

Compléments de cours en analyse numérique matricielle

M. Granger

1 Recherche de vecteurs propres et de valeurs propres

1.1 Méthode de la puissance itérée

1.1.1 Recherche de la valeur propre de plus grand module.

Considérons l'exemple suivant: la matrice $\begin{pmatrix} 10 & 0 \\ -9 & 1 \end{pmatrix}$ admet deux valeurs propres $\lambda = 10$ et $\mu = 1$, avec les vecteurs propres unitaires (pour la norme $\|\bullet\|_\infty$):

$$v_{10} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{et} \quad v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Considérons alors l'opération $v_k = A^k v_0$, de vecteur initial $v_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. On trouve

$$v_k = \begin{pmatrix} 10^k \\ -10^k + 2 \end{pmatrix}$$

On peut alors faire les constatations suivantes:

1. La direction limite pour v_k est un vecteur propre pour $\lambda = 10$. C'est à dire que

$$\frac{v_k}{\|v_k\|_\infty} = \begin{pmatrix} 1 \\ -1 + 2 \times 10^{-k} \end{pmatrix} \xrightarrow[k \rightarrow \infty]{} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

2. Le quotient $\frac{v_{k+1}(i)}{v_k(i)}$ est défini pour k assez grand et $\frac{v_{k+1}(i)}{v_k(i)} \xrightarrow[k \rightarrow \infty]{} \lambda = 10$

Cet exemple nous suggère l'algorithme suivant de recherche d'une valeur propre de module maximum. On fixe une norme $\|\bullet\|$ sur \mathbb{C}^n , et une matrice $A \in \mathcal{M}_{n,n}(\mathbb{C})$, et on désigne par $v(j)$ la j ème coordonnée d'un vecteur $v \in \mathbb{C}^n$:

ALGORITHME

1. Choisir $q_0 \in \mathbb{C}^n$ tel que $\|q_0\| = 1$.
2. $x_k = A.q_{k-1}$ et si $q_{k-1}(j) \neq 0$, $\lambda_k(j) = \frac{x_k(j)}{q_{k-1}(j)}$

3. Si $\gamma_k = \|x_k\| \neq 0$, on définit $q_k = \frac{x_k}{\gamma_k}$

Remarques:

- 1) Si $\gamma_k = 0$, l'algorithme s'arrête au cran k , ce qui n'arrive pas pour A inversible ni pour q_0 assez général (précisément, pourvu que A ne soit pas nilpotente, dès que q_0 est en dehors du sous espace caractéristique généralisé relatif à zéro)
- 2) q_k est unitaire donc pour chaque k : $\exists j, q_{k-1}(j) \neq 0$ autrement dit $\lambda_k(j)$ est défini. Si q_k a une limite sur la sphère unité, on peut choisir j indépendant de k , pour k assez grand.
- 3) $q_k = \frac{A^k q_0}{\|A^k q_0\|}$, par une récurrence immédiate sur k .

Théorème 1.1 *On suppose que A est diagonalisable avec une seule valeur propre notée λ_1 de module maximum, de multiplicité $p < n$. On ordonne les valeurs propres $\lambda_1 = \dots = \lambda_p, \lambda_{p+1}, \dots, \lambda_n$ par modules décroissants.*

1) *Pour q_0 bien choisi, précisément en dehors d'un sous espace vectoriel strict, la suite construite dans l'algorithme fournit, à la fois, une valeur approchée de λ_1 et un vecteur propre de la façon suivante:*

a) $\lim_{k \rightarrow \infty} (\frac{\overline{\lambda_1}}{|\lambda_1|})^k q_k = q$ existe et on a :

$$\|q\| = 1, \text{ et } q \in E_{\lambda_1}, \text{ sous espace propre associé à } \lambda_1$$

b) D'après a), il existe j tel que $\exists k_0, \forall k \geq k_0, q_k(j) \neq 0$, et alors :

$$\lambda_1 = \lim_{k \rightarrow \infty} \frac{x_{k+1}(j)}{q_k(j)}$$

2) L'erreur $\|(\frac{\overline{\lambda_1}}{|\lambda_1|})^k q_k - q\|$, est majorée par une série géométrique $Cste |\frac{\lambda_{p+1}}{\lambda_1}|^k$.

Démonstration On considère une base (u_1, \dots, u_n) de vecteurs propres, avec u_j associé à λ_j , et on note α_j , la jème coordonnée de q_0 :

$$q_0 = \alpha_1 u_1 + \dots + \alpha_n u_n$$

Dans la somme directe $\mathbb{C}^n = E_{\lambda_1} \oplus (\bigoplus_{\lambda_j \neq \lambda_1} E_{\lambda_j})$ la composante de q_0 sur E_{λ_1} est $u = \alpha_1 u_1 + \dots + \alpha_p u_p$ et on a

$$q_0 = u + \sum_{j=p+1}^n \alpha_j u_j.$$

La condition requise sur q_0 est $u \neq 0$, ce qui équivaut encore à $q_0 \notin \bigoplus_{\lambda_j \neq \lambda_1} E_{\lambda_j}$. Le vecteur $q_k = \frac{A^k q_0}{\|A^k q_0\|}$ est donc le vecteur unitaire associé à

$$\begin{aligned} A^k q_0 &= \lambda_1^k u + \sum_{j=p+1}^n \alpha_j \lambda_j^k u_j \\ &= \lambda_1^k (u + \sum_{j=p+1}^n \alpha_j (\frac{\lambda_j}{\lambda_1})^k u_j) \end{aligned}$$

Posons $e_k = \sum_{j=p+1}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^k u_j$, on a $\lim_{k \rightarrow \infty} e_k = 0$, et plus précisément du fait que pour tout $j \geq p+1$, $\lambda_j \leq \lambda_p + 1$, on déduit l'existence d'une constante C (en fait $C = \sum |\alpha_j| \|u_j\|$), telle que :

$$\|e_k\| \leq C \left| \frac{\lambda_{p+1}}{\lambda_1} \right|^k$$

On en tire

$$q_k = \frac{\lambda_1^k}{|\lambda_1|^k} \cdot \frac{u + e_k}{\|u + e_k\|} = \frac{|\lambda_1|^k}{\overline{\lambda_1}^k} \cdot \frac{u + e_k}{\|u + e_k\|}$$

donc

$$q'_k = \frac{\overline{\lambda_1}^k}{|\lambda_1|^k} q_k \xrightarrow[k \rightarrow \infty]{} q = \frac{u + e_k}{\|u + e_k\|}$$

Ceci démontre le premier résultat.

Pour le résultat sur la vitesse de convergence il suffit d'écrire: $q'_k - q = \frac{\|u\| \cdot (u + e_k) - \|u + e_k\| u}{\|u + e_k\| \cdot \|u\|}$, majoré en norme par $\frac{2 \cdot \|e_k\|}{\|u + e_k\|}$. \square

1.2 Généralisation

Si A est non diagonalisable mais avec une seule valeur propre, notée λ_1 , de module maximum, le résultat est encore valable. Ceci peut se voir sur les blocs de Jordan de A . On note :

$$J_q(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}$$

un bloc de Jordan de taille $q \times q$.

Cas 1 λ_1 est une valeur propre simple. La réduction de Jordan de A s'écrit alors:

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_1 & & 0 \\ & & & J_{q_2}(\lambda_2) & \\ 0 & & & & \ddots & \\ & & & & & J_{q_p}(\lambda_p) \end{pmatrix}$$

où $\lambda_2, \dots, \lambda_p$, sont les valeurs propres de A de modules tous $< |\lambda_1|$. On a encore une convergence vers zéro de $q'_k - q$, avec une vitesse de convergence similaire, puisque le reste est majorée par une série polynôme \times géométrique proportionnelle à $k^n |\frac{\lambda_2}{\lambda_1}|^k$. Les détails sont laissés en exercice.

Cela revient à considérer

$$P^{-1}A^kP = \begin{pmatrix} \lambda_1^k & & & & \\ & \ddots & & & \\ & & \lambda_1^k & & 0 \\ & & & J_{q_2}(\lambda_2)^k & \\ 0 & & & & \ddots & \\ & & & & & J_{q_p}(\lambda_p)^k \end{pmatrix}$$

et à voir que tous les coefficients de $\frac{J_{q_i}(\lambda_j)^k}{\lambda_1^k}$, sont de la forme $\frac{C_k^l \lambda_2^{k-\ell}}{\lambda_1^k}$, avec ℓ borne par n .

Cas général La convergence est beaucoup plus lente dès qu'il existe des blocs de Jordan de taille ≥ 2 pour λ_1 . Supposons pour simplifier les notations, tout en traitant essentiellement la question, que A est égale à un bloc de Jordan: $A = J_n(\lambda)$. Alors, la $(j+1)$ ième colonne de A^k est le vecteur:

$$x_k = A^k(e_{j+1}) = \begin{pmatrix} C_k^j \lambda_1^{k-j} \\ \vdots \\ C_k^1 \lambda_1^{k-1} \\ \lambda_1^k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

et on trouve encore que $q_k = \frac{x_k}{\|x_k\|}$, tend vers e_1 , vecteur propre.

Toutefois, $\frac{(q_{k+1})_1}{(q_k)_1} = \lambda_1 \frac{C_{k+1}^j}{C_k^j} = \frac{k+1}{k-j+1}$ tend vers λ_1 très lentement, le reste étant cette fois un $O(\frac{1}{k})$.

Remarques. 1) S'il existe plusieurs valeurs propres de module maximum il n'y a plus convergence. Par exemple, une matrice de rotation a deux valeurs propres $e^{i\theta}$ et $e^{-i\theta}$, et quel que soit $z \in \mathbb{C}^2$, dont la décomposition sur les espaces propres est $z = z_+ + z_-$, la direction du vecteur $Az = e^{in\theta}z_+ + e^{-in\theta}z_-$, ne tend vers aucune des deux directions propres.

2) (Sur le choix de q_0) Dans les problèmes concrets, où la valeur de λ_1 , n'est pas connu, il en est de même de l'espace des choix à exclure pour q_0 . Mais dans la pratique un choix au hasard contient toujours une composante propre sur A_{λ_1} .

3) Dans une matrice prise au hasard, selon la commande *rand*(n, n) de scilab par exemple, on est dans les hypothèses du théorème, l'espace des choix à exclure (matrice non diagonalisables ou avec des relations $|\lambda_i| = |\lambda_j|$) étant de mesure nulle.

4) Si on choisit $q_0 \in \bigoplus_{j \geq p+1} \mathbb{C}.u_j$, dans les notations de la démonstration du théorème 1.1, on trouverait de même, et en théorie la valeur propre λ_{p+1} . Dans la pratique, sauf cas simple où l'espace propre E_{λ_1} est connu explicitement, les erreurs d'arrondi sur q_0 , entraînent l'apparition d'une composante non nulle sur E_{λ_1} , qui finit toujours par l'emporter.

Pour une donnée initiale très proche de $E_{\lambda_{p+1}}$, on trouve une suite $k \mapsto \lambda_k(j)$, présentant un palier, en λ_2 , avant de converger vers λ_1 , comme dans la figure ci-dessous.

Voir dans le Lascaux Theodor des exemples de ce type ainsi que des exemples à convergence lente où $\frac{\lambda_2}{\lambda_1} \sim 0.99$.

1.2.1 La méthode de la puissance inverse

On peut d'abord remarquer que si A est inversible avec des valeurs propres rangées par ordre décroissant de façon que :

$$|\lambda_1| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n| > 0,$$

la méthode précédente appliquée à A^{-1} donne $\frac{1}{\lambda_n}$, comme plus grande valeur propre de A^{-1} , puis λ_n par inversion.

ALGORITHME pour la plus petite valeur propre.

- 1) Effectuer une fois pour toute une factorisation LU .
- 2) Résoudre $Ax_k = q_{k-1}$
- 3) Poursuivre le reste de l'algorithme comme dans le cas de la recherche de la plus grande valeur propre.

De même, la méthode appliquée à $(A - \mu I)^{-1}$, aboutit à $\frac{1}{\lambda - \mu}$, avec λ valeur propre la plus proche de μ .

1.2.2 Méthode de déflation

Lorsque les valeurs propres de A sont rangées par ordre de module croissant, en supposant pour simplifier ces modules tous distincts :

$$|\lambda_1| > \dots > |\lambda_n|$$

on peut chercher à calculer successivement $\lambda_1, \lambda_2, \dots$. A titre d'exemple nous donnons la méthode par soustraction appropriée pour les matrices hermitiennes. Dans ce cas $\lambda_i \in \mathbb{R}$.

Le principe est le suivant:

Soit q un vecteur propre de A pour la valeur propre λ_1 trouvé par exemple par la méthode de la puissance et qu'on supposera unitaire : $q' * q = 1$.

On note $B = A - \lambda_1 q \star q'$. Alors on remarque que $B \star q = 0$, et que si $v \in q^\perp$ est orthogonal à q , on a $B \star v = A \star v$. Comme A est diagonalisable dans une base orthonormée comme matrice hermitienne, on constate que B a les mêmes valeurs propres et vecteurs propres, à l'exception de λ_1 remplacé par zéro.

Ainsi de proche en proche on pourrait déterminer tout le spectre.

Malgré son apparente simplicité conceptuelle cette méthode a des limites car elle n'est pas stable à cause des erreurs d'arrondis.

Cette méthode est utilisée dans les problèmes où on ne souhaite calculer qu'un nombre limité de valeurs propres à partir des plus grandes.

2 Méthode de Jacobi

Il s'agit d'une méthode, assez ancienne, de réduction des matrices symétriques, efficace sur les matrices pleines.

Le principe est le suivant: on construit une suite de matrices orthogonales, Ω_k , et une suite de matrices semblables à A :

$$A_0 = A, \text{ et } A_{k+1} = {}^t\Omega_k A_k \Omega_k$$

la matrice Ω_k étant choisie à chaque étape pour annuler un terme non diagonal, et "se rapprocher" d'une matrice diagonale.

On utilise la matrice $\Omega = \Omega_{p,q}(\theta) = \omega_{i,j}$ définie par

$$\begin{cases} \omega_{i,j} = \delta_{i,j}, & \text{si } \{i,j\} \neq \{p,q\} , \\ \begin{pmatrix} \omega_{p,p} & \omega_{p,q} \\ \omega_{q,p} & \omega_{q,q} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \end{cases}$$

On appellera matrice orthogonale élémentaire une telle matrice.

Théorème 2.1 Soit $B = (b_{i,j})$ la matrice semblable à A , et symétrique:

$$B = \Omega^{-1} A \Omega = {}^t\Omega A \Omega$$

$$1) \sum b_{i,j}^2 = \sum a_{i,j}^2$$

2) Si $a_{p,q} \neq 0$, il existe un unique $\theta \in]-\frac{\pi}{4}, \frac{\pi}{4}] \setminus \{0\}$ tel que $b_{p,q} = 0$. Précisément θ est déterminé par l'égalité

$$\cotan\theta = \frac{a_{q,q} - a_{p,p}}{2a_{p,q}}$$

3) Avec la valeur de θ trouvée en 2), on a en plus:

$$\sum_{i=1}^n b_{i,i}^2 = \sum_{i=1}^n a_{i,i}^2 + 2a_{p,q}^2$$

Le dernier résultat montre que dans la somme fixe $\sum a_{i,j}^2$, la part de la diagonale augmente à chaque étape.

Démonstration 1) Un calcul élémentaire montre (exercice!) que $\sum a_{i,j}^2 = \text{tr } {}^tAA$. La propriété 1) se déduit alors du fait que les matrices tAA et tBB sont semblables, donc de même trace, conséquence de l'orthogonalité de Ω et du calcul suivant:

$${}^tBB = {}^t\Omega {}^tA\Omega {}^t\Omega A\Omega = {}^t\Omega ({}^tAA)\Omega$$

2) Le passage de B à A est un changement de base qui laisse fixe les e_j pour $j \notin \{p,q\}$, et qui dans l'espace engendré par $\{e_p, e_q\}$ est un changement de bases orthonormées directes. Les deux conséquences suivantes s'en déduisent:

- Si $\{i, j\} \cap \{p, q\} = \emptyset$, $b_{i,j} = a_{i,j}$.
- $\begin{pmatrix} b_{p,p} & b_{p,q} \\ b_{q,p} & b_{q,q} \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} a_{p,p} & a_{p,q} \\ a_{q,p} & a_{q,q} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$

où $c = \cos \theta$, $s = \sin \theta$. La dernière relation donne alors :

$$\begin{aligned} b_{p,q} &= b_{q,p} = c(a_{p,p}s + a_{p,q}c) - s(a_{q,p}s + a_{q,q}c) \\ &= cs(a_{p,p} - a_{q,q}) + a_{p,q}(c^2 - s^2) = \frac{\sin 2\theta}{2}(a_{p,p} - a_{q,q}) + \cos 2\theta a_{p,q} \end{aligned}$$

La condition requise pour annuler $b_{p,q}$ est donc:

$$\boxed{\frac{\cos 2\theta}{\sin 2\theta} = \frac{a_{q,q} - a_{p,p}}{2a_{p,q}}}$$

équation qui possède bien une solution unique avec $2\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}] \setminus \{0\}$.

3) Les sous-matrices 2×2 , $\begin{pmatrix} b_{p,p} & b_{p,q} \\ b_{q,p} & b_{q,q} \end{pmatrix}$ et $\begin{pmatrix} a_{p,p} & a_{p,q} \\ a_{q,p} & a_{q,q} \end{pmatrix}$, extraites de A et B par restriction à l'espace engendré par e_p et e_q , s'échangent aussi par une tranformation orthogonale élémentaire. On peut donc appliquer à cette matrice le résultat 1) ce qui s'écrit:

$$b_{p,p}^2 + b_{q,q}^2 = a_{p,p}^2 + a_{q,q}^2 + 2a_{p,q}^2$$

et donne l'égalité annoncée en 3). \square

On peut illustrer ce théorème par le schéma suivant

$$\left(\begin{array}{ccccccccc} & & & b_{1,p} & & & b_{1,q} & & \\ & = & & \times & = & & \times & = & \\ & & & \times & & & \times & = & \\ b_{p,1} & \times & \times & b_{p,p} & \times & \times & \times & 0 & \times & \times & b_{p,n} \\ & & & \times & & & \times & & & & \\ & = & & \times & = & & \times & = & & & \\ & & & \times & & & \times & & & & \\ b_{q,1} & \times & \times & 0 & \times & \times & \times & b_{q,q} & \times & \times & b_{q,n} \\ & & & \times & & & \times & & & & \\ & = & & \times & = & & \times & = & & & \\ & = & & b_{n,p} & = & & b_{n,q} & = & & & \end{array} \right)$$

Dans ce schéma, les termes contenus dans les cases avec un signe $=$ au centre sont inchangés par rapport aux $a_{i,j}$, alors que tous les $a_{i,j}$ dans lequel un au moins des indices est dans $\{p, q\}$ sont altérés par le transformation.

Voir dans Lascaux theodor, tome 2 p. , l'illustration suivante, aisée à programmer: on représente dans chaque position (i, j) la taille du coefficient $a_{i,j}^{(k)}$ de A_k à la k ème itération par un carré de coté $|a_{i,j}^{(k)}|$. L'aire totale de ces carrés est constante selon le résultat 1) du théorème

et selon le résultat 3), on observe une accumulation d'aire sur la diagonale, et sous réserve de convergence l'évanouissement progressif des autres carrés.

N.B. Prendre garde au fait que si à la k ième étape on a obtenu $b_{p,q} = 0$, une rotation ultérieure (avec par exemple un indice sélectionné) (p, q') peut faire resurgir in coefficient non nul d'indice (p, q) . Il n'est donc pas vrai que le processus converge en $\frac{n(n-1)}{2}$ étapes.

Dans ce qui suit on note $t = \tan \theta$. On rappelle que:

$$\cos 2\theta = \frac{1-t^2}{1+t^2}, \quad \sin 2\theta = \frac{2t}{1+t^2}, \quad \cot 2\theta = \frac{1-t^2}{2t}$$

La valeur de t issu du point 2) du théorème est donc la solution de module ≤ 1 (ou $t = 1$ dans le cas des racines $-1, +1$) de l'équation

$$t^2 + \frac{a_{q,q} - a_{p,p}}{a_{p,q}} t - 1 = 0, \quad t \in]0, 1] \quad (1)$$

ALGORITHME

- Sélectionner un indice (p, q) tel que $|a_{p,q}|$ soit maximum.
- Calculer t solution de l'équation 1.
- Faire $b_{p,q} = b_{q,p} = 0$, et pour $\{i, j\} \cap \{p, q\} \neq \emptyset$ avec $(p, q) \neq (i, j) \neq (q, p)$, calculer:

$$b_{i,j} = ({}^t\Omega A \Omega)_{i,j}$$

Commentaire sur l'algorithme 1) Il s'agit d'une description non formalisée de la k ième étape. A chaque étape on réaffecte conformément aux formules les variables $A(i, j)$.

2) Dans l'item 3) il est inutile de préciser que les autre $a_{i,j}$ sont inchangé.

3) On peut programmer une boucle du type pour k de 1 à n , ou en fonction d'une précision choisie avec une instruction conditionnelle: *for* $i \neq j$, *while* $|a_{i,j}| > 10^{-N}$.

4) Dans le calcul des $b_{i,j}$ à programmer selon le troisième item, il n'est pas utile de déterminer θ , puisque toutes les formules passent par des expressions algébriques de la variable t . En effet:

$$c = \cos t = \frac{1}{\sqrt{1+t^2}}, \quad \text{et } s = \sin t = ct = \frac{t}{\sqrt{1+t^2}}.$$

Variantes dans la stratégie de choix de l'indice (p, q) .

1) Au lieu de choisir le plus grand des $a_{p,q}$, on peut effectuer un balayage de tous les indices

$$(p, q) = (2, 1), \dots, (n, 1), (3, 2), \dots, \dots, (n, n-1)$$

Inconvénient: ne maximise pas à chaque étape le gain sur $\sum a_{i,i}^2$. Avantage: gain de temps de calcul consistant à éviter à chaque étape la recherche d'un max avec $n(n-1)/2$ comparaisons.

2) Même stratégie avec instruction contidionnelle d'omettre p, q si $|a_{p,q}| < 10^{-N}$. Elimine les étapes non rentables.

Théorème 2.2 Dans la méthode de Jacobi classique avec choix du coefficient maximum à chaque étape, la convergence a lieu: $\lim A_k = D$, matrice diagonale semblable à A

Démonstration On note $A_k = D_k + B_k$, le résultat de la k ième itération et D_k sa partie diagonale $D_k = \text{diag}(a_{1,1}^{(k)}, \dots, a_{n,n}^{(k)})$.

i) Montrons d'abord que la suite B_k tend vers zéro. Soit $\|\bullet\|_E$, la norme (qui n'est pas subordonnée ni même sous-multiplicative!) définie par $\|(a_{i,j})\|^2 = \sum a_{i,j}^2$. En vertu du point 1) du théorème 2.1, on a $\|A_k\| = \|A_{k+1}\|$, et si on pose

$$\epsilon_k = \|B_k\|^2 = \sum_{i \neq j} a_{i,j}^{(k)},$$

il s'agit d'établir le fait que $\epsilon_k \xrightarrow[k \rightarrow \infty]{} 0$.

Par le point 3) du théorème 2.1, on a :

$$0 \leq \epsilon_{k+1} = \epsilon_k - 2|a_{p_k, q_k}^{(k)}|^2 < \epsilon_k$$

Par le choix du coefficient maximum à l'étape k on a : $\forall (i, j), |a_{i,j}^{(k)}| \leq |a_{p_k, q_k}^{(k)}|$, et on obtient $\epsilon_k \leq n(n-1)|a_{p_k, q_k}^{(k)}|^2$. D'où

$$0 \leq \epsilon_{k+1} \leq \epsilon_k \left(1 - \frac{2}{n(n-1)}\right)$$

Par récurrence cela donne $\epsilon_k \leq \epsilon_0 \left(1 - \frac{2}{n(n-1)}\right)^k$ et montre que $B_k \xrightarrow[k \rightarrow \infty]{} 0$.

ii) On remarque ensuite que la suite D_k n'a qu'un nombre fini de valeurs d'adhérence. Soit en effet D une telle valeur d'adhérence et D_{k_p} , une suite extraite tendant vers D . Comme $B_k = A_k - D_k$ tend vers zéro on a aussi $\lim_{p \rightarrow \infty} A_{k_p} = D$, donc le polynôme caractéristique de D est (coefficient par coefficient) la limite de celui de A_{k_p} , qui est semblable à A_0 . Il en résulte que $P_D(X) := \det(XI - D) = \det(XI - A_0)$, est indépendant du choix de D . Les termes de la diagonale de D étant les racines de P_D sont les mêmes que les racines comptées avec multiplicités de $P_{A_0} = \det(XI - A_0)$ et cela donne au plus $n!$ choix possibles pour D (et même moins en cas de valeurs propres multiples).

iii) On remarque enfin que la suite $D_{k+1} - D_k$ tend vers zéro. En effet on peut extraire des calculs du théorème 2.1, l'évaluation suivante de la diagonale de $B - A$

$$\begin{aligned} b_{p,p} - a_{p,p} &= a_{p,p}c^2 + a_{q,q}s^2 - 2a_{p,q}cs - a_{p,p} = (a_{q,q} - a_{p,p})\sin^2 \theta - a_{p,q} \sin 2\theta \\ &= 2a_{p,q} \frac{\cos 2\theta}{\sin 2\theta} \sin^2 \theta - a_{p,q} \sin 2\theta = -\tan \theta \cdot a_{p,q} \end{aligned}$$

Du fait que $|\tan \theta| \leq 1$, on en tire $|b_{p,p} - a_{p,p}| \leq |a_{p,q}|$. Appliquée à la k ième itération de l'algorithme de Jacobi on déduit de ce calcul :

$$|a_{p_k, p_k}^{(k+1)} - a_{p_k, p_k}^{(k)}| \leq |a_{p_k, q_k}^{(k)}| \leq \|B_k\|_E$$

On a la même majoration pour $a_{q_k, q_k}^{(k+1)} - a_{q_k, q_k}^{(k)}$, qui est le seul autre coefficient non nul de $D_{k+1} - D_k$ et comme $\|B_k\|_E$ tend vers zéro cela donne le résultat voulu.

Pour conclure que A_k a une limite D , ou ce qui revient au même d'après le point i) que D_k a une limite D , il reste à appliquer le lemme de topologie suivant, laissé en exercice au lecteur, en remarquant aussi que la suite A_k est bornée (puisque la norme $\|A_k\|_E \geq \|D_k\|_E$ est constante égale à $\|A_0\|_E$) :

Lemme 2.3 Soit u_n une suite dans un espace vectoriel normé de dimension finie telle que

$$\lim_{k \rightarrow \infty} (u_{k+1} - u_k) = 0$$

et n'ayant qu'un nombre fini de valeurs d'adhérences. Alors la suite u_k est convergente ou tend en norme vers l'infini.

Indication sur le lemme: Soient l_1, \dots, l_q les valeurs d'adhérence de la suite u_n . On choisit $\epsilon > 0$ tel que quelque soit la paire d'indices (i, j) distincts on a $|l_i - l_j| \geq 3\epsilon$, et $R > 0$, tel que la boule fermée $\overline{B} = \overline{B}(0, R)$ contient toutes les boules $\overline{B}(l_i, 2\epsilon)$.

Considérons alors le compact $K = \overline{B}(0, R) \setminus B(l_i, \epsilon)$. On montre facilement qu'il existe n_0 , tel que $n \geq n_0 \Rightarrow u_n \notin K$. On en tire l'alternative :

$$\left[q = 1 \text{ et } \lim_{k \rightarrow \infty} u_k = l_1 \right] \quad \text{ou} \quad \left[q = 0 \text{ et } \lim_{k \rightarrow \infty} \|u_k\| = +\infty \right]$$

□