

DM n°06 — Projection orthogonale et méthode des moindres carrés

Orthogonalité et projection orthogonale · Moindres carrés et équations normales · Matrices de Gram · Ajustement affine · Approximation de fonctions

Trois mesures ne sont pas toujours exactement compatibles avec le modèle que l'on souhaite leur imposer. Il faut alors préciser ce que signifie « approcher au mieux ». À partir d'un ajustement élémentaire, on fera apparaître une propriété géométrique du résidu. Cette propriété conduira à la projection orthogonale, puis aux équations normales des moindres carrés, avant d'être retrouvée dans l'approximation de fonctions.

I Ajuster trois points

On considère les trois points

$$A(-1, 1), \quad B(0, 1), \quad C(1, 2).$$

On cherche à les approcher par le graphe d'une fonction affine

$$x \mapsto \alpha x + \beta.$$

- 1 Montrer qu'il n'existe aucun couple $(\alpha, \beta) \in \mathbb{R}^2$ tel que les trois points A, B, C appartiennent au graphe de $x \mapsto \alpha x + \beta$.
- 2 On mesure l'erreur d'un couple (α, β) par

$$S(\alpha, \beta) = (-\alpha + \beta - 1)^2 + (\beta - 1)^2 + (\alpha + \beta - 2)^2.$$

Déterminer le minimum de S sur $\mathbb{R}^2$ et le couple (α, β) pour lequel il est atteint.

- 3 Dans $\mathbb{R}^3$, muni du produit scalaire usuel, on pose

$$X = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad F = \text{Vect}(X, U).$$

- (a) Calculer $\alpha X + \beta U$ et montrer que

$$S(\alpha, \beta) = \|Y - (\alpha X + \beta U)\|^2.$$

En déduire que minimiser S revient à chercher, parmi les vecteurs de F , celui qui est le plus proche de Y .

- (b) Montrer que X et U sont linéairement indépendants. En déduire la dimension de F .
- (c) Pour le couple obtenu à la question 2, calculer

$$P = \alpha X + \beta U, \quad R = Y - P,$$

puis montrer que R est orthogonal à F .

4 Soit $Z \in F$.

- (a) Montrer que $Y - P$ et $P - Z$ sont orthogonaux. En déduire une relation entre $\|Y - Z\|^2$, $\|Y - P\|^2$ et $\|P - Z\|^2$.
- (b) Comparer $\|Y - Z\|$ et $\|Y - P\|$ lorsque Z parcourt F , et déterminer le cas d'égalité. Que peut-on en conclure pour le minimum de S ?

II Projection orthogonale

Dans cette partie, E est un espace vectoriel réel de dimension finie, muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. On note

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Une partie $F \subset E$ est un sous-espace vectoriel de E lorsque

$$0 \in F \quad \text{et} \quad \forall (u, v) \in F^2, \forall (\lambda, \mu) \in \mathbb{R}^2, \quad \lambda u + \mu v \in F.$$

Autrement dit, F contient le vecteur nul et est stable par combinaisons linéaires.

Soit désormais F un sous-espace vectoriel de E . On définit

$$F^\perp = \{x \in E ; \forall y \in F, \langle x, y \rangle = 0\}.$$

5 Montrer que $F^\perp$ est un sous-espace vectoriel de E , puis que

$$F \cap F^\perp = \{0\}.$$

6 On admet que tout sous-espace vectoriel de dimension finie possède une base orthonormée. Soit $(e_1, \dots, e_r)$ une base orthonormée de F . Pour $x \in E$, on pose

$$q = \sum_{i=1}^r \langle x, e_i \rangle e_i.$$

- (a) Montrer que $q \in F$ et que $x - q \in F^\perp$.
- (b) Montrer que toute décomposition $x = p + r'$, avec $p \in F$ et $r' \in F^\perp$, est unique. En déduire que, pour tout $x \in E$, il existe un unique $p \in F$ tel que $x - p \in F^\perp$. On note ce vecteur $P_F(x)$ et on l'appelle projeté orthogonal de x sur F .

7 Soit $x \in E$ et $p = P_F(x)$.

- (a) Pour tout $z \in F$, établir

$$\|x - z\|^2 = \|x - p\|^2 + \|p - z\|^2,$$

puis montrer que p est l'unique élément de F minimisant la distance à x .

- (b) Réciproquement, soit $p \in F$ tel que

$$\|x - p\| \leq \|x - z\| \quad \text{pour tout } z \in F.$$

Montrer que $x - p \in F^\perp$.

- (c) En déduire que, pour $p \in F$, les trois assertions suivantes sont équivalentes :

- $p = P_F(x)$;
- $x - p \in F^\perp$;
- p est l'unique élément de F minimisant la distance à x .

- 8 (a) Montrer que $P_F : E \rightarrow E$ est linéaire et que $P_F^2 = P_F$. Déterminer $\text{Im}(P_F)$ et $\ker(P_F)$.
- (b) Compléter une base orthonormée de F en une base orthonormée de E , puis donner la matrice de P_F dans cette base.
- (c) Montrer que, pour tous $x, y \in E$,

$$\langle P_F(x), y \rangle = \langle x, P_F(y) \rangle.$$

Montrer ensuite que $\|P_F(x)\| \leq \|x\|$ et caractériser le cas d'égalité.

III Le problème des moindres carrés

On munit désormais $\mathbb{R}^m$ et $\mathbb{R}^n$ de leurs produits scalaires usuels. Soient

$$A \in \mathcal{M}_{m,n}(\mathbb{R}), \quad y \in \mathbb{R}^m.$$

On cherche à minimiser sur $\mathbb{R}^n$ la fonction

$$\Phi(x) = \|Ax - y\|^2.$$

- 9 (a) Identifier l'ensemble des vecteurs Ax , lorsque x parcourt $\mathbb{R}^n$, et reformuler la minimisation de Φ comme un problème de meilleure approximation dans $\mathbb{R}^m$.
- (b) En utilisant la partie II, montrer que Φ admet toujours un minimum et que le vecteur ajusté Ax est le même pour tous les minimiseurs.
- (c) Si x_0 est un minimiseur, montrer que l'ensemble des minimiseurs est

$$x_0 + \ker A.$$

En déduire une condition nécessaire et suffisante d'unicité du paramètre minimisant x .

- 10 On rappelle que A^T désigne la transposée de A .

- (a) Montrer que, pour tous $x \in \mathbb{R}^n$ et $z \in \mathbb{R}^m$,

$$\langle Ax, z \rangle = \langle x, A^T z \rangle.$$

- (b) En déduire

$$(\text{Im } A)^\perp = \ker(A^T).$$

- (c) Montrer que x minimise Φ si et seulement si

$$A^T(Ax - y) = 0,$$

c'est-à-dire si et seulement si

$$A^T Ax = A^T y.$$

Ces relations sont appelées équations normales du problème des moindres carrés.

11 On note $a_1, \dots, a_n$ les colonnes de A .

(a) Montrer que

$$(A^T A)_{ij} = \langle a_i, a_j \rangle.$$

Ainsi, $A^T A$ est la matrice de Gram de la famille $(a_1, \dots, a_n)$.

(b) Montrer que

$$\ker(A^T A) = \ker A.$$

(c) En déduire l'équivalence entre les trois propriétés suivantes : $A^T A$ est inversible ; $\ker A = \{0\}$; les colonnes de A sont linéairement indépendantes.

(d) Sous ces hypothèses, déterminer l'unique minimiseur $\hat{x}$ et en déduire la matrice de la projection orthogonale sur $\text{Im } A$.

12 On considère

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(a) Déterminer tous les vecteurs $x \in \mathbb{R}^2$ minimisant $\|Ax - y\|^2$.

(b) Déterminer le vecteur ajusté Ax et le résidu $y - Ax$ pour un minimiseur.

(c) Écrire les équations normales. Vérifier qu'elles caractérisent bien tous les minimiseurs trouvés en (a), bien que $A^T A$ ne soit pas inversible.

IV Ajuster une droite à N observations

Soient $N \geq 2$ et $(x_1, y_1), \dots, (x_N, y_N)$ des points du plan. On suppose que les x_i ne sont pas tous égaux. On cherche la fonction affine $x \mapsto ax + b$ qui minimise

$$\sum_{i=1}^N (y_i - ax_i - b)^2.$$

On pose

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}.$$

13 (a) Construire une matrice $A \in \mathcal{M}_{N,2}(\mathbb{R})$ telle que

$$A \begin{pmatrix} a \\ b \end{pmatrix} = aX + b\mathbf{1}.$$

(b) Montrer que les deux colonnes de A sont linéairement indépendantes si et seulement si les x_i ne sont pas tous égaux.

(c) En déduire l'existence et l'unicité du couple (a, b) recherché.

14 On note

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i.$$

(a) Écrire les deux équations normales sous la forme de sommes portant sur les x_i et les y_i .

(b) Montrer que

$$b = \bar{y} - a\bar{x}.$$

(c) En déduire

$$a = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}, \quad b = \bar{y} - a\bar{x}.$$

(d) Montrer que la droite ajustée passe par le point $(\bar{x}, \bar{y})$.

15 On pose $X_c = X - \bar{x} \mathbf{1}$.

(a) Montrer que $X_c \perp \mathbf{1}$ et que

$$\text{Vect}(X, \mathbf{1}) = \text{Vect}(X_c, \mathbf{1}).$$

(b) À l'aide de $X_c \perp \mathbf{1}$, écrire directement la projection orthogonale de Y sur $\text{Vect}(X_c, \mathbf{1})$. En déduire les coefficients a et b obtenus à la question 14.

(c) Appliquer ce calcul aux trois points de la partie I pour retrouver le résultat de la question 2. Expliquer aussi, à partir de $X \perp U$, pourquoi le développement de $S(\alpha, \beta)$ ne contient aucun terme en $\alpha\beta$.

V Approximer des fonctions

On considère l'espace vectoriel

$$H = C([-1, 1], \mathbb{R})$$

et, pour $f, g \in H$,

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt.$$

16 (a) Montrer que cette formule définit un produit scalaire sur H .

On note

$$F = \{t \mapsto at + b; (a, b) \in \mathbb{R}^2\}.$$

(b) Montrer que F est de dimension 2. Justifier que, même si H n'est pas de dimension finie, la construction de la question 6 permet de projeter tout élément de H sur F .

- (c) Déterminer la fonction affine $p \in F$ qui minimise

$$\int_{-1}^1 (t^2 - p(t))^2 dt,$$

puis calculer la valeur minimale.

- 17 Soit maintenant H un espace vectoriel réel muni d'un produit scalaire et soit $(\varphi_0, \dots, \varphi_d)$ une famille linéairement indépendante de H . On pose

$$F = \text{Vect}(\varphi_0, \dots, \varphi_d).$$

Pour $f \in H$, on écrit son projeté sur F sous la forme

$$p = \sum_{j=0}^d c_j \varphi_j.$$

- (a) À partir de $f - p \in F^\perp$, montrer que les coefficients c_j vérifient le système

$$Gc = D, \quad G_{ij} = \langle \varphi_i, \varphi_j \rangle, \quad D_i = \langle f, \varphi_i \rangle.$$

- (b) Montrer que, pour tout $c = (c_0, \dots, c_d)^T$,

$$c^T Gc = \left\| \sum_{j=0}^d c_j \varphi_j \right\|^2.$$

En déduire que G est définie positive et inversible.

- 18 (a) Lorsque $(\varphi_0, \dots, \varphi_d)$ est orthonormée, déterminer directement les coefficients du projeté de f .
 (b) En prenant pour les φ_j les colonnes d'une matrice A , retrouver le système des équations normales.
 (c) Dans $C([-1, 1], \mathbb{R})$ muni du produit scalaire de la question 16, prendre $F = \mathbb{R}_d[X]$. Justifier que toute fonction continue possède une unique meilleure approximation polynomiale de degré au plus d pour la norme associée. Montrer enfin qu'une base orthonormée de F permet d'en obtenir les coefficients par de simples produits scalaires, contrairement à la base monomiale $(1, X, \dots, X^d)$.