

Ift 2421

Chapitre 7

Introduction

aux valeurs propres
et
aux vecteurs propres

Définition :

Si A est une matrice de $n \times n$, alors un vecteur non nul x est dit vecteur propre de A si

$$Ax = \lambda x$$

λ est appelé **valeur propre** de A ,
et x **vecteur propre** de A correspondant à λ .

Exemple :

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Si $\lambda > 1$ Dilatation.

Si $0 < \lambda < 1$ Contraction.

Si $\lambda < 0$ Changement de direction.

Calcul analytique des valeurs propres et des vecteurs propres

$$Ax = \lambda x$$

$$\Leftrightarrow$$

$$(\lambda I - A)x = 0$$

Nous avons une solution non nulle ssi

$$\det(A - \lambda I) = 0 \quad \text{Équation caractéristique de A.}$$

Les valeurs satisfaisant cette équation sont
les **valeurs propres** de A.

Exemple :

$$A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$$

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$$

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{bmatrix}$$

L'équation caractéristique de
A est $\lambda^2 - 3\lambda + 2 = 0$

Exemple : trouver les valeurs propres de:

$$A = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$$

et trouver les vecteurs propres de:

$$A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix}$$

$$\det \begin{bmatrix} 10-I & 0 & 0 \\ 1 & -3-I & -7 \\ 0 & 2 & 6-I \end{bmatrix}$$

Les valeurs propres sont: $\lambda_1=10$, $\lambda_2=4$ et $\lambda_3=-1$

$\lambda_1=10$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 10 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \begin{cases} 0x_1 & & = 0 \\ x_1 & -13x_2 & -7x_3 = 0 \\ & 2x_2 & -4x_3 = 0 \end{cases} \quad \begin{cases} x_1 = 1 \\ x_2 = \frac{2}{33} \\ x_3 = \frac{1}{33} \end{cases}$$

$\lambda_2=4$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \begin{cases} 6x_1 & & = 0 \\ x_1 & -7x_2 & -7x_3 = 0 \\ & 2x_2 & +2x_3 = 0 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = -1 \\ x_3 = 1 \end{cases}$$

$\lambda_3=-1$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \begin{cases} 11x_1 & & = 0 \\ x_1 & -2x_2 & -7x_3 = 0 \\ & 2x_2 & +7x_3 = 0 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = -\frac{1}{7} \\ x_3 = -\frac{2}{7} \end{cases}$$

La localisation des valeurs propres : Théorème de Gerschgorin

$$Ax = \lambda x$$

$$\sum_{j=1}^n a_{ij} x_j = \lambda x_i \quad \text{pour } i = 1 \text{ à } n.$$

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

Si nous choisissons x_i tel que $|x_i| = \max_j |x_j|$

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

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

Posons

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

C'est à dire la sommation des valeurs absolues des éléments de
la ligne i sauf a_{ii} .

La localisation des valeurs propres

Les disques de Guerschgorin sont définis par :

$$D_i = \left\{ \lambda \mid |\lambda - a_{ii}| \leq r_i \right\}, \quad i = 1 \text{ à } n.$$

Théorème :

Les valeurs propres de la matrice A sont éléments de l'union des disques D_i .

$$S = \bigcup_{i=1}^n D_i$$

Note :

1. On vérifie toutes les possibilités pour $i = 1$ à n qui pourraient vérifier $|x_i| = \max_j |x_j|$ pour chaque valeur (vecteur) propres.
2. Puisque les valeurs propres peuvent être complexes, nous obtenons des disques dans le plan complexe.
3. Si la matrice A est symétrique, les valeurs propres sont réelles et les disques deviennent de simples intervalles.
4. a_{ii} : centre du disque. r_i : rayon du disque.

Exemple :

$$A = \begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & 1 \\ -2 & 0 & -6 \end{bmatrix}$$

$$D_1 = \left\{ \lambda \mid |\lambda - a_{11}| \leq r_1 \right\}, \quad r_1 = \sum_{j=2}^3 |a_{1j}|.$$

$$|\lambda - 4| \leq 2$$

λ peut être complexe : $\lambda = \lambda_R + i\lambda_I$

$$\left| (\lambda_R - 4) + i(\lambda_I - 0) \right| \leq 2$$

$$\sqrt{(\lambda_R - 4)^2 + (\lambda_I)^2} \leq 2$$

$$(\lambda_R - 4)^2 + (\lambda_I - 0)^2 \leq 4$$

D_1 : centre = (4,0), rayon = 2

D_2 : centre = (1,0), rayon = 2

D_3 : centre = (-6,0), rayon = 2

La localisation des valeurs propres

Corollaires :

1. Les valeurs propres de la matrice A sont aussi éléments de l'union des disques D_i^T construit à partir de sa transposée A^T :

$$S^T = \bigcup_{i=1}^n D_i^T$$

Note : En effet, A et A^T ont les mêmes valeurs propres.

2. Les valeurs propres de la matrice A appartiennent donc à l'intersection de S et S^T :

$$\lambda \in (S \cap S^T)$$

3. Une borne supérieure pour la plus grande valeur propre est donc :

$$|\lambda| \leq \min \left\{ \max_{i=1 \dots n} \left[\sum_{j=1}^n |a_{ij}| \right], \max_{j=1 \dots n} \left[\sum_{i=1}^n |a_{ij}| \right] \right\}$$

Exemple (suite) : pouvons nous améliorer le résultat ?

$$A^T = \begin{bmatrix} 4 & 1 & -2 \\ -1 & 1 & 0 \\ 1 & 1 & -6 \end{bmatrix}$$

D_1^T : centre = (4,0), rayon = 3

D_2^T : centre = (1,0), rayon = 1

D_3^T : centre = (-6,0), rayon = 2

$$\lambda \in (S \cap S^T)$$

$$\lambda_1 = -5.7685$$

$$\lambda_2 = 3.4694$$

$$\lambda_3 = 1.2992$$

Note : corollaire 3 :

$$|\lambda| \leq \min\{\max[6,3,8], \max[7,2,8]\}$$

Démonstration du corollaire 3

$$Ax = \lambda x$$

$$\sum_{j=1}^n a_{ij} x_j = \lambda x_i \quad \text{pour } i = 1 \text{ à } n.$$

$$\lambda = \sum_{j=1}^n a_{ij} \frac{x_j}{x_i}$$

Il y a un i tel que : $|x_i| = \max_j |x_j|$

$$|\lambda| \leq \sum_{j=1}^n |a_{ij}| \left| \frac{x_j}{x_i} \right|$$

$$|\lambda| \leq \sum_{j=1}^n |a_{ij}|$$

Nous ne connaissons pas le i en question, nous choisissons le max des sommations des éléments des lignes.

$$|\lambda| \leq \max \sum_{j=1}^n |a_{ij}|$$

Si nous considérons A^T , la transposée de A :

$$|\lambda| \leq \max \sum_{i=1}^n |a_{ij}|$$

$$|\lambda| \leq \min \left\{ \max_{i=1..n} \left[\sum_{j=1}^n |a_{ij}| \right], \max_{j=1..n} \left[\sum_{i=1}^n |a_{ij}| \right] \right\}$$

Méthode des puissances

Supposons que nous cherchions la plus grande valeur propre de la matrice A ($n \times n$) et que cette matrice possède n vecteurs **linéairement indépendants** x_1, x_2, \dots, x_n associés respectivement aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ qui sont dans l'ordre.

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| > 0$$

La valeur propre λ_1 est dite **dominante**.

Tout vecteur x de R^n peut donc s'écrire :

$$x = \sum_{i=1}^n c_i x_i$$

puisque les vecteurs propres de A forment une base de R^n .

En multipliant ce vecteur par A , nous obtenons :

$$x^{(1)} = Ax^{(0)} = \sum_{i=1}^n c_i Ax_i = \sum_{i=1}^n c_i \lambda_i x_i$$

Si nous répétons cette opération k fois :

$$\begin{aligned} x^{(k)} &= A^k x^{(0)} = \sum_{i=1}^n c_i (\lambda_i)^k x_i \\ &= \lambda_1^k \left[c_1 x_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right] \end{aligned}$$

Or nous avons $\left| \frac{\lambda_i}{\lambda_1} \right| < 1$ pour $i \neq 1$

Donc le terme $\lambda_1^k c_1 x_1$ devient dominant dès que k est assez grand. Donc pour k assez grand, nous avons :

$$x^{(k)} \cong \lambda_1^k c_1 x_1$$

$$x^{(k+1)} \cong \lambda_1^{k+1} c_1 x_1 = \lambda_1 (\lambda_1^k c_1 x_1) \cong \lambda_1 x^{(k)}$$

$$x^{(k+1)} = Ax^{(k)} \cong \lambda_1 x^{(k)}$$

en choisissant 2 composantes i , nous obtenons :

$$\lambda_1 \cong \frac{x_i^{(k+1)}}{x_i^{(k)}}$$

- Nous obtenons donc la plus grande valeur propre et son vecteur propre en même temps.
- Dépassement de la capacité de l'ordinateur ?

Pour éviter de dépasser les capacités de l'ordinateur, nous avons intérêt à normaliser les vecteur itérés x^k à chaque étape ; nous pouvons par exemple ramener à 1 la plus grande composante (en valeur absolue) de x^k , en appliquant l'algorithme :

$$\begin{array}{ll} x^{(1)} = Ax^{(0)} & y^1 = \frac{1}{m_1} x^{(1)} \\ x^{(2)} = Ay^{(1)} & y^2 = \frac{1}{m_2} x^{(2)} \\ \vdots & \vdots \\ x^{(k)} = Ay^{(k-1)} & y^k = \frac{1}{m_k} x^{(k)} \end{array}$$

où m_k est la composante de x^k de module maximum.

Dans ces conditions, nous avons finalement, pour k assez grand :

$$y^k \cong y^{k+1} = \frac{1}{m_{k+1}} x^{(k+1)} = \frac{1}{m_{k+1}} Ay^k$$

donc

$$m_{k+1} y^k \cong Ay^k$$

Remarque :

$$\lim_{k \rightarrow \infty} m_k = \lambda_1 \qquad \lim_{k \rightarrow \infty} y^k = \vec{v}_1$$

Exemple :

$$A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \quad x^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 0.1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0.1 \\ 0 \end{bmatrix} = \begin{bmatrix} 10 \\ 0.7 \\ 0.2 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 0.07 \\ 0.02 \end{bmatrix}$$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0.07 \\ 0.02 \end{bmatrix} = \begin{bmatrix} 10 \\ 0.65 \\ 0.26 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 0.065 \\ 0.026 \end{bmatrix}$$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0.065 \\ 0.026 \end{bmatrix} = \begin{bmatrix} 10 \\ 0.623 \\ 0.286 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 0.0623 \\ 0.0286 \end{bmatrix}$$

$$\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0.0623 \\ 0.0286 \end{bmatrix} = \begin{bmatrix} 10 \\ 0.6129 \\ 0.2962 \end{bmatrix} = 10 \begin{bmatrix} 1 \\ 0.06129 \\ 0.02962 \end{bmatrix}$$

$$\lambda_1 = 10 \quad \vec{x}_1 = \begin{bmatrix} 1 \\ 2/33 \\ 1/33 \end{bmatrix}$$

Remarques :

1. Nous avons :

$$Ax = \lambda x$$

$$A^{-1} A x = A^{-1} \lambda x = \lambda A^{-1} x$$

$$x = \lambda A^{-1} x$$

$$A^{-1} x = \frac{1}{\lambda} x$$

Pour obtenir la plus petite valeur propre, nous pouvons donc utiliser la méthode des puissances sur A^{-1} .

2. Si $|\lambda_1| \approx |\lambda_2|$ ($\dots |\lambda_{n-1}| \approx |\lambda_n|$), alors la convergence sera très lente car $\frac{|\lambda_2|}{|\lambda_1|} \ll 1$ doit être vrai pour avoir une convergence rapide.

3. Le choix du vecteur initial influence beaucoup la rapidité de la convergence. Si le vecteur initial est près du vecteur propre \vec{x}_1 alors les coefficients c_i , $i \neq 1$, seront petit par rapport à c_1 .

$$x^{(k)} = A^k x^{(0)} = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2 + \dots + c_n \lambda_n^k x_n$$

Exemple :

$$A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} 0.1 & 0 & 0 \\ 0.15 & -1.5 & -1.75 \\ -0.05 & 0.5 & 0.75 \end{bmatrix} \quad x^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$A^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.1 \\ -3.1 \\ 1.2 \end{bmatrix} = -3.1 \begin{bmatrix} -0.0323 \\ 1 \\ -0.3871 \end{bmatrix}$$

$$A^{-1} y^1 = \begin{bmatrix} -0.0032 \\ -0.8274 \\ 0.2113 \end{bmatrix} = -0.8274 \begin{bmatrix} 0.0039 \\ 1 \\ -0.2554 \end{bmatrix}$$

$$A^{-1} y^2 = \begin{bmatrix} 0.004 \\ -1.0525 \\ 0.3083 \end{bmatrix} = -1.0525 \begin{bmatrix} -0.0004 \\ 1 \\ -0.2929 \end{bmatrix}$$

$$A^{-1} y^3 = \begin{bmatrix} 0.0000 \\ -0.9875 \\ 0.2803 \end{bmatrix} = -0.9875 \begin{bmatrix} 0 \\ 1 \\ -0.2839 \end{bmatrix}$$

Valeurs exactes :

$$\lambda_3 = -1 \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 1 \\ -0.2857 \end{bmatrix}$$

Comment obtenir la deuxième valeurs propres ?

Méthode de déflation

a. Calculer de λ_1 et \vec{x}_1 par la méthode des puissances.

b. Construire A^1 :

$$A^1 = A - \lambda_1 \frac{(\vec{x}_1 \vec{x}_1^T)}{\vec{x}_1^T \vec{x}_1}$$

c. Appliquer à nouveau la méthode des puissances sur A^1

$$\lambda_2 \text{ et } \vec{x}_2$$

Remarque :

A^1 possède les mêmes λ_i que A sauf $\lambda_1 = 0$

En effet :

$$\left(A - \lambda_1 \frac{(\vec{x}_1 \vec{x}_1^T)}{\vec{x}_1^T \vec{x}_1} \right) \vec{x}_1 = A\vec{x}_1 - \lambda_1 \frac{(\vec{x}_1 \vec{x}_1^T)}{\vec{x}_1^T \vec{x}_1} \vec{x}_1 = 0\vec{x}_1$$

$$\left(A - \lambda_1 \frac{(\vec{x}_1 \vec{x}_1^T)}{\vec{x}_1^T \vec{x}_1} \right)_{i \neq 1} \vec{x}_i = A\vec{x}_i - \lambda_1 \frac{(\vec{x}_1 \vec{x}_1^T)}{\vec{x}_1^T \vec{x}_1} \vec{x}_i = A\vec{x}_i = \lambda_i \vec{x}_i$$

Exemple :

$$A = \begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix}$$

$$A^1 = \begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix} - \frac{10}{1094} \begin{bmatrix} 1 \\ \frac{2}{33} \\ \frac{1}{33} \end{bmatrix} \begin{bmatrix} 1 & \frac{2}{33} & \frac{1}{33} \end{bmatrix}$$

$$A^1 = \begin{bmatrix} 0.0457 & -0.6033 & -0.3016 \\ 0.3967 & -3.3066 & -7.0183 \\ -0.3016 & 1.9817 & 5.9909 \end{bmatrix} \quad x^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.8593 \\ -9.6581 \\ 7.6709 \end{bmatrix} = -9.6581 \begin{bmatrix} 0.0890 \\ 1 \\ -0.7942 \end{bmatrix}$$

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^1 = \begin{bmatrix} -0.3596 \\ 2.5730 \\ -2.8033 \end{bmatrix} = -2.8033 \begin{bmatrix} 0.1283 \\ -0.9178 \\ 1 \end{bmatrix}$$

\vdots

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^k = \begin{bmatrix} 0.3010 \\ -3.9727 \\ 4.0001 \end{bmatrix} = 4.0001 \begin{bmatrix} 0.0752 \\ -0.9931 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 4.0001 \quad \vec{x}_2 = \begin{bmatrix} 0.0752 \\ -0.9931 \\ 1 \end{bmatrix}$$