J.-P. CROISILLE
24 oct 2012

La méthode des moindres carrés

[M.J.D. Powell : Approximation theory and methods]

C'est la méthode de base en théorie de l'approximation — Remonte à C.F. Gauss.

Problème d'approximation abstrait :

- Un "grand" espace de Hilbert $(\mathcal{B}, (\,,\,)_{\mathcal{B}})$

<u>exemple</u> : Espace des fonctions $L^2(0,1)$, muni de la norme $\|f\|_{L^2(0,1)} = \left(\int_0^1 f^2(x)\,dx\right)^{1/2}$

- Un sous-espace de dimension finie de $(\mathcal{B}, (\,,\,)_{\mathcal{B}})$ noté $\mathcal{A}$, $\mathcal{A} \subset \mathcal{B}$ —

<u>exemple</u> : $\mathcal{A} = \mathcal{P}_n[0,1]$, $(\dim \mathcal{A} = n+1)$

Problème : Pour $f \in \mathcal{B}$ fixé, résoudre le problème de "meilleure approximation" de $f \in \mathcal{B}$ à l'aide d'éléments de $\mathcal{A}$ — On cherche $p^* \in \mathcal{A}$ solution de

$$\min_{p \in \mathcal{A}} \|f - p\|_{\mathcal{B}}$$

<u>Théorème 1</u> : [Moindres carrés dans les espaces de Hilbert — Equation normale]

Soit $(\mathcal{B}, (\,,\,)_{\mathcal{B}})$ un espace de Hilbert et $\mathcal{A} \subset \mathcal{B}$ un sous-espace vect. dimension finie.

(i) Pour tout $f \in \mathcal{B}$, il existe un unique $p^* \in \mathcal{A}$ solution du problème "moindres carrés"

$$(M) \qquad \min_{p \in \mathcal{A}} \|f - p\|_{\mathcal{B}}$$

(ii) p^* est caractérisé par l'"équation normale" du problème

$$(f - p^*, p)_{\mathcal{B}} = 0 \quad \forall p \in \mathcal{A}$$

<u>dém</u>: ⓐ Montrons que l'équation normale caractérise une éventuelle solution :

1) Soit $e^* = F - p^*$ où p^* est solution de $\min_{p \in \mathcal{A}} \|f - p\|$ c.a.d.

$$\|f - p^*\|_B \leq \|f - p\|_B \quad \forall p \in \mathcal{A}$$

On note $\| \ \|_B = \| \ \|$.

Supposons que $(e^*, p_0)_B \neq 0$ pour un certain $p_0 \in \mathcal{A}$

Alors $\|F - p^* - \alpha p_0\|^2 = \underbrace{\|F - p^*\|^2 - 2\alpha (e^*, p_0) + \alpha^2 \|p_0\|^2}_{\varphi(\alpha)}$,

$\varphi(\alpha)$ = polynôme de degré 2 en α.

$\varphi'(\alpha) = -2(e^*, p_0) + 2\alpha \|p_0\|^2$ donc

$\varphi'(\alpha) = 0 \iff \alpha = \dfrac{(e^*, p_0)}{\|p_0\|^2}$ $\quad (p \neq 0$ car $(e^*, p)_B \neq 0)$

donc c'est $p^* + \alpha p_0 = p^* + \dfrac{(e^*, p_0)}{\|p_0\|^2} p_0$ qui est la meilleure approximation de f dans $\mathcal{A}$ et non p^*.

Donc $(e^*, p)_B = 0 \quad \forall p \in \mathcal{A}$.

Réciproquement si $(e^*, p) = 0 \quad \forall p \in \mathcal{A}$. Pour tout $q^* \in \mathcal{A}$, on a

$$\|F - q^*\|^2 - \|F - p^*\|^2 = \|q^*\|^2 - \|p^*\|^2 - 2(F, q^*) + 2(F, p^*)$$

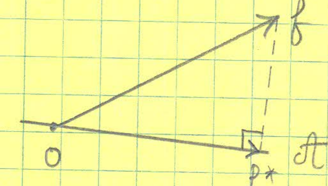
$$= \|q^* - p^*\|^2 + 2\underbrace{(F - p^*, p^* - q^*)}_{=0 \text{ par hypothèse.}}$$

donc $\|F - q^*\|^2 - \|F - p^*\|^2 \geq 0 \iff \|F - p^*\| \leq \|F - q^*\| \quad \forall q^* \in \mathcal{A}$

donc p^* est solution de (M)

Donc p^* est solution de (M) si et seulement si p^* est solution de

(EN) $\quad (f - p^*, p) = 0 \quad \forall p \in \mathcal{A}$.



(EN) s'appelle l'"équation normale" du problème

(b) Réciproquement (EN) possède une solution unique, car c'est un problème ~~de~~ linéaire dans $\mathcal{A}$, dim $\mathcal{A} < +\infty$, qui s'écrit

$$(p^*, p) = (f, p) \qquad \forall p \in \mathcal{A}$$

Si $f \equiv 0 \Rightarrow (p^*, p) = 0 \quad \forall p \in \mathcal{A}$; $p = p^*$ donne $p^* = 0$, donc le système linéaire a au plus une solution. Il en a donc exactement une seule. $\boxtimes$

Calcul explicite de p^* :

Soit $(\varphi_i)_{i=0,1\dots n}$ une base de $\mathcal{A}$.

Alors si $p^* = \sum_{j=0}^{n} c_j^* \varphi_j$, l'équation normale se traduit par

$$\left(\sum_{j=0}^{n} c_j^* \varphi_j ; \varphi_i\right) = (\varphi_i ; f) \qquad i = 0 \dots n \quad \text{, soit}$$

$$\sum_{j=0}^{n} (\varphi_j, \varphi_i)\, c_j^* = (\varphi_i ; f) \qquad i = 0 \dots n$$

En notant $X = \begin{bmatrix} c_0^* \\ c_1^* \\ \vdots \\ c_n^* \end{bmatrix} \in \mathbb{R}^{n+1}$; $A = A_{ij} = (\varphi_i, \varphi_j)$ et $b = \begin{bmatrix} (f, \varphi_0) \\ \vdots \\ (f, \varphi_n) \end{bmatrix}$

on a que X est solution de $Ax = b$,

Théorème 2 :

Si $\mathcal{A}$ est un sous-espace d'un espace de Hilbert engendré par $(\varphi_0, \varphi_1, \ldots \varphi_n)$. Si les fonctions $(\varphi_i)_{0 \leq i \leq n}$ sont mutuellement orthogonales

c.a.d. $(\varphi_i, \varphi_j) = 0 \quad \forall i \neq j$

alors pour tout $f \in \mathcal{B}$, la meilleure approximation de f par des éléments de $\mathcal{A}$ est

$$p^* = \sum_{j=0}^{n} \frac{(\varphi_j, f)}{\|\varphi_j\|^2} \varphi_j$$

dém : On a :

$$\sum_{j=0}^{n} (\varphi_i, \varphi_j) c_j = (\varphi_i, f_i)$$

donc

$$c_j = \frac{(\varphi_j, f)}{\|\varphi_j\|^2}$$

$\boxed{\varphi_{//}}$

Situation concrète pour l'approximation moindres carrés : le "grand" espace $\mathcal{B}$ n'est pas connu de façon exacte, mais seulement par des "mesures"

$(x_j, y_j)_{1 \leq j \leq m}$ avec m "grand" . On a $y_j \simeq f(x_j)$

et des "erreurs de mesure" sont commises.

L'espace approximant $\mathcal{A}$ va au contraire être de petite dimension . On peut souhaiter pondérer certaines mesures par les poids $w_j > 0$ reflétant le fait qu'elles sont entachées de plus ou moins d'incertitudes.

- w_j grand $\Rightarrow$ "bonnes" mesures
- w_j petit $\Rightarrow$ "mauvaises" mesures

L'espace $\mathcal{B}$ peut donc s'identifier à $\mathbb{R}^m$ muni de la norme

$$\|u\|_w = \left[\sum_{j=1}^{m} w_j u_j^2 \right]^{1/2}$$

C'est la contrepartie discrète de la norme

$$\left[\int_a^b w(t) F(t)^2 dt \right]^{1/2}, \quad w(t) > 0$$

Fonction "poids" sur $C^0[a,b]$.

À $p \in \mathcal{A}$, correspond $p(x_1) \dots p(x_m)$ des valeurs approximantes et le problème de moindres carrés consiste alors à minimiser

$$\|y - p(x)\|_{\mathcal{A}} = \left(\sum w_j \left(y_j - p(x_j) \right)^2 \right)^{1/2}$$

<u>Exemple fondamental:</u> La " régression linéaire"

$(x_j ; y_j)$ = mesures. On cherche $p(x) = ax+b$ t.q.

$$\varphi(a,b) = \left[\sum_{j=1}^{m} w_j \left(y_j - a x_j - b \right)^2 \right]^{1/2} \quad \text{minimum}$$

les valeurs de a et b sont données par l'équation normale de $\varphi(a,b) = \|y - ax - b\|_w^2$.

On a $\varphi(a,b) = \|y - ax - b\|_w^2$

où $(x,y)_w = \sum_{j=1}^{m} w_j x_j y_j$

Equation normale
$$\begin{cases} (y - ax - b, \mathbb{1})_w = 0 \\ (y - ax - b ; x)_w = 0 \end{cases}$$

c.a.d.
$$\begin{cases} a (x, \mathbb{1})_w + b (\mathbb{1}, \mathbb{1})_w = (\mathbb{1}, y)_w \\ a \|x\|_w^2 + b (\mathbb{1}, x)_w = (x, y)_w \end{cases}$$

où $x = (x_1 \dots x_m)$, $\mathbb{1} = (1, 1 \dots 1) \in \mathbb{R}^m$, $y = (y_1 \dots y_m) \in \mathbb{R}^m$

$$\begin{cases} a = \dfrac{(\mathbb{1},x)_w (\mathbb{1},y)_w - (x,y)_w (\mathbb{1},\mathbb{1})_w}{(x;\mathbb{1})_w^2 - \|x\|_w^2 (\mathbb{1},\mathbb{1})_w} = \dfrac{(x,y)_w (\mathbb{1},\mathbb{1})_w - (\mathbb{1},x)_w (\mathbb{1},y)_w}{\|x\|_w^2 (\mathbb{1},\mathbb{1})_w - (x,\mathbb{1})_w^2} \\[2ex] b = \dfrac{(\mathbb{1},x)_w (x,y)_w - \|x\|_w^2 (\mathbb{1},y)_w}{(x;\mathbb{1})_w^2 - \|x\|_w^2 (\mathbb{1},\mathbb{1})_w} = \dfrac{\|x\|_w^2 (\mathbb{1},y)_w - (\mathbb{1},x)_w (x,y)_w}{\|x\|_w^2 (\mathbb{1},\mathbb{1})_w - (x,\mathbb{1})_w^2} \end{cases}$$