Algèbre linéaire (partie 3)

Vincent Audigier

CNAM, Paris

STA101

Application linéaire

Une application f de $\mathbb{R}^I$ dans $\mathbb{R}^J$ est dite **linéaire** si elle vérifie pour tous les vecteurs u et v de $\mathbb{R}^I$ et tout scalaire k

$$f(u+v) = f(u) + f(v)$$
$$f(ku) = kf(u)$$

Quand f est à valeurs dans $\mathbb{R}^I$, on parle d'**endomorphisme**

Représentation matricielle

Soit $f : \mathbb{R}^I \to \mathbb{R}^J$ une application linéaire. On associe à $\mathbb{R}^I$ et $\mathbb{R}^J$ les bases $B_{\mathbb{R}^I} = (u_1, \ldots, u_I)$ et $B_{\mathbb{R}^J} = (v_1, \ldots, v_J)$. On peut représenter f sous la forme d'une matrice dont les colonnes sont les vecteurs $f(u_i)$ (pour $1 \leq i \leq I$) exprimés dans la base $B_{\mathbb{R}^J}$.

Ex : $J = I = 2$,
$B_{\mathbb{R}^I} = B_{\mathbb{R}^J} =$ bases canoniques

$$\begin{array}{ccc} f : \mathbb{R}^I & \to & \mathbb{R}^J \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto & \begin{pmatrix} 3x_1 \\ 2x_2 \end{pmatrix} \end{array}$$

$$Mat(f; B_{\mathbb{R}^I}, B_{\mathbb{R}^J}) = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

f peut être représentée par plusieurs matrices, mais cette représentation est unique pour 2 bases $B_{\mathbb{R}^I}, B_{\mathbb{R}^J}$ données.

Représentation matricielle

Soit $f : \mathbb{R}^I \to \mathbb{R}^J$ une application linéaire. On associe à $\mathbb{R}^I$ et $\mathbb{R}^J$ les bases $B_{\mathbb{R}^I} = (u_1, \ldots, u_I)$ et $B_{\mathbb{R}^J} = (v_1, \ldots, v_J)$. On peut représenter f sous la forme d'une matrice dont les colonnes sont les vecteurs $f(u_i)$ (pour $1 \leq i \leq I$) exprimés dans la base $B_{\mathbb{R}^J}$.

Ex : $J = I = 2$,
$B_{\mathbb{R}^I} = B_{\mathbb{R}^J} =$ bases canoniques

$$\begin{array}{rcl} f : \mathbb{R}^I & \to & \mathbb{R}^J \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto & \begin{pmatrix} 3x_1 \\ 2x_2 \end{pmatrix} \end{array}$$

$$Mat(f; B_{\mathbb{R}^I}, B_{\mathbb{R}^J}) = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

Soit $C = \left\{ \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \right\}$
une autre base de $\mathbb{R}^J$

$$Mat(f; B_{\mathbb{R}^I}, C) = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}$$

f peut être représentée par plusieurs matrices, mais cette représentation est unique pour 2 bases $B_{\mathbb{R}^I}, B_{\mathbb{R}^J}$ données.

Image d'un vecteur

Pour obtenir l'image dans la base $B_{\mathbb{R}^J}$ d'un vecteur $\begin{pmatrix} x_1 \\ \vdots \\ x_I \end{pmatrix}$ exprimé dans la base $B_{\mathbb{R}^I}$, il suffit de faire le produit

$$Mat(f; B_{\mathbb{R}^I}, B_{\mathbb{R}^J}) \begin{pmatrix} x_1 \\ \vdots \\ x_I \end{pmatrix}$$

Ex : $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ donne l'image par f du vecteur exprimé dans la base canonique de $\mathbb{R}^J$.

Intérêt de la représentation diagonale

Il est généralement préférable d'utiliser une représentation dans laquelle la matrice est diagonale, car les opérations sur les matrices diagonales sont plus faciles.

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_I \end{pmatrix}$$

$$D^{-1} = \begin{pmatrix} 1/d_1 & & \\ & \ddots & \\ & & 1/d_I \end{pmatrix}, \; D^p = \begin{pmatrix} d_1^p & & \\ & \ddots & \\ & & d_I^p \end{pmatrix}$$

Image d'une application linéaire

Soit f une application linéaire de $\mathbb{R}^I$ dans $\mathbb{R}^J$, on appelle **image** de f, l'ensemble des éléments de $\mathbb{R}^J$ qui s'écrivent $f(u)$ avec u dans $\mathbb{R}^I$.
Si A est la représentation matricielle de f, alors

$$Im(A) = \{x \in \mathbb{R}^J \text{ tel qu'il existe } u \in \mathbb{R}^I \text{ vérifiant } Au = x\}$$

- C'est l'espace vectoriel engendré par les colonnes de A
- C'est un sous-espace vectoriel de $\mathbb{R}^J$
- $dim(Im(A)) = Rang(A)$

Noyau d'une application linéaire

Le noyau d'une application linéaire $f : \mathbb{R}^I \to \mathbb{R}^J$ est l'ensemble des éléments de $\mathbb{R}^I$ tel que $f(u) = 0$

$$Ker(A) = \{u \in \mathbb{R}^I \text{ tel que } Au = 0\}$$

- C'est un sous-espace vectoriel de $\mathbb{R}^I$
- $dim(Ker(A)) = dim(\mathbb{R}^I) - Rang(A)$ (théorème du rang)

Vecteurs propres, valeurs propres, sous espaces propres

Un vecteur v de $\mathbb{R}^l$ est dit **vecteur propre** d'un endomorphisme f de $\mathbb{R}^l$, si v est non nul et $f(v)$ est colinéaire à v.

Dans ce cas, le scalaire λ tel que $f(v) = \lambda v$ est appelé **valeur propre** de f.

On appelle **sous-espace propre** associé à λ l'ensemble des vecteurs v tels que $f(v) = \lambda v$, c'est le noyau de $f - \lambda I_l$

Endomorphisme diagonalisable (1)

Un endomorphisme f est diagonalisable, s'il existe une base B de $\mathbb{R}^l$ dans laquelle la matrice de f est diagonale

Autrement dit, f est diagonalisable s'il existe une base formée de vecteurs propres de f, car si $f(v_1) = \lambda_1 v_1$, $f(v_2) = \lambda_2 v_2$,...

alors $Mat(f, B, B) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_l \end{pmatrix}$.

Endomorphisme diagonalisable (2)

Tous les endomorphismes ne sont pas diagonalisables.

Théorème : Si le polynôme caractéristique de f admet n racines distinctes, alors f est diagonalisable.

On appelle **polynôme caractéristique** de f,

$$\chi_f(\lambda) = det(f - \lambda I_I)$$

On note n le degré de ce polynôme en λ. Les valeurs propres de f sont racines de $\chi_f(\lambda)$

Exemple

Soit

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

la matrice représentant f dans la base canonique.

$$\begin{aligned}
\chi_f(\lambda) &= det(A - \lambda I_3) \\
&= det \begin{pmatrix} -\lambda & 2 & -1 \\ 3 & -2-\lambda & 0 \\ -2 & 2 & 1-\lambda \end{pmatrix} \\
&= -\lambda(-2-\lambda)(1-\lambda) - 3(2(1-\lambda)+2) - 2(-2-\lambda) \\
&= -\lambda^3 - \lambda^2 + 10\lambda - 8 \\
&= -(\lambda-1)(\lambda-2)(\lambda+4)
\end{aligned}$$

3 racines distinctes, polynôme d'ordre 3 $\Rightarrow$ diagonalisable

Exemple

Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$ et $\lambda_3 = -4$. On cherche les vecteurs propres v_1, v_2, v_3 associés. Par définition $f(v_1) = \lambda_1 v_1$ donc on résout en x, y, z

$$\begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

ce qui est équivaut à résoudre

$$\begin{pmatrix} -x + 2y - z \\ 3x - 3y \\ -2x + 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

On trouve $x = y = z$. Un vecteur propre est

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Exemple

On procède de même pour les deux autres valeurs propres. On trouve

$$v_2 = \begin{pmatrix} 4 \\ 3 \\ -2 \end{pmatrix} \text{ et } v_3 = \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}$$

Dans la base (v_1, v_2, v_3), la matrice A est donc diagonalisable et s'écrit $diag(1, 2, -4)$

Matrice de passage

Soit f est un endomorphisme de $\mathbb{R}^l$ et A sa représentation matricielle dans la base canonique. Soit D sa représentation matricielle dans une base de vecteurs propres.
Ex :

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix} \text{ et D = diag(1,2,-4)}$$

L'image d'un vecteur e par f s'obtient selon le produit Ae dans la base canonique ou selon $PDP^{-1}e$ où P est la matrice formée des vecteurs propres. P est appelée **matrice de passage** de la base de vecteurs propres à la base canonique.

P transforme un vecteur exprimé selon une nouvelle base en un vecteur d'anciennes coordonnées

Exemple

Dans l'exemple précédent, la matrice de passage est $P = \begin{pmatrix} 1 & 4 & 2 \\ 1 & 3 & -3 \\ 1 & -2 & 2 \end{pmatrix}$

On peut vérifier que $P^{-1} = \begin{pmatrix} 0 & 0.4 & 0.6 \\ 1/6 & 0 & -1/6 \\ 1/6 & -0.2 & 1/30 \end{pmatrix}$. Soit $e = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

On a

$$Ae = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}$$

$$\begin{aligned} PDP^{-1}e &= \begin{pmatrix} 1 & 4 & 2 \\ 1 & 3 & -3 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} 0 & 0.4 & 0.6 \\ 1/6 & 0 & -1/6 \\ 1/6 & -0.2 & 1/30 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} \end{aligned}$$

Application bilinéaire

Une application f de $\mathbb{R}^I \times \mathbb{R}^I$ dans $\mathbb{R}^J$ est dite **bilinéaire** si elle vérifie pour tous les vecteurs u, v, w de $\mathbb{R}^I$ et tout scalaire k

$$\begin{aligned} f(ku, v) &= kf(u, v) = f(u, kv) \\ f(u + w, v) &= f(u, v) + f(w, v) \\ f(u, v + w) &= f(u, v) + f(u, w) \end{aligned}$$

Représentation matricielle

Si $f : \mathbb{R}^I \times \mathbb{R}^I \to \mathbb{R}^J$ est bilinéaire et $\mathbb{R}^I$ est munie d'une base $B_{\mathbb{R}^I}$, alors il existe une matrice M telle que pour tous les couples (u, v) vecteurs colonnes de $\mathbb{R}^I$

$$f(u, v) = u^t M v$$

Ex : $< u, v >= \sum_{i=1}^{I} u_i v_i$, dans la base canonique.

$$< u, v >= (u_1, ..., u_I) \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_I \end{pmatrix}$$

Produit scalaire

Une matrice M définit un produit scalaire si elle est

- symétrique : $M = M^t$
- définie : $u'Mu = 0 \Leftrightarrow u = 0$
- positive : $u'Mu \geq 0$.

Une matrice est définie positive si et seulement si ses valeurs propres sont toutes strictement positives. Si certaines sont nulles, alors la matrice est dite semi-définie positive.

Ex :

$$< u, v >= (u_1, ..., u_l) \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_l \end{pmatrix}$$

est clairement un produit scalaire
M est appelée **métrique** de l'espace, elle **définie les distances** entre les vecteurs

Projection orthogonale sur un sous-espace

Soit $\mathcal{F}$ un sous-espace vectoriel de $\mathbb{R}^I$, la matrice de projection M-orthogonale sur $\mathcal{F}$ est la matrice $P_{\mathcal{F}}$ telle que pour tout e dans $\mathbb{R}^I$, $P_{\mathcal{F}}e \in \mathcal{F}$ et $< P_{\mathcal{F}}e; e - P_{\mathcal{F}}e >_M = 0$.

Propriétés

- $P_{\mathcal{F}}^2 = P_{\mathcal{F}}$ (matrice idempotente)
- $P'_{\mathcal{F}}M = MP_{\mathcal{F}}$ (M-symétrie)
- les valeurs propres de $P_{\mathcal{F}}$ sont 1 ou 0

Projection orthogonale sur un sous-espace

Soit $\mathcal{F}$ un sous-espace vectoriel de $\mathbb{R}^I$, la matrice de projection M-orthogonale sur $\mathcal{F}$ est la matrice $P_{\mathcal{F}}$ telle que pour tout e dans $\mathbb{R}^I$, $P_{\mathcal{F}}e \in \mathcal{F}$ et $< P_{\mathcal{F}}e; e - P_{\mathcal{F}}e >_M = 0$.

Propriétés

- $P_{\mathcal{F}}^2 = P_{\mathcal{F}}$ (matrice idempotente)
- $P'_{\mathcal{F}}M = MP_{\mathcal{F}}$ (M-symétrie)
- les valeurs propres de $P_{\mathcal{F}}$ sont 1 ou 0

Soit $e \in \mathbb{R}^I$, on a

$< P_{\mathcal{F}}e, e >_M = < e, P_{\mathcal{F}}e >_M$ pour tout e

$$\begin{aligned} &\Rightarrow \quad (P_{\mathcal{F}}e)^t Me = e^t MP_{\mathcal{F}}e \text{ pour tout } e \\ &\Rightarrow \quad e^t P_{\mathcal{F}}^t Me = e^t MP_{\mathcal{F}}e \text{ pour tout } e \\ &\Rightarrow \quad P_{\mathcal{F}}^t M = MP_{\mathcal{F}} \end{aligned}$$

Expression matricielle

Soit $X = (x_1, x_2, \ldots, x_d)$ une base de $\mathcal{F}$, alors le projecteur orthogonal sur $\mathcal{F}$ pour le produit scalaire $<;>_M$ est donné par

$$P_{\mathcal{F}} = X(X'MX)^{-1}X'M$$

Preuve

- Soit e dans $\mathbb{R}^l$, $< P_{\mathcal{F}}e; e - P_{\mathcal{F}}e >_M = 0$
- Comme $P_{\mathcal{F}}e \in \mathcal{F}$, il existe $b \in \mathbb{R}^d$ tel que $P_{\mathcal{F}}e = Xb$ où $X = (X_1, ...X_d)$ est une base de $\mathcal{F}$, X_i $(1 \leq i \leq d)$ est un vecteur colonne de la base et $1 \leq d \leq l$ la dimension de $\mathcal{F}$
- Donc $< Xb; e - Xb >_M = 0$
- De plus, pour tout vecteur X_i de la base, $< X_i; e - Xb >_M = 0$ (car $X_i \in \mathcal{F}$)

$$\begin{aligned}
&\Rightarrow \quad X_i^t M(e - Xb) = 0 \text{ pour tout } i \\
&\Rightarrow \quad X^t M(e - Xb) = 0 \text{ (ici 0 désigne le vecteur nul)} \\
&\Rightarrow \quad X^t Me = X^t MXb \\
&\Rightarrow \quad b = (X^t MX)^{-1} X^t Me
\end{aligned}$$

- d'où $P_{\mathcal{F}}e = Xb = X(X^t MX)^{-1} X^t Me$

Gradient

Soit

$$\begin{array}{rcl} f : \mathbb{R}^l & \rightarrow & \mathbb{R} \\ \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix} & \mapsto & A \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix} \end{array}$$

on appelle **gradient** de f le vecteur des dérivées partielles

$$grad\ f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_l} \right)$$

Ici,

$$grad\ f = \frac{\partial Ax}{\partial x} = \left(\frac{\partial \sum_{i=1}^{l} a_i x_i}{\partial x_1}, \ldots, \frac{\partial \sum_{i=1}^{l} a_i x_i}{\partial x_l} \right) = A$$

Dérivation d'une forme quadratique

Soit A une matrice carrée, on appelle **forme quadratique** associée à A, la fonction g définie par

$$\begin{aligned} g : \mathbb{R}^l &\rightarrow \mathbb{R} \\ \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix} &\mapsto \begin{pmatrix} x_1 & \dots & x_l \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix} \end{aligned}$$

$$grad\ g = \frac{\partial x' A x}{\partial x} = Ax + A'x$$

En particulier, pour A symétrique,

$$grad\ g = \frac{\partial x' A x}{\partial x} = 2Ax$$

Maximisation d'un quotient de formes quadratiques

Soient A et B deux matrices symétriques de même taille, B inversible, alors le rapport $\frac{x'Ax}{x'Bx}$ est maximal pour u vecteur propre de $B^{-1}A$ associée à sa plus grande valeur propre

Preuve

On dérive $\frac{x'Ax}{x'Bx}$ et on recherche le vecteur x qui annule sa dérivée :

$$\begin{aligned}
\frac{2Axx'Bx - x'Ax2Bx}{(x'Bx)^2} &= 0 \\
\Rightarrow \quad 2Axx'Bx &= x'Ax2Bx \\
\Rightarrow \quad x'Bx2Ax &= x'Ax2Bx \\
\Rightarrow \quad x'BxB^{-1}Ax &= x'AxB^{-1}Bx \\
\Rightarrow \quad B^{-1}Ax &= \frac{x'Ax}{x'Bx}x
\end{aligned}$$

Donc x est nécessairement vecteur propre de $B^{-1}A$ associé à la valeur propre $\frac{x'Ax}{x'Bx}$

Or maximiser $\frac{x'Ax}{x'Bx}$ c'est maximiser la valeur propre. Donc le vecteur recherché est le vecteur propre associé à la plus grande valeur propre.