Mathématiques pour l'informatique 2

Diagonalisation

Émilie Charlier

Université de Liège

Définition

Une matrice $A \in \mathbb{C}^{m \times m}$ est diagonalisable s'il existe une matrice inversible $S \in \mathbb{C}^{m \times m}$ telle que $S^{-1}AS$ est une matrice diagonale.

On dit qu'une telle matrice S diagonalise A.

Puissances de matrices diagonalisables

Lorsqu'une matrice est diagonalisable, il devient facile d'en calculer les puissances.

En effet, si $S^{-1}AS$ est une matrice diagonale $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, alors $\Delta^n = \mathrm{diag}(\lambda_1^n, \ldots, \lambda_m^n)$ et

$$\begin{aligned} A^n &= (S\Delta S^{-1})^n \\ &= S\Delta S^{-1} S\Delta S^{-1} \cdots S\Delta S^{-1} \\ &= S\Delta^n S^{-1} \\ &= S\,\mathrm{diag}(\lambda_1^n, \ldots, \lambda_m^n) S^{-1}. \end{aligned}$$

Convention

Dans ce chapitre, on fixe une matrice $A \in \mathbb{C}^{m \times m}$.

Valeur propre, vecteur propre et espace propre

Définition

Une valeur propre (vp) de A est un nombre complexe λ pour lequel il existe un vecteur $\boldsymbol{x} \in \mathbb{C}^m$ non nul tel que $A\boldsymbol{x} = \lambda\boldsymbol{x}$.

Un vecteur propre de A relatif à un nombre complexe λ est un vecteur $\boldsymbol{x} \in \mathbb{C}^m$ tel que $A\boldsymbol{x} = \lambda\boldsymbol{x}$.

Le sous-espace propre relatif à un nombre complexe λ est l'ensemble des vecteurs propres de A relatifs à λ :

$$E_\lambda(A) = \{\boldsymbol{x} \in \mathbb{C}^m \colon A\boldsymbol{x} = \lambda\boldsymbol{x}\}.$$

Remarque

λ est vp de A si et seulement si $E_\lambda(A) \neq \{\mathbf{0}\}$.

Multiplicité géométrique

On vérifie aisément que $E_\lambda(A)$ est un sev de $\mathbb{C}^m$, ce qui nous autorise à poser la définition suivante.

Définition

La multiplicité géométrique d'un nombre complexe λ en tant que vp de A est la dimension du sev $E_\lambda(A)$ sur $\mathbb{C}$.

Remarque

- La multiplicité géométrique est toujours inférieure ou égale à m.
- Si λ est vp de A, alors sa multiplicité géométrique est supérieure ou égale à 1.

Calcul de la multiplicité géométrique

Pour déterminer la multiplicité géométrique de $\lambda \in \mathbb{C}$, on cherche une base du sous-espace propre $E_\lambda(A)$ en résolvant l'équation matricielle $A\boldsymbol{x} = \lambda \boldsymbol{x}$.

Le résultat suivant nous donne une méthode de calcul alternative de la multiplicité géométrique.

Proposition

Pour tout $\lambda \in \mathbb{C}$,

$$E_\lambda(A) = \mathrm{Ker}(A - \lambda I_m) \qquad \text{et} \qquad \dim E_\lambda(A) = m - \mathrm{rg}(A - \lambda I_m).$$

Démonstration.

Pour tout $\boldsymbol{x} \in \mathbb{C}^m$, on a $A\boldsymbol{x} = \lambda \boldsymbol{x} \iff (A - \lambda I_m)\boldsymbol{x} = \mathbf{0}$.

On a donc $E_\lambda(A) = \mathrm{Ker}(A - \lambda I_m)$.

La deuxième partie découle du théorème du rang. $\square$

Proposition

Des vecteurs propres non nuls relatifs à des vp distinctes sont linéairement indépendants.

Démonstration

Soient $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p$ des vecteurs propres de A non nuls relatifs à des vp deux à deux distinctes $\lambda_1, \ldots, \lambda_p$.

Supposons que $\mu_1, \ldots, \mu_p \in \mathbb{C}$ sont tels que

$$\sum_{k=1}^{p} \mu_k \boldsymbol{x}_k = \mathbf{0}. \qquad (*)$$

Nous devons montrer que tous les μ_k sont nécessairement nuls.

Pour chaque $j \in \{1, \ldots, p\}$, on définit une matrice A_j par

$$A_j = (A - \lambda_1 I_m) \cdots (\widehat{A - \lambda_j I_m}) \cdots (A - \lambda_p I_m)$$

où $(\widehat{A - \lambda_j I_m})$ signifie qu'on omet le facteur $A - \lambda_j I_m$.

Pour tous $j, k \in \{1, \ldots, p\}$, on a

$$(A - \lambda_j I_m)\boldsymbol{x}_k = A\boldsymbol{x}_k - \lambda_j \boldsymbol{x}_k = (\lambda_k - \lambda_j)\boldsymbol{x}_k$$

et

$$A_j \boldsymbol{x}_k = (\lambda_k - \lambda_1) \cdots (\widehat{\lambda_k - \lambda_j}) \cdots (\lambda_k - \lambda_p)\boldsymbol{x}_k.$$

Comme les vp λ_k sont deux à deux distinctes et que les $\boldsymbol{x}_k$ sont des vecteurs non nuls, on en tire que $A_j \boldsymbol{x}_k = \boldsymbol{0}$ si et seulement si $j \neq k$.

Pour tout $j \in \{1, \ldots, p\}$, en multipliant les deux membres de l'égalité $(*)$ par A_j, on obtient que

$$A_j \sum_{k=1}^{p} \mu_k \boldsymbol{x}_k = \sum_{k=1}^{p} \mu_k A_j \boldsymbol{x}_k = \mu_j A_j \boldsymbol{x}_j = \boldsymbol{0}$$

et par conséquent que $\mu_j = 0$. $\square$

Polynôme caractéristique

Définition

Le polynôme caractéristique de A est le polynôme $\det(A - XI_m)$ de $\mathbb{C}[X]$.

Exemple Le polynôme caractéristique de la matrice

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 1 \\ 0 & i & 5 \end{pmatrix}$$

est

$$\begin{aligned} \det(A - XI_m) &= \begin{vmatrix} 1-X & 0 & 2 \\ 0 & 4-X & 1 \\ 0 & i & 5-X \end{vmatrix} \\ &= (1-X)(4-X)(5-X) - i(1-X) \\ &= -X^3 + 10X^2 - 29X + 20 - i + iX \\ &= -X^3 + 10X^2 - (29-i)X + 20 - i. \end{aligned}$$

Remarque

Le polynôme caractéristique d'une matrice carrée de taille m est toujours de degré m.

Plus précisément, le coefficient de X^m vaut toujours $(-1)^m$.

Ceci peut se voir en utilisant la multilinéarité du déterminant et en remarquant que

$$\det(A - XI_m) = \det\begin{pmatrix} C_1 - X\boldsymbol{e}_1 & \cdots & C_m - X\boldsymbol{e}_m \end{pmatrix}$$

où $C_1, \ldots, C_m$ désignent les colonnes de A et $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ désignent les matrices-colonnes unitaires.

Proposition

Un nombre complexe est vp de A si et seulement s'il est zéro du polynôme caractéristique de A.

Démonstration.

Un nombre complexe λ est vp de A si et seulement si le système linéaire homogène

$$(A - \lambda I_m)\boldsymbol{x} = \boldsymbol{0}$$

possède une solution $\boldsymbol{x}$ non nulle.

Ceci a lieu si et seulement si le système n'est pas de Cramer, c'est-à-dire si et seulement si la matrice $A - \lambda I_m$ n'est pas inversible.

Or une matrice est inversible si et seulement si son déterminant est non nul.

On obtient donc que λ est vp de A si et seulement si $\det(A - \lambda I_m) = 0$, c'est-à-dire si et seulement si λ annule le polynôme caractéristique de A. $\square$

Multiplicité algébrique

Définition

La multiplicité algébrique d'un nombre complexe λ est sa multiplicité en tant que zéro du polynôme caractéristique de A.

Remarque

Un nombre complexe est vp de A si et seulement si sa multiplicité algébrique vaut au moins 1.

Définition

Une vp simple est une vp dont la multiplicité algébrique vaut 1.

Remarque

Une matrice carrée de taille m possède donc exactement m vp répétées selon leurs multiplicités algébriques.

Autrement dit, si $\lambda_1, \dots, \lambda_p$ sont les vp de A et si $\alpha_1, \dots, \alpha_p$ en sont les multiplicités algébriques respectives, alors

$$\alpha_1 + \cdots + \alpha_p = m.$$

Matrices semblables

Proposition

Pour toute matrice inversible $S \in \mathbb{C}^{m\times m}$, les matrices A et $S^{-1}AS$ ont même polynôme caractéristique. En particulier, elles possèdent les mêmes vp avec les mêmes multiplicités algébriques.

Démonstration.

Nous avons successivement

$$\begin{aligned}
\det(S^{-1}AS - XI_m) &= \det(S^{-1}(A - XI_m)S) \\
&= \det(S^{-1})\det(A - XI_m)\det(S) \\
&= \det(S^{-1})\det(S)\det(A - XI_m) \\
&= \det(S^{-1}S)\det(A - XI_m) \\
&= \det(I_m)\det(A - XI_m) \\
&= \det(A - XI_m).
\end{aligned}$$

Ainsi, les matrices A et $S^{-1}AS$ possèdent le même polynôme caractéristique. Le cas particulier est immédiat. $\square$

Proposition

La multiplicité géométrique d'une vp est inférieure à sa multiplicité algébrique.

Démonstration

Soit λ une vp de A. Notons $d = \dim E_\lambda(A)$.

Alors il existe d vecteurs propres $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$ de A formant une base de $E_\lambda(A)$.

Puisque $\dim \mathbb{C}^m = m$, le théorème de complétion des bases nous dit qu'il existe $m - d$ vecteurs $\boldsymbol{x}_{d+1}, \ldots, \boldsymbol{x}_m$ qui complètent cette base pour former une base de $\mathbb{C}^m$.

Soit $S = (\boldsymbol{x}_1 \cdots \boldsymbol{x}_m)$ la matrice dont les colonnes sont les vecteurs $\boldsymbol{x}_i$.

Montrons que pour chaque $i \in \{1, \ldots, d\}$, la i-ième colonne de $S^{-1}AS$ est $\lambda \boldsymbol{e}_i$.

En effet, supposons que $i \in \{1, \ldots, d\}$ et que la i-ième colonne de $S^{-1}AS$ soit $\boldsymbol{c}_i = (\mu_1 \cdots \mu_m)^T$.

Alors
$$\lambda \boldsymbol{x}_i = A\boldsymbol{x}_i = S\boldsymbol{c}_i = \sum_{j=1}^{m} \mu_j \boldsymbol{x}_j.$$

Puisque les vecteurs $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ sont linéairement indépendants, on obtient que $\mu_i = \lambda$ et pour $j \neq i$, $\mu_j = 0$.

Ainsi,
$$S^{-1}AS = \begin{pmatrix} \lambda I_d & * \\ 0 & B \end{pmatrix}$$

où B est une matrice carrée de taille $m - d$ et $*$ représente des entrées quelconques.

On obtient que

$$\det(A - XI_m) = \det(S^{-1}AS - XI_m) = (\lambda - X)^d \det(B - XI_{m-d}).$$

La vp λ a donc une multiplicité algébrique au moins égale à d. $\square$

Déterminant, trace et vp

.

Définition

La trace de A est la somme de ses éléments diagonaux :

$$\mathrm{tr}(A) = \sum_{i=1}^{m} A_{ii}.$$

Proposition

Soient $\lambda_1, \ldots, \lambda_m$ les vp de A répétées selon leurs multiplicités algébriques. On a

$$\det(A) = \prod_{i=1}^{m} \lambda_i \quad \text{et} \quad \mathrm{tr}(A) = \sum_{i=1}^{m} \lambda_i.$$

Démonstration de la formule du déterminant

Par le théorème fondamental de l'algèbre et la remarque faite précédemment, nous savons que

$$\det(A - XI_m) = (-1)^m (X - \lambda_1) \cdots (X - \lambda_m) = (\lambda_1 - X) \cdots (\lambda_m - X).$$

En évaluant les deux membres en 0, on obtient que

$$\det(A) = \prod_{i=1}^{m} \lambda_i.$$

$\square$

Théorème de Cayley-Hamilton et polynôme minimal

Définition

Si
$$P = p_0 + p_1 X + p_2 X^2 + \cdots + p_d X^d,$$
alors
$$P(A) = p_0 I_m + p_1 A + p_2 A^2 + \cdots + p_d A^d.$$

Remarque

Le terme indépendant p_0 est vu comme étant $p_0 X^0$, et est donc remplacé par $p_0 A^0 = p_0 I_m$.

Exemple

Si

$$P = X^2 + 3X - 2 \qquad \text{et} \qquad A = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix},$$

alors

$$\begin{aligned} P(A) &= A^2 + 3A - 2I_2 \\ &= \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}^2 + 3\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 5 & 9 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 3 & 9 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 0 \\ 8 & 16 \end{pmatrix}. \end{aligned}$$

Théorème de Cayley-Hamilton (admis)

Toute matrice carrée complexe annule son polynôme caractéristique. Autrement dit, si $P = \det(A - XI_m)$, alors $P(A) = 0$.

Théorème de Cayley-Hamilton (admis)

Toute matrice carrée complexe annule son polynôme caractéristique. Autrement dit, si $P = \det(A - XI_m)$, alors $P(A) = 0$.

Remarque

Attention à ne pas faire l'erreur d'écrire

$$P(A) = \det(A - AI_m) = \det(0) = 0.$$

En effet, il faut d'abord calculer le polynôme $P = \det(A - XI_m)$ dans $\mathbb{C}[X]$ et ensuite de calculer $P(A)$ dans $\mathbb{C}^{m\times m}$ en substituant l'indéterminée X par la matrice A.

Exemple

Poursuivons l'exemple précédent. Le polynôme caractéristique est

$$\det(A - XI_2) = \begin{vmatrix} 2-X & 0 \\ 1 & 3-X \end{vmatrix} = (2-X)(3-X) = X^2 - 5X + 6.$$

On calcule

$$A^2 = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 5 & 9 \end{pmatrix}$$

et

$$\begin{aligned} A^2 - 5A + 6I_2 &= \begin{pmatrix} 4 & 0 \\ 5 & 9 \end{pmatrix} - 5 \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix} + 6 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4-10+6 & 0 \\ 5-5 & 9-15+6 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Définition

Le polynôme minimal de A est le polynôme de $\mathbb{C}[X] \setminus \{0\}$ de coefficient dominant 1 de plus petit degré qui est annulé par A.

Remarque

Si $P(A) = 0$ alors pour tout $c \in \mathbb{C}_0$, on a $(cP)(A) = cP(A) = 0$.

D'où l'importance de préciser que le polynôme minimal a un coefficient dominant égal à 1 dans la définition précédente.

Lemme

Le polynôme minimal de A divise tout polynôme de $\mathbb{C}[X]$ annulé par A.

Démonstration.

Notons π_A le polynôme minimal de A et soit $P \in \mathbb{C}[X]$ un polynôme annulé par A.

Notre but est de montrer que π_A divise P.

Soient donc Q et R les quotient et reste de la division euclidienne de P par π_A.

On a $P = Q\pi_A + R$.

En évaluant les deux membres de cette égalité en A, on obtient $0 = P(A) = (Q\pi_A + R)(A) = Q(A)\pi_A(A) + R(A) = R(A)$.

Donc R est un polynôme de degré strictement plus petit que celui de π_A et annulé par A.

Par définition de π_A, $R = 0$ et donc π_A divise P. $\square$

Proposition

Le polynôme minimal de A divise le polynôme caractéristique de A.

Démonstration.

C'est une conséquence directe du théorème de Cayley-Hamilton et du lemme précédent. $\square$

Lemme

Soient $P \in \mathbb{C}[X]$ et λ une vp de A.
Pour tout $\boldsymbol{x} \in E_\lambda$, on a $P(A)\boldsymbol{x} = P(\lambda)\boldsymbol{x}$.

Démonstration

D'abord, montrons par récurrence que pour tout $i \in \mathbb{N}$, on a $A^i\boldsymbol{x} = \lambda^i\boldsymbol{x}$.

Pour $i = 0$, on a bien $A^0\boldsymbol{x} = \boldsymbol{x} = \lambda^0\boldsymbol{x}$.
Ensuite, fixons $i \in \mathbb{N}$ et supposons que $A^i\boldsymbol{x} = \lambda^i\boldsymbol{x}$. Alors

$$A^{i+1}\boldsymbol{x} = A \cdot A^i\boldsymbol{x} = A \cdot \lambda^i\boldsymbol{x} = \lambda^i \cdot A\boldsymbol{x} = \lambda^i \cdot \lambda\boldsymbol{x} = \lambda^{i+1}\boldsymbol{x}.$$

Maintenant, si $P = p_0 + p_1X + p_2X^2 \cdots + p_dX^d$, alors

$$\begin{aligned}
P(A)\boldsymbol{x} &= (p_0I_m + p_1A + p_2A^2 \cdots + p_dA^d)\boldsymbol{x} \\
&= p_0I_m\boldsymbol{x} + p_1A\boldsymbol{x} + p_2A^2\boldsymbol{x} \cdots + p_dA^d\boldsymbol{x} \\
&= p_0\boldsymbol{x} + p_1\lambda\boldsymbol{x} + p_2\lambda^2\boldsymbol{x} \cdots + p_d\lambda^d\boldsymbol{x} \\
&= (p_0 + p_1\lambda + p_2\lambda^2 \cdots + p_d\lambda^d)\boldsymbol{x} \\
&= P(\lambda)\boldsymbol{x}.
\end{aligned}$$

$\square$

Exemple

Illustrons le lemme précédent avec la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

et le polynôme $P = X^3 + X^2 + X + 1$.

On calcule

$$\begin{aligned} \det(A - XI_2) &= \begin{vmatrix} 1-X & 3 \\ 3 & 1-X \end{vmatrix} \\ &= \begin{vmatrix} 4-X & 3 \\ 4-X & 1-X \end{vmatrix} \\ &= (4-X)(1-X-3) \\ &= (4-X)(-2-X). \end{aligned}$$

Les vp de A sont donc 4 et -2.

On a

$$A^2 = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} \quad \text{et} \quad A^3 = \begin{pmatrix} 28 & 36 \\ 36 & 28 \end{pmatrix}$$

et donc

$$P(A) = A^3 + A^2 + A + I = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix}.$$

Considérons d'abord la valeur propre 4. On peut vérifier que

$$E_4(A) = \left\{ \begin{pmatrix} a \\ a \end{pmatrix} : a \in \mathbb{C} \right\}.$$

On a

$$P(4) = 4^3 + 4^2 + 4^1 + 4^0 = 64 + 16 + 4 + 1 = 85$$

et pour tout $a \in \mathbb{C}$,

$$P(A) \begin{pmatrix} a \\ a \end{pmatrix} = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix} \begin{pmatrix} a \\ a \end{pmatrix} = \begin{pmatrix} 85a \\ 85a \end{pmatrix} = 85 \begin{pmatrix} a \\ a \end{pmatrix} = P(4) \begin{pmatrix} a \\ a \end{pmatrix}.$$

Considérons maintenant la valeur propre -2. On vérifie que

$$E_{-2}(A) = \left\{ \begin{pmatrix} a \\ -a \end{pmatrix} : a \in \mathbb{C} \right\}.$$

On a

$$P(-2) = (-2)^3 + (-2)^2 + (-2)^1 + (-2)^0 = -8 + 4 - 2 + 1 = -5$$

et pour tout $a \in \mathbb{C}$,

$$P(A) \begin{pmatrix} a \\ -a \end{pmatrix} = \begin{pmatrix} 40 & 45 \\ 45 & 40 \end{pmatrix} \begin{pmatrix} a \\ -a \end{pmatrix} = \begin{pmatrix} -5a \\ 5a \end{pmatrix} = -5 \begin{pmatrix} a \\ -a \end{pmatrix} = P(-2) \begin{pmatrix} a \\ -a \end{pmatrix}.$$

Proposition

Les zéros du polynôme minimal de A sont les vp de A.

Démonstration.

Notons π_A le polynôme minimal de A.

Puisque π_A divise le polynôme caractéristique de A, tout zéro de π_A est vp de A.

Il nous reste à montrer que la réciproque est vraie aussi, c'est-à-dire que toute vp de A est zéro de π_A.

Soit λ une vp de A et soit $\boldsymbol{x} \in E_\lambda \setminus \{\mathbf{0}\}$.

Par le lemme précédent, on a que $\pi_A(\lambda)\boldsymbol{x} = \pi_A(A)\boldsymbol{x} = \mathbf{0}$.

Mais puisque $\boldsymbol{x} \neq \mathbf{0}$, ceci implique que $\pi_A(\lambda) = 0$. $\square$

Exemple : calcul du polynôme minimal

Soit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de A est $(2-X)(1-X)^2$.

D'une part, $(X-2)(X-1)$ divise le polynôme minimal.

D'autre part, le polynôme minimal divise $(X-2)(X-1)^2$.

Puisque

$$(A-2I_3)(A-I_3) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

on conclut que le polynôme minimal est égal à
$(X-2)(X-1) = X^2 - 3X + 2$.

Corollaire

Si toutes les vp de A sont simples, alors les polynômes minimal et caractéristique de A coïncident (au signe près).

Démonstration.

Supposons que A possède m vp simples $\lambda_1, \ldots, \lambda_m$.

Alors le polynôme caractéristique de A est égal à $(\lambda_1 - X) \cdots (\lambda_m - X)$.

Nous savons que le polynôme $(X - \lambda_1) \cdots (X - \lambda_m)$ divise le polynôme minimal de A, qui lui-même divise le polynôme caractéristique de A.

D'où la conclusion. $\square$

Exemple

Soit

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

Le polynôme caractéristique de A est

$$(2-X)(1-X)(3-X) = -X^3 + 6X^2 - 11X + 6.$$

La matrice A possède donc trois vp simples.

Par le corollaire précédent, le polynôme minimal est égal à $X^3 - 6X^2 + 11X - 6$.

Critères de diagonalisation

La preuve du théorème suivant est aussi importante que son énoncé.

En effet, l'énoncé nous donne des critères pour déterminer si une matrice donnée (en supposant qu'on puisse en trouver les vp) est ou n'est pas diagonalisable.

Lorsque c'est le cas, la preuve du théorème nous donne une méthode effective de diagonalisation, c'est-à-dire de construction d'une matrice S diagonalisant A.

Théorème

Les conditions suivantes sont équivalentes.

1. A est diagonalisable.
2. A possède m vecteurs propres linéairement indépendants.
3. La somme des multiplicités géométriques des vp de A vaut m.
4. Les multiplicités algébriques et géométriques de chaque vp de A sont égales.

Lemme
Soit $S = (\boldsymbol{c}_1 \cdots \boldsymbol{c}_m)$ une matrice inversible et soient $\mu_1, \ldots, \mu_m \in \mathbb{C}$.
On a

$$S^{-1}AS = \operatorname{diag}(\mu_1, \ldots, \mu_m) \iff \forall j \in \{1, \ldots, m\},\ A\boldsymbol{c}_j = \mu_j \boldsymbol{c}_j.$$

Démonstration.
On a

$$S^{-1}AS = \operatorname{diag}(\mu_1, \ldots, \mu_m) \iff AS = S \operatorname{diag}(\mu_1, \ldots, \mu_m).$$

Pour tout $j \in \{1, \ldots, m\}$,

- la j-ième colonne de AS est $A\boldsymbol{c}_j$
- la j-ième colonne de $S \operatorname{diag}(\mu_1, \ldots, \mu_m)$ est $S\mu_j\boldsymbol{e}_j = \mu_j S\boldsymbol{e}_j = \mu_j \boldsymbol{c}_j$.

D'où la conclusion. $\square$

Démonstration du théorème

Soient $\lambda_1, \dots, \lambda_p$ les vp de A, soient $\alpha_1, \dots, \alpha_p$ leurs multiplicités algébriques respectives et soient $d_1, \dots, d_p$ leurs multiplicités géométriques respectives.

Le polynôme caractéristique étant de degré m, on a $\alpha_1 + \cdots + \alpha_p = m$.

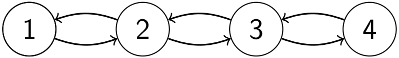

Figure – Étapes de la preuve

$1 \implies 2.$
Supposons que A soit diagonalisable par une matrice S.

Par le lemme, les colonnes de S sont des vecteurs propres de A.

Puisque S est inversible, ses colonnes sont linéairement indépendantes et A possède donc m vecteurs propres linéairement indépendants.

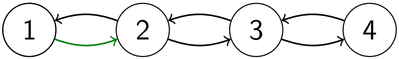

Figure – Étapes de la preuve

$2 \implies 1 \wedge 3.$

Supposons que A possède m vecteurs propres linéairement indépendants.

Notons ces vecteurs propres

$$\boldsymbol{x}_{1,1}, \dots, \boldsymbol{x}_{1,m_1}, \dots, \boldsymbol{x}_{p,1}, \dots, \boldsymbol{x}_{p,m_p}$$

où l'on suppose que pour chaque $i \in \{1, \dots, p\}$ et $\ell \in \{1, \dots, m_i\}$, $\boldsymbol{x}_{i,\ell}$ est un vecteur propre de A relatif à la vp λ_i.

On a donc $m_1 + \cdots + m_p = m$.

Soit

$$S = (\boldsymbol{x}_{1,1} \cdots \boldsymbol{x}_{1,m_1} \cdots \boldsymbol{x}_{p,1} \cdots \boldsymbol{x}_{p,m_p}).$$

Par le lemme,

$$S^{-1}AS = \mathrm{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{m_1 \text{ fois}}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{m_p \text{ fois}})$$

et donc A est diagonalisable. Le point 1 est montré.

De plus, on sait que pour tout $i \in \{1, \ldots, p\}$,

$$m_i \leq d_i \leq \alpha_i.$$

D'où

$$m = m_1 + \cdots + m_p \leq d_1 + \cdots + d_p \leq \alpha_1 + \cdots + \alpha_p = m$$

et donc $d_1 + \cdots + d_p = m$. Ainsi, le point 2 est montré aussi.

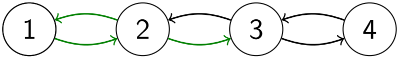

Figure – Étapes de la preuve

$4 \implies 3$ est immédiat.

$3 \implies 4$.
On montre la contraposée. S'il existe $i \in \{1, \dots, p\}$ tel que $d_i \neq \alpha_i$, alors $d_i < \alpha_i$ et $d_1 + \cdots + d_p < \alpha_1 + \cdots + \alpha_p = m$.

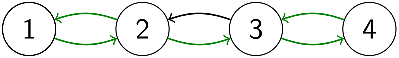

Figure – Étapes de la preuve

$3 \implies 2.$

Supposons que $d_1 + \cdots + d_p = m$.

Pour tout $i \in \{1, \ldots, p\}$, soit $(\boldsymbol{x}_{i,1}, \ldots, \boldsymbol{x}_{i,d_i})$ une base de $E_{\lambda_i}(A)$.

Il suffit de montrer que les m vecteurs $\boldsymbol{x}_{1,1}, \ldots, \boldsymbol{x}_{1,d_1}, \ldots, \boldsymbol{x}_{p,1}, \ldots, \boldsymbol{x}_{p,d_p}$ sont linéairement indépendants.

Soient $k_{1,1}, \ldots, k_{1,d_1}, \ldots, k_{p,1}, \ldots, k_{p,d_p} \in \mathbb{C}$ tels que

$$\underbrace{k_{1,1}\boldsymbol{x}_{1,1} + \cdots + k_{1,d_1}\boldsymbol{x}_{1,d_1}}_{=\boldsymbol{y}_1} + \cdots + \underbrace{k_{p,1}\boldsymbol{x}_{p,1} + \cdots + k_{p,d_p}\boldsymbol{x}_{p,d_p}}_{=\boldsymbol{y}_p} = \mathbf{0}.$$

Chaque vecteur $\boldsymbol{y}_i$ ainsi défini est un vecteur propre relatif à la valeur propre λ_i.

Puisque des vecteurs propres non nuls relatifs à des vp distinctes sont linéairement indépendants, on obtient qu'au moins un des vecteurs $\boldsymbol{y}_i$ est nul.

En itérant cet argument, on obtient que les vecteurs $\boldsymbol{y}_i$ sont en fait tous nuls.

Pour tout $i \in \{1, \ldots, p\}$, on a donc

$$\boldsymbol{y}_i = k_{i,1}\boldsymbol{x}_{i,1} + \cdots + k_{i,d_i}\boldsymbol{x}_{i,d_i} = \mathbf{0}.$$

Or, pour tout $i \in \{1, \ldots, p\}$, les vecteurs $\boldsymbol{x}_{i,1}, \ldots, \boldsymbol{x}_{i,d_i}$ sont linéairement indépendants, ce qui implique que $k_{i,1} = \cdots = k_{i,d_i} = 0$. $\square$

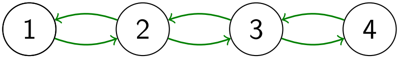

Figure – Étapes de la preuve

Corollaire

Si A possède uniquement des vp simples, alors A est diagonalisable.

Démonstration.

Dans ce cas, toutes les multiplicités algébriques et géométriques sont égales à 1. $\square$

Exemple

Soit

$$A = \begin{pmatrix} \frac{3}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Pour savoir si A est diagonalisable, on teste si les multiplicités géométriques et algébriques de chaque valeur propre coïncident.

Pour cela, il nous faut d'abord déterminer les zéros du polynôme caractéristique de A.

On a

$$\begin{aligned}
\det(A - XI_4) &= \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} - X & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & -X & 0 \\ 0 & 0 & 0 & \frac{1}{3} - X \end{vmatrix} \\
&= \left(\frac{1}{3} - X\right) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 \\ -\frac{1}{4} & -\frac{1}{4} - X & -1 \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix} \\
&\underset{L_2 \leftarrow L_2 + L_3}{=} \left(\frac{1}{3} - X\right) \begin{vmatrix} \frac{3}{4} - X & -\frac{45}{4} & 1 \\ 0 & -1 - X & -1 - X \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{3}-X\right)(-1-X)\begin{vmatrix} \frac{3}{4}-X & -\frac{45}{4} & 1 \\ 0 & 1 & 1 \\ \frac{1}{4} & -\frac{3}{4} & -X \end{vmatrix} \\
&\underset{C_3 \leftarrow C_3 - C_2}{=} \left(\frac{1}{3}-X\right)(-1-X)\begin{vmatrix} \frac{3}{4}-X & -\frac{45}{4} & \frac{49}{4} \\ 0 & 1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & \frac{3}{4}-X \end{vmatrix} \\
&= \left(\frac{1}{3}-X\right)(-1-X)\begin{vmatrix} \frac{3}{4}-X & \frac{49}{4} \\ \frac{1}{4} & \frac{3}{4}-X \end{vmatrix} \\
&= \left(\frac{1}{3}-X\right)(-1-X)\left(\left(\frac{3}{4}-X\right)^2-\frac{49}{16}\right) \\
&= \left(\frac{1}{3}-X\right)(-1-X)\left(\left(\frac{3}{4}-X\right)-\frac{7}{4}\right)\left(\left(\frac{3}{4}-X\right)+\frac{7}{4}\right) \\
&= \left(\frac{1}{3}-X\right)(-1-X)^2\left(\frac{5}{2}-X\right).
\end{aligned}
$$

Les vp de A sont donc $\frac{1}{3}, -1$ et $\frac{5}{2}$ de multiplicités algébriques respectives $1, 2$ et 1.

On en déduit que les multiplicités géométriques de $\frac{1}{3}$ et $\frac{5}{2}$ valent toutes deux 1.

Intéressons-nous à la multiplicité géométrique de -1.

A priori, celle-ci peut valoir 1 ou 2.

Pour tous $a, b, c, d \in \mathbb{C}$, on a

$$\begin{aligned}
& \begin{pmatrix} \frac{3}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = - \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \\
\iff \quad & \begin{pmatrix} \frac{7}{4} & -\frac{45}{4} & 1 & 0 \\ -\frac{1}{4} & \frac{3}{4} & -1 & 0 \\ \frac{1}{4} & -\frac{3}{4} & 1 & 0 \\ 0 & 0 & 0 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
\iff \quad & \left\{ \begin{array}{rcl} \frac{7}{4}a - \frac{45}{4}b + c & = & 0 \\ \frac{1}{4}a - \frac{3}{4}b + c & = & 0 \\ \frac{4}{3}d & = & 0 \end{array} \right. \\
\iff \quad & \left\{ \begin{array}{rcl} \frac{3}{2}a - \frac{21}{2}b & = & 0 \\ \frac{1}{4}a - \frac{3}{4}b + c & = & 0 \\ d & = & 0 \end{array} \right. \\
\iff \quad & \left\{ \begin{array}{rcl} a & = & 7b \\ c & = & -b \\ d & = & 0. \end{array} \right.
\end{aligned}$$

De ceci, on obtient que

$$\begin{aligned} E_{-1}(A) &= \left\{ \begin{pmatrix} 7b \\ b \\ -b \\ 0 \end{pmatrix} : b \in \mathbb{C} \right\} \\ &= \left\{ b \begin{pmatrix} 7 \\ 1 \\ -1 \\ 0 \end{pmatrix} : b \in \mathbb{C} \right\} \\ &= \left\rangle \begin{pmatrix} 7 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\langle . \end{aligned}$$

La multiplicité géométrique de -1 est $\dim E_{-1}(A) = 1$ et diffère de sa multiplicité algébrique, qui est 2.

On obtient que A n'est pas une matrice diagonalisable.

Exemple

Soit

$$A = \begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de A est

$$\begin{aligned} \det(A - XI_6) &= \begin{vmatrix} -X & i & 0 & 0 & 0 & 0 \\ i & -X & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-X & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-X & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-X & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-X \end{vmatrix} \\ &= (X^2+1)(1-X)^4 \\ &= (X-i)(X+i)(1-X)^4. \end{aligned}$$

Les vp de A sont donc $i, -i$ et 1 de multiplicités algébriques respectives $1, 1$ et 4.

On en déduit que les multiplicités géométriques de i et $-i$ valent toutes deux 1.

Intéressons-nous à la multiplicité géométrique de 1.

A priori, celle-ci est comprise entre 1 et 4.

Pour $a, b, c, d, e, f \in \mathbb{C}$, on a

$$\begin{aligned}
& \begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \\
\iff \quad & \begin{pmatrix} -1 & i & 0 & 0 & 0 & 0 \\ i & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
\iff \quad & \left\{ \begin{array}{rcl} -a + ib & = & 0 \\ ia - b & = & 0 \end{array} \right. \\
\iff \quad & \left\{ \begin{array}{rcl} a & = & 0 \\ b & = & 0 \end{array} \right.
\end{aligned}$$

De ceci, on obtient que

$$
\begin{aligned}
E_1(A) &= \left\{ \begin{pmatrix} 0 \\ 0 \\ c \\ d \\ e \\ f \end{pmatrix} : c, d, e, f \in \mathbb{C} \right\} \\
&= \left\{ c \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + e \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + f \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} : c, d, e, f \in \mathbb{C} \right\} \\
&= \left\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\langle .
\end{aligned}
$$

La multiplicité géométrique de 1 est dim $E_1(A) = 4$.

On obtient que A est une matrice diagonalisable.

Pour obtenir une matrice S qui diagonalise A, il suffit de trouver 6 vecteurs propres linéairement indépendants.

Des calculs qui précèdent, on sait déjà que

$$\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}$$

sont des vecteurs propres relatifs à la vp 1 linéairement indépendants.

Il nous faut encore trouver un vecteur propre relatif à la vp i et un autre relatif à la vp $-i$.

Pour $a, b, c, d, e, f \in \mathbb{C}$, on a

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = i \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \iff \left\{ \begin{array}{rcl} ib & = & ia \\ ia & = & ib \\ c & = & ic \\ d & = & id \\ e & = & ie \\ f & = & if \end{array} \right.$$

$$\iff \left\{ \begin{array}{l} a = b \\ c = d = e = f = 0 \end{array} \right.$$

et $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ est un vecteur propre relatif à la vp i.

De même, pour $a, b, c, d, e, f \in \mathbb{C}$, on a

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = -i \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \iff \left\{ \begin{array}{rcl} ib & = & -ia \\ ia & = & -ib \\ c & = & -ic \\ d & = & -id \\ e & = & -ie \\ f & = & -if \end{array} \right.$$

$$\iff \left\{ \begin{array}{l} b = -a \\ c = d = e = f = 0 \end{array} \right.$$

et $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ est un vecteur propre relatif à la vp i.

Les 6 vecteurs propres que nous avons obtenus sont nécessairement linéairement indépendants.

Dès lors, la matrice

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

est telle que

$$S^{-1}AS = \mathrm{diag}(1, 1, 1, 1, i, -i).$$

Quelques résultats supplémentaires sans preuve

Le premier est un autre critère de diagonalisation.

Proposition

Une matrice est diagonalisable si et seulement si son polynôme minimal ne possède que des zéros simples.

Nous mentionnons ensuite quelques propriétés de matrices particulières en lien avec leurs vp.

Rappelons que $A^* = \bar{A}^\intercal$ est la matrice adjointe de A.

Proposition

Supposons que A commute avec sa matrice adjointe, c'est-à-dire $A^*A = AA^*$ (une telle matrice est dite normale).

Alors

- $A^* = A$ (une telle matrice est dite hermitienne) si et seulement si toutes ses vp sont réelles ;
- $A^{-1} = A^*$ (une telle matrice est dite unitaire) si et seulement si ses vp de A sont toutes de module 1.

Le dernier résultat mentionné est utile notamment en théorie des graphes.

Rappelons que A est symétrique si $A^{\intercal} = A$.

Proposition

Toute matrice complexe symétrique est diagonalisable.

Remarque 1

Puisque la matrice d'adjacence d'un graphe non orienté est symétrique, cette proposition nous dit que tout graphe non orienté possède une matrice d'adjacence diagonalisable.

Remarque 2

Dans le cas d'une matrice symétrique, on peut même demander que la matrice S qui diagonalise A soit orthogonale, c'est-à-dire que $S^{-1} = S^{\intercal}$.

C'est évidemment une propriété très intéressante quand on sait combien il peut être fastidieux de calculer l'inverse d'une grande matrice (et la théorie des graphes fournit souvent de grandes matrices).