \\ \text{Pour } k = 0, 1, \dots \text{ calculer} \\ \quad y^{(k+1)} = Ax^{(k)} \\ \quad x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2} \\ \quad \lambda^{(k+1)} = (x^{(k)})^* Ax^{(k)} = (x^{(k)})^* y^{(k+1)} \end{array} \right.$$

Théorème 5.3

Soit A une matrice carrée d'ordre n , diagonalisable et de valeurs propres λ_i , $1 \leq i \leq n$, comptées avec leur ordre de multiplicité, et vérifiant :

$$0 \leq |\lambda_1| \leq \dots \leq |\lambda_{n-1}| < |\lambda_n|.$$

Soit $(u_i)_{1 \leq i \leq n}$ une base normée de vecteurs propres de A associés aux valeurs propres λ_i . Si $x^{(0)} = \sum_{i=1}^n \alpha_i u_i$ est tel que $\alpha_n \neq 0$, alors la méthode de la puissance converge au sens suivant :

$$(5.14) \quad \lim_{k \rightarrow +\infty} \lambda^{(k)} = \lambda_n$$

$$(5.15) \quad \lim_{k \rightarrow +\infty} (x^{(k)} - \beta^k \gamma u_n) = 0$$

où $\beta = \frac{\lambda_n}{|\lambda_n|}$ et γ est un nombre complexe de module 1.

Démonstration :

Montrons (5.15). Posons pour cela, $z^{(0)} = x^{(0)}$ et $z^{(k)} = A^k x^{(0)}$ pour $k \geq 1$. On a alors

$$(5.16) \quad x^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|_2} = \frac{Ax^{(k-1)}}{\|Ax^{(k-1)}\|_2} = \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|_2} = \frac{z^{(k)}}{\|z^{(k)}\|_2}.$$

Dans (5.16) la 3-ème égalité peut être démontrée par récurrence sur k .

Par ailleurs, on a

$$(5.17) \quad \begin{aligned} z^{(k)} = A^k x^{(0)} &= \sum_{i=1}^n \alpha_i A^k u_i = \sum_{i=1}^n \alpha_i \lambda_i^k u_i \\ &= \alpha_n \lambda_n^k u_n + \sum_{i=1}^{n-1} \alpha_i \lambda_i^k u_i. \end{aligned}$$

D'où

$$(5.18) \quad \frac{z^{(k)}}{\lambda_n^k} - \alpha_n u_n = \sum_{i=1}^{n-1} \alpha_i \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i.$$

Or, pour tout $i \in \{1, 2, \dots, n-1\}$, on a $|\lambda_i| < |\lambda_n|$. Et par suite, on obtient

$$(5.19) \quad \lim_{k \rightarrow +\infty} \frac{z^{(k)}}{\lambda_n^k} = \alpha_n u_n$$

et donc

$$(5.20) \quad \lim_{k \rightarrow +\infty} \frac{\|z^{(k)}\|_2}{|\lambda_n^k|} = |\alpha_n| \|u_n\|_2 = |\alpha_n|,$$

puisque u_n est de norme égale à 1. Utilisant l'hypothèse $\alpha_n \neq 0$, on obtient, d'après (5.19) et (5.20) :

$$(5.21) \quad \lim_{k \rightarrow +\infty} \frac{z^{(k)}}{\|z^{(k)}\|_2} \frac{|\lambda_n^k|}{\lambda_n^k} = \frac{\alpha_n}{|\alpha_n|} u_n.$$

En posant $\beta = \frac{\lambda_n}{|\lambda_n|}$ et $\gamma = \frac{\alpha_n}{|\alpha_n|}$, on obtient, grâce à (5.16) et (5.21) :

$$\lim_{k \rightarrow +\infty} (x^{(k)} - \beta^k \gamma u_n) = 0.$$

Ce qui prouve (5.15).

Prouvons (5.14). Utilisant (5.16) et (5.17), on obtient

$$\begin{aligned}
 (5.22) \quad x^{(k)} = \frac{z^{(k)}}{\|z^{(k)}\|_2} &= \frac{1}{\|z^{(k)}\|_2} \lambda_n^k \left(\alpha_n u_n + \sum_{i=1}^{n-1} \alpha_i \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i \right) \\
 &= \frac{\alpha_n \lambda_n^k}{\|z^{(k)}\|_2} u_n + \frac{\alpha_n \lambda_n^k}{\|z^{(k)}\|_2} \sum_{i=1}^{n-1} \frac{\alpha_i}{\alpha_n} \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i.
 \end{aligned}$$

Posons

$$(5.23) \quad \beta_n^k = \frac{\alpha_n \lambda_n^k}{\|z^{(k)}\|_2} \quad \text{et} \quad \beta_i^k = \beta_n^k \frac{\alpha_i}{\alpha_n} \left(\frac{\lambda_i}{\lambda_n} \right)^k, \quad \text{pour } i \in \{1, 2, \dots, n-1\}.$$

Utilisant (5.20), on obtient

$$(5.24) \quad \lim_{k \rightarrow +\infty} |\beta_n^k| = 1$$

et par suite, d'après (5.23), on a

$$(5.25) \quad \lim_{k \rightarrow +\infty} |\beta_i^k| = 0, \quad 1 \leq i \leq n-1$$

puisque $|\lambda_i| < |\lambda_n|$ pour tout $i \in \{1, 2, \dots, n-1\}$.

Par ailleurs, on a, d'après (5.22) et (5.23)

$$Ax^{(k)} = \sum_{i=1}^n \beta_i^k A u_i = \sum_{i=1}^n \beta_i^k \lambda_i u_i$$

et

$$(x^{(k)})^* = \sum_{j=1}^n \bar{\beta}_j^k u_j^*.$$

Il s'en suit, grâce à (5.13), que

$$\lambda^{(k+1)} \equiv (x^{(k)})^* A x^{(k)} = \sum_{i,j=1}^n \beta_i^k \bar{\beta}_j^k \lambda_i u_j^* u_i.$$

D'où

$$(5.26) \quad \lambda^{(k+1)} = |\beta_n^k|^2 \lambda_n \underbrace{\|u_n\|_2^2}_{=1} + \sum_{\substack{i,j=1 \\ (i,j) \neq (n,n)}}^n \beta_i^k \bar{\beta}_j^k \lambda_i u_j^* u_i.$$

En passant à la limite lorsque $k \rightarrow +\infty$, on obtient, grâce à (5.24) et (5.25)

$$\lim_{k \rightarrow +\infty} \lambda^{(k)} = \lambda_n.$$

Ce qui prouve (5.14) et achève la démonstration. ■

Définition 5.4

Le facteur de convergence d'une suite $(v^{(k)})_k$ qui converge vers v , dans un espace vectoriel normé, est le nombre défini par

$$\lim_{k \rightarrow +\infty} \frac{\|v^{(k+1)} - v\|}{\|v^{(k)} - v\|}.$$
■

Corollaire 5.1

Sous les hypothèses du théorème 5.3, le facteur de convergence des suites $\lambda^{(k)}$ et $\frac{z^{(k)}}{\lambda_n^k}$ générées par l'algorithme (5.13) de la méthode de la puissance, où $z^{(k)} = A^k x^{(0)}$, est égal à $\frac{|\lambda_{n-1}|}{|\lambda_n|}$.

Démonstration :

On a, d'après (5.18)

$$\frac{z^{(k)}}{\lambda_n^k} - \alpha_n u_n = \sum_{i=1}^{n-1} \alpha_i \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i.$$

Posons

$$e^{(k)} = \sum_{i=1}^{n-1} \alpha_i \left(\frac{\lambda_i}{\lambda_n} \right)^k u_i,$$

alors, d'après (5.19) et (5.18), $e^{(k)}$ tend vers 0. Puisque

$$\frac{|\lambda_i|}{|\lambda_n|} \leq \frac{|\lambda_{n-1}|}{|\lambda_n|} < 1, \quad 1 \leq i \leq n-1,$$

alors $e^{(k)}$ tend vers 0 comme $\frac{|\lambda_{n-1}|^k}{|\lambda_n|^k}$, et par suite on a

$$\lim_{k \rightarrow +\infty} \frac{\|e^{(k)}\|}{\|e^{(k-1)}\|} = \frac{|\lambda_{n-1}|}{|\lambda_n|}.$$

Ainsi, le facteur de convergence de la suite $\frac{z^{(k)}}{\lambda_n^k}$ est égal à $\frac{|\lambda_{n-1}|}{|\lambda_n|}$.

De même, utilisant (5.26), on a

$$\lambda^{(k+1)} = |\beta_n^k|^2 \lambda_n + \sum_{\substack{i,j=1 \\ (i,j) \neq (n,n)}}^n \beta_i^k \bar{\beta}_j^k \lambda_i u_j^* u_i,$$

et d'après (5.23) on a

$$|\beta_i^k| = |\beta_n^k| \frac{|\alpha_i|}{|\alpha_n|} \frac{|\lambda_i|^k}{|\lambda_n|^k},$$

où $|\beta_n^k| \rightarrow 1$ et $|\beta_i^k| \rightarrow 0$, lorsque $k \rightarrow +\infty$, comme $\frac{|\lambda_{n-1}|^k}{|\lambda_n|^k}$. Ainsi le facteur de convergence de la suite $\lambda^{(k)}$ est donné par

$$\lim_{k \rightarrow +\infty} \frac{|\lambda^{(k+1)} - \lambda_n|}{|\lambda^{(k)} - \lambda_n|} = \frac{|\lambda_{n-1}|}{|\lambda_n|}.$$

Ce qui prouve le résultat énoncé et achève la démonstration. ■

5.3 La méthode de la puissance inverse

C'est la méthode de la puissance appliquée à la matrice A^{-1} , qui permet d'obtenir la valeur propre de plus petit module de A et un vecteur propre qui lui est associé.

5.3.1 Recherche de la valeur propre de plus petit module

La valeur propre de plus grand module de A^{-1} est égale à l'inverse de la valeur propre de A de plus petit module, puisque

$$\max_i \frac{1}{|\lambda_i|} = \frac{1}{\min_i |\lambda_i|}.$$

On applique alors la méthode de la puissance à A^{-1} , puis on inverse la valeur propre trouvée, pour obtenir la valeur propre de A de plus petit module ainsi qu'un vecteur propre associé.

Cela conduit, en partant de (5.13), à l'algorithme suivant :

$$\left\{ \begin{array}{l} x^{(0)} \quad \text{donné dans } \mathbf{C}^n \text{ de norme } 1 \\ \text{Pour } k = 0, 1, \dots \text{ calculer} \\ \\ y^{(k+1)} = A^{-1}x^{(k)} \\ \\ x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2} \\ \\ \gamma^{(k+1)} = \frac{1}{(x^{(k)})^* A^{-1} x^{(k)}} = \frac{1}{(x^{(k)})^* y^{(k+1)}} \end{array} \right.$$

Dans la pratique, on ne calcule pas A^{-1} , mais on résout successivement les systèmes linéaires $Ay^{(k+1)} = x^{(k)}$, par exemple en ayant factorisé A une fois pour toute sous la forme du produit de deux matrices triangulaires : $A = LU$ (ou $PA = LU$ lorsque cela est nécessaire).

Algorithme de la méthode de la puissance inverse

$$(5.27) \quad \left\{ \begin{array}{l} x^{(0)} \quad \text{donné dans } \mathbf{C}^n \text{ de norme } 1 \\ \text{Pour } k = 0, 1, \dots \text{ calculer} \\ \\ y^{(k+1)} \text{ en résolvant le système } Ay^{(k+1)} = x^{(k)} \\ \\ x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2} \\ \\ \gamma^{(k+1)} = \frac{1}{(x^{(k)})^* y^{(k+1)}} \end{array} \right.$$

On a le résultat de convergence suivant.

Corollaire 5.2

Soit A une matrice diagonalisable de valeurs propres λ_i , $1 \leq i \leq n$, vérifiant

$$0 < |\lambda_1| < |\lambda_2| \leq |\lambda_3| \leq \dots \leq |\lambda_n|.$$

Soit $(u_i)_{1 \leq i \leq n}$ une base normée de vecteurs propres de A associés aux valeurs propres λ_i . Si $x^{(0)} = \sum_{i=1}^n \alpha_i u_i$ est tel que $\alpha_1 \neq 0$, alors la méthode de la puissance inverse converge :

$$\lim_{k \rightarrow +\infty} \gamma^{(k)} = \lambda_1.$$

Démonstration :

Il suffit d'appliquer le théorème 5.3 à la matrice A^{-1} . ■

Remarques :

1) Le facteur de convergence de la méthode de la puissance inverse est égal à

$$\frac{\frac{1}{|\lambda_2|}}{\frac{1}{|\lambda_1|}} = \frac{|\lambda_1|}{|\lambda_2|}$$

(il suffit d'appliquer le corollaire 5.1 à la matrice A^{-1}).

2) La "direction" du vecteur $x^{(k)}$, défini dans (5.27), converge vers une direction propre associée à la valeur propre λ_1 . ■

5.3.2 La méthode de la puissance inverse avec translation

Elle permet d'obtenir la valeur propre la plus proche d'un nombre donné. Soit μ un nombre donné. Considérons la matrice $(A - \mu I)$ obtenue à partir de A par la "translation $-\mu I$ " (*shift* en anglais). On applique l'algorithme (5.27) de la méthode de la puissance inverse à la matrice $(A - \mu I)$. Le plus grand module des valeurs propres de $(A - \mu I)^{-1}$ est $\frac{1}{\min_i |\lambda_i - \mu|}$, où les λ_i sont les valeurs propres de A .

Algorithme de la méthode de la puissance inverse avec translation

$$(5.28) \quad \left\{ \begin{array}{l} x^{(0)} \quad \text{donné dans } \mathbf{C}^n \text{ de norme } 1 \\ \text{Pour } k = 0, 1, \dots \text{ calculer} \\ \quad y^{(k+1)} \text{ en résolvant le système } (A - \mu I)y^{(k+1)} = x^{(k)} \\ \quad x^{(k+1)} = \frac{y^{(k+1)}}{\|y^{(k+1)}\|_2} \\ \quad \tilde{\gamma}^{(k+1)} = \mu + \frac{1}{(x^{(k)})^*(A - \mu I)^{-1}x^{(k)}} = \mu + \frac{1}{(x^{(k)})^*y^{(k+1)}} \end{array} \right.$$

On a le résultat de convergence suivant.

Corollaire 5.3

Soit A une matrice diagonalisable de valeurs propres λ_i , $1 \leq i \leq n$. Soit $(u_i)_{1 \leq i \leq n}$ une base normée de vecteurs propres de A associés aux valeurs propres λ_i . Si λ_{i_0} est la valeur propre de A la plus proche de μ avec de plus

$$|\lambda_{i_0} - \mu| < |\lambda_i - \mu|, \quad \text{pour } \lambda_i \in Sp(A) \setminus \{\lambda_{i_0}\},$$

et si $x^{(0)} = \alpha_{i_0}u_{i_0} + \sum_{\substack{i=1 \\ i \neq i_0}}^n \alpha_i u_i$ est tel que $\alpha_{i_0} \neq 0$, alors la méthode de la puissance inverse avec translation converge :

$$\lim_{k \rightarrow +\infty} \tilde{\gamma}^{(k)} = \lambda_{i_0}.$$

Démonstration :

Il suffit d'appliquer le théorème 5.3 à la matrice $(A - \mu I)^{-1}$. ■

Remarques :

1) Soit λ_{i_0} la valeur propre de A la plus proche de μ et soit λ_{i_1} la seconde plus proche valeur propre de μ . Le facteur de convergence de la méthode de la puissance inverse avec translation est alors égal à

$$\frac{|\mu - \lambda_{i_0}|}{|\mu - \lambda_{i_1}|}$$

(il suffit d'appliquer le corollaire 5.1 à la matrice $(A - \mu I)^{-1}$).

2) La "direction" du vecteur $x^{(k)}$, défini dans (5.28), converge vers une direction propre associée à la valeur propre λ_{i_0} (la plus proche de μ). ■

5.4 La méthode de Jacobi

Cette méthode s'applique aux matrices réelles et symétriques d'ordre n . Elle permet d'obtenir une approximation de toutes les valeurs propres et de tous les vecteurs propres.

Avant de décrire la méthode de Jacobi, nous allons donner son principe.

La matrice A étant symétrique, alors elle est diagonalisable. Et par suite, il existe une matrice orthogonale O telle que $O^T A O$ soit diagonale :

$$O^T A O = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

où les nombres λ_i , $1 \leq i \leq n$, sont les valeurs propres de la matrice A , comptées avec leur ordre de multiplicité. Rappelons que les vecteurs colonnes de la matrice O forment un ensemble orthonormal de vecteurs propres, le i -ème vecteur colonne étant un vecteur propre associé à la valeur propre λ_i .

Partant de la matrice $A_0 = A$, la méthode de Jacobi consiste à construire une suite $(Q_k)_k$ de matrices orthogonales, en s'arrangeant pour que la suite de matrices

$$A_{k+1} = Q_{k+1}^T A_k Q_{k+1},$$

encore symétriques, converge vers une matrice diagonale D ayant les mêmes valeurs propres que la matrice A . Notons au passage que toutes les matrices A_k admettent les mêmes valeurs propres qui sont celles de A , puisque

$$\begin{aligned} \det(A_{k+1} - \lambda I) &= \det(Q_{k+1}^T A_k Q_{k+1} - \lambda I) \\ (5.29) \qquad &= \det(Q_{k+1}^T (A_k - \lambda I) Q_{k+1}) \\ &= \det(A_k - \lambda I), \end{aligned}$$

car Q_{k+1} est orthogonale. Par ailleurs, on a :

$$\begin{aligned} A_k = Q_k^T A_{k-1} Q_k &= Q_k^T Q_{k-1}^T A_{k-2} Q_{k-1} Q_k = \dots = \\ (5.30) \qquad &= Q_k^T Q_{k-1}^T \dots Q_1^T A Q_1 Q_2 \dots Q_k, \end{aligned}$$

et si l'on pose

$$O_k = Q_1 Q_2 \dots Q_k,$$

alors on obtient :

$$A_k = O_k^T A O_k.$$

Ceci montre que, si en plus de la convergence de A_k vers une matrice diagonale D , la suite de matrices orthogonales $(O_k)_k$ converge vers une matrice orthogonale O , alors on a $D = O^T A O$, et les vecteurs colonnes de O forment une base orthonormée de vecteurs propres de A .

Le principe de chaque transformation :

$$A_k \longrightarrow A_{k+1} = Q_{k+1}^T A_k Q_{k+1}$$

est d'annuler un coefficient non diagonal de A_k , par exemple celui de plus grand module. À cause de la symétrie de A_k , en fait, on en annule deux. Reste à savoir comment choisir les matrices Q_k . Pour cela, commençons par donner le résultat suivant.

Théorème 5.4

Soient p et q deux entiers tels que $1 \leq p < q \leq n$ et $\theta \in \mathbb{R}$, auxquels on associe la matrice orthogonale

$$(5.31) \quad \Omega = \begin{matrix} & \begin{matrix} p & & q \\ \downarrow & & \downarrow \end{matrix} \\ \begin{matrix} p \rightarrow \\ \\ q \rightarrow \end{matrix} & \begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & 1 & & & & & & \\ & \cdot & \cdot & \cdot & \cos \theta & \cdot & \cdot & \cdot & \sin \theta & \cdot & \cdot & \cdot \\ & & & \cdot & 1 & & & & \cdot & & & \\ & & \circ & \cdot & & \cdot & & & \cdot & & \circ & \\ & & & \cdot & & & 1 & & \cdot & & & \\ & \cdot & \cdot & \cdot & -\sin \theta & \cdot & \cdot & \cdot & \cos \theta & \cdot & \cdot & \cdot \\ & & & \cdot & & & & & \cdot & 1 & & \\ & & \circ & \cdot & & \circ & & & \cdot & & \cdot & \\ & & & \cdot & & & & & \cdot & & & 1 \end{pmatrix} \end{matrix}$$

$$\begin{aligned} \Omega_{pp} &= \Omega_{qq} = \cos \theta, \quad \Omega_{pq} = -\Omega_{qp} = \sin \theta, \\ \Omega_{ii} &= 1, \text{ si } i \neq p \text{ et } i \neq q, \\ \Omega_{ij} &= 0, \text{ si } i \neq j, (i, j) \neq (p, q) \text{ et } (i, j) \neq (q, p). \end{aligned}$$

1) Si la matrice $A = (a_{ij})_{1 \leq i, j \leq n}$ est symétrique et si $B = \Omega^T A \Omega$, alors $B = (b_{ij})_{1 \leq i, j \leq n}$ est symétrique; et on a

$$(5.32) \quad \sum_{i, j=1}^n (b_{ij})^2 = \sum_{i, j=1}^n (a_{ij})^2.$$

2) Si $a_{pq} \neq 0$, alors il existe un unique $\theta \in]-\frac{\pi}{4}, 0[\cup]0, \frac{\pi}{4}[$ tel que $b_{pq} = 0$, et θ est déterminé par

$$(5.33) \quad \cotg 2\theta = \frac{a_{qq} - a_{pp}}{2a_{pq}}.$$

De plus, on a dans ce cas

$$(5.34) \quad \sum_{i=1}^n (b_{ii})^2 = \sum_{i=1}^n (a_{ii})^2 + 2(a_{pq})^2.$$

Démonstration :

1) Il est clair que la matrice $B = \Omega^T A \Omega$ est symétrique lorsque A est symétrique. Montrons (5.32). On a

$$\begin{aligned} (5.35) \quad \sum_{i, j=1}^n (b_{ij})^2 &= \sum_{i=1}^n \sum_{j=1}^n b_{ij} b_{ji}^t = \sum_{i=1}^n (BB^T)_{ii} = \text{tr}(BB^T) \\ &= \text{tr}(\Omega^T A \Omega \Omega^T A^T \Omega) = \text{tr}(\Omega^T A A^T \Omega) \end{aligned}$$

puisque Ω est orthogonale. Or pour deux matrices A et C quelconques, on a toujours $\text{tr}(AC) = \text{tr}(CA)$. Il s'en suit, grâce à (5.35), que

$$\sum_{i,j=1}^n (b_{ij})^2 = \text{tr}(\Omega^T A A^T \Omega) = \text{tr}(\Omega \Omega^T A A^T) = \text{tr}(A A^T).$$

Utilisant le fait que

$$\text{tr}(A A^T) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} a_{ji}^t,$$

et que A est symétrique, on obtient

$$\sum_{i,j=1}^n (b_{ij})^2 = \sum_{i,j=1}^n (a_{ij})^2.$$

Ce qui prouve (5.32).

2) Puisque $B = \Omega^T A \Omega$, alors utilisant la forme particulière de la matrice Ω , on obtient

$$(5.36) \quad \begin{pmatrix} b_{pp} & b_{pq} \\ b_{qp} & b_{qq} \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Or la matrice $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ étant orthogonale, alors le même raisonnement qu'en 1) nous conduit au résultat suivant :

$$(5.37) \quad b_{pp}^2 + b_{qq}^2 + 2b_{pq}^2 = a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2.$$

Par ailleurs, utilisant (5.36), on obtient

$$(5.38) \quad b_{pq} = b_{qp} = a_{pq} \cos 2\theta + \frac{a_{pp} - a_{qq}}{2} \sin 2\theta.$$

Si $a_{pq} \neq 0$ et si on choisit θ tel que

$$(5.39) \quad \cotg 2\theta = \frac{a_{qq} - a_{pp}}{2a_{pq}},$$

alors, d'après (5.38), on a $b_{pq} = 0$.

Utilisant la forme particulière de Ω , on obtient, pour $i \neq p$ et $i \neq q$, $a_{ii} = b_{ii}$. Et par suite, grâce à (5.37), on obtient, pour θ vérifiant (5.39) :

$$\sum_{i=1}^n b_{ii}^2 = \sum_{\substack{i=1 \\ i \neq p, q}}^n b_{ii}^2 + b_{pp}^2 + b_{qq}^2 = \sum_{\substack{i=1 \\ i \neq p, q}}^n a_{ii}^2 + a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2.$$

Ainsi, on a

$$\sum_{i=1}^n b_{ii}^2 = \sum_{i=1}^n a_{ii}^2 + 2a_{pq}^2.$$

Ce qui prouve (5.34).

Pour finir la démonstration, notons que l'existence de θ vérifiant (5.39) est unique dans $] -\frac{\pi}{4}, 0[\cup]0, \frac{\pi}{4}[$. ■

Commentaires :

1) Soit $B = \Omega^T A \Omega$, où Ω est la matrice donnée par (5.31). Ω n'est autre que la matrice de rotation d'angle θ dans le plan engendré par le p -ème et q -ème vecteur de la base canonique. Les coefficients b_{ij} de la matrice B sont donnés par :

$$(5.40) \quad \left\{ \begin{array}{ll} b_{ij} &= a_{ij}, \quad \text{pour } i \neq p, q \text{ et } j \neq p, q \\ b_{pj} &= a_{pj} \cos \theta - a_{qj} \sin \theta, \quad \text{pour } j \neq p, q \\ b_{iq} &= a_{ip} \sin \theta + a_{iq} \cos \theta, \quad \text{pour } i \neq p, q \\ b_{pp} &= a_{pp} \cos^2 \theta + a_{qq} \sin^2 \theta - a_{pq} \sin 2\theta \\ b_{qq} &= a_{pp} \sin^2 \theta + a_{qq} \cos^2 \theta + a_{pq} \sin 2\theta \\ b_{pq} &= a_{pq} \cos 2\theta + \frac{a_{pp} - a_{qq}}{2} \sin 2\theta \end{array} \right.$$

Si $a_{pq} \neq 0$, on choisit $\theta \in]-\frac{\pi}{4}, 0[\cup]0, \frac{\pi}{4}[$ tel que $b_{pq} = 0$, et θ est donné par $\cotg 2\theta = \frac{a_{qq} - a_{pp}}{2a_{pq}}$. Pour calculer les autres coefficients b_{ij} de la matrice $B = \Omega^T A \Omega$, pour ce choix de θ , on a besoin de calculer $\cos \theta$ et $\sin \theta$. En fait, il n'est pas nécessaire de calculer θ , seules interviennent les valeurs des fonctions trigonométriques en θ .

Posons $\xi = \frac{a_{qq} - a_{pp}}{2a_{pq}} = \cotg 2\theta$ et $t = \tg \theta$. Alors t est solution de l'équation

$$(5.41) \quad t^2 + 2\xi t - 1 = 0.$$

Ainsi, pour déterminer t , il suffit de résoudre l'équation (5.41) et de choisir la racine de plus petit module lorsque $\xi \neq 0$. Puisque $t = \tg \theta$, alors on a

$$\cos \theta = \frac{1}{\sqrt{1+t^2}} \quad \text{et} \quad \sin \theta = \frac{t}{\sqrt{1+t^2}} \quad (\text{si } \xi \neq 0).$$

En fait, si $\xi \neq 0$, on a $t = -\xi + (\text{signe } \xi) \sqrt{1+\xi^2}$ et si $\xi = 0$, on choisit $t = 1$.

Les formules (5.40) deviennent alors, en posant $c = \cos \theta = \frac{1}{\sqrt{1+t^2}}$ et $s = \sin \theta = \frac{t}{\sqrt{1+t^2}}$:

$$(5.42) \quad \left\{ \begin{array}{ll} b_{ij} &= a_{ij}, \quad \text{pour } i \neq p, q \text{ et } j \neq p, q \\ b_{pj} &= ca_{pj} - sa_{qj}, \quad \text{pour } j \neq p, q \\ b_{iq} &= sa_{ip} + ca_{iq}, \quad \text{pour } i \neq p, q \\ b_{pp} &= a_{pp} - ta_{pq} \\ b_{qq} &= a_{qq} + ta_{pq} \\ b_{pq} &= 0 \end{array} \right.$$

2) À chaque étape k de la méthode de Jacobi, on choisit donc p_k et q_k ($p_k < q_k$) tels que $(A_k)_{p_k q_k} \neq 0$. L'angle θ_k est donné par le théorème 5.4, i.e. vérifiant

$$\cotg 2\theta_k = \frac{(A_k)_{q_k q_k} - (A_k)_{p_k p_k}}{2(A_k)_{p_k q_k}},$$

et la matrice Q_{k+1} est alors la matrice Ω définie par (5.31), avec $p = p_k$, $q = q_k$ et $\theta = \theta_k$. ■

5.4.1 La méthode de Jacobi classique

La méthode de Jacobi classique consiste à choisir (p_k, q_k) avec $p_k < q_k$ tels que $(A_k)_{p_k q_k}$ soit un coefficient (non diagonal) de plus grand module de la matrice A_k :

$$(5.43) \quad |(A_k)_{p_k q_k}| = \max_{1 \leq i < j \leq n} |(A_k)_{ij}|.$$

Ainsi, partant de $A_0 = A$, une itération de la méthode de Jacobi classique consiste à :

$$(5.44) \quad \left\{ \begin{array}{l} \text{déterminer } (p_k, q_k) \text{ vérifiant (5.43)} ; \\ \text{déterminer la matrice orthogonale } Q_{k+1} \text{ définie par} \\ \quad Q_{k+1} = \Omega, \\ \text{où } \Omega \text{ est donnée par (5.31) avec } \theta = \theta_k, p = p_k \text{ et } q = q_k, \\ \text{et où } \theta_k \text{ est donnée par} \\ \quad \cotg 2\theta_k = \frac{(A_k)_{q_k q_k} - (A_k)_{p_k p_k}}{2(A_k)_{p_k q_k}}; \\ \text{calculer } A_{k+1} = Q_{k+1}^T A_k Q_{k+1}; \quad (\text{le coefficient } (A_{k+1})_{p_k q_k} = 0). \end{array} \right.$$

Pour calculer $A_{k+1} = Q_{k+1}^T A_k Q_{k+1}$ dans l'algorithme (5.44), on ne calcule pas explicitement la matrice Q_{k+1} , mais on utilise les formules (5.42) appliquées à $A = A_k$, $B = A_{k+1} = Q_{k+1}^T A_k Q_{k+1}$ et $\theta = \theta_k$.

Avant de donner un résultat de convergence de la méthode de Jacobi classique, nous donnons quelques résultats intermédiaires. Posons

$$(5.45) \quad A_k = D_k + B_k$$

où D_k est la matrice diagonale formée par les coefficients diagonaux de A_k .

Proposition 5.1

La matrice B_k donnée par (5.45), vérifie

$$(5.46) \quad \lim_{k \rightarrow +\infty} B_k = 0, \quad \text{dans } \mathcal{M}_n(\mathbb{R})$$

Démonstration :

On a $(B_k)_{ij} = (A_k)_{ij}$ si $i \neq j$, et $(B_k)_{ii} = 0$. Puisque $A_{k+1} = Q_{k+1}^T A_k Q_{k+1}$, où Q_{k+1} est définie par (5.44), alors, utilisant le théorème 5.4, on obtient

$$(5.47) \quad \begin{aligned} \sum_{\substack{i,j=1 \\ i \neq j}}^n (A_{k+1})_{ij}^2 &= \sum_{i,j=1}^n (A_{k+1})_{ij}^2 - \sum_{i=1}^n (A_{k+1})_{ii}^2 \\ &= \sum_{i,j=1}^n (A_k)_{ij}^2 - \sum_{i=1}^n (A_k)_{ii}^2 - 2(A_k)_{p_k q_k}^2, \end{aligned}$$

et par suite, on a

$$(5.48) \quad \sum_{\substack{i,j=1 \\ i \neq j}}^n (A_{k+1})_{ij}^2 = \sum_{\substack{i,j=1 \\ i \neq j}}^n (A_k)_{ij}^2 - 2(A_k)_{p_k q_k}^2.$$

Or, par construction, on a

$$|(A_k)_{ij}| \leq |(A_k)_{p_k q_k}|, \quad \text{pour } i \neq j,$$

d'où, grâce à (5.48), on obtient

$$\sum_{\substack{i,j=1 \\ i \neq j}}^n (A_{k+1})_{ij}^2 \leq \underbrace{\left(1 - \frac{2}{n^2 - n}\right)}_{=\alpha} \sum_{\substack{i,j=1 \\ i \neq j}}^n (A_k)_{ij}^2.$$

Comme $|\alpha| < 1$, alors on a nécessairement

$$\lim_{k \rightarrow +\infty} \sum_{\substack{i,j=1 \\ i \neq j}}^n (A_k)_{ij}^2 = 0.$$

Et par suite, on a

$$\lim_{k \rightarrow +\infty} \sum_{i,j=1}^n (B_k)_{ij}^2 = 0,$$

puisque $(B_k)_{ij} = (A_k)_{ij}$ pour $i \neq j$ et $(B_k)_{ii} = 0$. Ainsi, on a

$$\lim_{k \rightarrow +\infty} B_k = 0, \quad \text{dans } \mathcal{M}_n(\mathbb{R}),$$

puisque l'application qui à une matrice M associe $\left(\sum_{i,j=1}^n (M_{ij})^2\right)^{1/2}$ définit une norme matricielle sur $\mathcal{M}_n(\mathbb{R})$. Ce qui prouve (5.46) et achève la démonstration. ■

Lemme 5.2

Soit $(X, \|\cdot\|_X)$ un espace vectoriel normé de dimension finie. Soit $(x_k)_{k \in \mathbb{N}}$ une suite bornée dans X admettant un nombre fini de valeurs d'adhérences.

Si la suite $(x_k)_{k \in \mathbb{N}}$ vérifie $\lim_{k \rightarrow +\infty} \|x_{k+1} - x_k\|_X = 0$, alors elle est convergente.

Démonstration : À faire en exercice. ■

Théorème 5.5

Soit $\mathcal{S}_n$ l'ensemble des permutations de $\{1, 2, \dots, n\}$. Alors, il existe $\sigma \in \mathcal{S}_n$ telle que

$$\lim_{k \rightarrow +\infty} D_k = \text{diag}(\lambda_{\sigma(1)}, \lambda_{\sigma(2)}, \dots, \lambda_{\sigma(n)}).$$

Démonstration :

1) Soit $\|\cdot\|$ la norme matricielle définie, pour $M \in \mathcal{M}_n(\mathbb{R})$, par

$$\|M\| = \left(\sum_{i,j=1}^n (M_{ij})^2 \right)^{1/2}.$$

On a

$$(5.49) \quad \|D_k\|^2 = \sum_{i=1}^n (A_k)_{ii}^2 \leq \sum_{i,j=1}^n (A_k)_{ij}^2.$$

Or, d'après le théorème 5.4 appliqué à $\Omega = Q_k$, $A = A_{k-1}$ et $B = A_k = Q_k^T A_{k-1} Q_k$, on a

$$\sum_{i,j=1}^n (A_k)_{ij}^2 = \sum_{i,j=1}^n (A_{k-1})_{ij}^2.$$

Et par suite, grâce à (5.49), on a $\|D_k\|^2 \leq c$, où c est une constante réelle positive. Ainsi, la suite $(D_k)_k$ est bornée dans $(\mathcal{M}_n(\mathbb{R}), \|\cdot\|)$.

2) Montrons que la suite $(D_k)_k$ admet un nombre fini de valeurs d'adhérences. Soit D une valeur d'adhérence de la suite $(D_k)_k$. On a, d'après (5.45) :

$$\det(D_k + B_k - \lambda I) = \det(A_k - \lambda I),$$

et par suite, grâce à (5.29), on a

$$(5.50) \quad \det(D_k + B_k - \lambda I) = \det(A_k - \lambda I) = \det(A - \lambda I).$$

Soit $(D_{k'})_{k'}$ une sous-suite de $(D_k)_k$ qui converge vers D , alors, d'après (5.50), on a

$$\det(D_{k'} + B_{k'} - \lambda I) = \det(A - \lambda I).$$

En passant à la limite, lorsque $k' \rightarrow +\infty$, on obtient

$$\det(D - \lambda I) = \det(A - \lambda I),$$

puisque, d'après la proposition 5.1, $\lim_{k' \rightarrow +\infty} B_{k'} = 0$. Ainsi D est une matrice diagonale qui admet les mêmes valeurs propres que A . D'où, il existe $\sigma \in \mathcal{S}_n$, telle que $D = \text{diag}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)})$. Il existe donc au plus $n!$ ($= \text{card}\mathcal{S}_n$) valeurs d'adhérences possibles pour la suite $(D_k)_k$.

3) Montrons que

$$\lim_{k \rightarrow +\infty} (D_{k+1} - D_k) = 0, \quad \text{dans } (\mathcal{M}_n(\mathbb{R}), \|\cdot\|).$$

Utilisant (5.42) avec $B = Q_{k+1}^T A_k Q_{k+1}$, on obtient

$$(D_{k+1})_{ii} - (D_k)_{ii} = (A_{k+1})_{ii} - (A_k)_{ii} = \begin{cases} 0, & \text{si } i \neq p_k \text{ et } i \neq q_k \\ -\text{tg}\theta_k \times (A_k)_{p_k q_k}, & \text{si } i = p_k \\ \text{tg}\theta_k \times (A_k)_{p_k q_k}, & \text{si } i = q_k \end{cases}$$

Ainsi, on a

$$\max_{1 \leq i \leq n} |(D_{k+1})_{ii} - (D_k)_{ii}| \leq |(A_k)_{p_k q_k}|,$$

puisque $|\text{tg}\theta_k| \leq 1$ pour $\theta_k \in]-\frac{\pi}{4}, \frac{\pi}{4}[\setminus \{0\}$. Or, $(B_k)_{p_k q_k} = (A_k)_{p_k q_k}$ et $\lim_{k \rightarrow +\infty} B_k = 0$ dans $\mathcal{M}_n(\mathbb{R})$. Ainsi

$$\lim_{k \rightarrow +\infty} \|D_{k+1} - D_k\|_\infty = 0.$$

4) Utilisant les parties 1), 2) et 3) de cette démonstration et le lemme 5.2, on obtient

$$\exists \sigma \in \mathcal{S}_n, \quad \lim_{k \rightarrow +\infty} D_k = \text{diag}(\lambda_{\sigma(1)}, \lambda_{\sigma(2)}, \dots, \lambda_{\sigma(n)}).$$

Ce qui prouve le résultat énoncé et achève la démonstration. ■

Théorème 5.6 (Admis)

Si toutes les valeurs propres de A sont distinctes, alors la suite de matrices $(O_k)_k$, définies par $O_k = Q_1 Q_2 \cdots Q_k$, est convergente, et la matrice $O = \lim_{k \rightarrow +\infty} O_k$ est orthogonale et ses colonnes sont des vecteurs propres de A associés aux valeurs propres $\lambda_{\sigma(1)}, \lambda_{\sigma(2)}, \dots, \lambda_{\sigma(n)}$. ■

5.4.2 Variantes de la méthode de Jacobi

À chaque itération de la méthode de Jacobi, le coût pour chercher l'élément maximum hors diagonale est élevé. Pour éviter cela, on peut envisager les deux variantes suivantes :

1) La méthode de Jacobi cyclique :

Au lieu de choisir (p_k, q_k) tels que $|(A_k)_{p_k q_k}| = \max_{i \neq j} |(A_k)_{ij}|$, on décide de faire varier le couple (p_k, q_k) de façon systématique en prenant, par exemple, successivement les valeurs

$$(1, 2); (1, 3); \dots; (1, n) \text{ puis } (2, 3); (2, 4); \dots; (2, n); (3, 4); \dots$$

Lorsqu'on arrive à $(n-1, n)$, on recommence le cycle.

2) La méthode de Jacobi avec seuil :

On fixe une valeur $\varepsilon > 0$. On choisit alors le couple (p_k, q_k) en effectuant un balayage systématique comme pour la méthode de Jacobi cyclique, mais on n'effectue la rotation (i.e. la multiplication de A_k à gauche par Q_{k+1}^T et à droite par Q_{k+1}) que si $|(A_k)_{p_k q_k}| \geq \varepsilon$ (si $|(A_k)_{p_k q_k}| < \varepsilon$, on passe au couple suivant, ce qui revient à prendre $Q_{k+1} = I$).

Lorsque tous les éléments non diagonaux sont inférieurs à ε , on recommence le procédé en prenant une nouvelle valeur de ε inférieure à la précédente. On arrête le processus lorsque ε est de l'ordre de la précision désirée. ■

5.5 La méthode QR

La méthode QR s'applique au cas d'une matrice quelconque et donne toutes ses valeurs propres.

La méthode QR consiste à construire une suite de matrice $(A_k)_k$ par le procédé suivant :

$$(5.51) \quad \left\{ \begin{array}{l} A_1 = A; \\ \text{On factorise la matrice } A_k \text{ sous la forme QR} \\ \text{(par la méthode de Householder ou de Givens) :} \\ \quad A_k = Q_k R_k \\ \text{où } Q_k \text{ est orthogonale et } R_k \text{ est triangulaire supérieure;} \\ \text{On forme la matrice } A_{k+1} = R_k Q_k. \end{array} \right.$$

Propriétés 5.1

Les matrices $(A_k)_k$ construites dans l'algorithme (5.51), sont toutes semblables à la matrice de départ A par des transformations orthogonales.

Démonstration :

On a $A_{k+1} = R_k Q_k = Q_k^T A_k Q_k$. Comme Q_k est orthogonale, alors A_{k+1} et A_k sont semblables. Posons

$$\tilde{Q}_k = Q_1 Q_2 \cdots Q_k \quad \text{et} \quad \tilde{R}_k = R_1 R_2 \cdots R_k.$$

On a alors

$$A_{k+1} = \tilde{Q}_k^T A \tilde{Q}_k.$$

Ce qui prouve que A_{k+1} est semblable à A , puisque $\tilde{Q}_k$ est orthogonale. Et par suite A_{k+1} admet les mêmes valeurs propres que A . ■

Donnons à présent un résultat de convergence de la méthode QR .

Théorème 5.7 (Admis)

Supposons que la matrice A a des valeurs propres simples vérifiant :

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0.$$

Soit S une matrice telle que

$$S^{-1}AS = D,$$

où la matrice D est diagonale avec :

$$d_{ii} = \lambda_i, \quad 1 \leq i \leq n.$$

Supposons que la matrice S^{-1} se factorise sous la forme

$$S^{-1} = LU,$$

avec L triangulaire inférieure à diagonale unité et U triangulaire supérieure.

Alors la méthode QR converge au sens suivant :

$$\lim_{k \rightarrow +\infty} (A_k)_{ij} = 0, \quad 1 \leq j < i \leq n$$

$$\lim_{k \rightarrow +\infty} (A_k)_{ii} = \lambda_i, \quad 1 \leq i \leq n$$

■

Remarque : Dans la pratique, on n'applique pas la méthode QR à la matrice A , mais à une matrice $\tilde{A} = (\tilde{a}_{ij})_{1 \leq i, j \leq n}$ semblable à A telle que $\tilde{A}$ soit de Hessenberg supérieure, i.e. $\tilde{a}_{ij} = 0$ pour $1 \leq j \leq i - 2$ (c.f. séries d'exercices). ■

5.6 La méthode de Givens-Householder (bissection)

C'est une méthode particulièrement bien adaptée à la recherche de valeurs propres sélectionnées d'une matrice **symétrique**, par exemple toutes les valeurs propres situées dans un intervalle déterminé à l'avance, ou bien les valeurs propres d'un rang donné (en les supposant ordonnées par ordre croissant), etc Par contre, cette méthode ne permet pas le calcul des vecteurs propres.

La méthode de Givens-Householder (ou bissection) comprend deux étapes :

– **1ère étape : tridiagonalisation**

Étant donnée une matrice symétrique A , on détermine une matrice orthogonale P (d'un type particulier) telle que la matrice (symétrique) $P^T A P$ soit tridiagonale. Cette étape, qui ne nécessite qu'un nombre fini d'opérations élémentaires, constitue la méthode de Householder de réduction d'une matrice symétrique à la forme tridiagonale.

– **2ème étape : bissection**

On est ainsi ramené au calcul des valeurs propres d'une matrice **symétrique tridiagonale**, qui s'effectue par la méthode de Givens, appelée encore méthode de la bissection.

5.6.1 Tridiagonalisation d'une matrice symétrique (Householder)

Soit donc $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice symétrique d'ordre n , qu'on écrit sous la forme :

$$A = \left(\begin{array}{c|c} a_{11} & a_1^t \\ \hline a_1 & \tilde{A}_1 \end{array} \right) \quad \text{où } a_1 \in \mathbb{R}^{n-1} \text{ et } \tilde{A}_1 \in \mathcal{M}_{n-1}(\mathbb{R}).$$

Décrivons la première étape de la méthode de Householder.

Si $a_1 \neq 0$, alors il existe une matrice de Householder élémentaire $\tilde{H}_1$ (symétrique et orthogonale) dans $\mathcal{M}_{n-1}(\mathbb{R})$ telle que l'on ait :

$$\tilde{H}_1 a_1 = \alpha e^{(1)} = (\alpha, 0, \dots, 0)^t \in \mathbb{R}^{n-1}, \quad \text{où } \alpha \in \mathbb{R}.$$

Soit H_1 la matrice de $\mathcal{M}_n(\mathbb{R})$ définie par :

$$H_1 = \left(\begin{array}{c|c} 1 & O \\ \hline O & \tilde{H}_1 \end{array} \right)$$

On a alors :

$$\begin{aligned} H_1^T A H_1 &= \left(\begin{array}{c|c} 1 & O \\ \hline O & \tilde{H}_1^T = \tilde{H}_1 \end{array} \right) \left(\begin{array}{c|c} a_{11} & a_1^t \\ \hline a_1 & \tilde{A}_1 \end{array} \right) \left(\begin{array}{c|c} 1 & O \\ \hline O & \tilde{H}_1 \end{array} \right) \\ &= \left(\begin{array}{c|c} a_{11} & a_1^t \\ \hline \tilde{H}_1 a_1 & \tilde{H}_1 \tilde{A}_1 \end{array} \right) \left(\begin{array}{c|c} 1 & O \\ \hline O & \tilde{H}_1 \end{array} \right) = \left(\begin{array}{c|c} a_{11} & a_1^t \tilde{H}_1 = (\alpha e^{(1)})^t \\ \hline \alpha & \tilde{H}_1 \tilde{A}_1 \tilde{H}_1 \\ 0 & \\ \vdots & \\ 0 & \end{array} \right) \end{aligned}$$

D'où l'on obtient :

$$H_1^T A H_1 = H_1 A H_1 = \left(\begin{array}{c|ccc} a_{11} & \alpha & 0 & \dots & 0 \\ \hline \alpha & & & & \\ 0 & & \tilde{H}_1 \tilde{A}_1 \tilde{H}_1 & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right)$$

Notons A_2 cette matrice : $A_2 = H_1 A H_1$. En notant $A_1 = A$ et en suivant la même démarche que ci-dessus, on détermine de proche en proche $(n-2)$ matrices symétriques et orthogonales $H_1, H_2, \dots, H_{n-2}$ de $\mathcal{M}_n(\mathbb{R})$, telles que les matrices symétriques :

$$A_k = H_{k-1}^T A_{k-1} H_{k-1} = (H_1 H_2 \dots H_{k-1})^T A (H_1 H_2 \dots H_{k-1}), \quad 2 \leq k \leq n-1$$

soient de la forme :

$$A_k = \left(\begin{array}{ccc|c|c} \times & \times & \circ & & \\ \times & \cdot & \cdot & & \circ \\ & \cdot & \cdot & \times & \\ \circ & & \times & \times & a_k^t \\ \hline & \circ & & a_k & \tilde{A}_k \end{array} \right)$$

où $\tilde{A}_k \in \mathcal{M}_{n-k}(\mathbb{R})$, et a_k est le vecteur de $\mathbb{R}^{n-k}$ de composantes les éléments a_{ik}^k , $k+1 \leq i \leq n$, de la matrice $A_k = (a_{ij}^k)_{1 \leq i, j \leq n}$. De la sorte, la matrice

$$A_{n-1} = (H_1 H_2 \cdots H_{n-2})^T A (H_1 H_2 \cdots H_{n-2})$$

est tridiagonale et est semblable à A , puisque la matrice $P = H_1 H_2 \cdots H_{n-2}$ est orthogonale.

Décrivons à présent la k -ème étape de la méthode de Householder qui consiste à transformer la matrice A_k en la matrice $A_{k+1} = H_k^T A_k H_k$. Chaque transformation $A_k \mapsto A_{k+1}$ est effectuée à l'aide d'une matrice H_k de la forme :

$$H_k = \left(\begin{array}{c|c} I_k & O \\ \hline O & \tilde{H}_k \end{array} \right)$$

où $I_k \in \mathcal{M}_k(\mathbb{R})$ désigne la matrice identité et $\tilde{H}_k \in \mathcal{M}_{n-k}(\mathbb{R})$ est la matrice de Householder élémentaire vérifiant :

$$\tilde{H}_k a_k = (\alpha, 0, \dots, 0)^t \in \mathbb{R}^{n-k}, \quad \alpha \in \mathbb{R}.$$

[Si le vecteur a_k est nul, on choisit $\tilde{H}_k = I$, matrice identité de $\mathcal{M}_{n-k}(\mathbb{R})$].

Utilisant les propriétés de la multiplication par blocs des matrices, on obtient :

$$A_{k+1} = H_k^T A_k H_k = \left(\begin{array}{ccc|c} \times & \times & \circ & \circ \\ \times & \cdot & \cdot & \circ \\ & \cdot & \cdot & \times \\ \hline \circ & & \times & \times \\ \hline & \circ & \boxed{\tilde{H}_k a_k} & \boxed{(\tilde{H}_k a_k)^t} \\ & & \boxed{\tilde{H}_k \tilde{A}_k \tilde{H}_k} & \end{array} \right)$$

Plus précisément, on a

$$A_{k+1} = \left(\begin{array}{ccc|c} \times & \times & \circ & \circ \\ \times & \cdot & \cdot & \circ \\ & \cdot & \cdot & \times \\ \hline \circ & & \times & \times \\ \hline & \circ & \boxed{\alpha} & \boxed{0} \\ & & \boxed{0} & \boxed{0} \\ & & \boxed{0} & \boxed{0} \\ & & \boxed{\tilde{H}_k \tilde{A}_k \tilde{H}_k} & \end{array} \right)$$

avec $a_{ij}^{k+1} = a_{ij}^k$ pour $1 \leq i, j \leq k$.

À la fin de la $(n-2)$ -ème étape, on obtient une matrice A_{n-1} tridiagonale et semblable à la matrice A :

$$A_{n-1} = (H_1 H_2 \cdots H_{n-2})^T A (H_1 H_2 \cdots H_{n-2}).$$

Récapitulons ce que l'on vient de faire :

Théorème 5.8

Étant donnée une matrice symétrique A d'ordre n , il existe une matrice orthogonale P , produit de $(n-2)$ matrices de Householder, telle que la matrice $P^T A P$ soit tridiagonale et semblable à A . ■

5.6.2 La Méthode de la bisection

Cette méthode permet la recherche de valeurs propres d'une matrice $B \in \mathcal{M}_n(\mathbb{R})$ **symétrique et tridiagonale** :

$$B = \begin{pmatrix} b_1 & c_1 & & & \\ c_1 & b_2 & c_2 & & O \\ & c_2 & \cdot & \cdot & \\ O & & \cdot & \cdot & c_{n-1} \\ & & & c_{n-1} & b_n \end{pmatrix}$$

[La matrice B jouera le rôle de la matrice A_{n-1} obtenue à l'étape de tridiagonalisation].

On observe tout d'abord que si l'un des éléments c_i est nul, $1 \leq i \leq n-1$, alors la matrice B se décompose par blocs en deux sous-matrices tridiagonales :

$$B = \left(\begin{array}{ccccc|ccccc} b_1 & c_1 & & & O & & & & \\ c_1 & \cdot & & & \cdot & & & & \circ \\ & \cdot & & & c_{i-1} & & & & \\ O & & c_{i-1} & & b_i & c_i=0 & & & \\ & & & c_i=0 & & b_{i+1} & c_{i+1} & & O \\ & & & & & c_{i+1} & \cdot & \cdot & \\ & & & & & & \cdot & \cdot & c_{n-1} \\ & & & & & & O & c_{n-1} & b_n \end{array} \right) = \left(\begin{array}{c|c} C & O \\ \hline O & D \end{array} \right),$$

et par suite les valeurs propres de B sont celles de C ou de D , puisque $\text{Sp}(B) = \text{Sp}(C) \cup \text{Sp}(D)$, où $\text{Sp}(B)$ désigne le spectre de la matrice B . Ainsi, dans le cas où l'un des éléments $c_i = 0$, on est ramené à la recherche de valeurs propres de matrices de plus petites tailles.

On peut donc supposer, sans restreindre la généralité, que pour tout $i = 1, \dots, n-1$, $c_i \neq 0$.

Soit B_i la sous-matrice d'ordre i de B , donnée par :

$$B_i = \begin{pmatrix} b_1 & c_1 & & & \\ c_1 & b_2 & c_2 & & O \\ & c_2 & \cdot & \cdot & \\ O & & \cdot & \cdot & c_{i-1} \\ & & & c_{i-1} & b_i \end{pmatrix}, \quad 2 \leq i \leq n$$

et soit $p_i(\lambda)$ le polynôme caractéristique de B_i , $2 \leq i \leq n$:

$$p_i(\lambda) = \det(B_i - \lambda I_i), \quad \text{où } I_i \in \mathcal{M}_i(\mathbb{R}).$$

Les valeurs propres de B sont alors les racines de p_n .

Nous commençons par établir que les racines des polynômes p_i ont des propriétés "d'emboîtement" illustrées au théorème suivant.

Théorème 5.9

On pose

$$p_0(\lambda) = 1 \quad \text{et} \quad p_1(\lambda) = b_1 - \lambda.$$

Alors, les polynômes caractéristiques p_i des sous-matrices B_i de B vérifient la relation :

$$(5.52) \quad p_i(\lambda) = (b_i - \lambda)p_{i-1}(\lambda) - c_{i-1}^2 p_{i-2}(\lambda), \quad 2 \leq i \leq n$$

et les propriétés suivantes :

$$(5.53) \quad \begin{cases} \lim_{\lambda \rightarrow -\infty} p_i(\lambda) = +\infty \\ \lim_{\lambda \rightarrow +\infty} p_i(\lambda) = \begin{cases} +\infty, & \text{si } i \text{ est pair} \\ -\infty, & \text{si } i \text{ est impair} \end{cases} \end{cases} \quad 1 \leq i \leq n$$

$$(5.54) \quad \text{Si } p_i(\lambda_0) = 0, \quad \text{alors } p_{i-1}(\lambda_0)p_{i+1}(\lambda_0) < 0, \quad 1 \leq i \leq n-1.$$

$$(5.55) \quad \text{Le polynôme } p_i \text{ possède } i \text{ racines réelles distinctes, qui séparent les } (i+1) \text{ racines du polynôme } p_{i+1}, 1 \leq i \leq n-1.$$

Démonstration :

1) Montrons (5.52). Il suffit pour cela de développer le déterminant de la matrice $(B_i - \lambda I_i)$ par rapport à la dernière ligne (ou colonne).

2) Montrons (5.53). Pour cela, nous allons montrer par récurrence que le coefficient du monôme λ^i dans $p_i(\lambda)$ est égal à $(-1)^i$, $1 \leq i \leq n$. Il est clair que cette propriété est vérifiée pour $i = 1$. Supposons alors que jusqu'à l'ordre $i-1$, le coefficient de λ^k dans p_k est égal à $(-1)^k$, $1 \leq k \leq i-1$. Utilisant alors la relation (5.52), on vérifie que le coefficient de λ^i dans p_i est égal à $(-1)^i$. Et par suite (5.53) est vérifiée.

3) Montrons (5.54). Soit donc λ_0 un zéro de p_i , $1 \leq i \leq n-1$. Utilisant alors (5.52), on obtient :

$$p_{i+1}(\lambda_0) = (b_{i+1} - \lambda_0) \underbrace{p_i(\lambda_0)}_{=0} - c_i^2 p_{i-1}(\lambda_0) = -c_i^2 p_{i-1}(\lambda_0)$$

Comme $c_i \neq 0$, alors deux cas peuvent se produire : soit

$$p_{i+1}(\lambda_0)p_{i-1}(\lambda_0) < 0$$

et donc (5.54) est vérifiée ; soit

$$(5.56) \quad p_{i+1}(\lambda_0) = p_{i-1}(\lambda_0) = 0.$$

Dans ce deuxième cas, utilisant (5.52) :

$$\underbrace{p_i(\lambda_0)}_{=0} = (b_i - \lambda_0) \underbrace{p_{i-1}(\lambda_0)}_{=0} - c_{i-1}^2 p_{i-2}(\lambda_0)$$

on obtient $p_{i-2}(\lambda_0) = 0$, puisque $c_{i-1} \neq 0$. Il s'en suit, grâce à (5.52), que

$$p_i(\lambda_0) = p_{i-1}(\lambda_0) = \dots = p_0(\lambda_0) = 0,$$

ce qui contredit le fait que $p_0(\lambda_0) = 1$. Ainsi le deuxième cas (5.56) ne peut pas avoir lieu. Ainsi si $p_i(\lambda_0) = 0$, alors nécessairement $p_{i-1}(\lambda_0)p_{i+1}(\lambda_0) < 0$, ce qui prouve (5.54).

4) Montrons (5.55). Il est clair, que les racines de p_i sont toutes réelles, puisque la matrice B_i est symétrique. Le fait qu'elles sont toutes distinctes découle de ce que l'on va montrer, à savoir que les zéros de p_i séparent strictement les zéros de p_{i+1} , ce qui se traduit par :

$$\lambda_1^{i+1} < \lambda_1^i < \lambda_2^{i+1} < \lambda_2^i < \dots < \lambda_i^{i+1} < \lambda_i^i < \lambda_{i+1}^{i+1}$$

où $\lambda_1^i, \dots, \lambda_i^i$ sont les zéros de p_i et $\lambda_1^{i+1}, \dots, \lambda_{i+1}^{i+1}$ ceux de p_{i+1} . Montrons ce résultat par récurrence.

On a $p_1(\lambda_1^1) = 0$ ($\lambda_1^1 = b_1$). D'où, d'après (5.54), on obtient

$$(5.57) \quad \underbrace{p_0(\lambda_1^1)}_{=1} p_2(\lambda_1^1) < 0.$$

On a, d'après (5.52) :

$$\begin{aligned} p_2(\lambda) &= (b_2 - \lambda)p_1(\lambda) - c_1^2 p_0(\lambda) \\ &= (b_2 - \lambda)p_1(\lambda) - c_1^2 \\ &= (b_2 - \lambda)(b_1 - \lambda) - c_1^2, \end{aligned}$$

on en déduit le tableau de variation suivant :

λ	$-\infty$	λ_1^2	λ_1^1	λ_2^2	$+\infty$
p_1	$+\infty$	$+$	0	$-$	$-\infty$
p_2	$+\infty$	0	<0 (*)	0	$+\infty$

(*) d'après (5.57)

Ainsi, $\lambda_1^2 < \lambda_1^1 < \lambda_2^2$, i.e. la racine de p_1 sépare strictement les deux racines de p_2 . Ce qui prouve que le résultat est vérifié à l'ordre 1.

Supposons que le résultat est vrai jusqu'à l'ordre i , i.e. que les zéros de p_i séparent ceux de p_{i+1} :

$$\lambda_1^{i+1} < \lambda_1^i < \lambda_2^{i+1} < \dots < \lambda_i^{i+1} < \lambda_i^i < \lambda_{i+1}^{i+1}$$

et démontrons le à l'ordre $i+1$, i.e. que les zéros de p_{i+1} séparent ceux de p_{i+2} :

$$\lambda_1^{i+2} < \lambda_1^{i+1} < \lambda_2^{i+2} < \dots < \lambda_{i+1}^{i+2} < \lambda_{i+1}^{i+1} < \lambda_{i+2}^{i+2}$$

Pour cela, donnons le tableau de variation de p_i , p_{i+1} et p_{i+2} en se basant sur l'hypothèse de récurrence et sur les propriétés (5.52), (5.53) et (5.54).

λ	$-\infty$	λ_1^{i+1}	λ_1^i	λ_2^{i+1}	$\dots$	λ_i^{i+1}	λ_i^i	λ_{i+1}^{i+1}	$+\infty$		
p_i	$+\infty$	$ $	$ $	$ $	$ $	$ $	$+$	$-$	$-\infty^{(1)}$		
	$+$	$\oplus$	$+$	0	$-$	$\ominus$	$\dots$	$ $	0		
	$ $	$ $	$ $	$ $	$ $	$ $	$-$	$+$	$\oplus$	$+$	$+\infty^{(2)}$
p_{i+1}	$+\infty$	$ $	$ $	$ $	$ $	$ $	$ $	$ $	$+$	$+\infty^{(1)}$	
	$+$	0	$-$	0	$\dots$	0	$ $	0	$ $	$+$	$+\infty^{(1)}$
	$ $	$ $	$ $	$ $	$ $	$ $	$ $	$ $	$ $	$-$	$-\infty^{(2)}$
p_{i+2}	$+\infty$	$ $	$ $	$ $	$ $	$ $	$ $	$\oplus$	$-\infty^{(1)}$		
	$\mathbf{0}^{(3)}$	$\ominus$	$\mathbf{0}$	$\oplus$	$\dots$	$ $	$ $	$\mathbf{0}$	$ $	$\mathbf{0}$	$-\infty^{(1)}$
	$+\infty$	$ $	$ $	$ $	$ $	$ $	$ $	$\ominus$	$+\infty^{(2)}$		
	λ_1^{i+2}	$ $	λ_2^{i+2}	$ $	$\dots$	$ $	$ $	$ $	λ_{i+2}^{i+2}		

(1) : cas où i est impair, (2) : cas où i est pair.

(3) : Montrons que p_{i+2} admet un zéro dans l'intervalle $] -\infty, \lambda_1^{i+1}[$.

Puisque λ_1^{i+1} est un zéro de p_{i+1} , alors, d'après (5.54), on a $p_i(\lambda_1^{i+1})p_{i+2}(\lambda_1^{i+1}) < 0$. Or, d'après le tableau de variation ci-dessus, on a $p_i(\lambda_1^{i+1}) > 0$, et par suite on a $p_{i+2}(\lambda_1^{i+1}) < 0$. Il s'en suit, grâce à (5.53), que p_{i+2} admet un zéro, noté λ_1^{i+2} , dans l'intervalle $] -\infty, \lambda_1^{i+1}[$.

En utilisant le même raisonnement que ci-dessus, on montre que p_{i+2} admet $(i+2)$ zéros qui sont séparés par les $(i+1)$ zéros de p_{i+1} , ce qui prouve (5.55) et achève la démonstration. ■

Remarque : Une suite de polynômes vérifiant les conditions (5.53), (5.54) et (5.55) du théorème 5.9 précédent est appelée *suite de Sturm*.

La méthode de Givens (ou de bissection) repose sur une propriété tout à fait remarquable d'une telle suite, qui est de permettre un calcul immédiat du nombre des racines $< \mu$, $\mu \in \mathbb{R}$ donné, de chacun des polynômes de la suite, comme le montre le théorème suivant.

Théorème 5.10

Soit i un entier appartenant à $\{1, 2, \dots, n\}$. Étant donné un nombre $\mu \in \mathbb{R}$, on pose

$$\text{sgn}(p_i(\mu)) = \begin{cases} \text{signe de } p_i(\mu), & \text{si } p_i(\mu) \neq 0 \\ \text{signe de } p_{i-1}(\mu), & \text{si } p_i(\mu) = 0. \end{cases}$$

Soient $E(i, \mu)$ l'ensemble ordonné, de $(i+1)$ signes, défini par :

$$E(i, \mu) = \{+, \text{sgn}(p_1(\mu)), \text{sgn}(p_2(\mu)), \dots, \text{sgn}(p_i(\mu))\}$$

et $N(i, \mu)$ le nombre de changement de signes entre éléments consécutifs de l'ensemble $E(i, \mu)$.

On a alors

$$(5.58) \quad N(i, \mu) = \text{nombre de zéros de } p_i \text{ qui sont } < \mu.$$

Démonstration : Vérifions d'abord que l'expression $\text{sgn}(p_i(\mu))$ est bien définie dans tous les cas. Si $p_i(\mu) = 0$, alors, d'après (5.54), on a $p_{i-1}(\mu) \neq 0$. Ainsi $\text{sgn}(p_i(\mu))$ est bien défini.

Nous allons démontrer le résultat (5.58) par récurrence. Pour $i = 1$, on a $p_1(\mu) = b_1 - \mu$. Deux cas peuvent alors se produire (rappelons que b_1 est le zéro de p_1) :

– si $b_1 < \mu$, alors, d'après (5.53), $p_1(\mu) < 0$, d'où $E(1, \mu) = \{+, -\}$. Et par suite, $N(1, \mu) = 1$;

– si $b_1 \geq \mu$, alors $p_1(\mu) > 0$, d'où $E(1, \mu) = \{+, +\}$. Et par suite, $N(1, \mu) = 0$.

Ainsi dans les deux cas, $N(1, \mu)$ est égal au nombre de zéros de p_1 qui sont $< \mu$.

Supposons la propriété vraie pour les entiers $1, 2, \dots, i$ et montrons qu'elle est encore vraie pour l'entier $i+1$. Soient

$$\lambda_1^i < \lambda_2^i < \dots < \lambda_i^i \quad \text{et} \quad \lambda_1^{i+1} < \lambda_2^{i+1} < \dots < \lambda_{i+1}^{i+1}$$

les racines, rangées par ordre croissant, des polynômes p_i et p_{i+1} respectivement. On a, par hypothèse de récurrence, $N(i, \mu) = \text{nombre de zéros de } p_i \text{ qui sont } < \mu$:

$$\lambda_1^i < \lambda_2^i < \dots < \lambda_{N(i, \mu)}^i < \mu \leq \lambda_{N(i, \mu)+1}^i < \dots < \lambda_i^i.$$

Utilisant (5.55), on a :

$$\lambda_{N(i, \mu)}^i < \lambda_{N(i, \mu)+1}^{i+1} < \lambda_{N(i, \mu)+1}^i.$$

Il suffit alors d'examiner les deux cas suivants :

- **1er cas :** $\lambda_{N(i, \mu)}^i < \mu \leq \lambda_{N(i, \mu)+1}^{i+1}$

Dans ce cas, $\text{sgn}(p_{i+1}(\mu)) = \text{sgn}(p_i(\mu))$ d'après le tableau de variation ci-dessus si $\mu < \lambda_{N(i,\mu)+1}^{i+1}$. Si $\mu = \lambda_{N(i,\mu)+1}^{i+1}$ alors $\text{sgn}(p_{i+1}(\mu)) = \text{sgn}(p_i(\mu))$ par construction. Et par suite on a $N(i, \mu) = N(i+1, \mu)$ et donc (5.58) est vérifiée à l'ordre $i+1$.

- **2ème cas :** $\lambda_{N(i,\mu)+1}^{i+1} < \mu \leq \lambda_{N(i,\mu)+1}^i$

Dans ce cas $\text{sgn}(p_{i+1}(\mu)) = -\text{sgn}(p_i(\mu))$ d'après le tableau de variation ci-dessus. D'où $N(i+1, \mu) = N(i, \mu) + 1$ et donc (5.58) est vérifiée à l'ordre $i+1$.

Ainsi (5.58) est vérifiée pour tout entier i . ■

Application : On souhaite calculer la valeur propre de B de rang j_0 :

$$\lambda_1 < \lambda_2 < \dots < \lambda_{j_0} < \dots < \lambda_n.$$

On part d'un intervalle $[\alpha_0, \beta_0]$ qui contient la valeur propre que l'on cherche, par exemple :

$$[\alpha_0, \beta_0] = [-\|B\|, \|B\|]$$

où $\|\cdot\|$ est une norme matricielle. Soit alors

$$\mu_0 = \frac{\alpha_0 + \beta_0}{2}$$

Utilisant le théorème 5.10, on a :

$$\begin{array}{ll} \text{Si } N(n, \mu_0) \geq j_0 & \text{alors } \lambda_{j_0} \in [\alpha_0, \mu_0], \quad \beta_1 = \mu_0 \\ \text{Si } N(n, \mu_0) < j_0 & \text{alors } \lambda_{j_0} \in [\mu_0, \alpha_0], \quad \alpha_1 = \mu_0 \end{array}$$

de telle sorte que $\lambda_{j_0} \in [\alpha_1, \beta_1]$. Ainsi, on construit une suite d'intervalles $[\alpha_k, \beta_k]$ tels que $\lambda_{j_0} \in [\alpha_k, \beta_k]$. On arrête alors les itérations lorsque $|\alpha_k - \beta_k| \leq \varepsilon$, où ε est un paramètre assez petit, et on prend comme approximation de λ_{j_0} la valeur de $(\alpha_k + \beta_k)/2$. ■