

CORRIGÉ

PSI · CORRIGÉ

CHAPITRES

DM n°02 — Reconstruire un signal à partir de mesures imparfaites

Polynômes · Matrices et systèmes ·
Algèbre linéaire · Produit scalaire ·
Moindres carrés

Le problème part d'une reconstruction exacte, puis examine ce qui change lorsque les mesures ne sont plus compatibles avec le modèle choisi. Le point central est géométrique : à la meilleure reconstruction, l'erreur restante est orthogonale à toutes les variations que le modèle permet de produire.

I Reconstruire à partir de mesures exactes

1

On écrit

$$s(t) = \alpha + \beta t + \gamma t^2.$$

La mesure en 0 donne immédiatement

$$\alpha = 10.$$

Les deux autres mesures donnent

$$10 - \beta + \gamma = 13, \quad 10 + \beta + \gamma = 11,$$

soit

$$-\beta + \gamma = 3, \quad \beta + \gamma = 1.$$

En additionnant ces deux égalités,

$$2\gamma = 4,$$

donc $\gamma = 2$, puis $\beta = -1$. Ainsi

$$s(t) = 10 - t + 2t^2.$$

2

(a) Supposons

$$A(t_1, t_2, t_3)x = 0, \quad x = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

Le polynôme

$$p(t) = \alpha + \beta t + \gamma t^2$$

s'annule alors en t_1, t_2, t_3 . Ces trois nombres étant distincts et p étant de degré au plus 2, on obtient nécessairement $p = 0$. Ainsi $x = 0$ et l'application $x \mapsto A(t_1, t_2, t_3)x$ est injective.

Elle est linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$. En dimension finie, une application linéaire injective entre deux espaces de même dimension est bijective. Pour trois valeurs mesurées quelconques, il existe donc un unique paramètre x qui les reproduit :

trois mesures prises en trois instants distincts déterminent un unique signal du modèle.

(b) Avec une seule mesure en t_1 , le polynôme non nul

$$q(t) = t - t_1$$

s'annule au point mesuré. Si s convient, alors $s + \lambda q$ convient encore pour tout $\lambda \in \mathbb{R}$.

Avec deux mesures prises en deux instants distincts t_1, t_2 , on peut utiliser

$$q(t) = (t - t_1)(t - t_2).$$

Ce polynôme est non nul, de degré 2, et s'annule aux deux instants. Là encore, si s convient, tous les signaux

$$s + \lambda q$$

conviennent également. Une ou deux mesures ne suffisent donc pas, en général, à assurer l'unicité.

(c) Trois mesures exactes en trois instants distincts ont déjà fixé un unique signal du modèle. Une quatrième mesure parfaitement compatible ne modifie donc pas ce signal : elle apporte une vérification supplémentaire de l'adéquation des données au modèle.

II Quand les mesures deviennent incompatibles

3

Les trois mesures centrales donnent

$$s(-1) = 11, \quad s(0) = 10, \quad s(1) = 13.$$

Comme à la question 1, $s(0) = 10$ impose $\alpha = 10$. Les deux autres égalités donnent

$$-\beta + \gamma = 1, \quad \beta + \gamma = 3.$$

On obtient

$$\beta = 1, \quad \gamma = 2.$$

Le seul polynôme du modèle compatible avec ces trois mesures est donc

$$s(t) = 10 + t + 2t^2.$$

Mais

$$s(2) = 10 + 2 + 8 = 20 \neq 15.$$

Il ne peut donc pas reproduire la mesure en $t = 2$. Par conséquent

$$Ax = y \text{ n'a pas de solution.}$$

4

(a) Une somme algébrique peut masquer de grandes erreurs. Par exemple, deux composantes du résidu peuvent être respectivement égales à 5 et -5 : leur somme est nulle alors que les deux écarts sont importants.

La somme des composantes ne mesure donc pas correctement la taille globale du résidu.

(b) Les composantes de $Ax - y$ sont

$$\alpha - 2\beta + 4\gamma - 21,$$

$$\alpha - \beta + \gamma - 11,$$

$$\alpha - 10,$$

$$\alpha + \beta + \gamma - 13,$$

$$\alpha + 2\beta + 4\gamma - 15.$$

Ainsi

$$\begin{aligned} J(x) &= (\alpha - 2\beta + 4\gamma - 21)^2 + (\alpha - \beta + \gamma - 11)^2 \\ &\quad + (\alpha - 10)^2 + (\alpha + \beta + \gamma - 13)^2 + (\alpha + 2\beta + 4\gamma - 15)^2. \end{aligned}$$

Chaque écart intervient par son carré : son signe ne permet plus de compenser un autre écart. Minimiser J revient donc à chercher le signal dont les cinq valeurs prédites sont, au sens de la distance euclidienne dans $\mathbb{R}^5$, aussi proches que possible des cinq mesures.

III Caractériser la meilleure reconstruction

5

(a) Fixons $h \in \mathbb{R}^3$ et considérons les paramètres

$$\hat{x} + \lambda h, \quad \lambda \in \mathbb{R}.$$

Comme $\hat{x}$ minimise J , la fonction

$$\lambda \mapsto J(\hat{x} + \lambda h)$$

atteint son minimum en $\lambda = 0$.

Or

$$r(\hat{x} + \lambda h) = A(\hat{x} + \lambda h) - y = r(\hat{x}) + \lambda Ah.$$

Donc

$$\begin{aligned} J(\hat{x} + \lambda h) &= \|r(\hat{x}) + \lambda Ah\|^2 \\ &= J(\hat{x}) + 2\lambda \langle r(\hat{x}), Ah \rangle + \lambda^2 \|Ah\|^2. \end{aligned}$$

Si $Ah = 0$, l'égalité recherchée est immédiate. Si $Ah \neq 0$, le membre de droite est un trinôme du second degré de coefficient dominant strictement positif. Pour qu'il soit minimal en $\lambda = 0$, son terme linéaire doit être nul. Dans tous les cas,

$$\langle r(\hat{x}), Ah \rangle = 0 \quad \text{pour tout } h \in \mathbb{R}^3.$$

(b) Pour tout $h \in \mathbb{R}^3$,

$$\langle r(\hat{x}), Ah \rangle = h^T A^T r(\hat{x}).$$

Cette quantité étant nulle pour tout h , on a

$$A^T r(\hat{x}) = 0.$$

Comme $r(\hat{x}) = A\hat{x} - y$,

$$A^T(A\hat{x} - y) = 0.$$

Ainsi

$$A^T A \hat{x} = A^T y.$$

Ces relations sont les équations normales du problème.

6

(a) Supposons

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

Pour $h \in \mathbb{R}^3$,

$$\begin{aligned} J(x+h) &= \|Ax - y + Ah\|^2 \\ &= J(x) + 2\langle Ax - y, Ah \rangle + \|Ah\|^2. \end{aligned}$$

Or

$$\langle Ax - y, Ah \rangle = h^T A^T(Ax - y) = 0.$$

Par conséquent

$$J(x+h) = J(x) + \|Ah\|^2 \geq J(x).$$

Le paramètre x minimise donc J .

(b) Si A est injective, alors pour tout $h \neq 0$,

$$Ah \neq 0,$$

donc

$$\|Ah\|^2 > 0.$$

La relation précédente devient

$$J(x+h) > J(x).$$

Ainsi, lorsque A est injective, tout paramètre vérifiant les équations normales est l'unique minimiseur de J .

7

(a) Supposons $Ax = 0$. Le polynôme

$$p(t) = \alpha + \beta t + \gamma t^2$$

s'annule alors aux cinq instants distincts $-2, -1, 0, 1, 2$. Un polynôme de degré au plus 2 possédant cinq racines distinctes est nécessairement nul. Ainsi $x = 0$ et

A est injective.

(b) Soit $B \in \mathcal{M}_{m,n}(\mathbb{R})$ injective. Supposons

$$B^T B u = 0.$$

En multipliant à gauche par u^T ,

$$0 = u^T B^T B u = (Bu)^T (Bu) = \|Bu\|^2.$$

Donc $Bu = 0$. L'injectivité de B impose $u = 0$. Ainsi

$$\ker(B^T B) = \{0\}.$$

La matrice $B^T B$ est carrée de taille n , donc

$B^T B$ est inversible.

(c) Dans notre problème, A est injective. La matrice $A^T A$ est donc inversible, et les équations normales possèdent l'unique solution

$$\hat{x} = (A^T A)^{-1} A^T y.$$

La question 6 montre que cette solution est précisément l'unique minimiseur de J . L'existence et l'unicité de la meilleure reconstruction sont donc établies.

IV Calculer puis interpréter la reconstruction

8

(a) Les symétries des instants $-2, -1, 0, 1, 2$ donnent

$$\sum t_i = 0, \quad \sum t_i^2 = 10, \quad \sum t_i^3 = 0, \quad \sum t_i^4 = 34.$$

On obtient donc

$$A^T A = \begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}.$$

D'autre part,

$$\sum y_i = 70,$$

$$\sum t_i y_i = -2 \times 21 - 11 + 13 + 2 \times 15 = -10,$$

et

$$\sum t_i^2 y_i = 4 \times 21 + 11 + 13 + 4 \times 15 = 168.$$

Ainsi

$$A^T y = \begin{pmatrix} 70 \\ -10 \\ 168 \end{pmatrix}.$$

(b) Les équations normales sont

$$\begin{cases} 5\alpha + 10\gamma = 70, \\ 10\beta = -10, \\ 10\alpha + 34\gamma = 168. \end{cases}$$

On obtient immédiatement

$$\beta = -1.$$

La première équation donne

$$\alpha + 2\gamma = 14.$$

En la combinant avec la troisième,

$$10(14 - 2\gamma) + 34\gamma = 168,$$

donc

$$14\gamma = 28, \quad \gamma = 2, \quad \alpha = 10.$$

Par conséquent

$$\hat{x} = \begin{pmatrix} 10 \\ -1 \\ 2 \end{pmatrix}$$

et

$$\hat{s}(t) = 10 - t + 2t^2.$$

(c) Les valeurs prédites sont

$$A\hat{x} = \begin{pmatrix} 20 \\ 13 \\ 10 \\ 11 \\ 16 \end{pmatrix}.$$

Donc

$$r(\hat{x}) = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

La valeur minimale est

$$J(\hat{x}) = 1 + 4 + 0 + 4 + 1 = \boxed{10}.$$

9

(a) Tout vecteur de $E = \text{Im}(A)$ s'écrit Ah avec $h \in \mathbb{R}^3$. La question 5 donne précisément

$$\langle r(\hat{x}), Ah \rangle = 0$$

pour tout h . Ainsi

$$\boxed{r(\hat{x}) \perp E.}$$

(b) Soit $v \in E$. Comme $A\hat{x} \in E$, il existe $h \in \mathbb{R}^3$ tel que

$$v = A\hat{x} + Ah.$$

Or

$$y - A\hat{x} = -r(\hat{x})$$

est orthogonal à E , donc en particulier à Ah . Par conséquent

$$\begin{aligned} \|y - v\|^2 &= \|y - A\hat{x} - Ah\|^2 \\ &= \|y - A\hat{x}\|^2 + \|Ah\|^2. \end{aligned}$$

On a donc

$$\|y - v\| \geq \|y - A\hat{x}\|.$$

L'égalité n'est possible que si $Ah = 0$, c'est-à-dire si $v = A\hat{x}$. Ainsi $A\hat{x}$ est l'unique point de E le plus proche de y . De plus,

$$y - A\hat{x} \perp E.$$

D'après la définition donnée dans l'énoncé,

$$\boxed{A\hat{x} \text{ est la projection orthogonale de } y \text{ sur } E.}$$

10

Pour

$$x_c = \begin{pmatrix} 10 \\ 1 \\ 2 \end{pmatrix},$$

le signal est

$$s_c(t) = 10 + t + 2t^2.$$

Aux cinq instants,

$$Ax_c = \begin{pmatrix} 16 \\ 11 \\ 10 \\ 13 \\ 20 \end{pmatrix}.$$

Ainsi

$$Ax_c - y = \begin{pmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 5 \end{pmatrix}$$

et

$$J(x_c) = 25 + 25 = 50.$$

Or

$$J(\hat{x}) = 10.$$

Le signal s_c reproduit exactement trois mesures, mais au prix de deux écarts importants sur les mesures extrêmes. Les moindres carrés ne cherchent pas à rendre certaines erreurs nulles : ils répartissent les écarts de façon à minimiser leur somme quadratique sur l'ensemble des données.

V Une formule qui vaut pour tout modèle linéaire injectif

11

(a) Soit

$$B \in \mathcal{M}_{m,n}(\mathbb{R})$$

injective. D'après la question 7(b), la matrice $B^T B$ est inversible.

Posons

$$\hat{u} = (B^T B)^{-1} B^T z.$$

Alors

$$B^T (B\hat{u} - z) = 0.$$

Le raisonnement de la question 6, valable sans changement dans $\mathbb{R}^n$, montre que $\hat{u}$ minimise $\|Bu - z\|^2$. L'injectivité de B assure l'unicité. Ainsi

$$\boxed{\hat{u} = (B^\top B)^{-1} B^\top z.}$$

(b) Par définition,

$$R = (B^\top B)^{-1} B^\top.$$

Donc

$$RB = (B^\top B)^{-1} B^\top B = I_n.$$

Ainsi

$$\boxed{RB = I_n.}$$

Si les données sont exactement de la forme

$$z = Bu_0,$$

alors leur reconstruction vaut

$$Rz = RBu_0 = u_0.$$

La formule retrouve donc exactement le paramètre ayant produit les données.

(c) Posons

$$P = BR.$$

Alors

$$P^2 = BRBR = B(RB)R = BI_n R = BR = P.$$

Donc

$$\boxed{P^2 = P.}$$

Comme $P = BR$, tout vecteur de l'image de P appartient à l'image de B :

$$\text{Im}(P) \subset \text{Im}(B).$$

Réciproquement, si $v \in \text{Im}(B)$, il existe u tel que $v = Bu$. Alors

$$Pv = BRBu = B(RB)u = Bu = v.$$

Ainsi $v \in \text{Im}(P)$, d'où

$$\boxed{\text{Im}(P) = \text{Im}(B).}$$

(d) Pour $z \in \mathbb{R}^m$, on a déjà

$$Pz \in \text{Im}(B).$$

Calculons

$$\begin{aligned} B^T(z - Pz) &= B^Tz - B^TBRz \\ &= B^Tz - B^TB(B^TB)^{-1}B^Tz \\ &= 0. \end{aligned}$$

Pour tout $u \in \mathbb{R}^n$,

$$\langle z - Pz, Bu \rangle = u^T B^T(z - Pz) = 0.$$

Ainsi

$$z - Pz \perp \text{Im}(B).$$

Le vecteur Pz appartient à $\text{Im}(B)$ et l'écart $z - Pz$ lui est orthogonal. Par définition,

Pz est la projection orthogonale de z sur $\text{Im}(B)$.

La formule des moindres carrés n'est donc pas seulement une recette de calcul : elle réalise la projection orthogonale des données sur l'ensemble des valeurs que le modèle peut produire.

SOURCES

Gilbert Strang — Projection Matrices and Least Squares. [Lien direct](#)

NIST/SEMATECH — Linear Least Squares Regression. [Lien direct](#)