

Cours no 4
Le 24 février 2012

Mathématiques appliquées et numériques

Licence 3, Dpt Géosciences
Année 2011-2012, 2e semestre

Présentation synthétique du cours

Janvier – Juin 2012

Cours donné en 3^e année de
Licence de Sciences de la planète Terre
par Michael Ghil et Jean Roux
TD par Mohamadou Diallo
École normale supérieure, Paris

Quatrième cours

Algèbre linéaire I. Valeurs et vecteurs propres d'une matrice : théorie et méthodes numériques

C'est un cours particulièrement utile à la résolution des systèmes d'équations différentielles linéaires.

Il est à peine utile de noter que le calcul des valeurs et vecteurs propres se pose dans de nombreux problèmes de la physique et en particulier dans les Sciences de la Terre. Les complications qui se présentent sont bien plus grandes que pour la résolution des systèmes linéaires. Tout d'abord il n'existe pas de méthode directe : toutes les méthodes sont itératives ; ceci est clair puisqu'il s'agit, comme on va le voir, de calculer les zéros du polynôme caractéristique. En raison des grandes difficultés attachées à ce problème, ce chapitre ne doit être considéré que comme une introduction au sujet.

4.1 Valeurs propres et vecteurs propres d'une matrice

De grande importance dans l'étude des matrices $A \in \mathbb{R}^{n,n}$ sont les vecteurs spéciaux (les vecteurs propres ci-dessous) dont les directions sont inchangées lorsqu'ils sont multipliés par A . Un scalaire réel ou complexe λ tel que

$$Ax = \lambda x, \quad x \neq 0, \quad (4.1.1)$$

est appelé une *valeur propre* de A et x un vecteur propre (associé à λ) de A . Il se peut évidemment qu'une matrice réelle ait une valeur propre complexe (voir l'exemple de la matrice A à la fin de ce paragraphe). Notons aussi qu'en réalité les valeurs propres complexes apparaissent par paires complexes conjuguées, ce sera démontré ci-dessous. Si λ est complexe le vecteur propre associé x est complexe, car s'il était réel, on aurait Ax réel et λx complexe, ce qui est impossible ; de plus si $Ax = \lambda x$ avec $\lambda \in \mathbb{C}$ et nécessairement x complexe (on vient de le démontrer), on a (en prenant le conjugué des deux membres) $\overline{Ax} = \overline{\lambda x}$, soit, puisque A est réelle par hypothèse, $A\bar{x} = \bar{\lambda}\bar{x}$. Ce qui signifie que $\bar{\lambda}$ est aussi valeur propre de A associée au vecteur propre \bar{x} , qui est le complexe conjugué du vecteur propre x associé à λ .

Lorsqu'une valeur propre est connue, la détermination du vecteur propre associé revient à résoudre le système linéaire homogène $(A - \lambda I)x = 0$. Il est évident que si x est vecteur propre, alors αx est aussi vecteur propre pour tout scalaire $\alpha \neq 0$ (le vecteur propre est donc défini à un scalaire près non nul). On peut donc choisir α tel que $\|x\| = 1$. Il suit que λ est une valeur propre si et seulement si le système linéaire homogène $(A - \lambda I)x = 0$ a une solution non triviale $x \neq 0$, ou, de façon équivalente, si et seulement si la matrice $A - \lambda I$ est singulière. Par conséquent les valeurs

propres satisfont à l'équation $p_A(\lambda) = \det(A - \lambda I) = 0$, où $p_A(\lambda)$ est un *polynôme caractéristique* de degré n si la matrice A est d'ordre n . Les *valeurs propres* $\lambda_i = \lambda_i(A)$, $1 \leq i \leq n$, sont les n racines, distinctes ou confondues, réelles et/ou imaginaires de $p_A(\lambda)$. Maintenant il est facile de voir que si λ est une racine complexe de $p_A(\lambda)$ où A est une matrice réelle, alors $\bar{\lambda}$ est aussi racine. Ce polynôme s'écrit évidemment $p_A(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n$ où les coefficients a_i sont réels. Soit $\lambda_0 \in \mathbb{C}$ une racine de $p_A(\lambda)$, on a : $0 = p_A(\lambda_0) = a_0 + a_1\lambda_0 + \cdots + a_n(\lambda_0)^n$. Prenons le complexe conjugué de tous les membres, on a : $\bar{0} = \overline{p_A(\lambda_0)} = a_0 + a_1\bar{\lambda}_0 + \cdots + a_n(\bar{\lambda}_0)^n$, ou encore $0 = a_0 + a_1\bar{\lambda}_0 + \cdots + a_n(\bar{\lambda}_0)^n$, ce qui prouve que $\bar{\lambda}_0$ est aussi racine. Ce qui démontre l'assertion.

Le calcul "à la main" des valeurs et vecteurs propres peut être parfois un peu pénible, même s'il s'agit toujours de trouver les racines d'un polynôme (ce n'est possible que pour n de l'ordre de 2 ou 3) et de résoudre un petit système linéaire algébrique. L'emploi d'algorithmes numériques est donc usuellement nécessaire.

Nous avons le

Théorème 4.1.1. *Si λ_1 et λ_2 sont des valeurs propres telles que $\lambda_1 \neq \lambda_2$, alors les vecteurs propres correspondants sont linéairement indépendants.*

Démonstration : Soient $Av_1 = \lambda_1v_1$ et $Av_2 = \lambda_2v_2$. Raisonnons par l'absurde en supposant que v_1 et v_2 sont linéairement dépendants, c'est-à-dire qu'il existe une constante $k \neq 0$ telle que $v_2 = kv_1$. Alors $A(kv_1) = \lambda_2(kv_1)$ soit $Av_1 = \lambda_2v_1$, en soustrayant de la relation $Av_1 = \lambda_1v_1$ il vient $(\lambda_2 - \lambda_1)v_1 = 0$ avec $v_1 \neq 0$, on a donc $\lambda_1 = \lambda_2$ ce qui est absurde. \square

Il peut exister deux valeurs propres identiques sur la diagonale associées à deux vecteurs propres indépendants. Par exemple soit la matrice I_2 suivante $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, le nombre 1 est une valeur propre double et on peut choisir $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ comme vecteurs propres indépendants. On distingue la *multiplicité algébrique* d'une valeur propre - ici $\lambda = 1$ est une valeur propre de multiplicité algébrique 2 - de sa *multiplicité géométrique* qui est la dimension de l'espace propre associé. Dans le cas de la matrice I_2 les deux multiplicités sont les mêmes. Par contre la matrice $B = \begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix}$ a pour valeur propre μ de multiplicité algébrique 2, mais sa multiplicité géométrique est 1 (exercice).

Si la multiplicité algébrique d'une valeur propre est égale à un, la valeur propre est dite *simple*, elle est dite *multiple* dans le cas contraire.

Si $\lambda \in \mathbb{C}$ est valeur propre, alors $\bar{\lambda}$ est valeur propre. En dimension 2 l'exemple typique est celui de la matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Il est impossible,

en dimension 2, que la matrice D se présente sous la forme $D = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$, car i serait racine double du polynôme caractéristique $(\lambda - i)^2 = \lambda^2 - 2i\lambda - 1$ qui n'est pas à coefficients réels comme on le suppose. Par contre il peut arriver (il est facile d'exhiber un exemple naïf), par exemple en dimension 4, que $D = \begin{pmatrix} i & & & \\ & -i & 0 & \\ & 0 & i & \\ & & & -i \end{pmatrix}$, les valeurs propres doubles apparaissant par paires conjuguées.

4.2 Réduction d'une matrice

Soit un espace vectoriel X de dimension finie n . Soit $\mathcal{A} : X \rightarrow X$ un opérateur linéaire représenté par la matrice A dans une base (e_i) de X . Si (f_i) est une autre base de X , le même opérateur est représenté dans cette base (f_i) par la matrice

$$A' = F^{-1}AF$$

où F est la matrice régulière dont le $j^{\text{ème}}$ vecteur colonne représente le vecteur (f_j) dans la base (e_i) . La matrice F est dite la *matrice de changement de base*. La matrice A' est dite *semblable* à A .

Le même opérateur linéaire \mathcal{A} est ainsi représenté par différentes matrices selon la base choisie. Le problème se pose donc de trouver une base dans laquelle la matrice représentant l'opérateur \mathcal{A} soit "particulièrement simple"-on réduit la matrice à sa forme la plus simple possible. Le meilleur des cas est celui où il existe une matrice **régulière** P (la matrice de changement de base) telle que la matrice $P^{-1}AP$ soit diagonale. On a alors

$$P^{-1}AP = \begin{pmatrix} d_1 & & & \\ & d_2 & & 0 \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}. \quad (4.2.1)$$

Or il est évident que si λ_i est une valeur propre de A , elle est aussi (c'est une vérification élémentaire) une valeur propre de $D = P^{-1}AP$ et réciproquement ; le spectre de cette dernière matrice étant évidemment l'ensemble des d_i . Si donc la matrice A est diagonalisable, la matrice D est constituée que des valeurs propres de A , c'est-à-dire que $D = \text{diag}(\lambda_i)$. On vérifie aussi que si y_i est un vecteur propre de D correspondant à λ_i ($Dy_i = \lambda_i y_i$), alors Py_i est un vecteur propre de A correspondant à λ_i : en effet de l'égalité $AP = PD$ on tire $A(Py_i) = P(Dy_i) = \lambda_i(Py_i)$, ce qui démontre l'assertion.

Si la matrice est diagonalisable les λ_i ne sont pas nécessairement tous différents, mais il existe toujours n vecteurs propres linéairement indépendants. La réciproque est vraie : si une matrice possède n vecteurs propres

linéairement indépendants - ce qui revient à dire que la matrice P précédente est régulière - alors elle est diagonalisable. Si toutes les valeurs propres sont distinctes alors la matrice est diagonalisable, cela suit du fait que les vecteurs propres associés à des valeurs propres distinctes sont linéairement indépendants. Mais il se peut que, la matrice A étant diagonalisable, on ait $\lambda_1 = \lambda_2$ avec x_1 indépendant de x_2 (prendre, par exemple, la matrice identité I_2 de l'exemple du paragraphe précédent).

Toute matrice n'est pas diagonalisable, mais une large classe de matrices possède la propriété de diagonalisation. elle l'est lorsqu'elle est *normale*, i.e. lorsque la matrice réelle A permute avec sa transposée : $A(A^T) = (A^T)A$, ce qui est vérifié si, par exemple, la matrice est symétrique ou orthogonale. Alors il existe une matrice C régulière (invertible) telle que $C^{-1}AC = D$, où la matrice D est diagonale et constituée des valeurs propres de A . En effet démontrons que les matrices A et D ont les mêmes valeurs propres. On a $\det(D - \lambda I) = \det(C^{-1}AC - \lambda I) = \det(C^{-1}(A - \lambda C I C^{-1})C)$, soit $\det(D - \lambda I) = \det(C^{-1}) \det(A - \lambda I) \det(C)$. Comme le déterminant de l'inverse d'une matrice est l'inverse du déterminant de la matrice, on a $\det(D - \lambda I) = \det(A - \lambda I)$, les matrices A et D ont mêmes spectres. En effet démontrons que les matrices A et D ont les mêmes valeurs propres. On a $\det(D - \lambda I) = \det(C^{-1}AC - \lambda I) = \det(C^{-1}(A - \lambda C I C^{-1})C)$, soit $\det(D - \lambda I) = \det(C^{-1}) \det(A - \lambda I) \det(C)$. Comme le déterminant de l'inverse d'une matrice est l'inverse du déterminant de la matrice, on a $\det(D - \lambda I) = \det(A - \lambda I)$, les matrices A et D ont mêmes spectres.

Il existe des matrices qui ne sont pas diagonalisables, mais on peut toujours "au moins" réduire toute matrice A à sa *forme de Jordan*, qui est d'ailleurs diagonale si A est diagonalisable. La mise sous forme de Jordan sera utile au cours 6.

On a le remarquable résultat algébrique suivant :

Théorème 4.2.1. "Théorème" : Soit A une matrice $n \times n$ quelconque. Il existe une matrice régulière C qui réduit A à sa forme de Jordan suivante :

$$J = C^{-1}AC = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_r \end{pmatrix} \quad (4.2.2)$$

où chaque bloc de Jordan J_l , d'ordre n_l , est associé à une valeur propre λ_l de multiplicité algébrique n_l , pour $1 \leq l \leq r$ est associé à, ce que l'on appelle, un sous-espace propre généralisé (ou encore un sous-espace invariant

irréductible). Et où J_l est une matrice triangulaire supérieure de la forme :

$$J_l = \lambda_l \quad \text{si } n_l = 1, \quad \text{ou} \quad J_l = \begin{pmatrix} \lambda_l & 1 & 0 & \dots & \dots & 0 \\ & \lambda_l & 1 & 0 & \dots & 0 \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & 0 \\ & & & & \ddots & 1 \\ & & & & & \lambda_l \end{pmatrix} \quad \text{si } n_l \geq 2 \quad (4.2.3)$$

on a $\sum_{l=1}^r n_l = n$. Chaque bloc de Jordan est associé à un vecteur propre, mais aussi, si $n_l \geq 2$, à un sous-espace propre généralisé. Si λ_k est une valeur propre de A de multiplicité algébrique n_k on peut avoir plusieurs blocs de Jordan, par exemple J_{l_1} et J_{l_2} associés à λ_k de multiplicités algébriques respectives n_{l_1} et n_{l_2} avec $n_{l_1} + n_{l_2} = n_k$: dans ce cas λ_k est associée à deux vecteurs propres linéairement indépendants (la multiplicité géométrique de λ_k est deux), les deux sous-espaces propres généralisés correspondants sont indépendants.

Commentaires : La matrice A est diagonalisable si et seulement si les sous-matrices J_l sont d'ordre 1 (c'est-à-dire que $n_l = 1$, $1 \leq l \leq r = n$).

Nous mettons le mot Théorème entre guillemets, pour indiquer que nous en donnons seulement l'idée (voir le livre pour plus de précisions).

On montre que le nombre de sous-matrices du type J_l correspondant à la même valeur propre coïncide avec le nombre de sous-espaces propres généralisés distincts associés à cette même valeur propre, c'est-à-dire est égal à la multiplicité géométrique de la valeur propre. Cette matrice en forme normale de Jordan est unique à l'ordre des blocs de Jordan près.

S'il existe un bloc de Jordan associé à $\lambda_i \in \mathbb{C}$ d'ordre n_i , il existe un bloc de Jordan associé à $\bar{\lambda}_i$ de même ordre, et les blocs de Jordan associés à λ_i et $\bar{\lambda}_i$ se correspondent 2 à 2.

Il est facile de voir qu'il n'y a qu'un seul vecteur propre (à une constante multiplicative près) attaché à un bloc de Jordan défini par (4.2.3). En effet soit $u = (u_1, u_2, \dots, u_{n_l})^T$ un vecteur propre associé à la valeur propre λ_l . Posons $\lambda = \lambda_l$. Exprimons $J_l u = \lambda u$ composante par composante, il vient

$$\begin{cases} \lambda u_1 + u_2 = \lambda u_1 \\ \lambda u_2 + u_3 = \lambda u_2 \\ \vdots \\ \lambda u_{n_l} = \lambda u_{n_l} \end{cases} \quad . \quad \text{En partant de la première relation il vient } u_2 = 0, \text{ puis}$$

par la deuxième $u_3 = 0$, ainsi de suite jusqu'à l'avant-dernière relation qui donne $u_{n_l-1} = 0$. La dernière équation est alors nécessairement satisfaite. Finalement on a $u = (u_1, 0, \dots, 0)^T$ avec $u_1 \neq 0$ car u est un vecteur propre ; comme tout vecteur propre est défini à une constante près on peut écrire $u = \alpha e_1$ avec α constante, où e_1 est le premier vecteur de la base canonique

de \mathbb{R}^{n_l} .

4.2.1 Valeur propre dominante simple ou multiple. Unicité de la valeur propre dominante

Soit une matrice diagonalisable ou non, on appelle valeur propre *dominante* la valeur propre la plus grande en module.

Notons que, évidemment, la valeur propre dominante peut ne pas être unique. Par exemple si deux valeurs propres dominantes sont réelles et opposées : $\lambda_1 = -\lambda_2 > 0$, λ_1 et $-\lambda_2 > 0$ sont deux valeurs propres distinctes dominantes de même module et on a $\lambda_1 > |\lambda_3| \geq |\lambda_4| \geq \dots$; ou encore si on a une paire dominante de valeurs propres complexes conjuguées : $\lambda_1 \in \mathbb{C}$, $\overline{\lambda_1}$ sont deux valeurs propres distinctes dominantes et $|\lambda_1| = |\overline{\lambda_1}| > |\lambda_3| \geq \dots$.

La valeur propre dominante λ_1 est simple si $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, autrement dit si sa multiplicité algébrique est égale à 1, elle est multiple si, par exemple, dans le cas d'une multiplicité algébrique égale à 2, $\lambda_1 = \lambda_2$ et $|\lambda_1| > |\lambda_3| \geq \dots \geq |\lambda_n|$.

Il faut donc distinguer les notions de multiplicité algébrique d'une valeur propre et d'unicité de la valeur propre dominante. Cette dominance peut être unique pour une valeur propre algébriquement multiple. A contrario, des valeurs propres simples distinctes peuvent être associées à une valeur propre dominante multiple.

Nous allons seulement considérer, dans ce cours, le cas où la valeur propre dominante est unique. Remarquons qu'elle est alors nécessairement réelle, car si elle était complexe, sa complexe conjuguée serait aussi valeur propre dominante et cela contredirait l'hypothèse d'unicité.

4.3 Calcul numérique des valeurs et vecteurs propres. Petit panorama des méthodes.

Il est à peine utile de noter que le calcul des valeurs et vecteurs propres se pose dans de nombreux problèmes de la physique et en particulier dans les Sciences de la Terre. Les complications qui se présentent sont bien plus grandes que pour la résolution des systèmes linéaires. Tout d'abord il n'existe pas de méthode directe : toutes les méthodes sont itératives ; ceci est clair puisqu'il s'agit, en quelque sorte, de calculer les zéros du polynôme caractéristique. En raison des grandes difficultés attachées à ce problème, ce qui suit ne doit être considéré que comme une introduction au sujet.

Dressons d'abord un catalogue non exhaustif des méthodes relatives au calcul : trouver $x \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$ tels que

$$Ax = \lambda x \tag{4.3.1}$$

où $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. On sait qu'une matrice réelle peut avoir des valeurs propres complexes auxquelles on associe nécessairement des vecteurs propres complexes.

(A) Calcul d'une partie du spectre

- Calcul du plus grand (resp. petit) module $\rho(A) = \max_i |\lambda_i|$ (resp. $\min_i |\lambda_i|$) des valeurs propres (que la valeur propre soit simple ou non) par la méthode de la puissance (resp. puissance inverse).
- Estimation des valeurs propres par "paquets" : méthode de Lanczos, méthode des itérations d'un sous espace. C'est hors programme.

(B) Cas d'une matrice symétrique

On démontre que les valeurs propres d'une telle matrice sont réelles (théorème 2.18 page 47 du livre).

- Méthode de Jacobi. On calcule tout le spectre et les vecteurs propres associés. Elle s'applique, en particulier, lorsque la matrice est pleine. C'est hors programme.
- Méthode de Givens-Householder particulièrement adaptée au calcul de valeurs propres sélectionnées. Cette méthode calcule le spectre mais pas les vecteurs propres qui, s'ils doivent être calculés, sont estimés par une autre méthode. Cette méthode de bisection opère en deux étapes
 1. Tridiagonalisation de la matrice A (par une méthode de Householder - voir cours 7),
 2. Calcul du spectre de la matrice symétrique tridiagonale obtenue (semblable à A) en utilisant la méthode de Givens appelée aussi méthode de bisection.

(C) Cas d'une matrice non symétrique

- Méthode LR de Rutishauser. Calcul de tout le spectre. Méthode basée sur la décomposition $A = LU$ - voir cours 7. Cette méthode est étudiée dans les compléments du livre.

Il est utile, avant tout calcul de valeurs propres, d'avoir une idée de leur localisation afin de valider les calculs. Cela peut s'estimer, par exemple, par le théorème de Gerschgorin-Hadamard - voir livre. On sait aussi que le rayon spectral $\rho(A)$ vérifie $\rho(A) \leq \|A\|$ pour toute norme matricielle - voir livre.

4.4 Méthode de la puissance

4.4.1 Exemple

Soit la matrice

$$A = \begin{pmatrix} 10 & 0 \\ -9 & 1 \end{pmatrix}$$

de valeurs propres 10 et 1 auxquelles sont associées respectivement les vecteurs propres $u_1 = (1, -1)^T$ et $u_2 = (0, 1)^T$.

Soit un vecteur initial $x_0 = (2, 1)^T$ et successivement la suite $x_1 = Ax_0, x_2 = Ax_1, \dots, x_{k+1} = Ax_k, \dots$. On obtient le tableau suivant où on indique les valeurs des deux composantes des itérés successifs

x_0	x_1	x_2	x_3	x_4	x_5
2	20	200	2000	20000	200000
1	-17	-197	-1997	-19997	-199997

Constats : i) La direction de x_k tend vers celle du vecteur propre $(1, -1)^T$ associé à la valeur propre dominante.

ii) On a presque $x_{k+1} = 10x_k$; puisque $x_{k+1} = Ax_k$ cela qualifie la valeur 10 comme valeur propre de la matrice A .

Explication : On a $x_k = Ax_{k-1} = \dots = A^k x_0$. Calculons A^k . On a $A = SDS^{-1}$ où

$$D = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ et } S^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Dès lors $A^k = SD^k S^{-1}$ où

$$D^k = \begin{pmatrix} 10^k & 0 \\ 0 & 1 \end{pmatrix};$$

il vient donc

$$A^k = \begin{pmatrix} 10^k & 0 \\ -10^k + 1 & 1 \end{pmatrix}, \text{ d'où } x_k = \begin{pmatrix} 2 \cdot 10^k \\ -2 \cdot 10^k + 3 \end{pmatrix} = 2 \cdot 10^k \begin{pmatrix} 1 \\ -1 + \frac{3}{2 \cdot 10^k} \end{pmatrix}.$$

De cette analyse il ressort que :

1. $x_k \xrightarrow[k \rightarrow \infty]{} u_1$
2. le rapport de deux composantes successives de même indice du vecteur itéré x_k tend vers la valeur propre de plus grand module, en effet :

$$\lim_{k \rightarrow \infty} \frac{2 \cdot 10^k}{2 \cdot 10^{k-1}} = \lim_{k \rightarrow \infty} \frac{-2 \cdot 10^k + 3}{-2 \cdot 10^{k-1} + 3} = 10,$$

on en déduit que pour toute norme

$$\forall x_0, \lim_{k \rightarrow \infty} \frac{\|x_k\|}{\|x_{k-1}\|} = 10.$$

Remarque 4.4.1. *Sur cet exemple les composantes de x_k varient comme 10^k , ce qui entraîne les risques de dépassement de capacité machine ! Comme un vecteur propre est défini à un coefficient de proportionnalité près, il est naturel de le normaliser (à chaque itération par exemple).*

On imagine alors, grâce à cet exemple, l'algorithme suivant.

4.4.2 Algorithme de la puissance

Soit A une matrice réelle d'ordre n . À chaque itération précédente, on normalise x_k par une norme quelconque $\|\cdot\|$ de \mathbb{C}^n , il semble qu'il n'y ait pas d'inconvénient à choisir a priori la norme euclidienne, plus facile à calculer que la norme du maximum, car n'exigeant pas des comparaisons.

On choisit donc la norme euclidienne et on considère l'algorithme suivant:

1. choix de $q_0 \in \mathbb{C}^n$ tel que $\|q_0\| = 1$,
2. pour $k = 1, 2, \dots$ calculer

$$\left\{ \begin{array}{l} \bullet x_k = Aq_{k-1} \\ \bullet \lambda^k(j) = \frac{x_k(j)}{q_{k-1}(j)} \quad j = 1, 2, \dots, n \\ \bullet \gamma_k = \|x_k\| \\ \bullet q_k = \frac{x_k}{\gamma_k} \end{array} \right. ; \quad (4.4.1)$$

on note par $x_k(j)$ la j^e composante du vecteur x_k et par $\lambda^k(j)$ la valeur propre associée à $x_k(j)$.

Le critère d'arrêt de l'algorithme est fourni par la stabilisation de $\lambda^k(j)$ autour d'une même valeur.

4.4.3 Théorème de convergence

Afin d'étudier la convergence des suites définies par l'algorithme, posons une définition.

Définition 4.4.1. *Le facteur de convergence d'une suite (z_k) qui converge vers z est le nombre défini par*

$$\lim_{k \rightarrow \infty} \frac{\|z_{k+1} - z\|}{\|z_k - z\|}.$$

Remarque 4.4.2. *On voit que plus le facteur de convergence est petit plus la convergence de la suite (z_k) vers z est asymptotiquement rapide.*

Nous donnons ci-après un théorème de convergence de la méthode de la puissance valide dans le cas où la matrice est diagonalisable et où la valeur propre dominante est réelle et unique (ce qui n'exclut pas le fait qu'elle puisse être multiple algébriquement)

Voir le livre et ses compléments en ligne pour les autres cas.

Théorème 4.4.1. *Soit une matrice A carrée d'ordre n , diagonalisable, dont la valeur propre λ_1 de plus grand module est réelle et unique (éventuellement multiple).*

Si le vecteur q_0 n'est pas orthogonal au sous-espace propre à gauche associé à λ_1 , alors la suite définie par

* $q_0 \in \mathbb{C}^n$ tel que $\|q_0\| = 1$,

* pour $k = 1, 2, \dots$, $x_k = Aq_{k-1}$, $q_k = \frac{x_k}{\|x_k\|}$ c'est la normalisation,

vérifie

i) $\lim_{k \rightarrow \infty} \frac{|\lambda_1|^k}{\lambda_1^k} q_k$ est un vecteur propre à droite de norme unité associé à λ_1 ,

ii) $\lim_{k \rightarrow \infty} \|Aq_k\| = |\lambda_1|$,

iii) $\lim_{k \rightarrow \infty} \frac{x_{k+1}(j)}{q_k(j)} = \lambda_1$, pour $1 \leq j \leq n$ si $q_k(j) \neq 0$.

Le facteur de convergence de chacune des suites est $|\frac{\lambda_{p+1}}{\lambda_1}|$ où p est la multiplicité de la valeur propre λ_1 .

N.B. : Le vecteur propre usuel est le vecteur propre à droite. Le vecteur propre est toujours défini à un coefficient multiplicatif près, et on ne peut parler de convergence vers un vecteur propre qu'à un facteur multiplicatif près.

Preuve : Voir livre. \square

Remarque 4.4.3. Ce théorème est donné pour une matrice carrée quelconque - sinon le théorème serait trop restrictif - ; ce qui nécessite, pour la preuve mathématique, l'introduction dans l'hypothèse portant sur q_0 , de la notion de sous-espace propre à gauche. Si la matrice est, par exemple, symétrique (dans le cas réel), un vecteur propre à gauche est un vecteur y satisfaisant à l'équation $y^T A = \lambda y^T$ (voir livre pour des compléments) et c'est aussi un vecteur propre à droite. Mais toutes les matrices réelles ne sont pas symétriques ! Évidemment, en pratique, un vecteur initial q_0 quelconque vérifiant $\|q_0\| = 1$ vérifie presque toujours d'être non orthogonal au vecteur propre à gauche v_1 associée à λ_1 ; i.e. que l'on a presque toujours, en pratique, $(q_0, v_1) \neq 0$ où $(.,.)$ est le produit scalaire de \mathbb{C}^n .

4.5 Méthode de la puissance inverse

C'est la méthode de la puissance appliquée, soit à la matrice A^{-1} (on suppose évidemment son existence) qui permet d'obtenir la valeur propre de plus petit module de A et un vecteur propre associé, soit à la matrice $(A - \mu I)^{-1}$ qui permet d'obtenir la valeur propre la plus proche de μ et un vecteur propre associé.. On remarque que si on utilisait la méthode de la puissance à la matrice $(A - \mu I)$, on convergerait vers la valeur propre $\lambda_i - \mu$ telle que $|\lambda_i - \mu| = \max_j |\lambda_j - \mu|$, cette valeur propre λ_i n'est évidemment pas la plus proche de μ !

4.5.1 Recherche de la valeur propre de plus petit module

Soit λ_i une valeur propre de A , alors $\frac{1}{\lambda_i}$ est valeur propre de A^{-1} . Comme la valeur propre de plus grand module de A^{-1} est $\max_i |\frac{1}{\lambda_i}| = \frac{1}{\min_i |\lambda_i|}$, la valeur propre de plus petit module de A est bien égale à l'inverse de la valeur propre de plus grand module de A^{-1} .

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

En reprenant l'algorithme de la puissance, pour calculer la valeur propre de plus grand module de A^{-1} , il faut faire l'itération $x_k = A^{-1}q_{k-1}$; il semble donc qu'il est nécessaire de connaître la matrice A^{-1} . Mais dans la pratique on ne calcule pas A^{-1} (cela coûte trop cher !). En fait, on résout les systèmes $Ax_k = q_{k-1}$ en factorisant la matrice A *une fois pour toute* sous la forme du produit de 2 matrices triangulaires $A = LU$ au départ de l'algorithme - voir le cours 6 pour cette factorisation. L'intérêt est que les résolutions des systèmes triangulaires sont très rapides, et qu'il suffit de changer le second membre à chaque itération pour calculer x_k . Dès lors la recherche de la valeur propre de plus petit module devient rapide elle aussi.

4.5.2 Recherche de la valeur propre la plus proche d'un nombre donné

Soit μ un nombre donné quelconque. On applique l'algorithme précédent à la matrice $(A - \mu I)^{-1}$. 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 . En prenant donc l'inverse du nombre calculé ainsi, on obtient $\min_i |\lambda_i - \mu|$ ce qui donne la valeur propre la plus proche de μ - voir livre pour des informations complémentaires.

4.6 Méthode de bisection ou méthode de Givens-Householder

Elle opère dans le cas où la matrice B est symétrique, rappelons que B est alors diagonalisable. La méthode de bisection que nous allons considérer a pour objectif la recherche soit d'une valeur propre de rang m donné (le rang étant défini par l'ordre dans lequel elles sont rangées de façon décroissante (ou croissante) par module), soit de tout le spectre, ceci pour une *matrice symétrique* d'ordre n .

Notons enfin que la méthode de bisection permet d'être sûr, sans en oublier, de calculer toutes les valeurs propres appartenant à un intervalle donné.

4.6.1 Algorithme général

Elle se met en œuvre en deux étapes successives.

Première étape : Triangulation de la matrice. Étant donnée une matrice B symétrique, on détermine par la méthode de Householder (voir cours 7) une matrice orthogonale P telle que la matrice $P^T B P = A$ soit tridiagonale. La matrice A ainsi obtenue sera aussi nécessairement symétrique.

On est ainsi ramené au problème du calcul des valeurs propres d'une matrice symétrique tridiagonale A , semblable à B , donc ayant même spectre que la matrice B .

Nous allons maintenant décrire, dans une deuxième étape, une méthode (dite de bisection) qui calcule le spectre d'une telle matrice A .

Deuxième étape : C'est la Méthode de bisection proprement dite. Soit donc la matrice symétrique et tridiagonale obtenue après la première étape

$$A = \begin{pmatrix} c_1 & b_1 & & 0 \\ b_1 & c_2 & b_2 & \\ & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & b_{n-1} \\ 0 & & & b_{n-1} & c_n \end{pmatrix} \quad (4.6.1)$$

Pour simplifier on suppose tous les b_i non nuls. Avec les coefficients de la matrice A définissons la suite des polynômes $p_i(\lambda)$, pour $i = 0, 1, 2, \dots, n$, par

$$\begin{aligned} p_0(\lambda) &= 1 \\ p_1(\lambda) &= c_1 - \lambda \\ &\dots \\ p_i(\lambda) &= (c_i - \lambda)p_{i-1}(\lambda) - b_{i-1}^2 p_{i-2}(\lambda), \quad 2 \leq i \leq n, \end{aligned}$$

On vérifie facilement que $p_n(\lambda) = \det(A - \lambda I)$ (polynôme caractéristique de la matrice A). On démontre (voir livre) que le polynôme p_i a i racines distinctes qui séparent strictement les racines du polynôme p_{i+1} , (pour $1 \leq i \leq (n-1)$). Il en résulte que les zéros des polynômes successifs ont l'allure indiquée sur la figure 4.1 suivante

Pour $\lambda \in \mathbb{R}$ donné et un indice i , $1 \leq i \leq n$, posons

$$\operatorname{sgn} p_i(\lambda) = \begin{cases} \operatorname{sgn} p_i(\lambda) & \text{si } p_i(\lambda) \neq 0 \\ \operatorname{sgn} p_{i-1}(\lambda) & \text{si } p_i(\lambda) = 0 \end{cases} \quad (4.6.2)$$

Considérons l'ensemble

$$E_i(\lambda) = \{+, \operatorname{sgn} p_1(\lambda), \operatorname{sgn} p_2(\lambda), \dots, \operatorname{sgn} p_{i-1}(\lambda), \operatorname{sgn} p_i(\lambda)\}$$

défini pour tout i , $1 \leq i \leq n$.

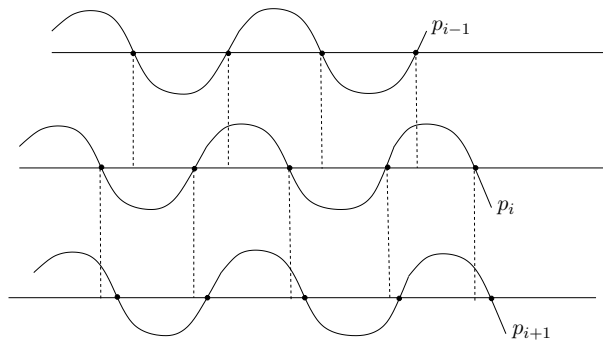


Figure 4.1:

Appelons $\sigma_i(\lambda)$ le nombre de paires consécutives de même signe dans $E_i(\lambda)$. Par exemple si $E_{10}(\lambda) = \{\overbrace{+, +}^1, -, +, \overbrace{-, -}^2, +, \overbrace{-, -, -}^3, +\}$, on a $\sigma_{10}(\lambda) = 4$.

On a alors le résultat (surprenant !) suivant

Théorème 4.6.1. *Pour $\lambda \in \mathbb{R}$ donné, le nombre $\sigma_n(\lambda)$ est égal au nombre de racines de $p_n(\lambda)$ qui sont supérieures ou égales à λ .*

Démonstration : Admise (voir livre). Elle utilise essentiellement la propriété précédente de séparation des racines. \square

4.6.2 Algorithme de bisection

On va calculer toutes les valeurs propres de la matrice tridiagonale et symétrique A grâce au théorème 4.6.1, en sachant, par les remarques précédentes sur les racines, que les valeurs propres sont toutes distinctes. Rangeons-les dans l'ordre décroissant (on sait que la matrice étant symétrique les valeurs propres sont réelles, et l'on peut donc les ordonner)

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{n-1} > \lambda_n.$$

Même plus, cet algorithme va nous permettre de calculer une valeur propre d'ordre i donné dans ce rangement. Cherchons donc à calculer λ_i .

Choix d'un intervalle de confiance On sait que pour toute norme matricielle on a $\rho(A) \leq \|A\|$. Il est dès lors facile de se donner un intervalle de confiance $[a_0, b_0]$ tel que $\lambda_i \in [a_0, b_0]$ pour i fixé mais quelconque tel que $1 \leq i \leq n$. Il suffit de choisir $b_0 = \|A\|_1$ ou $b_0 = \|A\|_\infty$ car ces normes sont faciles à calculer, et de prendre $a_0 = -b_0$. Alors toutes les valeurs propres appartiennent à l'intervalle $[a_0, b_0]$.

Examen d'une suite d'intervalles $[a_k, b_k]$ tels que $\lambda_i \in [a_k, b_k]$ Partons de l'intervalle $[a_0, b_0]$ et soit c_0 le milieu de cet intervalle ($c_0 = \frac{a_0+b_0}{2}$). On calcule alors $\sigma_n(c_0)$. On remarque qu'il faut calculer les valeurs des polynômes $p_1(\lambda), p_2(\lambda), \dots, p_n(\lambda)$ pour $\lambda = c_0$ pour pouvoir déterminer leurs signes en ce point : on peut (on doit !) utiliser la formule de récurrence les définissant pour calculer les $p_i(c_0)$. Alors de deux choses l'une :

- i) si $\sigma_n(c_0) \geq i$, d'après le théorème 4.6.1 cela signifie qu'il existe plus de i racines supérieures ou égales à c_0 . D'après le rangement des valeurs propres cela signifie que $\lambda_i \in [c_0, b_0]$,
- ii) si $\sigma_n(c_0) < i$ alors, a contrario, $\lambda_i \in [a_0, c_0[\subset [a_0, c_0]$.

De sorte que $\lambda_i \in [a_1, b_1]$ avec $[a_1, b_1] = [c_0, b_0]$ dans le cas i) et $[a_1, b_1] = [a_0, c_0]$ dans le cas ii). On recommence le procédé avec le milieu c_1 de $[a_1, b_1]$ et ainsi de suite. On construit ainsi une suite d'intervalles emboîtés $[a_k, b_k]$ par division en deux des intervalles successifs (c'est de là que vient le mot bisection donné à la méthode) tels que $\lambda_i \in [a_k, b_k]$ pour tout k avec

$$b_k - a_k = \frac{b_0 - a_0}{2^k} \xrightarrow[k \rightarrow \infty]{} 0.$$

On arrête le processus dès que $b_k - a_k < \varepsilon$, $\varepsilon > 0$ donné ($\varepsilon \simeq 10^{-4}$ par exemple), et on décide alors que $\lambda_i = c_k = \frac{b_k + a_k}{2}$. \square

Calcul du vecteur propre associé Une fois la valeur propre d'ordre i calculée par la méthode de bisection, on peut déterminer le vecteur propre associé par la méthode de la puissance itérée inverse (voir livre).