Algèbre linéaire et applications aux sciences des données

17 décembre 2024

Dernière semaine !

- Aujourd'hui : Dernier résultat important + TD
- Vendredi : Dernière application SVD + derniers exos + copies

ENDOMORPHISMES AUTO-ADJOINTS

Rappel : Si (E,q) est un espace euclidien de dimension n et que $\langle \cdot,\cdot \rangle$ est le produit scalaire associé à q, alors $\forall f \in \mathcal{L}(E)$, il existe un unique $f^* \in \mathcal{L}(E)$ tel que $\langle f(x), y \rangle = \langle x, f^*(y) \rangle \quad \forall (x,y) \in E^2$.

f^* s'appelle l'adjoint de f.

Si $A \in \mathbb{R}^{n \times n}$ est la matrice de f dans une base orthonormée, alors la matrice de f^* dans cette base est A^T.

Cas particulier : Si $E = \mathbb{R}^n$, $(Ax)^T y = x^T (A^T y) \quad \forall (x,y) \in (\mathbb{R}^n)^2$

On confond $x \in \mathbb{R}^n$ et $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$ $\quad u^T v = \sum_{i=1}^{n} u_i v_i$ produit scalaire canonique

Ex) $E = \mathbb{R}^3$ muni de la structure canonique d'espace euclidien

$q(x) = x_1^2 + x_2^2 + x_3^2$ et $\langle x, y \rangle = \sum_{i=1}^{3} x_i y_i = x^T y$

$$f : x \longmapsto \begin{bmatrix} 2x_1 + 3x_2 \\ x_1 + 2x_2 \\ x_1 + x_2 + x_3 \end{bmatrix} \in \mathcal{L}(E)$$

La matrice de f dans la base canonique (orthonormée) est $A = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

La matrice de f^* dans cette base est $A^T = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Par conséquent, l'adjoint est défini par

$$f^* : x \in \mathbb{R}^3,\ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \longmapsto f(x) = A^T x = \begin{bmatrix} 2x_1 + x_2 + x_3 \\ 3x_1 + 2x_2 + x_3 \\ x_3 \end{bmatrix}$$

Définition : On dit que $f \in \mathcal{L}(E)$ est auto-adjoint si $f^* = f$

→ Si A est la matrice de f dans une base orthonormée, alors A est symétrique.

→ Réciproquement, si A est la matrice de f dans une base orthonormée et si A est symétrique, alors f est auto-adjoint.

Remarque : Une matrice symétrique définit aussi une forme quadratique (dans une certaine base).

Théorème (dit "théorème spectral")

Soit $f \in \mathcal{L}(E)$ auto-adjoint.

Alors f est diagonalisable dans une base orthonormée

$\Leftrightarrow$ il existe une base orthonormée dans laquelle la matrice de f est diagonale

Corollaire : Toute matrice symétrique réelle est diagonalisable dans une base orthonormée.

$\Leftrightarrow$ Pour toute matrice $A \in \mathbb{R}^{n \times n}$, $A^T = A$, il existe $P \in \mathbb{R}^{n\times n}$ orthogonale $(PP^T = P^TP = I_n)$

telle que $A = P\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}P^T$ avec $\lambda_1, \dots, \lambda_n$ les valeurs propres réelles de A

$\rightarrow$ les colonnes de P forment une base orthonormée de vecteurs propres de A.

Lemme 1 : Pour tout $f \in \mathcal{L}(E)$ auto-adjoint, le polynôme caractéristique de f est scindé sur $\mathbb{R}$ (que des racines réelles)

$\Longleftrightarrow$ Toute matrice $A \in \mathbb{R}^{n\times n}$ symétrique ne possède que des valeurs propres réelles (pas forcément distinctes)

Preuve (dans le cas matriciel)

Supposons que A ait une valeur propre complexe. $\lambda \in \mathbb{C}$

Alors $\exists x \in \mathbb{C}^n$, $x \neq 0_{\mathbb{C}^n}$ tel que $Ax = \lambda x$

$\lambda = a+ib$
$(a,b) \in \mathbb{R}^2$
$\bar{\lambda} = a - ib$

Dans ce cas, $\bar{\lambda}$ (conjugué de λ) est aussi valeur propre de A et on a $A\bar{x} = \bar{\lambda}\bar{x}$

(admis, provient de la définition du polynôme caractéristique)

$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{bmatrix}$

On va montrer que $\lambda = \bar{\lambda} \Longleftrightarrow \lambda \in \mathbb{R}$

D'une part, $(Ax)^T\bar{x} = (\lambda x)^T \bar{x} = \lambda x^T \bar{x}$

D'autre part, $(Ax)^T\bar{x} \underset{\text{adjoint}}{=} x^T(A^T\bar{x}) \underset{A \text{ symétrique}}{=} x^T A \bar{x} = x^T(\bar{\lambda}\bar{x})$

$= \bar{\lambda} x^T \bar{x}$

$(AB)^T = B^T A^T$

On en déduit que $\lambda x^T \bar{x} = \bar{\lambda} x^T \bar{x}$

$a+ib \neq 0 \Leftrightarrow (a+ib)(a-ib) = a^2+b^2 \neq 0$

Comme $x \neq 0_{\mathbb{C}^n}$, on a $x^T \bar{x} \neq 0$,

d'où $\lambda = \bar{\lambda} \Leftrightarrow \lambda \in \mathbb{R}$

Lemme 2 Soit $f \in \mathcal{L}(E)$ auto-adjoint.

Alors les sous-espaces propres de f sont orthogonaux 2 à 2.

Preuve Soient λ et μ deux valeurs propre (réelles) distinctes de f, et soient E_λ et E_μ les sous-espaces propres associés

$E_\lambda = \{ x \in E \mid f(x) = \lambda x \}$ (vecteurs propres + vecteur nul)

$E_\mu = \{ y \in E \mid f(y) = \mu y \}$

Pour tous $(x,y) \in E_\lambda \times E_\mu$,

$\langle f(x), y \rangle = \langle \lambda x, y \rangle$ car x est vecteur propre

$\langle f(x), y \rangle = \langle x, f(y) \rangle$ car f est auto-adjoint

$\langle x, f(y) \rangle = \langle x, \mu y \rangle$ car y est vecteur propre

On en déduit que

$\lambda \langle x,y \rangle = \langle \lambda x, y \rangle = \langle x, \mu y \rangle = \mu \langle x,y \rangle$

$\Leftrightarrow (\lambda - \mu) \langle x,y \rangle = 0$

Comme $\lambda \neq \mu$ par hypothèse, on en déduit que $\langle x,y \rangle = 0$ $\forall x \in E_\lambda, \forall y \in E_\mu$

et donc E_λ et E_μ sont des sous-espaces orthogonaux.

Démonstration du théorème (par récurrence sur $n = \dim(E)$)

$\underline{n=1}$: Toute matrice de taille 1×1 est scalaire, donc diagonalisable dans n'importe quelle base

M scalaire $\Leftrightarrow M = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

Supposons que le théorème soit vrai en dimension n. Montrons qu'il est vrai en dimension $n+1$.

Soit $f \in \mathcal{L}(E)$ auto-adjoint, $\dim(E) = n+1$

D'après le lemme 1, f possède des valeurs propres réelles

Soit $\lambda \in \mathbb{R}$ une valeur propre de f.

Si λ est la seule valeur propre de f, alors f se représente par une matrice scalaire $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$ dans n'importe quelle base et on a donc le résultat.

Plus généralement, soit $x \in F$ un vecteur propre de f associé à λ. Alors si on note $F = \mathrm{vect}\{x\} = \{u \in E \mid u = \alpha x, \alpha \in \mathbb{R}\}$

on sait que $E = F \oplus F^\perp$

Soit g la restriction de f à $F^\perp$

$$g : \begin{array}{l} F^\perp \longrightarrow F^\perp \\ y \longmapsto g(y) = f(y) \end{array}$$

Alors $\forall (y,z) \in (F^\perp)^2$, $\langle g(y), z\rangle = \langle f(y), z\rangle = \langle y, f(z)\rangle = \langle y, g(z)\rangle$

Donc g est auto-adjoint sur un espace de dimension n.

Par l'hypothèse de récurrence, il existe une base orthonormée de vecteurs propres de g, qu'on note $(v_1, \dots, v_n)$

En posant $v_{n+1} = \frac{x}{\|x\|}$ ($\|x\| \neq 0$ car x vecteur propre),

Norme induite par le produit scalaire $\|x\| = \sqrt{\langle x, x \rangle}$

on obtient une base orthonormée de E $(v_1, \dots, v_n, v_{n+1})$

$$\forall\, 1 \leq i \leq j \leq n,\ \langle v_i, v_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} \qquad (v_1, \dots, v_n) \text{ orthonormée}$$

$$\forall\, i \in \{1, \dots, n\},\ \langle v_{n+1}, v_i \rangle = 0 \quad \text{car } v_i \in F^\perp \text{ et } v_{n+1} \in F$$

$$\langle v_{n+1}, v_{n+1} \rangle = \left\langle \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle = \frac{\langle x, x \rangle}{\|x\|^2} = 1$$

De plus, tous les v_i sont des vecteurs propres

$$\forall\, i \in \{1, \dots, n\},\ f(v_i) = g(v_i) = \lambda_i v_i$$ car v_i vecteur propre de g par hypothèse de récurrence

et $$f(v_{n+1}) = f\left(\frac{x}{\|x\|}\right) = \frac{1}{\|x\|} f(x) = \lambda \frac{x}{\|x\|} = \lambda v_{n+1}$$ car x vecteur propre de f.

On en conclut qu'il existe bien une base orthonormée de vecteurs propres pour f.

Corollaire Soit $A \in \mathbb{R}^{n \times n}$ symétrique. Alors

- Si $A \succeq 0$ (A semi-définie positive), alors A possède n valeurs propres (réelles) positives ou nulles.
- Si $A \succ 0$ (A définie positive), alors A possède n

valeurs propres strictement positives.

- Si $A \preceq 0$ (A semi-définie négative), A possède n valeurs propres négatives ou nulles
- Si $A \prec 0$ (A définie négative), A possède n valeurs propres strictement négatives.

Application (Bonus !)

Soit $f : \mathbb{R} \to \mathbb{R}$ deux fois dérivable.

Soit $x \in \mathbb{R}$ tel que $f''(x) = 0$

Alors si $f''(x) > 0$, x est un minimum de f.

Si on considère $f : \mathbb{R}^n \to \mathbb{R}$, alors

$\forall x \in \mathbb{R}^n$ tel que $\frac{\partial f}{\partial x_i}(x) = 0 \quad \forall i = 1..n$

et tel que $\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right] \in \mathbb{R}^{n \times n}$ est

définie positive, x est un minimum de f.