Diagonalisation

La diagonalisation est une opération fondamentale des matrices. Nous allons énoncer des conditions qui déterminent exactement quand une matrice est diagonalisable. Nous reprenons pas à pas les notions du chapitre « Valeurs propres, vecteurs propres », mais du point de vue plus théorique des applications linéaires.

Notations.
Dans ce chapitre, E est un $\mathbb{K}$-espace vectoriel. $\mathbb{K}$ est un corps. Dans les exemples de ce chapitre, $\mathbb{K}$ sera $\mathbb{R}$ ou $\mathbb{C}$. Sauf mention contraire, E sera de dimension finie.

1. Valeurs propres, vecteurs propres

Commençons par définir les valeurs et les vecteurs propres d'une application linéaire. Il est important d'avoir d'abord compris le chapitre « Valeurs propres, vecteurs propres » des matrices.

1.1. Définitions

Rappel. $f : E \to E$ est appelé un ***endomorphisme*** si f est une application linéaire de E dans lui-même. Autrement dit, pour tout $v \in E$, $f(v) \in E$ et, en plus, pour tous $u, v \in E$ et tout $\alpha \in \mathbb{K}$:

$$f(u+v) = f(u) + f(v) \quad \text{et} \quad f(\alpha v) = \alpha f(v)$$

Définition 1.
Soit $f : E \to E$ un endomorphisme.
- $\lambda \in \mathbb{K}$ est dite ***valeur propre*** de l'endomorphisme f s'il existe un vecteur non nul $v \in E$ tel que
$$\boxed{f(v) = \lambda v.}$$
- Le vecteur v est alors appelé ***vecteur propre*** de f, associé à la valeur propre λ.
- Le ***spectre*** de f est l'ensemble des valeurs propres de f. Notation : $\mathrm{sp}(f)$ (ou $\mathrm{sp}_{\mathbb{K}}(f)$ si on veut préciser le corps de base).

Si v est un vecteur propre alors, pour tout $\alpha \in \mathbb{K}^*$, αv est aussi un vecteur propre.
Ces définitions sont bien sûr compatibles avec celles pour les matrices. Soit $A \in M_n(\mathbb{K})$. Soit $f : \mathbb{K}^n \to \mathbb{K}^n$ l'application linéaire définie par $f(v) = Av$ (où v est considéré comme un vecteur colonne). Alors les valeurs propres (et les vecteurs propres) de f sont celles de A.

1.2. Exemples

La principale source d'exemples provient des matrices et nous renvoyons encore une fois au chapitre « Valeurs propres, vecteurs propres ».

Exemple 1.
Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$f(x,y,z) = \big(-2x-2y+2z, -3x-y+3z, -x+y+z\big).$$

1. Écriture en terme de matrice. L'application f s'écrit aussi $f(X) = AX$ avec :
$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{et} \qquad A = \begin{pmatrix} -2 & -2 & 2 \\ -3 & -1 & 3 \\ -1 & 1 & 1 \end{pmatrix}$$
2. Le vecteur $v_1 = (1,1,0)$ est vecteur propre.
En effet, $f(1,1,0) = (-4,-4,0)$, autrement dit $f(v_1) = -4v_1$. Ainsi v_1 est un vecteur propre associé à la valeur propre $\lambda_1 = -4$.
Si on préfère faire les calculs avec les matrices, on considère v_1 comme un vecteur colonne et on calcule $Av_1 = -4v_1$.
3. $\lambda_2 = 2$ est valeur propre.
Pour le prouver, il s'agit de trouver un vecteur non nul dans $\mathrm{Ker}(f - \lambda_2 \,\mathrm{id}_{\mathbb{R}^3})$ pour $\lambda_2 = 2$. Pour cela, on calcule $A - \lambda_2 I_3$:
$$A - 2I_3 = \begin{pmatrix} -4 & -2 & 2 \\ -3 & -3 & 3 \\ -1 & 1 & -1 \end{pmatrix}$$
On trouve que $v_2 = (0,1,1)$ est dans le noyau de $A - 2I_3$, c'est-à-dire $(A - 2I_3)v_2$ est le vecteur nul. En d'autres termes, $v_2 \in \mathrm{Ker}(f - \lambda_2 \,\mathrm{id}_{\mathbb{R}^3})$, c'est-à-dire $f(v_2) - 2v_2 = 0$, donc $f(v_2) = 2v_2$.
Bilan : v_2 est un vecteur propre associé à la valeur propre $\lambda_2 = 2$.
4. $\lambda_3 = 0$ est valeur propre.
On peut faire juste comme au-dessus et trouver que $v_3 = (1,0,1)$ vérifie $f(v_3) = (0,0,0)$. Ainsi $f(v_3) = 0 \cdot v_3$. Bilan : v_3 est vecteur propre associé à la valeur propre $\lambda_3 = 0$.
5. On a trouvé 3 valeurs propres, et il ne peut y en avoir plus car la matrice A est de taille 3×3.
Conclusion : $\mathrm{sp}(f) = \{-4, 2, 0\}$.

Exemple 2.
Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ l'application linéaire définie par $(x_1, \ldots, x_{n-1}, x_n) \mapsto (x_1, \ldots, x_{n-1}, 0)$. Géométriquement, f est une projection sur $\mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$. Notons $e_1 = (1,0,0,\ldots)$, $e_2 = (0,1,0,\ldots)$,..., $e_n = (0,\ldots,0,1)$ les n vecteurs de la base canonique de $\mathbb{R}^n$.
Alors

$$f(e_1) = e_1 \quad f(e_2) = e_2 \quad \ldots \quad f(e_{n-1}) = e_{n-1} \quad \text{et} \quad f(e_n) = 0.$$

Ainsi $e_1, \ldots, e_{n-1}$ sont des vecteurs propres associés à la valeur propre 1. Et e_n est un vecteur propre associé à la valeur propre 0. Conclusion : $\mathrm{sp}(f) = \{0, 1\}$.

Voici d'autres exemples plus théoriques.

Exemple 3.

1. Soit $E = \mathbb{R}_n[X]$ l'espace des polynômes de degré $\leqslant n$. Soit $d : E \to E, P(X) \mapsto P'(X)$ l'application de dérivation. Pour des raisons de degré,
$$P' = \lambda P \qquad \Longrightarrow \qquad \lambda = 0 \quad \text{et} \quad P \text{ constant}$$
De plus, tout polynôme constant non nul est un vecteur propre de d, de valeur propre associée 0 ; donc $\operatorname{sp}(d) = \{0\}$.
2. (Cet exemple est en dimension infinie.) Soit $E = \mathscr{C}^\infty(\mathbb{R})$ l'espace des fonctions infiniment dérivables de $\mathbb{R}$ dans $\mathbb{R}$. Soit $d : E \to E$, $\phi \mapsto \phi'$ l'application de dérivation.
Pour tout $\lambda \in \mathbb{R}$, définissons la fonction
$$e_\lambda : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \exp(\lambda x).$$
On a $e_\lambda' = \lambda e_\lambda$, donc chaque fonction e_λ est un vecteur propre de d de valeur propre associée λ. Ici, $\operatorname{sp}(d) = \mathbb{R}$.

1.3. Sous-espaces propres

Cherchons une autre écriture de la relation de colinéarité définissant les vecteurs propres :
$$\begin{aligned} f(v) = \lambda v &\iff f(v) - \lambda v = 0 \\ &\iff (f - \lambda \operatorname{id}_E)(v) = 0 \\ &\iff v \in \operatorname{Ker}(f - \lambda \operatorname{id}_E) \end{aligned}$$
D'où la définition :

Définition 2.
Soit f un endomorphisme de E. Soit $\lambda \in \mathbb{K}$. Le *sous-espace propre* associé à λ est le sous-espace vectoriel E_λ défini par :
$$\boxed{E_\lambda = \operatorname{Ker}(f - \lambda \operatorname{id}_E)}$$

On notera aussi ce sous-espace $E_\lambda(f)$ si on souhaite signaler sa dépendance vis-à-vis de l'endomorphisme f.
Autrement dit :
$$\boxed{E_\lambda = \left\{ v \in E \mid f(v) = \lambda v \right\}}$$

C'est le sous-espace vectoriel de E constitué des vecteurs propres de f associés à la valeur propre λ, auquel on ajoute le vecteur nul. Être valeur propre, c'est donc exactement avoir un sous-espace propre non trivial :
$$\lambda \text{ valeur propre} \iff E_\lambda \neq \{0\}$$

Remarque.
Plaçons-nous dans le cas où E est de dimension finie.
- Si λ est une valeur propre de f, alors le sous-espace propre associé E_λ est de dimension $\geqslant 1$.
- Le sous-espace propre E_λ est stable par f, c'est-à-dire $f(E_\lambda) \subset E_\lambda$. En effet :
$$v \in \operatorname{Ker}(f - \lambda \operatorname{id}_E) \implies f(f(v)) = f(\lambda v) = \lambda f(v) \implies f(v) \in \operatorname{Ker}(f - \lambda \operatorname{id}_E)$$

Théorème 1.
Soit f un endomorphisme de E. Soient $\lambda_1, \ldots, \lambda_r$ des valeurs propres **distinctes** *de f. Alors les sous-espaces propres associés $E_{\lambda_1}, \ldots, E_{\lambda_r}$ sont en somme directe.*

On retrouve un résultat déjà prouvé dans le cas des matrices :

Corollaire 1.
Soient $\lambda_1, \ldots, \lambda_r$ des valeurs propres distinctes de f et, pour $1 \leqslant i \leqslant r$, soit v_i un vecteur propre associé à λ_i. Alors les v_i sont linéairement indépendants.

Cela implique que le nombre de valeurs propres est $\leqslant \dim E$.
Avant de lire les exemples et la preuve de ce théorème, lire si besoin la section suivante sur les sommes directes.

Exemple 4.
Reprenons l'exemple 1 avec $f : \mathbb{R}^3 \to \mathbb{R}^3$ définie par
$$f(x, y, z) = \big(-2x - 2y + 2z, -3x - y + 3z, -x + y + z\big).$$
Nous avions trouvé les valeurs propres et les vecteurs propres associés suivants :
$$\lambda_1 = -4 \quad v_1 = (1, 1, 0) \qquad \lambda_2 = 2 \quad v_2 = (0, 1, 1) \qquad \lambda_3 = 0 \quad v_3 = (1, 0, 1)$$
Par le corollaire 1, (v_1, v_2, v_3) forme une famille libre de $\mathbb{R}^3$ (ce que l'on vérifie par un calcul direct). Mais trois vecteurs indépendants de $\mathbb{R}^3$ forment automatiquement une base. Conclusion : (v_1, v_2, v_3) est une base de vecteurs propres de $\mathbb{R}^3$.
Ce que l'on peut aussi écrire :
$$\mathbb{R}^3 = \mathbb{R}v_1 \oplus \mathbb{R}v_2 \oplus \mathbb{R}v_3$$
ou encore
$$\mathbb{R}^3 = E_{-4} \oplus E_2 \oplus E_0.$$

Exemple 5.
Reprenons l'exemple 2, avec $f : \mathbb{R}^n \to \mathbb{R}^n$ définie par $(x_1, \ldots, x_{n-1}, x_n) \mapsto (x_1, \ldots, x_{n-1}, 0)$.
Nous avions trouvé deux valeurs propres 0 et 1.
Pour la valeur propre 0, nous avions un seul vecteur propre $e_n = (0, \ldots, 0, 1)$, ainsi $E_0 = \mathbb{R}e_n$.
Pour la valeur propre 1, nous avions trouvé $n-1$ vecteurs propres linéairement indépendants $e_1 = (1, 0, 0, \ldots)$, $e_2 = (0, 1, 0, \ldots)$,..., $e_{n-1} = (0, \ldots, 0, 1, 0)$. Plus précisément,
$E_1 = \mathrm{Ker}(f - \mathrm{id}_{\mathbb{R}^n}) = \mathrm{Vect}(e_1, \ldots, e_{n-1}) = \big\{(x_1, x_2, \ldots, x_{n-1}, 0) \in \mathbb{R}^n \mid x_1, \ldots, x_{n-1} \in \mathbb{R}\big\} = \mathbb{R}^{n-1} \times \{0\}$.
Nous avons bien
$$\mathbb{R}^n = E_0 \oplus E_1 = \big(\mathbb{R}e_n\big) \oplus \big(\mathbb{R}^{n-1} \times \{0\}\big).$$

Preuve du théorème 1. Pour chaque $1 \leqslant i \leqslant r$, soit $v_i \in E_{\lambda_i}$. On suppose $v_1 + \cdots + v_r = 0$, et nous allons montrer par récurrence qu'alors $v_1 = 0$, $v_2 = 0$,..., $v_r = 0$.
Si $r = 1$, c'est vérifié. Fixons $r \geqslant 2$ et supposons notre assertion vraie pour les familles de $r-1$ vecteurs. Soit une famille qui vérifie
$$v_1 + v_2 + \cdots + v_{r-1} + v_r = 0. \tag{1}$$
Par composition par l'application linéaire f,
$$f(v_1) + f(v_2) + \cdots + f(v_{r-1}) + f(v_r) = 0.$$
Mais comme $v_i \in E_{\lambda_i}$ alors $f(v_i) = \lambda_i v_i$ et donc :
$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{r-1} v_{r-1} + \lambda_r v_r = 0. \tag{2}$$
À partir des équations (1) et (2), on calcule l'expression (2) $- \lambda_r$(1) :
$$(\lambda_1 - \lambda_r)v_1 + (\lambda_2 - \lambda_r)v_2 + \cdots + (\lambda_{r-1} - \lambda_r)v_{r-1} = 0$$

(le vecteur v_r n'apparaît plus dans cette expression). On applique l'hypothèse de récurrence à la famille de $n-1$ vecteurs $(\lambda_1-\lambda_r)v_1, \ldots,(\lambda_{r-1}-\lambda_r)v_{r-1}$, ce qui implique que tous ces vecteurs sont nuls :

$$(\lambda_1-\lambda_r)v_1=0 \qquad \ldots \qquad (\lambda_{r-1}-\lambda_r)v_{r-1}=0$$

Comme les valeurs propres sont distinctes, alors $\lambda_i-\lambda_r \neq 0$ (pour $i=1,\ldots,r-1$). Ainsi

$$v_1=0 \qquad \ldots \qquad v_{r-1}=0.$$

L'équation (1) implique en plus

$$v_r=0.$$

Cela termine la récurrence. □

1.4. Rappels sur les sommes directes

Il faut bien comprendre le vocabulaire suivant. On commence par le cas de deux sous-espaces.

Définition 3.
Soient E_1, E_2 deux sous-espaces vectoriels d'un espace vectoriel E.

- La *somme* de E_1 et de E_2 est
$$E_1+E_2=\left\{v_1+v_2 \mid v_1 \in E_1 \text{ et } v_2 \in E_2\right\}.$$
- On dit que E_1 et E_2 sont en *somme directe* si $E_1 \cap E_2=\{0\}$.
- On dit que E_1 et E_2 sont en *somme directe dans E* si $E_1+E_2=E$ et $E_1 \cap E_2=\{0\}$. On note alors $E=E_1 \oplus E_2$.

Cela se généralise à plusieurs sous-espaces.

Définition 4.
Soient $E_1, E_2, \ldots, E_r$ des sous-espaces vectoriels d'un espace vectoriel E.

- La *somme* de $E_1, E_2, \ldots, E_r$ est
$$E_1+E_2+\cdots+E_r=\left\{v_1+v_2+\cdots+v_r \mid v_1 \in E_1, v_2 \in E_2, \ldots, v_r \in E_r\right\}.$$
- On dit que $E_1, E_2, \ldots, E_r$ sont en *somme directe* si
$$\forall v_1 \in E_1, \ldots, \forall v_r \in E_r \qquad v_1+\cdots+v_r=0 \implies v_1=0, \ldots, v_r=0.$$
- On dit que $E_1, E_2, \ldots, E_r$ sont en *somme directe dans E* s'ils sont en somme directe et que $E_1+E_2+\cdots+E_r=E$. On note alors $E=E_1 \oplus E_2 \oplus \cdots \oplus E_r$.

Exemple 6.

- Si $(v_1, \ldots, v_n)$ est une famille libre de E, alors les droites $\mathbb{K}v_1, \ldots, \mathbb{K}v_n$ sont en somme directe.
- Si $(v_1, \ldots, v_n)$ est une base de E, alors les droites $\mathbb{K}v_1, \ldots, \mathbb{K}v_n$ sont en somme directe dans E : $E=\mathbb{K}v_1 \oplus \cdots \oplus \mathbb{K}v_n$.

La notion de somme directe généralise celle de base :

Proposition 1.
Les sous-espaces vectoriels $E_1, \ldots, E_r$ sont en somme directe si et seulement si, pour chaque $v \in E_1+\cdots+E_r$, il existe $v_i \in E_i$ unique $(1 \leqslant i \leqslant r)$ tel que

$$v=v_1+v_2+\cdots+v_r.$$

En particulier, $E = E_1 \oplus \cdots \oplus E_r$ si et seulement si, pour tout $v \in E$, il existe un unique $v_i \in E_i$ tel que

$$v = v_1 + v_2 + \cdots + v_r.$$

Voici une autre application : si $E = E_1 \oplus \cdots \oplus E_r$ et si $\mathscr{B}_i$ est une base de E_i (pour $1 \leqslant i \leqslant r$) alors $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_r$ est une base de E.

Il est facile de calculer la dimension d'une somme directe :

Proposition 2.
Les sous-espaces vectoriels $E_1, \ldots, E_r$ sont en somme directe si et seulement si

$$\dim(E_1 + \cdots + E_r) = \dim E_1 + \cdots + \dim E_r.$$

En particulier, si $E = E_1 + \cdots + E_r$, alors les sous-espaces vectoriels $E_1, \ldots, E_r$ sont en somme directe dans E si et seulement si

$$\dim E = \dim E_1 + \cdots + \dim E_r.$$

Mini-exercices.

1. Soit $f : E \to E$ un endomorphisme. Quel est le lien entre l'assertion « f injective » et les valeurs propres de f ? Si E est de dimension finie, que peut-on dire de plus ?
2. Soit $f : E \to E$ un endomorphisme. Dire si les assertions suivantes sont vraies ou fausses. Justifier.
   (a) Si λ_1 et λ_2 sont valeurs propres, alors $\lambda_1 + \lambda_2$ aussi.
   (b) Si v_1 et v_2 sont vecteurs propres, alors $v_1 + v_2$ aussi.
   (c) Si λ est valeur propre, alors $\mu \cdot \lambda$ aussi (pour $\mu \in \mathbb{K}^*$).
   (d) Si v est vecteur propre, alors $\mu \cdot v$ aussi (pour $\mu \in \mathbb{K}^*$).
3. Soient $f, g : E \to E$ deux endomorphismes. Dire si les assertions suivantes sont vraies ou fausses. Justifier.
   (a) Si λ est valeur propre pour f et pour g, alors λ est valeur propre pour $f + g$.
   (b) Si v est vecteur propre pour f et pour g, alors v est vecteur propre pour $f + g$.
   (c) Si λ est valeur propre pour f, alors $\mu \cdot \lambda$ est valeur propre pour $\mu \cdot f$ (pour $\mu \in \mathbb{K}^*$).
   (d) Si v est vecteur propre pour f, alors $\mu \cdot v$ est vecteur propre pour $\mu \cdot f$ (pour $\mu \in \mathbb{K}^*$).
4. Montrer (sans utiliser le cours) que si λ et μ sont deux valeurs propres distinctes d'un endomorphisme $f : E \to E$ alors $E_\lambda \cap E_\mu = \{0\}$.
5. Montrer que si $f : E \to E$ est un endomorphisme vérifiant $f^2 = f$ (c'est-à-dire, pour tout $x \in E$, $f(f(x)) = f(x)$) alors $E_0 = \operatorname{Ker} f$ et $E_1 = \operatorname{Im} f$.

2. Polynôme caractéristique

Le polynôme caractéristique permet de trouver facilement les valeurs propres. Encore une fois, le chapitre « Valeurs propres, vecteurs propres » sur les matrices fournit de nombreux exemples.

2.1. Polynôme caractéristique

Définition 5.
Soit $f : E \to E$ un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie n. Soit $A \in M_n(\mathbb{K})$ la matrice de f dans une base $\mathcal{B}$.
Le ***polynôme caractéristique*** de f est égal au polynôme caractéristique de la matrice A :

$$\boxed{\chi_f(X) = \chi_A(X) = \det(A - XI_n).}$$

Le polynôme caractéristique est indépendant de la matrice A (et du choix de la base $\mathcal{B}$). En effet, si B est la matrice du même endomorphisme f mais dans une autre base $\mathcal{B}'$, alors on sait qu'il existe $P \in M_n(\mathbb{K})$ inversible telle que $B = P^{-1}AP$. On écrit :

$$B - XI_n = P^{-1}(A - XI_n)P.$$

Alors,

$$\chi_B(X) = \det(B - XI_n) = \frac{1}{\det(P)} \cdot \det(A - XI_n) \cdot \det(P) = \det(A - XI_n) = \chi_A(X).$$

2.2. Caractérisation des valeurs propres

Proposition 3.

$$\boxed{\lambda \text{ valeur propre de } f \quad \iff \quad \chi_f(\lambda) = 0}$$

Voyons une autre formulation. Soit $f : E \to E$. Soit $A \in M_n(\mathbb{K})$ sa matrice dans une base $\mathcal{B}$. Soit $\lambda \in \mathbb{K}$. Alors :

$$\lambda \text{ valeur propre de } f \iff \det(A - \lambda I_n) = 0$$

Démonstration.

$$\begin{aligned}
\lambda \text{ est une valeur propre de } f &\iff \exists v \in E \setminus \{0\}, \quad f(v) = \lambda v \\
&\iff \operatorname{Ker}(f - \lambda \operatorname{id}_E) \neq \{0\} \\
&\iff f - \lambda \operatorname{id}_E \text{ n'est pas injective} \\
&\iff f - \lambda \operatorname{id}_E \text{ n'est pas bijective} \\
&\iff A - \lambda I_n \text{ n'est pas inversible} \\
&\iff \det(A - \lambda I_n) = 0 \\
&\iff \chi_f(\lambda) = 0
\end{aligned}$$

Noter que l'équivalence entre « $f - \lambda \operatorname{id}_E$ non injective » et « $f - \lambda \operatorname{id}_E$ non bijective » repose sur le fait que : (a) $f - \lambda \operatorname{id}_E$ est un endomorphisme (il va de E dans lui-même) et (b) E est de dimension finie. □

Exemple 7.
Si D est la matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

alors $\chi_D(X) = (\lambda_1 - X) \cdots (\lambda_n - X)$ et donc les λ_i sont les racines de $\chi_D(X)$ et aussi les valeurs propres de D.

Exemple 8.
Soit E un $\mathbb{C}$-espace vectoriel de dimension finie. Soit $f : E \to E$ une *symétrie*, c'est-à-dire un endomorphisme qui vérifie $f^2 = -f$. Montrons que le polynôme caractéristique est de la forme $\chi_f(X) = \pm X^a (X+1)^b$ avec $a, b \geqslant 0$.
Pour cela, cherchons quelle peut être une valeur propre de f. Soit $\lambda \in \mathbb{K}$ une valeur propre, et soit $v \in E \setminus \{0\}$ un vecteur propre associé. Alors :

$$\begin{aligned} f(v) = \lambda v &\implies f\big(f(v)\big) = f(\lambda v) \\ &\implies -f(v) = \lambda f(v) \qquad \text{car } f^2 = -f \\ &\implies -\lambda v = \lambda^2 v \qquad \text{car } v \text{ vecteur propre} \\ &\implies -\lambda = \lambda^2 \qquad \text{car } v \text{ non nul} \\ &\implies \lambda(\lambda+1) = 0 \\ &\implies \lambda = 0 \quad \text{ou} \quad \lambda = -1 \end{aligned}$$

Conséquence : les seules valeurs propres possibles sont 0 ou -1. Par la proposition 3, les seules racines possibles de $\chi_f(X)$ sont 0 et -1. Donc $\chi_f(X) = \alpha X^a (X+1)^b$ où $\alpha \in \mathbb{C}^*$, $a, b \geqslant 0$. Nous verrons juste après que le coefficient dominant est ± 1. Ainsi $\chi_f(X) = \pm X^a (X+1)^b$.

2.3. Coefficients du polynôme caractéristique

Proposition 4.
Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Soit $f : E \to E$ un endomorphisme. Soit A la matrice de f dans une base $\mathcal{B}$. Le polynôme caractéristique de f est de degré n et vérifie :

$$\chi_f(X) = (-1)^n X^n + (-1)^{n-1} (\operatorname{tr} A) X^{n-1} + \cdots + \det A.$$

Si f admet n valeurs propres, qui sont donc toutes les racines de $\chi_f(X)$, alors de l'égalité

$$\chi_f(X) = (-1)^n \prod_{i=1}^{n} (X - \lambda_i)$$

on en déduit :

$$\boxed{\text{La somme des valeurs propres vaut } \operatorname{tr} A.}$$

$$\boxed{\text{Le produit des valeurs propres vaut } \det A.}$$

Preuve de la proposition 4. Si $A=(a_{ij})_{1\leqslant i,j\leqslant n}$ est la matrice de f, on a

$$\chi_f(X)=\det(A-XI_n)=\begin{vmatrix} a_{11}-X & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-X & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-X \end{vmatrix}.$$

Par la définition du déterminant :

$$\chi_f(X)=\sum_{\sigma\in\mathscr{S}_n}\epsilon(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n} \qquad \text{où} \quad b_{ij}=a_{ij} \ \text{ si } \ i\neq j \quad \text{ et } \quad b_{ii}=a_{ii}-X$$

On met à part la permutation identité :

$$\chi_f(X)=(a_{11}-X)\cdots(a_{nn}-X)+\sum_{\sigma\in\mathscr{S}_n,\sigma\neq\mathrm{id}}\epsilon(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n}.$$

Or, si $\sigma\neq\mathrm{id}$, il y a au plus $n-2$ entiers k tels que $\sigma(k)=k$, et donc le polynôme

$$\sum_{\sigma\in\mathscr{S}_n,\sigma\neq\mathrm{id}}\epsilon(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n}$$

est de degré au plus $n-2$.
Conclusion :

- Le polynôme $\chi_f(X)$ est de degré n.
- Les termes de degré n et $n-1$ proviennent du produit
$$(a_{11}-X)\cdots(a_{nn}-X)=(-1)^nX^n+(-1)^{n-1}(\mathrm{tr}A)X^{n-1}+\cdots$$
- Le terme constant, quant à lui, est donné par $\chi_f(0)=\det A$.

□

2.4. Exemples et applications

Voyons quelques applications du polynôme caractéristique :

- Si E est un $\mathbb{K}$-espace vectoriel de dimension n, alors tout endomorphisme $f:E\to E$ admet au plus n valeurs propres. En effet, le polynôme caractéristique de f est un polynôme de degré n, donc admet au plus n racines dans $\mathbb{K}$.
- Si E est un $\mathbb{C}$-espace vectoriel, alors tout endomorphisme $f:E\to E$ admet au moins une valeur propre. En effet, le polynôme caractéristique de f est un polynôme complexe non constant donc admet (au moins) une racine $\lambda\in\mathbb{C}$. Alors λ est une valeur propre de f.

Exemple 9.
Soit $A\in M_n(\mathbb{R})$ une matrice. Alors A possède un sous-espace invariant de dimension 1 ou 2.

Démonstration. Considérons la matrice $A\in M_n(\mathbb{R})$ comme une matrice de $M_n(\mathbb{C})$. Alors A possède une valeur propre $\lambda=a+\mathrm{i}\,b\in\mathbb{C}$ (a,b réels), et un vecteur propre associé $Z=X+\mathrm{i}\,Y\in\mathbb{C}^n\setminus\{0\}$ où $X,Y\in\mathbb{R}^n$.
Alors :

$$\begin{aligned} AZ=\lambda Z &\implies A(X+\mathrm{i}\,Y)=(a+\mathrm{i}\,b)(X+\mathrm{i}\,Y)\\ &\implies AX+\mathrm{i}AY=(aX-bY)+\mathrm{i}(bX+aY)\\ &\implies \left\{\begin{array}{rcl} AX &=& aX-bY\\ AY &=& bX+aY\end{array}\right.\end{aligned}$$

En particulier, AX et AY appartiennent à $\mathrm{Vect}(X,Y)$, donc le sous-espace (réel) $\mathrm{Vect}(X,Y)$ est stable par A. Or X ou Y n'est pas nul, donc $\mathrm{Vect}(X,Y)$ est de dimension 1 ou 2. □

Exemple 10.
Soit f un endomorphisme de E. Si E est de dimension n et si le polynôme caractéristique $\chi_f(X) \in \mathbb{K}[X]$ admet n racines distinctes dans $\mathbb{K}$, alors il existe une base de E formée de vecteurs propres de f.

Démonstration. Soient $\lambda_1,\ldots,\lambda_n \in \mathbb{K}$ les n racines distinctes de $\chi_f(X)$. Ce sont aussi n valeurs propres de f. Soient $v_1,\ldots,v_n$ des vecteurs propres associés. Par le corollaire 1, la famille $(v_1,\ldots,v_n)$ est une famille libre de E. C'est donc une famille libre à n éléments dans un espace vectoriel de dimension n : cela implique que $(v_1,\ldots,v_n)$ est une base de E. □

Mini-exercices.

1. Calculer le polynôme caractéristique d'une matrice triangulaire.
2. Trouver une application linéaire $f : \mathbb{R}^2 \to \mathbb{R}^2$ qui n'admet aucune valeur propre réelle. Montrer que les valeurs propres complexes d'un tel endomorphisme f seront toujours conjuguées.
3. Calculer le polynôme caractéristique de $A = \begin{pmatrix} -1 & \alpha+1 & 0 \\ 1 & \alpha & 1 \\ 3 & -\alpha-1 & 2 \end{pmatrix}$ en fonction de $\alpha \in \mathbb{R}$. Montrer que -1 est valeur propre et en déduire les autres valeurs propres. Quelle est la multiplicité de chaque valeur propre ? Trouver un vecteur propre pour chaque valeur propre.
4. Soit E un $\mathbb{C}$-espace vectoriel de dimension n. Soit $f : E \to E$ un endomorphisme tel que f^n soit l'application nulle (c'est-à-dire, pour tout $x \in E$, $f \circ f \circ \cdots \circ f(x) = 0$). Si λ est une valeur propre de f, que peut valoir λ ? En déduire le polynôme caractéristique de f.

3. Diagonalisation

Dans le chapitre « Valeurs propres, vecteurs propres », nous avions énoncé un critère qui permet de diagonaliser certaines matrices. Ici nous allons énoncer un critère plus fort : nous trouvons des conditions qui sont exactement équivalentes à ce qu'une matrice soit diagonalisable.

3.1. Endomorphisme diagonalisable

Définition 6.
On dit qu'un endomorphisme $f : E \to E$ est ***diagonalisable*** s'il existe une base de E formée de vecteurs propres de f.

Rappelons que :

Définition 7.
Soit A une matrice de $M_n(\mathbb{K})$. On dit que A est ***diagonalisable*** sur $\mathbb{K}$ s'il existe une matrice $P \in M_n(\mathbb{K})$ inversible telle que $P^{-1}AP$ soit diagonale.

Bien sûr, les deux définitions sont cohérentes :

Proposition 5.
Si A est la matrice de f dans une base $\mathcal{B}$ quelconque alors :

$$\boxed{f \text{ diagonalisable} \quad \Longleftrightarrow \quad A \text{ diagonalisable}}$$

Cette proposition est facile, mais il faut bien comprendre ce lien.

Démonstration.
- $\Longrightarrow$. Soit f un endomorphisme diagonalisable.
  - Si D est la matrice de f dans la base $(v_1, \ldots, v_n)$ formée de vecteurs propres, alors D est une matrice diagonale. En effet, comme $f(v_i) = \lambda_i v_i$, la matrice D est diagonale et le i-ème coefficient de la diagonale est λ_i.
  - Si A est la matrice de f dans une base $\mathscr{B}$ quelconque, alors A est semblable à la matrice D ci-dessus. Il existe donc P inversible telle que $D = P^{-1}AP$ soit diagonale.
- $\Longleftarrow$. Soit A une matrice diagonalisable.
  L'endomorphisme f, considéré comme une application $f : \mathbb{K}^n \to \mathbb{K}^n$, s'écrit $f(X) = AX$ où les coordonnées de X s'expriment dans la base canonique $(Y_1, \ldots, Y_n)$:
  $$Y_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \qquad Y_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \qquad \cdots$$
  Soit P une matrice telle que $D = P^{-1}AP$ soit une matrice diagonale. Notons $\lambda_1, \ldots, \lambda_n$ les coefficients de la diagonale. Notons $(X_1, \ldots, X_n)$ les vecteurs colonnes de P. Ils s'obtiennent aussi comme $X_i = PY_i$. Montrons que X_i est un vecteur propre de f, associé à la valeur propre λ_i :
  $$f(X_i) = AX_i = (PDP^{-1})(PY_i) = PDY_i = P(\lambda_i Y_i) = \lambda_i(PY_i) = \lambda_i X_i.$$
  Comme P est inversible, alors $(X_1, \ldots, X_n)$ est une base de vecteurs propres.

□

Exemple 11 (Projection).
On suppose que $E = F \oplus G$ avec F et G deux sous-espaces vectoriels de E. N'importe quel $v \in E$ se décompose de façon unique en $v = x + y$ avec $x \in F$, $y \in G$. La projection sur F suivant G est l'endomorphisme de E défini par :

$$p : \begin{array}{ccc} E & \longrightarrow & E \\ v = x + y & \longmapsto & x \end{array}$$

- Pour $v = x \in F$, on a $p(x) = x$; ces x sont les vecteurs propres pour la valeur propre 1 :
  $$F = \operatorname{Ker}(p - \mathrm{id}_E) = E_1(p).$$
- Pour $v = y \in G$, on a $p(y) = 0$; ces y sont les vecteurs propres pour la valeur propre 0 :
  $$G = \operatorname{Ker} p = E_0(p).$$
- Comme $E = F \oplus G$, alors l'union d'une base de vecteurs propres de $E_1(p)$ et d'une base de vecteurs propres de $E_0(p)$ forme une base de vecteurs propres de E.
- Conclusion : p est diagonalisable.

Exemple 12 (Réflexion).
On suppose encore que $E = F \oplus G$. On définit la réflexion par rapport à F suivant G par :

$$r : \begin{array}{ccc} E & \longrightarrow & E \\ v = x + y & \longmapsto & x - y \end{array}$$

De façon semblable à l'exemple précédent, on montre que r est diagonalisable avec

$$F = \operatorname{Ker}(r - \mathrm{id}_E) = E_1(r) \qquad \text{et} \qquad G = \operatorname{Ker}(r + \mathrm{id}_E) = E_{-1}(r).$$

Proposition 6.
Si f est un endomorphisme de E et si on note $\lambda_1, \ldots, \lambda_r$ ses valeurs propres distinctes alors :
$$f \text{ est diagonalisable} \iff E = \operatorname{Ker}(f - \lambda_1 \operatorname{id}_E) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_E).$$
Autrement dit, f est diagonalisable si et seulement si E est somme directe des sous-espaces propres de f.

Notons que, dans les deux exemples précédents (la projection et la symétrie), l'espace vectoriel E est bien la somme directe des deux seuls sous-espaces propres.

Démonstration.

- $\Longrightarrow$. Par le théorème 1, les sous-espaces propres de f sont en somme directe. Notons $F = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$. Comme f est supposé diagonalisable, alors il existe une base de E formée de vecteurs propres de f. Mais ces vecteurs propres sont aussi des éléments de F. Ainsi F contient une base de E. On en conclut que $E = F = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.
- $\Longleftarrow$. Par hypothèse, les sous-espaces propres sont en somme directe dans E. On choisit une base pour chacun des sous-espaces propres $E_\lambda = \operatorname{Ker}(f - \lambda \operatorname{id}_E)$. Les vecteurs de chacune de ces bases sont des vecteurs propres de f. L'union de ces bases est une base de E formée de vecteurs propres de f, donc f est diagonalisable.

□

3.2. Rappels sur les polynômes

Rappelons quelques définitions. Soit $P(X) \in \mathbb{K}[X]$ un polynôme.

- $\lambda \in \mathbb{K}$ est ***racine*** de P si $P(\lambda) = 0$.
- $\lambda \in \mathbb{K}$ est racine de P si et seulement si $P(X) = (X - \lambda)Q(X)$ pour un polynôme $Q \in \mathbb{K}[X]$.
- La ***multiplicité*** de $\lambda \in \mathbb{K}$ dans P est le plus grand entier m tel que $P(X) = (X - \lambda)^m Q(X)$ pour un polynôme $Q \in \mathbb{K}[X]$.

Notation. On note $m(\lambda)$ la multiplicité de λ comme racine de P.

- Une racine de multiplicité 1 est une ***racine simple***.
- Une racine de multiplicité 2 est une ***racine double***...
- Si λ n'est pas racine de P, on posera $m(\lambda) = 0$.

Exemple 13.

- $P(X) = (X-2)^3(X^2 + X + 1) \in \mathbb{R}[X]$ admet 2 comme racine, et sa multiplicité est 3.
- Le même polynôme considéré cette fois dans $\mathbb{C}[X]$ s'écrit
$$P(X) = (X-2)^3 \left(X + \frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2} \right) \left(X + \frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2} \right).$$
Les racines complexes $-\frac{1}{2} \pm \mathrm{i}\frac{\sqrt{3}}{2}$ sont chacune de multiplicité 1.

Exemple 14.
Si $\lambda_1, \ldots, \lambda_r$ sont deux à deux distincts et si
$$P(X) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r},$$
alors m_i est la multiplicité de λ_i dans $P(X)$, pour tout i $(1 \leqslant i \leqslant r)$.

Définition 8.
Un polynôme $P(X) \in \mathbb{K}[X]$ est ***scindé*** sur $\mathbb{K}$ s'il s'écrit
$$P(X) = a_n(X - \lambda_1) \cdots (X - \lambda_n)$$

pour certains $\lambda_i \in \mathbb{K}$ et un $a_n \in \mathbb{K}^*$.

Souvent, on regroupe les racines égales et on écrit :

$$P(X) = a_n(X-\lambda_1)^{m(\lambda_1)} \cdots (X-\lambda_r)^{m(\lambda_r)}$$

avec les λ_i deux à deux distinctes et leurs multiplicités $m(\lambda_i) \geqslant 1$.

Exemple 15.

- Le polynôme $P(X) = (X-2)^3(X^2+X+1) \in \mathbb{R}[X]$ n'est pas scindé sur $\mathbb{R}$, car X^2+X+1 n'a pas de racine réelle. Par contre, il est scindé sur $\mathbb{C}$ (voir le commentaire ci-dessous).
- Le polynôme $P(X) = X^2+4X-3$ est scindé sur $\mathbb{R}$ car ses racines sont les réels $\lambda_1 = -2-\sqrt{7}$, $\lambda_2 = -2+\sqrt{7}$. Il s'écrit donc aussi $P(X) = (X-\lambda_1)(X-\lambda_2)$.

Quelques commentaires importants :

- Pour un polynôme $P \in \mathbb{K}[X]$ non nul, on a :
$$P \text{ scindé sur } \mathbb{K} \quad \Longleftrightarrow \quad \sum_{\lambda \text{ racine de } P} m(\lambda) = \deg P$$
- D'après le théorème de d'Alembert-Gauss :
$$\boxed{\text{Tous les polynômes sont scindés lorsque le corps de base est } \mathbb{C}.}$$
- Et donc, si $\mathbb{K} = \mathbb{C}$, on a toujours :
$$\deg P = \sum_{\lambda \text{ racine de } P} m(\lambda).$$

3.3. Diagonalisation

Nous allons énoncer un critère simple qui caractérise si un endomorphisme est diagonalisable ou pas. Ce critère se base sur le polynôme caractéristique et la dimension des sous-espaces propres, pour lesquels on établit un premier lien dans la proposition suivante. Les preuves seront faites dans la section 3.6.

Proposition 7.

Soient f un endomorphisme de E et χ_f son polynôme caractéristique. Soit λ une valeur propre de f, de multiplicité $m(\lambda)$ comme racine de χ_f, et soit E_λ le sous-espace propre associé. Alors on a

$$\boxed{1 \leqslant \dim E_\lambda \leqslant m(\lambda).}$$

Énonçons maintenant le théorème principal de ce chapitre. C'est un critère pour savoir si un endomorphisme – ou une matrice – est diagonalisable. Contrairement aux critères précédents, il s'agit ici d'une équivalence.

Théorème 2.

Soit $f : E \to E$ un endomorphisme. Alors :

$$\boxed{f \text{ est diagonalisable sur } \mathbb{K} \quad \Longleftrightarrow \quad \begin{cases} i) \ \chi_f(X) \text{ est scindé sur } \mathbb{K} \\ \text{et} \\ ii) \text{ pour toute valeur propre } \lambda \text{ de } f, \\ \quad m(\lambda) = \dim \operatorname{Ker}(f - \lambda \operatorname{id}_E). \end{cases}}$$

Voici une autre reformulation pour mieux comprendre ce théorème.
Soit $f : E \to E$. L'endomorphisme f est diagonalisable si et seulement si le polynôme caractéristique de f, $\chi_f(X)$, est scindé sur $\mathbb{K}$ et si, pour chacune des racines λ, la multiplicité de λ est égale à la dimension du sous-espace propre $E_\lambda = \mathrm{Ker}(f - \lambda\,\mathrm{id}_E)$.

Bien évidemment, il faut savoir transcrire ce théorème en termes de matrices :
Soit $A \in M_n(K)$. Alors :

$$A \text{ est diagonalisable sur } \mathbb{K} \quad \iff \quad \begin{cases} \text{i) } \chi_A(X) \text{ est scindé sur } \mathbb{K} \\ \text{ et} \\ \text{ii) pour toute valeur propre } \lambda \text{ de } A, \\ \quad m(\lambda) = \dim \mathrm{Ker}(A - \lambda I_n). \end{cases}$$

Corollaire 2.
Si le polynôme $\chi_f(X)$ (resp. $\chi_A(X)$) est scindé et si les racines sont simples, alors f (resp. A) est diagonalisable.

En effet, dans ce cas, la multiplicité $m(\lambda)$ vaut 1 pour chaque valeur. Par la proposition 7, on a $1 \leqslant \dim E_\lambda \leqslant m(\lambda)$, donc la dimension de chaque sous-espace propre est aussi 1. Par le théorème 2, l'endomorphisme (ou la matrice) est diagonalisable.

3.4. Exemples

Exemple 16.
Toute matrice réelle 2×2 symétrique $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ est diagonalisable sur $\mathbb{R}$.
La trace vaut $\mathrm{tr}\,A = a + d$, le déterminant vaut $\det A = ad - b^2$. On utilise la formule de la proposition 4 pour en déduire, sans calculs, que le polynôme caractéristique est :

$$\chi_A(X) = X^2 - \mathrm{tr}(A)X + \det(A) = X^2 - (a+d)X + ad - b^2.$$

Sans les calculer, montrons que $\chi_A(X)$ admet deux racines réelles. On calcule le discriminant de l'équation du second degré donnée par $\chi_A(X) = 0$:

$$\Delta = (a+d)^2 - 4(ad - b^2) = a^2 + d^2 - 2ad + 4b^2 = (a-d)^2 + 4b^2. \tag{3}$$

Cela prouve que $\Delta \geqslant 0$. Ainsi les deux racines λ_1 et λ_2 du polynôme caractéristique sont réelles.
Conclusion :

- Si $\Delta > 0$ alors ces deux racines sont réelles et distinctes. Ainsi $\chi_A(X) = (X - \lambda_1)(X - \lambda_2)$ est scindé à racines simples, donc la matrice A est diagonalisable.
- Si $\Delta = 0$ alors, par l'équation (3), on a $(a-d)^2 = 0$ et $b^2 = 0$. Donc $a = d$ et $b = 0$. La matrice A est une matrice diagonale (donc diagonalisable !).

Exemple 17.
La matrice de permutation circulaire

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 & 1 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \in M_n(\mathbb{C})$$

est diagonalisable sur $\mathbb{C}$.

En effet :

- son polynôme caractéristique est $\chi_A(X) = (-1)^n(X^n - 1)$ (voir le chapitre « Valeurs propres, vecteurs propres », section « Matrice compagnon »),
- les valeurs propres sont les racines n-ièmes de l'unité :
$$1, e^{\mathrm{i}\frac{2\pi}{n}}, \ldots, e^{\mathrm{i}\frac{2(n-1)\pi}{n}}$$
- les racines sont simples,
- le polynôme caractéristique est bien sûr scindé sur $\mathbb{C}$,
- par le corollaire 2, la matrice A est donc diagonalisable.

Exercice : Trouver une base de vecteurs propres.

Exemple 18.
Soit $n \geqslant 2$. Soit la matrice
$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \in M_n(\mathbb{K})$$
définie par des 1 au-dessus de la diagonale. Cette matrice n'est jamais diagonalisable ! En effet :

- Le polynôme caractéristique de A est $\chi_A(X) = (-1)^n X^n$ (la matrice est triangulaire, de diagonale nulle). Donc $\lambda = 0$ est la seule valeur propre et $m(0) = n$.
- Par contre, $E_0 = \operatorname{Ker}(A - \lambda I_n) = \operatorname{Ker} A$ est de dimension $\dim \operatorname{Ker} A < n$ car A n'est pas la matrice nulle.
- Comme $\dim E_0 < m(0)$ alors, par le théorème 2, A n'est pas diagonalisable.

3.5. Diagonaliser

Diagonaliser une matrice $A \in M_n(\mathbb{K})$ signifie trouver, si elles existent, $P \in M_n(\mathbb{K})$ inversible et $D \in M_n(\mathbb{K})$ diagonale telles que
$$A = PDP^{-1}.$$

Soit $A \in M_n(\mathbb{K})$ une matrice carrée $n \times n$. Pour la diagonaliser :

1. On calcule d'abord son polynôme caractéristique $\chi_A(X)$.
2. On cherche les racines de $\chi_A(X)$: ce sont les valeurs propres de A. Si $\chi_A(X)$ n'est pas scindé sur $\mathbb{K}$, alors A n'est pas diagonalisable.
3. Pour chaque valeur propre λ de A, on cherche une base de $\operatorname{Ker}(A - \lambda I_n)$, c'est-à-dire on cherche une base de l'espace des solutions du système
$$AX = \lambda X.$$
4. Si, pour toute valeur propre λ de A, $\dim \operatorname{Ker}(A - \lambda I_n) = m(\lambda)$, alors A est diagonalisable. Sinon elle n'est pas diagonalisable.
5. Dans le cas diagonalisable, la réunion des bases des sous-espaces propres forme une base de vecteurs propres. Ainsi, si P est la matrice dont les vecteurs colonnes sont ces vecteurs propres, alors $D = P^{-1}AP$ est une matrice diagonale dont les éléments diagonaux sont les valeurs propres λ de A, chacune apparaissant $m(\lambda)$ fois.

On renvoie une dernière fois au chapitre « Valeurs propres, vecteurs propres » pour des exemples de diagonalisation.

Exemple 19.
Soit

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

Démontrons que A est diagonalisable sur $\mathbb{R}$ et trouvons une matrice P telle que $P^{-1}AP$ soit diagonale.

1. Commençons par calculer le polynôme caractéristique de A :
$$\chi_A(X) = \det(A - XI_3) = \begin{vmatrix} 1-X & 0 & 0 \\ 0 & 1-X & 0 \\ 1 & -1 & 2-X \end{vmatrix} = (1-X)^2(2-X)$$
2. Les racines du polynôme caractéristique sont les réels 1 avec la multiplicité $m(1) = 2$, et 2 avec la multiplicité $m(2) = 1$. On remarque de plus que le polynôme est scindé sur $\mathbb{R}$.
3. Déterminons les sous-espaces propres associés.
   - Soit E_1 le sous-espace propre associé à la valeur propre double 1 : $E_1 = \operatorname{Ker}(A - I_3) = \{X \in \mathbb{R}^3 \mid A \cdot X = X\}$. Si on note $X = \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ alors :
   $$X \in E_1 \iff AX = X \iff \left\{ \begin{array}{rcl} x & = & x \\ y & = & y \\ x - y + 2z & = & z \end{array} \right. \iff x - y + z = 0$$
   $E_1 = \left\{ \left(\begin{smallmatrix} x \\ y \\ -x+y \end{smallmatrix}\right) \mid x \in \mathbb{R}, y \in \mathbb{R} \right\}$ est donc un plan vectoriel dont, par exemple, les vecteurs $X_1 = \left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)$ et $X_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$ forment une base.
   - Soit E_2 le sous-espace propre associé à la valeur propre simple 2 : $E_2 = \operatorname{Ker}(A - 2I_3) = \{X \in \mathbb{R}^3 \mid A \cdot X = 2X\}$. Alors :
   $$X \in E_2 \iff AX = 2X \iff \left\{ \begin{array}{rcl} x & = & 2x \\ y & = & 2y \\ x - y + 2z & = & 2z \end{array} \right. \iff x = 0 \text{ et } y = 0$$
   $E_2 = \left\{ \left(\begin{smallmatrix} 0 \\ 0 \\ z \end{smallmatrix}\right) \mid z \in \mathbb{R} \right\}$ est donc une droite vectorielle, dont le vecteur $X_3 = \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right)$ est une base.
4. Les dimensions des sous-espaces propres sont égales aux multiplicités des valeurs propres correspondantes : $\dim E_1 = 2 = m(1)$, $\dim E_2 = 1 = m(2)$. La matrice A est donc diagonalisable.
5. Dans la base (X_1, X_2, X_3), l'endomorphisme représenté par A (dans la base canonique) a pour matrice
$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$
Autrement dit, si on note P la matrice de passage dont les vecteurs colonnes sont X_1, X_2 et X_3, c'est-à-dire
$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix},$$
alors $P^{-1}AP = D$.

3.6. Preuves

Il nous reste à prouver la proposition 7 et le théorème 2 de la section 3.3. Rappelons l'énoncé de la proposition 7.

Proposition (Proposition 7).
Soient f un endomorphisme de E et χ_f son polynôme caractéristique. Soient λ une valeur propre de f, de multiplicité $m(\lambda)$ comme racine de χ_f, et E_λ le sous-espace propre associé. Alors on a

$$1 \leqslant \dim E_\lambda \leqslant m(\lambda).$$

Démonstration. Tout d'abord, par définition d'une valeur propre et d'un sous-espace propre, on a $\dim E_\lambda \geqslant 1$. Notons $p = \dim E_\lambda$ et $(e_1, \dots, e_p)$ une base de E_λ. On complète cette base en une base $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ de E. Dans cette base, la matrice de f est de la forme

$$A = \left(\begin{array}{c|c} \lambda I_p & C \\ \hline 0 & B \end{array}\right).$$

En effet, pour chaque $1 \leqslant i \leqslant p$, on a $f(e_i) = \lambda e_i$. Maintenant, en calculant le déterminant d'une matrice triangulaire par blocs :

$$\det(A - XI_n) = \det((\lambda - X)I_p) \cdot \det(B - XI_{n-p}) = (\lambda - X)^p \det(B - XI_{n-p}).$$

Cela prouve que $(\lambda - X)^p$ divise $\chi_f(X)$ et donc, par définition de la multiplicité d'une racine, on a $m(\lambda) \geqslant p$. $\square$

Passons à la preuve du théorème 2.

Théorème (Théorème 2).
Soit $f : E \to E$ un endomorphisme. Alors :

$$f \text{ est diagonalisable sur } \mathbb{K} \iff \begin{cases} i)\ \chi_f(X) \text{ est scindé sur } \mathbb{K} \\ \text{et} \\ ii)\ \text{pour toute valeur propre } \lambda \text{ de } f, \\ \quad m(\lambda) = \dim \operatorname{Ker}(f - \lambda \operatorname{id}_E). \end{cases}$$

Démonstration.

- $\Longrightarrow$. Supposons f diagonalisable et notons $\lambda_1, \dots, \lambda_r$ ses valeurs propres et $m(\lambda_1), \dots, m(\lambda_r)$ leurs multiplicités respectives dans $\chi_f(X)$. Comme f est diagonalisable, alors il existe une base $\mathcal{B}$ dans laquelle la matrice de f est une matrice diagonale D. Notons n_i le nombre de fois où λ_i apparaît dans la diagonale de D. On a alors

$$\chi_f(X) = \chi_D(X) = \prod_{i=1}^{r} (\lambda_i - X)^{n_i}.$$

  Cela prouve que $\chi_f(X)$ est scindé sur $\mathbb{K}$ et que $m(\lambda_i) = n_i$ pour tout $1 \leqslant i \leqslant r$.
  Comme D est diagonale, pour tout $1 \leqslant i \leqslant r$, il existe n_i vecteurs v de la base $\mathcal{B}$ de E tels que $f(v) = \lambda_i v$. Il existe donc n_i vecteurs linéairement indépendants dans E_{λ_i}, d'où $\dim E_{\lambda_i} \geqslant n_i$. Mais on sait que $n_i = m(\lambda_i)$, donc $\dim E_{\lambda_i} \geqslant m(\lambda_i)$. Enfin, on a démontré dans la proposition 7 que $\dim E_{\lambda_i} \leqslant m(\lambda_i)$, d'où l'égalité.
- $\Longleftarrow$. On suppose que $\chi_f(X)$ est scindé sur $\mathbb{K}$ et que, pour toute racine λ_i (avec $1 \leqslant i \leqslant r$), on a $\dim E_{\lambda_i} = m(\lambda_i)$. En particulier, on a

$$\chi_f(X) = \prod_{i=1}^{r} (\lambda_i - X)^{m(\lambda_i)}.$$

Notons $F = E_{\lambda_1} + \cdots + E_{\lambda_r}$. On sait que les sous-espaces propres sont en somme directe d'après le théorème 1, donc $F = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$. Ainsi, par la proposition 2, $\dim F = \sum_{i=1}^r \dim E_{\lambda_i} = \sum_{i=1}^r m(\lambda_i) = \deg \chi_f = \dim E$. On en conclut que $F \subset E$ et $\dim F = \dim E$, d'où $F = E$.
Pour chaque $1 \leqslant i \leqslant r$, on note $\mathscr{B}_i$ une base de E_{λ_i}. Soit $\mathscr{B} = \bigcup_{i=1}^r \mathscr{B}_i$. Alors $\mathscr{B}$ est une base de E (puisque c'est une base de F). Les vecteurs de E_{λ_i} sont des vecteurs propres. Ainsi, il existe une base de E formée de vecteurs propres de f, ce qui prouve que f est diagonalisable.

□

Mini-exercices.

1. Montrer que si λ est racine simple du polynôme caractéristique alors $\dim E_\lambda = 1$. Que peut-on dire pour une racine double ?
2. Soit $A = \left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right)$. Calculer le polynôme caractéristique de A. En déduire les valeurs propres. Déterminer une base de chaque sous-espace propre. La matrice A est-elle diagonalisable ? Généraliser au cas d'une matrice de taille $n \times n$ dont tous les coefficients sont 1.
3. Soit
$$A = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in M_n(\mathbb{R}).$$
Calculer le polynôme caractéristique de A, ses valeurs propres, leur multiplicité et la dimension des sous-espaces propres. A est-elle diagonalisable ?
4. Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E. On suppose qu'il existe un sous-espace F de E laissé stable par f. Notons $\chi_{f|F}$ le polynôme caractéristique de la restriction à F. Montrer alors que $\chi_{f|F}(X)$ divise $\chi_f(X)$ dans $\mathbb{K}[X]$. *Indication :* s'inspirer de la preuve de la proposition 7.

Auteurs du chapitre
D'après un cours de Sandra Delaunay et un cours d'Alexis Tchoudjem.
Revu et augmenté par Arnaud Bodin.
Relu par Stéphanie Bodin et Vianney Combet.