Ch. I. • VALEURS PROPRES ET VECTEURS PROPRES D'UNE MATRICE.
• DIAGONALISATION D'UNE MATRICE.

1. De l'utilité de la diagonalisation d'une matrice.

2. Valeurs propres, vecteurs propres d'une matrice : généralités.

3. Valeurs propres, vecteurs propres d'une matrice : propriétés élémentaires.
   - 3.1 Cas des matrices (2,2)
   - 3.2 Cas des matrices (3,3)
   - 3.3 Cas général
   - 3.4 Cas des matrices diagonales, triangulaires
   - 3.5 Spectres de matrices semblables

4. Le théorème de CAYLEY-HAMILTON.

5. Diagonalisable ou pas diagonalisable ? là est la question...

6. Un mot des applications.

Jean-Baptiste HIRIART-URRUTY
Mathématiques UFR MIG
Université Paul Sabatier (Toulouse III)
118, route de Narbonne
31062 TOULOUSE Cedex 4 - France

1. De l'utilité de la diagonalisation d'une matrice.

Dans l'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées à coefficients dans $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$), les matrices diagonales sont assurément les plus simples...
Même si $A \in \mathcal{M}_n(\mathbb{K})$ n'est pas diagonale, on peut peut-être s'y ramener. Ainsi, on dit que A est diagonalisable si elle est semblable à une matrice diagonale, c'est-à-dire s'il existe une matrice P inversible telle que $P^{-1}AP$ soit diagonale :

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{bmatrix} = \mathrm{diag}(\lambda_1, \dots, \lambda_n) =: D. \qquad (1)$$

Ceci est le point de vue matriciel (essentiellement le nôtre dans ce Cours).
Le point de vue "application linéaire" est le suivant. Soit E un espace vectoriel de dimension n, muni d'une base $\mathcal{B} = \{e_1, \dots, e_n\}$, f une application linéaire de E de E (on dit que f est un endomorphisme de E) dont la représentation matricielle dans la base $\mathcal{B}$ est A. Dire que f est diagonalisable signifie qu'il existe une (autre) base $\mathcal{B}'$ de E telle que la représentation matricielle de f dans la base $\mathcal{B}'$ soit une matrice diagonale.

Malheureusement, toutes les matrices ne sont pas diagonalisables, et notre objectif dans ce chapitre est de détecter quand une matrice est diagonalisable et comment la diagonaliser.

A quoi sert d'être arrivé à une situation comme (1) ? Comme on l'a déjà dit, D est une matrice on ne peut plus simple, avec

$$\mathrm{tr}\, D = \lambda_1 + \lambda_2 + \dots + \lambda_n, \quad \mathrm{dét}\, A = \lambda_1 \lambda_2 \cdots \lambda_n. \qquad (2)$$

Deux résultats simples (qu'il faut savoir redémontrer) :

– Si B est semblable à A (c'est-à-dire ?), alors $\mathrm{tr}\, B = \mathrm{tr}\, A$. (3)

En effet, si $B = P^{-1}AP$,

$$\mathrm{tr}\, B = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}(PP^{-1}A) \quad [\text{car } \mathrm{tr}(UV) = \mathrm{tr}(VU)]$$
$$= \mathrm{tr}\, A.$$

– Si B est semblable à A, alors $\mathrm{dét}\, B = \mathrm{dét}\, A$. (4)

En effet, si $B = P^{-1}AP$,

$$\mathrm{dét}\, B = \mathrm{dét}(P^{-1})\, \mathrm{dét}\, A\, \mathrm{dét}\, P \quad [\text{car } \mathrm{dét}\, UV = \mathrm{dét}\, U \cdot \mathrm{dét}\, V]$$
$$= \mathrm{dét}\, A \quad \left[\text{car } \mathrm{dét}(P^{-1}) = \frac{1}{\mathrm{dét}\, P}\right].$$

A quoi sert d'arriver à une situation telle que celle décrite en (1) ?

Réponses :

— Des éléments spécifiques associés à A (comme le déterminant) sont inchangés en passant à D et plus faciles à calculer sur D.

— Les calculs relatifs à A s'en trouvent simplifiés, car ces mêmes calculs sur D sont simples. Par exemple :

• Calcul de puissances de A, A^p où p est un entier naturel.

Si $P^{-1}AP = D$, c'est que $A = PDP^{-1}$, et il est aisé de démontrer par récurrence sur p, que

$$A^p = P D^p P^{-1}. \qquad (5)$$

Si on dispose d'une telle matrice P (et de son inverse P^{-1}), sachant que $D^p = \operatorname{diag}(\lambda_1^p, \lambda_2^p, \ldots, \lambda_m^p)$, l'expression de A^p via (5) s'en trouve facilitée.

• A (semblable à $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$) est inversible si et seulement si $\det A = \det D = \lambda_1 \lambda_2 \cdots \lambda_m \neq 0$ (c'est-à-dire aucun des λ_i n'est nul). Dans ce cas,

$$\begin{aligned} A^{-1} &= (P D P^{-1}) \\ &= P D^{-1} P \qquad [\text{car } (UV)^{-1} = V^{-1} U^{-1}]. \end{aligned}$$

En fait, (5) est valable pour tout exposant p qui serait un entier relatif.

• Calcul de la matrice exp A.

Une notion qui interviendra en Automatique est celle d'exponentielle d'une matrice (carrée) A :

$$\begin{aligned} e^A \text{ (ou } \exp A) &:= \sum_{p=0}^{+\infty} \frac{A^p}{p!} = I_m + A + \frac{A^2}{2} + \frac{A^3}{6} + \cdots \\ &= \lim_{N \to +\infty} \sum_{p=0}^{N} \frac{A^p}{p!}. \end{aligned}$$

Grâce à des propriétés comme celle de (5) et la continuité de l'application $A \mapsto \exp A$, on démontre que si $P^{-1}AP = D$, alors ($A = PDP^{-1}$) et

$$\exp A = P (\exp D) P^{-1} \qquad [\text{"l'exponentielle passe à travers"}] \qquad (6)$$

Comme $\exp D = \operatorname{diag}(e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_m})$ est de calcul immédiat, on voit

l'intérêt d'une expression comme (6).
La propriété la plus importante de la fonction "exponentielle de matrices" est : si A et B commutent (i.e., $AB=BA$), alors $(\exp A)(\exp B) = \exp(A+B)$.

2. Valeurs propres, vecteurs propres d'une matrice : généralités.

Soit $A \in \mathcal{M}_n(\mathbb{R})$ (mais A pourra être dans $\mathcal{M}_n(\mathbb{C})$). A représente une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ ($\mathbb{R}^n$ est muni de sa base canonique $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, où $\vec{e}_i = (0,\dots,0,1,0,\dots,0)$, 1 à la i-ème place).
Si $\vec{v} \in \mathbb{R}^n$, $A\vec{v}$ désigne l'image de $\vec{v}$ par A. Comme $\vec{v}$ est de composantes $(x_1, x_2, \dots, x_n)$ dans la base canonique de $\mathbb{R}^n$), on écrit

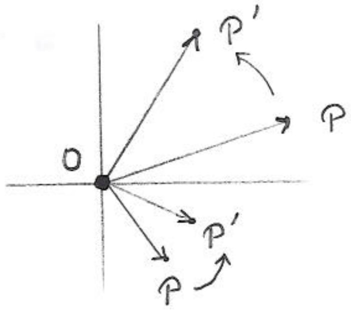
$$A\vec{v} \quad \text{ou} \quad AX = \begin{bmatrix} & & \\ & a_{ij} & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} * \\ * \\ \vdots \\ * \end{bmatrix}$$

composantes du vecteur $A\vec{v}$

pour l'image de $\vec{v}$ par A.

Ex. 1
$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$ représente une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

$$\vec{v} = \overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \overrightarrow{OP'} = A\vec{v} = A.\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y \end{bmatrix}. \qquad (7)$$



Définition 1. On dit que le scalaire λ est une valeur propre de A s'il existe un vecteur $\vec{v}$ non nul tel que
$$A\vec{v} = \lambda\vec{v}. \qquad (7)$$
$\vec{v}$ est alors appelé un vecteur propre de A associé à la valeur propre λ.

On comprend qu'on ait demandé à $\vec{v}$ de ne pas être nul car on a toujours $A\vec{0} = \lambda\vec{0}$ $(=\vec{0})$.
On n'a pas dit que A avait des valeurs propres !... Lorsque $A \in \mathcal{M}_n(\mathbb{R})$, il est normal de chercher des λ réels et des $\vec{v} \in \mathbb{R}^n$ (non nuls) pour lesquels (7) a lieu. Mais cela n'est pas toujours possible ; essayez dans l'exemple du milieu de page de trouver

des $\lambda \in \mathbb{R}$ et des $\vec{v} \in \mathbb{R}^2$ non nuls tels que $A\vec{v} = \lambda\vec{v}$ ... Il n'y en a pas. La raison géométrique en est simple : ce que demande (7) est que l'image de $\vec{v}$ par A et $\vec{v}$ lui-même soient colinéaires ; or dans l'exemple en question, $\overrightarrow{OP'}$ n'est jamais colinéaire avec $\overrightarrow{OP}$.
Donc, si on n'y arrive pas dans $\mathbb{R}$, on passe dans $\mathbb{C}$: $A \in \mathcal{M}_n(\mathbb{R})$ (a fortiori $A \in \mathcal{M}_n(\mathbb{C})$) représente aussi une application linéaire de $\mathbb{C}^n$ dans $\mathbb{C}^n$. On cherche alors $\lambda \in \mathbb{C}$ et $\vec{v} \in \mathbb{C}^n$ non nul tels que $A\vec{v} = \lambda\vec{v}$; on verra que ce qui n'était pas toujours possible dans le royaume des réels le sera toujours dans celui des complexes.

> "Le plus court chemin entre deux vérités dans le domaine réel passe par le domaine complexe" J. HADAMARD (1865 - 1963)

> " Mais Monsieur, avec les réels c'est déjà assez complexe..."
> Une étudiante d'une promotion précédente.

<u>Caractérisation matricielle.</u>

On est à la recherche de scalaires λ et de $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ non nul tels que $AX = \lambda X$, c'est-à-dire encore

$$(A - \lambda I_n)\, X = 0. \tag{8}$$

Or on sait depuis l'étude des systèmes linéaires (cf. Cours de L1) que le système linéaire (8) a des solutions X non nulles si et seulement si $A - \lambda I_n$ est singulière, c'est-à-dire de déterminant nul.
La clé de notre affaire réside donc en

$$\underline{P_A(\lambda) := \text{dét}\,(A - \lambda I_n)},$$

c'est ce qu'on appelle le <u>polynôme caractéristique</u> de A.
D'une manière claire :

- $P_A(\lambda)$ est un polynôme de degré n, à coefficients réels si $A \in \mathcal{M}_n(\mathbb{R})$, à coefficients complexes si $A \in \mathcal{M}_n(\mathbb{C})$. Il a toujours la forme suivante :

$$P_A(\lambda) = \underbrace{(-1)^n \lambda^n} + ////\,\lambda^{n-1} + \cdots ////\,\lambda^k + \cdots + ////\,\lambda + \underbrace{\text{dét}\,A}_{\text{terme constant}} \tag{9}$$

λ^n si n est pair, $n = 2$ par exemple ;
$-\lambda^n$ si n est impair, $n = 3$ par exemple

• D'après ce qui a été dit plus haut,

$$\boxed{(\lambda \text{ valeur propre de } A) \Leftrightarrow (\lambda \text{ racine du polynôme } P_A, \text{ c'est-à-dire } P_A(\lambda) = 0).} \quad (10)$$

On comprend donc qu'il faudra accepter des λ complexes même si $P_A(\lambda)$ est à coefficients réels (lorsque $A \in \mathcal{M}_n(\mathbb{R})$ par exemple).

• Si l'on accepte de travailler dans $\mathbb{C}$, <u>on peut toujours factoriser complètement $P_A(\lambda)$</u> :

$$P_A(\lambda) = -(\lambda-\lambda_1)^{\alpha_1}(\lambda-\lambda_2)^{\alpha_2}\cdots\cdots(\lambda-\lambda_p)^{\alpha_p},$$

où

λ_1 (éventuellement complexe) est racine d'ordre α_1 (de $P_A(\lambda)$)
$\vdots$
λ_p ——————————————— α_p,

et $\alpha_1 + \alpha_2 + \dots + \alpha_p = n$ (au total il y a n racines si on compte les racines λ_i avec leurs multiplicités α_i).

Lorsque $A \in \mathcal{M}_n(\mathbb{R})$, et que donc $P_A(\lambda)$ est à coefficients réels, ses valeurs propres sont

soit réelles (éventuellement aucune d'entre elle),

soit complexes, auquel cas elles vont deux par deux : λ_i de multiplicité α_i, et $\bar{\lambda}_i$ de même multiplicité α_i.

<u>Ex. 2</u> • <u>$n=2$</u>. La matrice $A \in \mathcal{M}_2(\mathbb{R})$ a

soit deux valeurs propres réelles (éventuellement confondues),

soit deux valeurs propres complexes conjuguées.

• <u>$n=3$</u>. La matrice $A \in \mathcal{M}_3(\mathbb{R})$ a

toujours une valeur propre réelle (tiens, pourquoi ?)

deux autres valeurs propres réelles ou deux autres valeurs propres complexes conjuguées.

<u>Définition 2</u> Si λ est une valeur propre de A, on appelle <u>sous-espace propre (de A) associé à λ</u> l'ensemble suivant

$$\underline{E_\lambda := \{ \vec{v} \mid A\vec{v} = \lambda \vec{v} \}} \quad (11)$$

(ensemble des vecteurs propres de A associés à λ, qu'on a complété avec le vecteur nul $\vec{0}$).

E_λ est un sous-espace vectoriel de $\mathbb{R}^n$ (ou $\mathbb{C}^n$ suivant le cas) de dimension ≥ 1 (puisque, par définition, il contient un vecteur non nul et donc la droite vectorielle portée par ce vecteur).
En écriture matricielle, déterminer E_λ consiste à déterminer les $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ solutions du système linéaire (8), à savoir $(A - \lambda I_n) X = 0$.

En pratique, comment procède t-on ?

Etape 1. Calcul du polynôme caractéristique $P_A(\lambda)$ de A.

Etape 2. Calcul des valeurs propres de A, c'est-à-dire des racines du polynôme $P_A(\lambda)$.

Etape 3. Pour chaque valeur propre λ, détermination du sous-espace propre E_λ associé en résolvant le système linéaire $(A - \lambda I_n) X = 0$.

Illustrons cela sur des exemples simples, avec $n = 2$ ou $n = 3$.

Ex. 3 Soit

$$A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

On a :

$$P_A(\lambda) = \text{dét}(A - \lambda I_2) = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = \lambda^2 + 7\lambda + 6.$$

Les valeurs propres de A sont les racines de P_A, c'est-à-dire :

$$\lambda_1 = -1, \quad \lambda_2 = -6.$$

Ces racines étant simples, on parlera de valeurs propres simples.
Déterminons les sous-espaces propres associés. Par définition,

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \in E_{-1} \iff (A + I_2)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

soit

$$\begin{cases} -4x + 2y = 0, \\ 2x - y = 0. \end{cases} \qquad (12)$$

Les deux équations de ce système linéaire sont les mêmes. Donc les solutions de (12) sont $(x, 2x)$, $x \in \mathbb{R}$. Ainsi

$$E_{-1} = \left\{ \vec{v} = \begin{pmatrix} x \\ 2x \end{pmatrix} \mid x \in \mathbb{R} \right\}, \text{ droite vectorielle de } \mathbb{R}^2.$$

On procède de même pour déterminer E_{-6}.

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \in E_{-6} \Longleftrightarrow (A + 6 I_2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

soit

$$\begin{cases} x + 2y = 0, \\ 2x + 4y = 0. \end{cases}$$

Par suite,

$$E_{-6} = \left\{ \vec{v} = \begin{pmatrix} x \\ -\frac{x}{2} \end{pmatrix} \mid x \in \mathbb{R} \right\},$$

E_1 est la droite vectorielle de $\mathbb{R}^2$ engendrée par le vecteur $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

<u>Ex. 4</u> Soit

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

Ici, $P_A(\lambda) = \text{dét}(A - \lambda I_3) = \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & 2 & -\lambda \end{vmatrix} = -\lambda^3 - \lambda^2 + 21\lambda + 45$

$$= -(\lambda - 5)(\lambda + 3)^2.$$

$\lambda_1 = 5$, racine simple de P_A, est valeur propre simple de A;
$\lambda_2 = -3$, racine double de P_A, est dite <u>valeur propre double</u> de A.
Ensuite,

$$E_5 = \left\{ \vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid (A - 5 I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ 2x \\ -x \end{pmatrix} \mid x \in \mathbb{R} \right\},$$ droite vectorielle de $\mathbb{R}^3$ dirigée par le vecteur $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.

Quant à E_{-3}, nous avons :

$$\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-3} \Longleftrightarrow (A + 3 I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ce qui se réduit à une seule équation : $x + 2y - 3z = 0$. Ceci est l'équation d'un plan vectoriel (sous-espace vectoriel de dimension 2). Deux vecteurs linéairement indépendants engendrant ce plan sont, par exemple,

$$\vec{v}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ et } \vec{v}_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Ici, la dimension de E_{λ_2} est égale à la multiplicité de la racine λ_2.

Ex. 5. Soit
$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

On a : $P_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2$, de sorte que $\lambda = 0$ est valeur propre double (la seule). Ensuite,
$$E_0 = \left\{ \vec{v} = \begin{pmatrix} x \\ y \end{pmatrix} \mid A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathrm{Ker}\, A$$
$$= \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}.$$
Dans cet exemple, la dimension de E_λ est strictement inférieure à la multiplicité de la racine λ.

Ex. 6 Soit
$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$$
On a :
$$P_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Il n'y a pas de racine réelle pour P_A, donc pas de valeur propre réelle de A. On raisonne donc dans $\mathbb{C}$ en considérant $A \in \mathcal{M}_2(\mathbb{C})$. Il y a deux racines complexes conjuguées de P_A, qui sont
$$\lambda_1 = i \quad \text{et} \quad \lambda_2 = \overline{\lambda_1} = -i.$$
Ensuite,
$$E_i = \left\{ \vec{v} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2 \mid (A - iI_2) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$
d'où :
$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in E_i \iff z_2 = i z_1,$$
de sorte que
$$E_i \text{ est la droite vectorielle engendrée par } \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

Pour déterminer E_{-i}, inutile de tout recommencer... En effet, puisque A est à coefficients réels,
$$(A - \overline{\lambda_1} I_2) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff (A - \lambda_1 I_2) \begin{pmatrix} \overline{z_1} \\ \overline{z_2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{(d'accord ?)}$$
Ainsi, $E_{\overline{\lambda_1}}$ est la droite vectorielle engendrée par $\overline{\begin{pmatrix} 1 \\ i \end{pmatrix}} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$.

Ex. 7 (Ex. 1 repris). A n'a pas de valeurs propres réelles, seulement deux valeurs propres complexes conjuguées
$$\lambda_1 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2} \quad \text{et} \quad \lambda_2 = \overline{\lambda_1} = \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}.$$

Une remarque générale (illustrée par les 2 exemples précédents).
Si

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathcal{M}_2(\mathbb{R}),$$ matrice antisymétrique donc,
alors les valeurs propres de A sont

$$\lambda_1 = a + ib \text{ et } \lambda_2 = \overline{\lambda}_1 = a - ib. \qquad (13)$$

Définition - Propriété 3 On appelle <u>multiplicité algébrique</u> de la valeur propre λ_i de A la multiplicité de λ_i comme racine du polynôme caractéristique P_A de A (multiplicité notée α_i)
On appelle <u>multiplicité géométrique</u> de la valeur propre λ_i de A la dimension du sous-espace propre E_{λ_i} associé à λ_i (dimension notée d_i).
<u>Propriété</u> :

$$\underline{d_i \leq \alpha_i}. \qquad (14)$$

Dans l'exemple 5 (page 8), α vaut 2 alors que d vaut 1.
Naturellement, si $\alpha_i = 1$ (i.e., λ_i est racine simple de P_A), alors $d_i = 1$ (car $d_i \geq 1$ toujours).

3. Valeurs propres, vecteurs propres d'une matrice : propriétés élémentaires

3.1 Cas des matrices (2,2).

Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ Alors, $P_A(\lambda) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - (a+d)\lambda + (ad - bc)$,

$$\underline{P_A(\lambda) = \lambda^2 - \operatorname{tr} A\, \lambda + \text{dét}\, A}. \qquad (15)$$

Si λ_1 et λ_2 sont les racines de P_A (éventuelle une racine double, éventuellement deux racines complexes),

$$P_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2,$$

de sorte que

$$\underline{\lambda_1 + \lambda_2 = \operatorname{tr} A}, \quad \underline{\lambda_1 \lambda_2 = \text{dét}\, A}. \qquad (16)$$

Avec ce résultat, on peut chaque fois vérifier si le calcul des valeurs propres λ_1 et λ_2 de A est correct ou pas.

3.2 Cas des matrices (3,3).

Soit $A \in \mathcal{M}_3(\mathbb{C})$. On a :

$$\left.\begin{aligned} P_A(\lambda) = \text{dét}(A-\lambda I_3) &= -\lambda^3 + /\!/\!/ \lambda^2 + /\!/\!/ \lambda + \text{dét}\, A \\ &= -(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3), \end{aligned}\right\} \quad (17)$$

où $\lambda_1, \lambda_2, \lambda_3$ désignent les valeurs propres (éventuellement complexes) de A.

A l'aide de (17), on peut être plus précis sur les coefficients de P_A.

Si $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, les calculs du déterminant de $A - \lambda I_3$ conduisent à

$$\underline{P_A(\lambda) = -\lambda^3 + \text{tr}\, A\ \lambda^2 - (A_{11} + A_{22} + A_{33})\lambda + \text{dét}\, A}, \quad (18)$$

où A_{ii} désigne le cofacteur associé à a_{ii} (déterminant de la matrice (2,2) où on a enlevé la i-ème ligne et la i-ème colonne ; exemple $A_{11} = a_{22}a_{33} - a_{23}a_{32}$).

Comme par ailleurs

$$P_A(\lambda) = -\lambda^3 + (\lambda_1+\lambda_2+\lambda_3)\lambda^2 - (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)\lambda + \lambda_1\lambda_2\lambda_3,$$

on a :

$$\left.\begin{aligned} &\underline{\text{tr}\, A = \lambda_1 + \lambda_2 + \lambda_3}, \quad \underline{\text{dét}\, A = \lambda_1\lambda_2\lambda_3} \\ &A_{11} + A_{22} + A_{33} = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \end{aligned}\right\} \quad (19)$$

On ne retiendra que l'essentiel de (19), i.e. sa première ligne.

Si $A \in \mathcal{M}_3(\mathbb{R})$, P_A (à coefficients réels) a automatiquement une racine réelle au moins (tiens, pourquoi ?), plus deux autres racines qui peuvent être réelles ou complexes conjuguées.

3.3 Si $n = 4$, $A \in \mathcal{M}_4(\mathbb{C})$,

$$P_A(\lambda) = \lambda^4 - (\text{tr}\, A)\lambda^3 + \dots\dots + \text{dét}\, A.$$

Si $n = 5$, $A \in \mathcal{M}_5(\mathbb{C})$,

$$P_A(\lambda) = -\lambda^5 + (\text{tr}\, A)\lambda^4 + \dots\dots + \text{dét}\, A$$

De manière générale, si $A \in \mathcal{M}_n(\mathbb{C})$, on retiendra pour

$$\underline{P_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\, \text{tr}\, A\ \lambda^{n-1} + \dots\dots + \text{dét}\, A} \quad (20)$$

que
- le terme de plus haut degré est $(-1)^n$,
- le terme suivant est $(-1)^{n-1}\, \text{tr}\, A$,
- le terme constant est toujours $\text{dét}\, A$.

Si on écrit $P_A(\lambda)$ sous le format suivant

$$P_A(\lambda) = (-1)^n + (-1)^{n-1} \square \, \lambda^{n-1} + \dots (-1)^k \square \, \lambda^k + \dots \square ,$$

les coefficients $\square$ qui apparaissent sont appelés les <u>invariants principaux</u> de A (notion utile en Mécanique notamment). Ainsi,

- pour $n = 2$, les deux invariants principaux de $A \in \mathcal{M}_2(\mathbb{C})$ sont $\operatorname{tr} A$ et $\det A$;
- pour $n = 3$, les trois invariants principaux de $A \in \mathcal{M}_3(\mathbb{C})$ sont
$$\operatorname{tr} A, \quad A_{11} + A_{22} + A_{33} = \text{trace de la matrice des cofacteurs de } A, \text{ et } \det A.$$

3.4 <u>Cas des matrices diagonales, triangulaires.</u>

Si $A = \operatorname{diag}(a_{11}, a_{22}, \dots, a_{nn})$, $A - \lambda I_n = \operatorname{diag}(a_{11} - \lambda, a_{22} - \lambda, \dots, a_{nn} - \lambda)$, de sorte que

$$P_A(\lambda) = \det(A - \lambda I_n) = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda).$$

Les valeurs propres de A, c'est-à-dire les racines de P_A, sont donc $a_{11}, a_{22}, \dots, a_{nn}$.

Plus intéressant est qu'il en est de même pour les matrices triangulaires.

Si

$$A = \begin{bmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}, \quad P_A(\lambda) = \det(A - \lambda I_n) = (a_{11} - \lambda) \dots (a_{nn} - \lambda),$$

il n'y a aucun calcul à faire : les valeurs propres de A sont ses termes diagonaux !

Encore plus intéressant est le cas des matrices diagonales par blocs ou même triangulaires par blocs (des matrices structurées ainsi interviennent souvent dans les calculs de Sciences de l'ingénieur). Soit donc

$$A = \begin{bmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_p \end{bmatrix}.$$

$A - \lambda I_n$ a la même structure, et $P_A(\lambda) = P_{A_1}(\lambda) \, P_{A_2}(\lambda) \dots P_{A_p}(\lambda)$. Les valeurs propres de A sont obtenues en réunissant les valeurs propres de $A_1, A_2, \dots, A_p$.

On désigne par <u>spectre</u> de A l'ensemble des valeurs propres de A. Ainsi,

$$\text{spectre } A = (\text{spectre } A_1) \cup (\text{spectre } A_2) \dots \cup (\text{spectre } A_p).$$

C'est comme si chaque A_i avait une vie autonome et qu'on assemblait ensuite le tout.

3.5 Spectres de matrices semblables.

Théorème 4. Si les matrices A et B sont semblables, alors elles ont les mêmes valeurs propres.

Démonstration. Puisque A et B sont semblables, il existe P inversible telle que $B = P^{-1}AP$. On a donc

$$\begin{aligned} P_B(\lambda) := \text{dét}(B - \lambda I_n) &= \text{dét}(P^{-1}AP - \lambda I_n) \\ &= \text{dét}\, P^{-1}(A - \lambda I_n)P \qquad [\text{astuce } P^{-1}P = I_n\,!] \\ &= \text{dét}(A - \lambda I_n) \\ &= P_A(\lambda). \end{aligned}$$

Les polynômes caractéristiques de A et B sont les mêmes, donc leurs racines aussi : les valeurs propres de A et B sont les mêmes.

Autre démonstration. On sait que λ est valeur propre de B si et seulement si il existe $\vec{v}$ non nul tel que

$$B\vec{v} = \lambda\vec{v}, \text{ soit } (P^{-1}AP)(\vec{v}) = \lambda\vec{v}.$$

Ceci se transforme en

$$A\,(P\vec{v}) = \lambda\,(P\vec{v})$$

Avec le vecteur $\vec{w} := P\vec{v}$, non nul car P est inversible, on a ainsi la traduction du fait que λ est valeur propre de A. On a même démontré une correspondance supplémentaire : $\vec{v}$ est un vecteur propre de B associé à la valeur propre λ si et seulement si $\vec{w} := P\vec{v}$ est un vecteur propre de A associé à la valeur propre λ. □

Remarque. Dans la démonstration au-dessus, on a montré que les coefficients de P_A et P_B sont les mêmes. Donc, si A et B sont semblables, les invariants principaux de A et B sont les mêmes, notamment la trace et le déterminant (ce que nous savions déjà, cf. bas de la page 1).

3.6 Valeurs propres de matrices inversibles, de matrices singulières.

Théorème 5. Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors :

$$(A \text{ inversible}) \iff (0 \text{ n'est pas valeur propre de } A)$$

$$\left(\text{ou bien, } \begin{pmatrix} A \text{ singulière} \\ \text{c'est-à-dire non inversible} \end{pmatrix} \iff (0 \text{ est valeur propre de } A)\right).$$

Si A est inversible, alors

$$\text{spectre } A^{-1} = \left\{ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n} \;\middle|\; \lambda_1, \dots, \lambda_n \in \text{spectre } A \right\}.$$

Démonstration. Par définition de $P_A(\lambda)$, $P_A(0) = \text{dét} A$. Donc

$$(P_A(0) \neq 0) \iff (A \text{ est inversible}),$$

et la première assertion dit que 0 n'est pas valeur propre de A.
Supposons A inversible. Par définition,

$$P_{A^{-1}}(\mu) = \text{dét}(A^{-1} - \mu I_n), \quad P_A(\lambda) = \text{dét}(A - \lambda I_n),$$

d'où :

$$(\lambda \text{ valeur propre de } A) \iff (P_A(\lambda) = 0)$$

$$\iff \frac{(-1)^n}{\lambda^n} \cdot \frac{1}{\text{dét} A} \cdot \text{dét}(A - \lambda I_n) = 0 \quad [\text{car } \lambda \neq 0, \text{ dét} A \neq 0]$$

$$\iff \frac{1}{\text{dét} A} \cdot \text{dét}\left(I_n - \frac{A}{\lambda}\right) = 0 \quad [\text{règles de calcul sur les déterminants}]$$

$$\iff \frac{1}{\text{dét} A} \cdot \text{dét}\left[A\left(A^{-1} - \frac{1}{\lambda} I_n\right)\right] = 0 \quad [\text{astuce } A A^{-1} = I_n]$$

$$\iff \text{dét}\left(A^{-1} - \frac{1}{\lambda} I_n\right) = 0,$$

c'est-à-dire $P_{A^{-1}}\left(\frac{1}{\lambda}\right) = 0$, ou encore : $\frac{1}{\lambda}$ est valeur propre de A^{-1}!

On aurait pu aussi invoquer la définition première de valeur propre : μ est valeur propre de A^{-1} si et seulement si il existe $\vec{w}$ non nul tel que

$$A^{-1} \vec{w} = \mu \vec{w}, \quad \text{soit encore} \quad \frac{1}{\mu} \vec{w} = A^{-1} \vec{w}$$

(car, A^{-1} étant inversible (eh oui !), $\mu \neq 0$ nécessairement).

Mises en garde !

- Même si on connaît les valeurs propres de A et de B, on ne peut en déduire (en général) les valeurs propres de $A + B$ (sauf cas particuliers comme celui où $B = r I_n$) ni celles de AB.
- N'avoir que des valeurs propres nulles (spectre $A = \{0\}$) ne signifie aucunement que la matrice A est nulle. Exemple, $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ n'a que 0 comme valeur propre (laquelle est double).

4. Le théorème de CAYLEY-HAMILTON

Le polynome caractéristique P_A de la matrice A fournit une information supplémentaire sur A, en fait une relation sur certaines puissances de A : en remplaçant λ par A dans $P_A(\lambda)$, on obtient la matrice nulle.

> Théorème de CAYLEY-HAMILTON Soit $A \in \mathcal{M}_n(\mathbb{K})$ de polynôme caractéristique P_A. Alors :
> $$P_A(A) = 0 \quad \text{(la matrice nulle).} \tag{21}$$

Ce résultat fut démontré par W. HAMILTON puis A. CAYLEY dans les années 1850, mais seulement pour des cas particuliers de n. La première démonstration générale est due à G. FROBENIUS en 1878.

Voyons tout de suite ce qui dit (21) pour $n = 2$: si $A \in \mathcal{M}_2(\mathbb{K})$, alors $P_A(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A$, de sorte que

$$A^2 - (\operatorname{tr} A)A + (\det A)I_2 = 0 \quad \left(= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right).$$

Cela permet d'avoir $A^2 = (\operatorname{tr} A)A - (\det A)I_2$, puis

$$\begin{aligned} A^3 = A A^2 &= (\operatorname{tr} A)A^2 - (\det A)A \\ &= \operatorname{tr} A\left[(\operatorname{tr} A)A - (\det A)I_2\right] - (\det A)A \\ &= \left[(\operatorname{tr} A)^2 - (\det A)\right]A - (\operatorname{tr} A)(\det A)I_2, \end{aligned} \tag{22}$$

etc. pour $A^4, \ldots, A^p, \ldots$

Ex. 8 Utilisation de (21) dans le calcul d'une puissance A^p de A.

Pour $p > n$, faisons la division euclidienne de x^p par $P_A(x)$:

$$x^p = P_A(x)\,Q(x) + R(x) \quad [Q \text{ pour quotient, } R \text{ pour reste}],$$

où $R(x)$ est de degré $< n$. En remplaçant x par la matrice A, il s'ensuit la relation matricielle

$$A^p = P_A(A)\,Q(A) + R(A).$$

Par suite, puisque $P_A(A) = 0$, on a $\underline{A^p = R(A)}$.

Si l'on reprend l'exemple de $n = 2$ traité plus haut

$$x^3 = \left[x^2 - (\operatorname{tr} A)x + \det A\right]Q(x) \underbrace{- (\det A)x + (\operatorname{tr} A)^2 x - (\operatorname{tr} A)(\det A)}_{R(x)},$$

d'où $A^3 = R(A)$, ce qui redonne évidemment (22).

Ex. 9 Utilisation de (21) dans le calcul de A^{-1}.

Soit $A \in \mathcal{M}_n(\mathbb{K})$ inversible, c'est-à-dire telle que $\det A \neq 0$. De la relation (21), c'est-à-dire de

$$(-1)^n A^n + /\!/\!/\!/\, A^{n-1} + \cdots + /\!/\!/\!/\, A^k + \cdots + (\det A)\, I_n = 0,$$

on tire

$$\left[(-1)^n A^{n-1} + /\!/\!/\!/\, A^{n-2} + \cdots + /\!/\!/\!/\, A^{k-1} + \cdots\right] A = -(\det A)\, I_n,$$

d'où

$$A^{-1} = -\frac{1}{\det A}\left[(-1)^n A^{n-1} + /\!/\!/\!/\, A^{n-2} + \cdots + /\!/\!/\!/\, A^{k-1} + \cdots\right]. \tag{23}$$

La formule (23) n'est pas à retenir, seulement la <u>méthode</u> pour y arriver.

Ex. 10 Utilisation de (21) dans le calcul de $\exp A$.

Soit

$$A = \begin{bmatrix} 8 & -1 & -5 \\ -2 & 3 & 1 \\ 4 & -1 & 1 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

On demande de calculer $\exp(tA) = \sum_{p=0}^{+\infty} \frac{t^p A^p}{p!}$ pour tout $t \in \mathbb{R}$.

Une attaque frontale du calcul de A^p risque de nous faire transpirer... et sans assurance d'aboutir. Contournons le calcul.

Le polynôme caractéristique de A est $P_A(\lambda) = (2-\lambda)(4-\lambda)^2$. D'après le théorème de CAYLEY-HAMILTON, $P_A(A)$ est la matrice nulle, c'est-à-dire

$$(A - 2I_3)(A - 4I_3)^2 = 0.$$

En posant $B = A - 4I_3$, on a donc $(B + 2I_2)B^2 = 0$, d'où $B^3 = -2B^2$. Un raisonnement par récurrence simple conduit à

$$B^p = (-2)^{p-2} B^2 \text{ pour tout entier } p \geqslant 2.$$

Le calcul de $\exp(tB)$, $t \in \mathbb{R}$, devient du coup facile :

$$\begin{aligned} \exp(tB) &= I_3 + tB + \sum_{p=2}^{+\infty} \frac{t^p(-2)^p}{p!} B^2 \\ &= I_3 + tB + \frac{1}{4}\left(e^{-2t} - 1 + 2t\right)B^2 \quad \left[\text{en se souvenant que } e^x = \sum_{p=0}^{+\infty} \frac{x^p}{p!}\right]. \end{aligned}$$

Comme $tA = 4tI_3 + tB$ et que tI_3 et tB commutent,

$$\begin{aligned} \exp(tA) &= \underbrace{\exp(4tI_3)\exp(tB)} \\ &= e^{4t} \cdot \exp(tB) = e^{4t} I_3 + t e^{4t} B + \frac{1}{4}\left(e^{2t} - e^{4t} + 2te^{4t}\right) B^2. \end{aligned}$$

Il n'y a plus qu'à y remplacer B par $A - 4I_3$ et B^2 par son expression calculée (calcul facile).

Démonstration du théorème de CAYLEY-HAMILTON. La démonstration que nous avons choisi de présenter est intéressante par son aspect "calcul matriciel".

Par définition,

$$P_A(\lambda) = (-1)^n \lambda^n + \cdots\cdots + //// \lambda^k + \cdots\cdots + \det A.$$

Posons $B(\lambda) := [\operatorname{cof}(A - \lambda I_n)]^T$; $B(\lambda)$ est la transposée de la matrice des cofacteurs de $A - \lambda I_n$. Par construction, les coefficients de $B(\lambda)$ sont des expressions polynomiales de λ, de degré $\leq n-1$ (c'est le seul point délicat à observer).

On traduit cela matriciellement en écrivant

$$B(\lambda) = \lambda^{n-1} B_{n-1} + \cdots + \lambda^k B_k + \cdots + B_1 \lambda + B_0, \tag{24}$$

où $B_{n-1}, \ldots, B_0$ sont des matrices ad hoc (on a mis ce qu'il faut pour que (24) lieu).

La propriété fondamentale de la matrice des cofacteurs nous permet d'écrire :

$$(A - \lambda I_n)\, B(\lambda) = \det(A - \lambda I_n)\, I_n, \tag{25}$$

soit, en développant à gauche,

$$\left.\begin{aligned} &\lambda^{n-1} A B_{n-1} + \cdots + \lambda^k A B_k + \cdots + A B_1 \lambda + A B_0 \\ &- \lambda^n B_{n-1} - \lambda^{n-1} B_{n-2} \cdots - \lambda^{k+1} B_k - \lambda^k B_{k-1} \cdots - \lambda^2 B_1 - \lambda B_0 \\ &= (-1)^n \lambda^n I_n \cdots\cdots + //// \lambda^k I_n + \cdots + (\det A)\, I_n \end{aligned}\right\} \tag{26}$$

Ce que nous avons en (26) est une relation matricielle, valable quel que soit $\lambda \in \mathbb{K}$. On identifie donc les termes de même puissance de λ, de sorte que :

$$\begin{aligned} -B_{n-1} &= (-1)^n I_n \\ -B_{n-2} + A B_{n-1} &= //// \, I_n \quad (//// = \text{coefficient de } \lambda^{n-1} \text{ dans } P_A(\lambda)) \\ &\cdots\cdots \\ -B_0 + A B_1 &= //// \, I_n \quad (//// = \text{coefficient de } \lambda \text{ dans } P_A(\lambda)) \\ A B_0 &= (\det A)\, I_n \quad (\det A \text{ est le terme constant dans } P_A(\lambda)). \end{aligned}$$

Multiplions ces équations, respectivement, par $A^n, A^{n-1}, \ldots, A, I_n$. On a alors :

$$\begin{aligned} -A^n B_{n-1} &= (-1)^n A^n \\ -A^{n-1} B_{n-2} + A^n B_{n-1} &= //// \, A^{n-1} \\ &\cdots\cdots \\ -A B_0 + A^2 B_1 &= //// \, A \\ A B_0 &= (\det A)\, I_n. \end{aligned} \tag{27}$$

Additionnons ces équations membre à membre : à gauche de (27) se produit un télescopage $(-A^n B_{n-1} + A^n B_{n-1})$, $(-A^{n-1} B_{n-2} + A^{n-1} B_{n-2})$, ..., $(-AB_0 + AB_0)$, (il ne reste plus rien !), tandis qu'à droite on retrouve $(-1)^n A^n + \ldots A^{n-1} + \ldots + (\det A) I_n$, c'est-à-dire $P_A(A)$.

On a bien démontré que $P_A(A) = 0$. □

5. Diagonalisable ou pas diagonalisable ? là est la question...

Supposons que nous soyons arrivés à la situation suivante : les valeurs propres λ_i de A ont été calculées, les sous-espaces propres correspondants déterminés ainsi que les dimensions de ces sous-espaces propres :

$$\lambda_1, \lambda_2, \ldots\ldots, \lambda_p$$

$$E_{\lambda_1}, E_{\lambda_2}, \ldots\ldots, E_{\lambda_p} \quad (E_{\lambda_k} \subset \mathbb{R}^n \text{ ou } \mathbb{C}^n \text{ suivant les cas})$$

$$d_1 = \dim E_{\lambda_1}, \ldots\ldots, d_p = \dim E_{\lambda_p}$$

On rappelle que $d_k \leq \alpha_k$ (α_k étant l'ordre de multiplicité de la racine λ_k de P_A), et donc

$$d_1 + d_2 + \ldots\ldots + d_p \leq n. \tag{28}$$

- Ensuite, première observation : "la vie à l'intérieur de chaque sous-espace E_{λ_k} est autonome", c'est-à-dire

si $\vec{v} \in E_{\lambda_k}$, son image $A\vec{v}$ par A est encore dans E_{λ_k}

si $\{\vec{v}_1, \ldots, \vec{v}_{d_k}\}$ est une base de E_{λ_k}, la matrice de l'application linéaire (représentée par A) restreinte à E_{λ_k} serait

$$\underbrace{\begin{bmatrix} \lambda_k & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}}_{d_k} \left.\vphantom{\begin{bmatrix} \lambda_k \\ \\ 0 \end{bmatrix}}\right\} d_k \tag{29}$$

- Deuxième observation : des vecteurs propres correspondant à des valeurs propres distinctes sont linéairement indépendants.

Théorème 6. Soit $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ des vecteurs propres correspondant à des valeurs propres <u>distinctes</u> $\lambda_1, \lambda_2, \ldots, \lambda_m$. Alors les vecteurs $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ sont linéairement indépendants.

Démonstration. Raisonnons par l'absurde. Supposons que les $\vec{v}_k$ soient linéairement dépendants, et déroulons un raisonnement qui nous conduira à une contradiction. Soit $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_s$ un ensemble "minimal" de vecteurs pour lesquels le théorème est en défaut, c'est-à-dire (quitte à réindexer les $\vec{v}_i$) :

$$\left.\begin{array}{ll} \vec{v}_1, \vec{v}_2, \dots, \vec{v}_s & \text{sont linéairement dépendants,} \\ \vec{v}_2, \dots, \vec{v}_s & \text{sont linéairement indépendants.} \end{array}\right\} \quad (30)$$

(juste un cran en dessous) (vu ?)

Puisque $\vec{v}_1 \neq 0$, $s \geq 2$. De (30), il vient

$$\vec{v}_1 = a_2 \vec{v}_2 + \dots + a_s \vec{v}_s, \tag{31}$$

où les a_i ne sont pas tous nuls (on est d'accord ?).

En appliquant A aux deux membres de (31), il vient :

$$A\vec{v}_1 = a_2 A \vec{v}_2 + \dots + a_s A \vec{v}_s.$$

Comme $\vec{v}_i$ est vecteur propre de A associé à la valeur propre λ_i, $A\vec{v}_i = \lambda \vec{v}_i$.

D'où :

$$\lambda_1 \vec{v}_1 = \lambda_2 \vec{v}_2 + \dots + \lambda_s \vec{v}_s \tag{32}$$

La comparaison de (31) multipliée par λ_1 et de (32) conduit à :

$$a_2(\lambda_1 - \lambda_2)\vec{v}_2 + a_3(\lambda_1 - \lambda_3)\vec{v}_3 + \dots + a_s(\lambda_1 - \lambda_s)\vec{v}_s = 0.$$

Puisque les vecteurs $\vec{v}_2, \dots, \vec{v}_s$ sont linéairement indépendants, les coefficients de la combinaison linéaire au-dessus sont tous nuls :

$$a_2(\lambda_1 - \lambda_2) = 0, \; a_3(\lambda_1 - \lambda_3) = 0, \dots, a_s(\lambda_1 - \lambda_s) = 0.$$

Les λ_i étant distincts deux à deux, on a $\underline{a_2 = a_3 = \dots = a_s = 0}$. Ceci entre en contradiction avec le fait que "les a_i ne sont pas tous nuls".

Conclusion : l'hypothèse de départ est absurde, c'est son contraire qui est vrai, à savoir que les vecteurs $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ sont linéairement indépendants. □

Le résultat de cette affaire est que les sous-espaces propres $E_{\lambda_1}, \dots, E_{\lambda_p}$ (associés aux valeurs propres distinctes $\lambda_1, \dots, \lambda_p$) sont "en somme directe", c'est-à-dire

$$E_{\lambda_1} \cap E_{\lambda_2} = \{0\}, \; (E_{\lambda_1} + E_{\lambda_2}) \cap E_{\lambda_3} = \{0\}, \dots, (E_{\lambda_1} + \dots + E_{\lambda_{p-1}}) \cap E_{\lambda_p} = \{0\};$$

on écrit

$$E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p} \quad (\subset \mathbb{R}^n \text{ ou } \mathbb{C}^n).$$

La dimension de ce sous-espace $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$ est $d_1 + d_2 + \dots + d_p$, et on construit une base de $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$ en assemblant des bases de $E_{\lambda_1}, \dots, E_{\lambda_p}$:

$$(\text{base de } E_{\lambda_1}) \cup (\text{base de } E_{\lambda_2}) \cup \dots \cup (\text{base de } E_{\lambda_p}) = \text{base de } E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}.$$

Mais il se peut que $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ "ne remplisse pas" $\mathbb{R}^n$ (ou $\mathbb{C}^n$) tout entier, c'est-à-dire que $d_1 + d_2 + \cdots + d_p < n$. Tout le problème de la diagonalisation se trouve là.

Théorème 7. A est diagonalisable si et seulement si
$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p} = \mathbb{R}^n \text{ (ou } \mathbb{C}^n),$$
ce qui est équivalent au fait que
$$d_1 + d_2 + \cdots + d_p = n. \qquad (33)$$

Du coup, il y a au moins un cas où A est manifestement diagonalisable, c'est celui où toutes ses valeurs propres sont simples.

Corollaire 8. Si $P_A(\lambda) = -(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$, où les λ_i sont distincts (on a ainsi n racines simples λ_i du polynôme $P_A(\lambda)$), alors A est diagonalisable.

Beware..., ceci n'est qu'une condition suffisante de diagonalisabilité.

Diagonalisation : marche à suivre.

Etape 1. Calculer les valeurs propres de A et déterminer les sous-espaces propres associés.

Etape 2. Pour chaque sous-espace propre E_{λ_i} (associé à la valeur propre λ_i), déterminer la dimension d_i de E_{λ_i} (c'est-à-dire, en fait, une base de E_{λ_i}).

Etape 3. Test de diagonalisabilité :

(α) Si $d_1 + d_2 + \cdots + d_p < n$, A n'est pas diagonalisable ;

(β) Si $d_1 + d_2 + \cdots + d_p = n$, A est diagonalisable.

Etape 4 (dans le cas (β) seulement).
Soit $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ une base de $\mathbb{R}^n$ (ou $\mathbb{C}^n$) constituée de vecteurs propres de A (par assemblage des bases des sous-espaces propres E_{λ_i}), et soit P la matrice dont les colonnes sont les composantes des vecteurs $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ (ce sera notre matrice de passage). Alors
$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

<u>Ex 10.</u> Soit $A = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$. On a $P_A(\lambda) = \det(A - \lambda I_2) = (\lambda - 4)^2$.

Ainsi, $\lambda_1 = 4$ est la seule valeur propre de A (c'est une valeur propre réelle, double). Sous-espace propre associé :

$$E_4 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid (A - 4I_2)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x - y = 0 \right\}.$$

$E_4 = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\}$ est une droite vectorielle de $\mathbb{R}^2$, dirigée par le vecteur $\vec{v} = (1,1)$ par exemple, de dimension 1 donc. Avec les vecteurs propres de A, on ne peut constituer une base de $\mathbb{R}^2$, donc A n'est pas diagonalisable.

<u>Ex 11.</u> Soit $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$. On a $P_A(\lambda) = (\lambda - 1)(\lambda - 6)$.

A, qui a deux valeurs propres simples ($\lambda_1 = 1$ et $\lambda_2 = 6$), est diagonalisable.

Diagonalisons-là effectivement.

Pour $\lambda_1 = 1$, $E_1 = \left\{ X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid (A - 1I_2)X = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} \mid x \in \mathbb{R} \right\}$. Une base en est $\{\vec{v_1}\}$, avec $\vec{v_1} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ par exemple.

Pour $\lambda_2 = 6$, $E_6 = \left\{ X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid (A - 6I_2)X = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 4x \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\}$. Une base en est $\{\vec{v_2}\}$, avec $\vec{v_2} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ par exemple.

En conséquence, une matrice P qui va servir à diagonaliser A est

$$P = \begin{array}{c} \begin{matrix} \vec{v_1} & \vec{v_2} \end{matrix} \\ \begin{bmatrix} 1 & 4 \\ -1 & 1 \end{bmatrix} \end{array}.$$

Par suite, nous savons que nous aurons

$$P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2).$$

Vous pouvez le vérifier si vous n'êtes pas convaincu(e) !

<u>Ex 12.</u> Soit $A = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix}$. On a $P_A(\lambda) = \lambda^2 + 1$.

Donc, A n'a pas de valeurs propres réelles... Pas question de diagonaliser A, du moins en ne considérant que des matrices à coefficients réels.

Mais A peut être considérée comme élément de $\mathcal{M}_2(\mathbb{C})$. Ainsi, P_A a deux racines simples i et $-i$, c'est-à-dire que A a deux valeurs propres complexes qui sont $\lambda_1 = i$ et $\lambda_2 = \overline{\lambda_1} = -i$. On est donc assuré de l'existence de P inversible (à coefficients complexes) telle que

$$P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2).$$

Pour déterminer P, on procède comme cela a été indiqué : détermination de $E_i \subset \mathbb{C}^2$, de $E_{-i} \subset \mathbb{C}^2$, etc.

<u>Ex 13</u>. Soit $A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \in \mathcal{M}_3(\mathbb{R})$. On trouve $P_A(\lambda) = -(\lambda-1)(\lambda+2)^2$.

A est diagonalisable si et seulement si le sous-espace propre E_{-2} associé à la valeur propre $\lambda_2 = -2$ est de dimension 2 (il pourrait être de dimension 1).
E_{-2} est d'équation $x+y+z=0$ (après calculs). C'est un plan vectoriel dont une base est, par exemple, $\{\vec{v_1}, \vec{v_2}\}$

$$\vec{v_1} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{v_2} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$ (2 vecteurs linéairement indépendants) de E_{-2}.

Pour E_1 on trouve la droite vectorielle d'équation $\begin{cases} x-2y+z=0 \\ y-z=0 \end{cases}$; un vecteur générateur en est $\vec{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Avec $P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix}$, on est assuré de diagonaliser A ; de manière plus précise

$$P^{-1} A P = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \operatorname{diag}(-2,-2,1).$$

6. <u>Un mot des applications.</u>

Les applications des notions de valeurs propres, de vecteurs propres, de la diagonalisation de matrices sont nombreuses et variées dans les Sciences de l'ingénieur :

- Mécanique des vibrations (qui est à l'origine des notions de valeurs et vecteurs propres), oscillations électriques, réseaux électriques, etc.
- Etude des systèmes évolutifs discrets $X_{k+1} = A X_k$, $X_k \in \mathbb{R}^n$, $A \in \mathcal{M}_n(\mathbb{R})$.
- Etude des systèmes différentiels à coefficients constants

$$\begin{cases} x'(t) = a\,x(t) + b\,y(t) \\ y'(t) = c\,x(t) + d\,y(t). \end{cases}$$

- Statistique. Etc

Et quand on ne peut pas diagonaliser A ? Il existe d'autres types de "réductions" de A, moins restrictives que la diagonalisation, donnons-en un exemple.

Pour toute matrice A, il existe P inversible (mais à coefficients complexes) telle que

$$P^{-1} A P = \begin{bmatrix} \bullet & & \\ & \ddots & \\ 0 & & \bullet \end{bmatrix} = T \quad \text{matrice triangulaire.}$$

On peut donc toujours triangulariser (on dit aussi trigonaliser) A.
L'intérêt, entre autres, est que dans T "on voit le spectre de A", puisqu'il s'agit des éléments diagonaux de T ($\bullet$ = valeurs propres de T = valeurs propres de A).

Entendu à un cadre scientifique d'une entreprise à qui on demandait de rapporter sur un volumineux document : "Je n'ai pas tout lu..., j'ai diagonalisé et j'en ai retenu l'essentiel".