

Chapitre 04: Calcul des valeurs propres

Le problème de la détermination des valeurs propres d'une matrice est en général beaucoup plus difficile que celui de la résolution des systèmes linéaires.

De nombreux algorithmes ont été développés pour le calcul des valeurs propres. Ces méthodes sont en fait assez semblables aux méthodes de résolutions des systèmes linéaires. Dans le cadre de ce cours, nous étudions tout d'abord la méthode de la puissance itérée qui permet de calculer la valeur propre de plus grand module d'une matrice réelle et nous étudions également, d'autres méthodes connues et qui sont : la méthode de Jacobi, Givens-Householder et la méthode dite QR. Ces dernières méthodes reposent sur le principe suivant : On construit une suite $(A_k)_{k \in \mathbb{N}}$ de matrices semblables à A et donc ayant les mêmes valeurs propres de A et qui converge vers une matrice plus simple (diagonale, triangulaire) encore semblable à A et donc les valeurs propres cherchées sont les termes diagonaux de la matrice limite.

4.1 Méthode de puissance itérée :

La méthode de la puissance itérée consiste à calculer la suite

$$x^{(k+1)} = \frac{A x^{(k)}}{\|A x^{(k)}\|}, k \in \mathbb{N} \dots (4.1)$$

qui converge vers un vecteur propre associé à λ_1 , la valeur propre de plus grand module de A .

D'abord, on peut démontrer par récurrence que pour tout $k \geq 1$.

$$x^{(k)} = \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|}.$$

Supposons que $A \in M_n(\mathbb{R})$ soit diagonalisable. On note $\lambda_1, \lambda_2, \dots, \lambda_n$ ses valeurs propres dans $\mathbb{C}$ et on suppose qu'elles satisfont

$$|\lambda_1| > |\lambda_i|, \quad i=2, \dots, n.$$

On appellera alors λ_1 la valeur propre dominante de A .

Soit $x^{(0)} \in \mathbb{R}^n$ un vecteur quelconque. $x^{(0)}$ peut être décomposé dans une base de vecteurs propres sous la forme

$$x^{(0)} = e_1 + e_2 + \dots + e_n$$

ou $e_i, i=1, \dots, n$ appartient à l'espace propre associé à la valeur propre λ_i .

Nous avons alors,

$$A x^{(0)} = \lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n$$

Par récurrence,

$$A^k x^{(0)} = \lambda_1^k e_1 + \lambda_2^k e_2 + \dots + \lambda_n^k e_n$$

$$\begin{aligned} \text{D'où} \quad A^k x^{(0)} &= \lambda_1^k \left(e_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^k e_2 + \dots + \left(\frac{\lambda_n}{\lambda_1} \right)^k e_n \right) \\ &= \lambda_1^k z_k \end{aligned}$$

Comme $\left| \frac{\lambda_i}{\lambda_1} \right| < 1, \forall i=2, \dots, n$, alors $\left(\frac{\lambda_i}{\lambda_1} \right)^k \xrightarrow{k \rightarrow \infty} 0$ et par suite $z_k \xrightarrow{k \rightarrow \infty} e_1$.

De plus, on a

$$x^{(k)} = \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|} = \frac{\lambda_1^k z_k}{\|\lambda_1^k z_k\|} = \frac{\lambda_1^k z_k}{|\lambda_1^k| \|z_k\|} = \text{signe}(\lambda_1) \cdot \frac{z_k}{\|z_k\|}$$

Le signe de λ_1^k est positif si k est pair et du même signe que λ_1 si k est impair.

Posons $u_1 = \frac{e_1}{\|e_1\|}$, alors on a $\frac{z_k}{\|z_k\|} \xrightarrow{k \rightarrow \infty} u_1$. Ainsi,

$$x^{(2k)} \xrightarrow{k \rightarrow \infty} u_1 \text{ et } x^{(2k+1)} \xrightarrow{k \rightarrow \infty} \text{signe}(\lambda_1) u_1.$$

Nous avons alors montré que les sous-suites $(x^{(2k)})_{k \in \mathbb{N}}$ et $(x^{(2k+1)})_{k \in \mathbb{N}}$ sont convergentes vers un vecteur propre associé à la valeur propre λ_1 .
Donc, on a le théorème suivant :

Théorème : Soit A une matrice diagonalisable de valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ satisfaisant

$$|\lambda_1| > |\lambda_i|, \quad i = 2, \dots, n$$

Soit la suite de vecteurs $(x^{(k)})_{k \in \mathbb{N}}$ définie par (4.1). Si $x^{(0)}$ est de composante non nulle dans la direction des vecteurs propres associés à λ_1 . Alors, la suite $(x^{(2k)})_{k \in \mathbb{N}}$ et $(x^{(2k+1)})_{k \in \mathbb{N}}$ convergent vers un vecteur propre de norme 1, associé à la valeur propre λ_1 .

$\leadsto$ Maintenant, pour calculer une approximation de λ_1 , sachant qu'on a déjà obtenu une approximation d'un vecteur propre associé à λ_1 , on peut utiliser le quotient de Rayleigh comme suit.

Définition : (Quotient de Rayleigh)

Soit $A \in M_n(\mathbb{C})$ et $x \in \mathbb{C}^n$, alors le quotient de Rayleigh de A relativement au vecteur $x \neq 0$ est défini par :

$$R(A, x) = \frac{x^* A x}{x^* x} = \frac{\langle Ax, x \rangle}{x^* x}$$

Remarque : $R(A, x) = \frac{x^T A x}{x^T x}$, si $A \in M_n(\mathbb{R})$.

Proposition : si x est un vecteur propre associé à la valeur propre λ , alors $R(A, x) = \lambda$.

Preuve: $R(A, x) = \frac{x^* A x}{x^* x} = \frac{x^* dx}{x^* x} = d \frac{x^* x}{x^* x} = d.$

En résumé, une fois un vecteur propre obtenu par la méthode de la puissance itérée, on en déduit immédiatement la valeur propre associée en calculant le quotient de Rayleigh.

Remarque: (méthode de la puissance inverse)

Pour calculer la plus petite valeur propre, on applique la méthode de la puissance itérée à la matrice A^{-1} comme les valeurs propres de A^{-1} sont les inverses de celles de A .

4.2 Méthode de Jacobi

La méthode de Jacobi s'applique aux matrices symétriques réelles

Principe: On construit une suite de matrices $(A_k)_{k \in \mathbb{N}}$ avec $A_0 = A = (a_{ij}) \in M_n(\mathbb{R})$

On passe de A_k à A_{k+1} par une transformation telle que

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

où Ω_k sont des matrices de rotation de type

$$\Omega_k = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cos \theta_k & \cdots & \sin \theta_k & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ 0 & \cdots & -\sin \theta_k & \cdots & \cos \theta_k & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$\downarrow$ $p^{\text{ième}}$ colonne $\downarrow$ $q^{\text{ième}}$ colonne

$\leftarrow$ $p^{\text{ième}}$ ligne
 $\leftarrow$ $q^{\text{ième}}$ ligne

Il est clair que les matrices (A_k) sont semblables à A , et on a :

$$A_{k+1} = \begin{pmatrix} \varrho_k & \varrho_{k-1} & \dots & \varrho_1 \end{pmatrix} A \begin{pmatrix} \varrho_k & \varrho_{k-1} & \dots & \varrho_1 \end{pmatrix}^T$$

On veut que $A_k = (a_{ij}^{(k)})$ tend vers une matrice diagonale D

quand $k \rightarrow +\infty$. Donc, on s'arrange pour annuler deux éléments

symétriques hors diagonaux $a_{pq}^{(k)}$ et $a_{qp}^{(k)}$, ou plus précisément,

on choisit ϱ_k et (p_k, q_k) qui minimisent la somme des carrés des éléments hors diagonaux de la matrice A_{k+1} .

Théorème

(i) Si $A = (a_{ij})$ symétrique, alors la matrice $B = \varrho^T A \varrho = (b_{ij})$

est symétrique et vérifie :

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

(ii) Si $a_{pq} \neq 0$, il existe une unique valeur de $\theta \in]-\frac{\pi}{4}, 0] \cup]0, \frac{\pi}{4}[$ telle que $b_{pq} = 0$, c'est la solution de l'équation

$$\cot(2\theta) = \frac{a_{qq} - a_{pp}}{2a_{pq}}$$

et on a

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

Théorème (convergence de la méthode de Jacobi)

La suite (A_k) de matrices obtenues par la méthode de Jacobi est convergente et on a

$$\lim_{k \rightarrow +\infty} A_k = D$$

où D est une matrice diagonale tel que $\text{Spec}(D) = \text{Spec}(A)$.

Proposition: On suppose que toutes les valeurs propres de A sont distinctes. Alors, la matrice

$$U_k = I_1 \cdot I_2 \cdots I_k$$

converge vers une matrice orthogonale U dont les vecteurs colonnes sont des vecteurs propres de la matrice A .

Remarque:

(1) Il ya plusieurs manières de choisir les couples (p_k, q_k) . La plus classique est de prendre

$$|a_{p_k q_k}^{(k)}| = \max_{i \neq j} |a_{ij}^{(k)}|$$

et θ_k est choisie tel que $b_{p_k q_k} = 0$.

(2) En pratique, on évite de calculer θ_k pour déterminer les lignes et les colonnes de A_{k+1} . On utilise un calcul reposant sur les relations algébriques suivantes:

$$R = \frac{a_{q_k q_k}^{(k)} - a_{p_k p_k}^{(k)}}{2 a_{p_k q_k}^{(k)}} \quad \text{et} \quad t = \tan \theta_k$$

et la valeur de t est calculée comme la racine de plus petit module de

$$t^2 + 2Rt - 1 = 0, \quad t \in]-1, 0[\cup]0, 1[$$

Puis on calcule $c = \frac{1}{\sqrt{1+t^2}}$ et $s = t \cdot c$

Enfin, on calcule les formules suivantes:

$$(k+1) \left\{ \begin{array}{l} a_{p_k i}^{(k+1)} = c a_{p_k i}^{(k)} - s a_{q_k i}^{(k)}, \text{ si } i \neq p_k, q_k \\ a_{q_k i}^{(k+1)} = s a_{p_k i}^{(k)} + c a_{q_k i}^{(k)}, \text{ si } i \neq p_k, q_k \end{array} \right. \quad \left| \quad \begin{array}{l} a_{p_k p_k}^{(k+1)} = a_{p_k p_k}^{(k)} - t a_{p_k q_k}^{(k)} \\ a_{q_k q_k}^{(k+1)} = a_{q_k q_k}^{(k)} + t a_{p_k q_k}^{(k)} \\ a_{p_k q_k}^{(k+1)} = a_{q_k p_k}^{(k+1)} = 0. \end{array} \right.$$

4.3 La méthode QR

4.3.1 Transformation de Householder et forme de Hessenberg

Définition: (Matrice de Householder)

On appelle matrice de Householder, une matrice de la forme

$$H_v = \begin{cases} I_n - \frac{2vv^T}{\|v\|_2^2} = I_n - \frac{vv^T}{\beta}, \text{ avec } \beta = \frac{1}{2}\|v\|_2^2, \text{ si } v \in \mathbb{R}^n \setminus \{0\} \\ I_n, \text{ si } v = 0_{\mathbb{R}^n} \end{cases}$$

Remarque: 1- Il est clair que toute matrice de Householder est symétrique ($H_v^T = H_v$) et orthogonale ($H_v^{-1} = H_v^T$).

2- Si $v \in \mathbb{C}^n$, on remplace v^T par v^* dans la définition.

Exemple: Soit $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^2$, on a $v^T = (1, 2)$

$$\|v\|_2^2 = 1^2 + 2^2 = 1 + 4 = 5$$

$$H_v = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} (1, 2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

Théorème (4.3.1) Soit $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, avec $\sum_{i=2}^n |a_i| \neq 0$. Alors, il existe une matrice de Householder H_v telle que $H_v a$ ait ses $(n-1)$ dernières composantes nulles. Plus précisément, si $\{e_1, e_2, \dots, e_n\}$ est la base canonique de $\mathbb{R}^n$. On a

$$H_v a = -k e_2, \quad H_v a = k e_1$$

avec $k = \text{signe}(a_1) \cdot \|a\|_2$, $v = a + k e_1$, $\beta = a - k e_1$.

Remarque (Application du Th^m dans les techniques de factorisation)

Si a est un vecteur colonne d'une matrice A , on fait apparaître

des zéros dans A en multipliant A par H_v puisque la colonne a sera remplacée par la colonne $H_v a$ qui a des zéro sauf la première composante.

Remarque: En pratique, on calcule $\|a\|_2$, puis $v = a \mp \|a\|_2 e_1$ et on choisit "+" ou "-" pour que $\beta = \frac{1}{2} \|v\|_2^2$ soit le plus grand possible.

Definition: (Matrice de Hessenberg).

Une matrice de Hessenberg (supérieure) $H = (a_{ij})_{1 \leq i, j \leq n}$ est une matrice telle que tous les termes au dessous de sa sous-diagonale sont nuls, c'à d,

$$a_{ij} = 0, \text{ pour } j < i-1.$$

Elle est évidemment triangulaire (supérieure) si tous les coefficients de sa sous-diagonale sont nuls.

Example

Example

diagonale

Rows-diagonal

$$\begin{pmatrix} 1 & 4 & 2 & 3 \\ 3 & 4 & 1 & 7 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

est une matrice de Hessenberg.

$\leadsto$ Le principe de la mise sous forme de Hessenberg d'une matrice $A \in M_n(\mathbb{R})$ est de déterminer des matrices de Householder $H_1, H_2, \dots, H_{n-2}$ telles que $A_{n-1} = H_{n-2} \dots H_1 A H_1 \dots H_{n-2}$ soit une matrice de Hessenberg semblable à A et donc elle a les mêmes valeurs propres que A .

Supposons pour $k \geq 2$:

$$\lambda_{k-1} =$$

$(k-1)^{\text{ième}}$ ligne

$A_{k-1} = \begin{pmatrix} a_{11}^{(k-1)} & a_{12}^{(k-1)} & \dots & a_{1n}^{(k-1)} \\ a_{21}^{(k-1)} & a_{22}^{(k-1)} & \dots & a_{2n}^{(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$

$(k-1)^{\text{ème}} \text{ ligne} \rightarrow$

$(k-1)^{\text{ème}} \text{ colonne}$

le vecteur α

Considérons le vecteur

$$a = \left(a_{k,k-1}^{(k-1)}, a_{k+1,k-1}^{(k-1)}, \dots, a_{n,k-1}^{(k-1)} \right) \in \mathbb{R}^{n-k+1}$$

et soit H_α une matrice de Householder de dimension $(n-k+1)$ associée au vecteur a (Par le Théorème 4.3.1).

Posons alors

$$H_{k-1} = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_\alpha \end{pmatrix}, \quad I_{k-1} \in \Pi_{k-1}(\mathbb{R})$$

la matrice unité

Si on note

$$A_{k-1} = \begin{pmatrix} A_{k-1}^{(1)} & A_{k-1}^{(2)} \\ A_{k-1}^{(3)} & A_{k-1}^{(4)} \end{pmatrix}$$

En multipliant par blocs, on obtient

$$H_{k-1} A_{k-1} = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_\alpha \end{pmatrix} \begin{pmatrix} A_{k-1}^{(1)} & A_{k-1}^{(2)} \\ A_{k-1}^{(3)} & A_{k-1}^{(4)} \end{pmatrix} = \begin{pmatrix} A_{k-1}^{(1)} & A_{k-1}^{(2)} \\ H_\alpha A_{k-1}^{(3)} & H_\alpha A_{k-1}^{(4)} \end{pmatrix}$$

et par suite:

$$A_k = H_{k-1} A_{k-1} H_{k-1} = \begin{pmatrix} A_{k-1}^{(1)} & A_{k-1}^{(2)} \\ H_\alpha A_{k-1}^{(3)} & H_\alpha A_{k-1}^{(4)} \end{pmatrix} \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_\alpha \end{pmatrix} = \begin{pmatrix} A_{k-1}^{(1)} & A_{k-1}^{(2)} H_\alpha \\ H_\alpha A_{k-1}^{(3)} & H_\alpha A_{k-1}^{(4)} H_\alpha \end{pmatrix}$$

Or, $H_\alpha A_{k-1}^{(3)}$ est de la forme

$$H_\alpha A_{k-1}^{(3)} = \begin{pmatrix} 0 & 0 & \dots & \star \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

une seule
composante
non nulle

c'est à dire A_k a la même forme que A_{k-1} . Donc, par récurrence

On obtient que A_{n-1} est une matrice de Hessenberg.

4.3.2 La factorisation QR

Soit $A = (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{R})$. La factorisation QR de la matrice A

consiste à trouver une matrice orthogonale Q et une matrice triangulaire supérieure R telles que $A = QR$, en utilisant les matrices de Householder.

Plus précisément, on va construire une suite de $(n-1)$ matrices de Householder

telles que $H_{n-1} H_{n-2} \dots H_2 H_1 A = R$, où R est une matrice triangulaire

supérieure, et la construction de telles matrices repose sur le

théorème (4.3.1). Supposons que $A_k = H_{k-1} H_{k-2} \dots H_1 A$, $k \geq 1$, est

de la forme

$$A_k = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & \dots & a_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(k)} \end{pmatrix}$$

← k^{ème} ligne

le vecteur a

↑
k^{ème} colonne

Posons $a = (a_{kk}^{(k)}, a_{k+1,k}^{(k)}, \dots, a_{nk}^{(k)}) \in \mathbb{R}^{n-k+1}$

Soit H_0 une matrice de Hessenberg de dimension $n-k+1$ associée au vecteur a selon le théorème (4.3.1). Posons

$$H_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_0 \end{pmatrix}, \quad I_{k-1} \in M_{k-1}(\mathbb{R}) \text{ est la matrice identité}$$

Notons $A_k = \begin{pmatrix} A_k^{(1)} & A_k^{(2)} \\ 0 & A_k^{(3)} \end{pmatrix}$

La décomposition par blocs correspondants à A_k . Par multiplication, on trouve

$$A_{k+1} = H_k A_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_0 \end{pmatrix} \begin{pmatrix} A_k^{(1)} & A_k^{(2)} \\ 0 & A_k^{(3)} \end{pmatrix}$$

$$= \begin{pmatrix} A_k^{(1)} & A_k^{(2)} \\ 0 & H_0 A_k^{(3)} \end{pmatrix}$$

Or, la matrice $H_0 A_k^{(3)}$ est de la forme

$$\begin{pmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \times & \cdots & \times \end{pmatrix}$$

Alors A_{k+1} a la même forme que A_k . Par itération jusqu'à $k=n-1$,

on obtient une matrice triangulaire qu'on note R , et alors

$$R = H_{n-1} H_{n-2} \cdots H_1 A \quad \text{et} \quad Q = (H_{n-1} H_{n-2} \cdots H_1)^{-1} = H_1 H_2 \cdots H_{n-1}$$

qui est la factorisation voulue.

4.3.3 La méthode QR:

La méthode QR est la méthode la plus couramment utilisée pour trouver toutes les valeurs propre d'une matrice quelconque, notamment non symétrique.

Soit $A_1 = A$ une matrice quelconque, on écrit alors sa factorisation

QR, soit $A_1 = Q_1 R_1$. Puis, on forme $A_2 = R_1 Q_1$ et on écrit sa

factorisation QR, on obtient $A_2 = Q_2 R_2$. Ensuite, on forme $A_3 = R_2 Q_2$

puis on écrit sa factorisation QR, on trouve $A_3 = Q_3 R_3$.

A l'étape k , on écrit la factorisation QR de $A_k = R_{k-1} Q_{k-1}$, on obtient $A_k = Q_k R_k$, et ainsi de suite ..., on trouve alors une suite de matrices (A_k) semblables à A .

$$A_1 = A = Q_1 R_1$$

$$A_2 = R_1 Q_1 = Q_1^{-1} A Q_1$$

⋮

$$A_{k+1} = R_k Q_k = Q_k^{-1} A_k Q_k = \dots = (Q_1 Q_2 \dots Q_k)^{-1} A (Q_1 Q_2 \dots Q_k)$$

Remarque: Dans la pratique avant d'appliquer la méthode QR à une matrice A , on commence par mettre la matrice A sous forme d'une matrice de Hessenberg semblable à A . L'intérêt de cette étape est que la suite des matrices A_k de la méthode QR d'une matrice de Hessenberg restent sous la forme de Hessenberg ce qui réduit le temps de calcul.