Chapitre 6

Approximation par des polynômes

Sommaire

6.1 Théorèmes généraux

Soit E un espace vectoriel normé (e.v.n.) sur $\mathbb{R}$ ou $\mathbb{C}$, muni d'une norme $\|\cdot\|$.

Théorème 6.1 *Si M est un sous-espace vectoriel* ***de dimension finie*** *de E, alors pour tout g dans E, il existe au moins un y dans M tel que $\|g-y\| = \inf\limits_{x\in M} \|g-x\|$.*

Démonstration On se fixe x_0 dans M. On définit $K = \{x \in M, \|g-x\| \leq \|g-x_0\|\}$. Alors $\inf_M = \inf_K$. La fonction $x \mapsto \|g-x\|$ est une fonction continue sur K compact, elle admet une borne inférieure d'après le théorème de Weirstrass. ■

Théorème 6.2 (Théorème de projection dans un Hilbert) *Soit E est un espace de Hilbert muni du produit scalaire $(\cdot,\cdot)$, $\|\cdot\|$ est la norme associée. Soit M un* *sous-espace vectoriel de E, alors pour tout g dans E, il existe* **un unique** *y dans M tel que $\|g-y\| = \inf_{x\in M} \|g-x\|$. On le note $P_M g$. C'est la projection de g sur M, caractérisée par*

$$\forall z \in M, \quad (g - P_M g, z) = 0$$

Démonstration Vu en L2. ■

On va appliquer ces théorèmes à $M = \mathbf{P}_n$. D'abord un résultat d'approximation

Théorème 6.3 (Théorème de Weierstrass) *Si $(a,b]$ est compact, toute fonction continue sur $(a,b]$ peut être approchée uniformément par des polynômes, ou encore l'espace des polynômes est dense dans $\mathcal{C}^0((a,b])$ pour la norme de L^∞.*

Démonstration La démonstration est un peu longue, elle s'appuie sur les polynômes de Bernestein

$$B_n(f) = \sum_{k=0}^{n} \binom{n}{k} f(\frac{k}{n}) x^k (1-x)^{n-k}$$

■

L'équivalent pour les fonctions périodiques est très utile :

Théorème 6.4 *Soit f une fonction continue de période 2π. Alors il existe des coefficients réels $a_0, \cdots, a_n, \cdots$ et $b_1, \cdots, b_n, \cdots$ tels que*

$$S_n(t) = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt)$$

converge uniformément vers f sur $\mathbb{R}$.

6.2 Polynômes orthogonaux, moindres carrés

L'espace $L^2(a,b)$ des fonctions de carré intégrable sur (a,b) est un espace de Hilbert pour le produit scalaire $(f,g) = \int_a^b f(x)g(x)\,dx$. Il contient les polynômes.

Définition 6.1 *On dit qu'une suite de polynômes $p_0, \cdots, p_n, \cdots$ forme une suite de polynômes orthogonaux si*

- *$d^\circ p_n = n$ pour tout n,*
- *$(p_i, p_j) = 0$ pour $i \neq j$.*

Il existe une unique suite, à une constante multiplicative près. Exemple : sur $[-1,1]$, les polynômes de Legendre sont définis par $L_n(1) = 1$. Ils sont définis également par la formule de récurrence

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - n\,L_{n-1}(x), \quad L_0 = 1,\, L_1 = x.$$

Ils sont aussi solution de l'équation différentielle

$$(1-x^2)y\text{''} - 2xy' + n(n+1)y = 0.$$

Leur norme est égale à $2/(2n+1)$.

Soit maintenant f une fonction de $L^2(a,b)$. Par le théorème de projection, il existe un unique P_n dans $\mathbf{P}_n$, projection de f sur $\mathbf{P}_n$. Décomposons le sur la base des polynômes orthogonaux p_j :

$$P_n = \sum_{j=0}^{n} \alpha_j^n p_j.$$

Par la caractérisation de la projection, on doit avoir pour $0 \leq j \leq n$:

$$(f, p_j) = (Q_n, p_j) = \alpha_j^n \|p_j\|^2$$

et donc α_j^n ne dépend pas de n et

$$\alpha_j^n = \alpha_j = \frac{(f, p_j)}{\|p_j\|^2}$$

Théorème 6.5 *Soit f une fonction de $L^2(a,b)$.*

1. *Pour tout n positif, il existe un unique polynôme $P \in \mathbf{P}_n$ tel que $\|f - P_n\| = \inf_{P\in\mathbf{P}_n} \|f - P\|$. Il est donné par*

$$P_n = \sum_{k=0}^{n} \frac{(f, p_j)}{\|p_j\|^2} p_j$$

2. *Si de plus $[a,b]$ est compact et f est continue, alors P_n tend vers f dans L^2 et*

$$\|f\|^2 = \sum_{k=0}^{\infty} \frac{(f, p_j)^2}{\|p_j\|^2}$$

6.3 Moindres carrés discrets

On se place maintenant dans $E = \mathbb{R}^N$. On se donne N points x_i, N mesures f_i, et on cherche $p \in \mathbf{P}_{n-1}$ qui "approche" les f_i aux points x_i. Il est clair que si $N >> n$, il n'existe en général pas de polynôme qui passe par tous les points. On va alors chercher à passer "le plus près possible", c'est-à-dire à minimiser la distance entre les $p(x_i)$ et les f_i : on cherche donc p_{n-1} qui minimise $\sum_{i=1}^N |p_{n-1}(x_i) - f_i|^2$. On cherche p_n sous la forme

$$p_{n-1}(x) = \sum_{k=0}^{n-1} a_k x^k,$$

c'est-à-dire qu'on cherche les a_k, et et on minimise $\sum_{i=1}^N | \sum_{k=0}^{n-1} a_k x_i^k - f_i|^2$. Notons A la matrice des $a_{ik} = x_i^{k-1}$, $1 \leq i \leq N$, $1 \leq k \leq n$, $y = (a_0, \cdots, a_{n-1})$, $b = (f_1, \cdots, f_N)$.

$$Trouver\ y \in \mathbb{R}^n,\ \|Ay - b\| = \inf_{z \in \mathbb{R}^n} \|Az - b\| \tag{6.1}$$

Théorème 6.6 *Soit A une matrice $N \times n$. Le problème de minimisation* (6.1) *admet une solution, caractérisée par $A^T Ay = A^T b$. La solution est unique si et seulement si A est injective (i.e.* $\mathrm{rg\,A = n}$*).*

Ce sont les équations normales. Pour résoudre le problème de moindres carrés, on n'a donc qu'à résoudre un système linéaire de taille n. Mais ce problème est mal conditionné. Un exemple :

$$A = \begin{pmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$$

alors

$$A = \begin{pmatrix} 1+\varepsilon^2 & 1 & 1 \\ 1 & 1+\varepsilon^2 & 1 \\ 1 & 1 & 1+\varepsilon^2 \end{pmatrix}$$

Définition 6.2 *Les valeurs singulières de A sont les racines carrées positives des n valeurs propres de $A^T A$.*

Théorème 6.7 *Soit A une matrice $N \times n$ avec $N \geq n$. Alors il existe 2 matrices* *orthogonales U $(N\times N)$ et V $(n\times n)$, et un matrice Σ de taille $N\times n$*

$$\Sigma = \left(\begin{array}{ccc} \sigma_1 & & \\ & \sigma_2 & 0 \\ 0 & & \ddots \\ & & \sigma_n \\ \hline & 0 & \\ & & \end{array}\right)$$

telles que $A = U\Sigma V^T$.

Démonstration

$$V^T(A^TA)V = \mathrm{diag}(\sigma_i^2)$$

Soit c_j le j-ème vecteur de AV. On a

$$c_i^T c_j = \sigma_i^2 \delta_{ij}$$

On ordonne les σ_i : $\sigma_1, \cdots, \sigma_r$ non nulles. Donc $c_i \equiv 0$ pour $i > r$. On pose

$$u_j = \frac{c_j}{\sigma_j}$$

pour $j \leq r$. Ils forment un système orthonormé, qu'on complète en une base orthonormée de $\mathbb{R}^N$. Soit U la matrice des u_j. ■

En corollaire, le rang de A est égal au nombre de valeurs singulières de A.

On a $A = \sum \sigma_i u_i v_i^T$, $A^TA = \sum \sigma_i^2 v_i v_i^T$. On en déduit que $u_1, \cdots, u_r$ forment une base de ImA, $v_{r+1}, \cdots, v_n$ une base de $\ker A$, $v_1, \cdots, v_r$ une base de $\mathrm{ImA^T} = (\ker \mathrm{A})^{\perp}$.

On appelle maintenant pseudo-inverse de Σ la matrice

$$\Sigma^\dagger = \left(\begin{array}{cccccc|c} 1/\sigma_1 & & 0 & & & & \\ & 1/\sigma_2 & & & & & \\ & & \ddots & & 0 & & 0 \\ & & & 1/\sigma_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \end{array}\right)$$

On définit alors la pseudo-inverse de A par

$$A^\dagger = V\Sigma^\dagger U^T$$

On a maintenant $A^\dagger = \sum 1/\sigma_i v_i u_i^T$. On en déduit que $AA^\dagger = \sum u_i u_i^T$ est la matrice de la projection orthogonale sur ImA et $A^\dagger A = \sum v_i v_i^T$ la matrice de la projection orthogonale sur $\mathrm{ImA}^{\mathrm{T}}$. Revenons à notre système de moindres carrés. Les équations normales ont maintenant une interprétation agréable. Si b n'appartient pas à ImA, nous le projetons sur ImA en $AA^\dagger b$ et nous résolvons alors $Ax = AA^\dagger b$, ou encore $A(x - A^\dagger b) = 0$, ou $x - A^\dagger b \in \ker A$, ou $x - A^\dagger b = (I - A^\dagger A)w$ pour un quelconque w.

Théorème 6.8 *La solution générale du problème de moindres carrés discrets s'écrit*

$$y = A^\dagger b + (I - A^\dagger A)w$$

Si A est de rang n, il y a une solution unique. Si $\mathrm{rgA} < \mathrm{n}$, *l'ensemble des solutions est un espace vectoriel de dimension $n - r$.*

Dans le deuxième cas, on choisit dans l'ensemble des solutions celle de norme minimale, i.e. on cherche $w \in \ker A$ tel que

$$\|A^\dagger b + w\| = \inf_{z \in \ker A} \|A^\dagger b + z\|$$

ce qui revient à projeter $-A^\dagger b$ sur $\ker A$.

6.4 Régression linéaire

On a mesuré, sur N individus, $n+1$ variables $Y, X_1, \cdots, X_n$. Appelons $\bar{x}_{ij}$ la mesure de la variable X_j sur l'individu i, et $\bar{y}_i$ celle de la variable Y. On cherche à reconstituer la loi de Y à partir de celles des X_i, supposées linéairement indépendantes, par une formule linéaire

$$Y = b_0 + b_1 X_1 + \cdots + b_n X_n$$

Soit y le vecteur des $y_i = b_0 + b_1 \bar{x}_{i1} + \cdots + b_n \bar{x}_{in}$, et $\bar{y}$ le vecteur de composantes $\bar{y}_i$. On cherche $b = (b_0, \cdots, b_n)$ qui minimise la norme de $y - \bar{y}$. EN notant $\bar{X}$ la matrice

$$\bar{X} = \begin{pmatrix} 1 & \bar{x}_{11} & \cdots & \bar{x}_{1in} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{x}_{i1} & \cdots & \bar{x}_{in} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{x}_{N1} & \cdots & \bar{x}_{Nn} \end{pmatrix}$$

on est ramenés à minimiser $\|\bar{X}b - \bar{y}\|$.

6.5 Résolution des équations normales

6.5.1 Méthode de Cholewski

Supposons la matrice A injective. La matrice $B = A^T A$ est alors symétrique définie positive. On écrit sa décomposition de Cholewski : il existe une unique matrice S triangulaire supérieure à coefficients diagonaux strictement positifs, telle que $B = S^T S$. On résout alors successivement les deux problèmes triangulaires.

Malheureusement, les équations normales ne sont pas bien conditionnées, et la décomposition n'est valable que si A est injective. On va donc faire différemment.

6.5.2 Décomposition QR

Toujours en supposant la matrice A injective, écrivons successivement

$$\begin{gathered} S^T S = A^T A \\ \left(S^T\right)^{-1} A^T A S^{-1} = I \\ \left(AS^{-1}\right)^T (AS^{-1}) = I \end{gathered}$$

ce qui montre que la matrice $Q_1 = AS^{-1}$ est orthogonale. On peut aussi écrire

$$A = Q_1 S$$

Augmentons Q_1 (de taille $N \times n$) par une matrice Q_2 en une matrice carrée de taille N : $Q = (Q_1|Q_2)$, alors nous pouvons écrire l'égalité précédente comme

$$A = (Q_1|Q_2)\begin{pmatrix} S \\ 0 \end{pmatrix} = QR$$

En fait cette décomposition peut être obtenue par d'autre moyen (voir plus bas), et ne nécessite pas que A soit injective. Puisque la matrice Q est orthogonale on a pour tout z, d'après le théorème de Pythagore,

$$\|Az - b\|^2 = \|Q^T(Az - b)\|^2 = \|Rz - Q^T b\|^2 = \|Sz - (Q^T b)_1\|^2 + \|(Q^T b)_2\|^2$$

Si R est inversible, c'est-à-dire si le rang r de A est égal à n, l'équation $Sz = (Q^T b)_1$ a une seule solution y, et le minimum est atteint pour y :

$$\inf_z \|Az - b\| = \|AS^{-1}(Q^T b)_1\| = \|(Q^T b)_1\|.$$

Si $r < n$, notons que $\ker A = \ker S$: il y a une infinité de solution, comme mentionné dans le théorème 6.8. Pour en trouver une, nous effectuons une factorisation QR de S^T, sous la forme

$$S^T = PV$$

où P^T est une matrice orthogonale $n \times n$ et V de taille $n \times r$ triangulaire supérieure de rang r

$$V = \left(\begin{array}{cccc} v_1 & \times & \times & \times \\ & v_2 & \times & \times \\ 0 & & \ddots & \times \\ & & & v_r \\ \hline \\ & & 0 & \\ \\ \end{array} \right) = \left(\begin{array}{c} \bar{V}^T \\ \hline 0_{n-r,r} \end{array} \right)$$

La matrice $\bar{V}$ est donc triangulaire inférieure de rang r. D'où

$$S = V^T P = (\bar{V} | 0_{r,n-r})\, P$$

Décomposons $\tilde{z} = Pz$ sous la forme

$$\tilde{z} = \left(\begin{array}{c} \tilde{z}_1 \\ \hline \tilde{z}_2 \end{array} \right), \quad \tilde{z}_1 \in \mathbb{R}^r,\ \tilde{z}_2 \in \mathbb{R}^{n-r}.$$

Donc

$$Sz = \left(\begin{array}{c} \bar{V}\tilde{z}_1 \\ \hline 0_{r,n-r} \end{array} \right)$$

et

$$\|Sz - (Q^T b)_1\|^2 = \|\bar{V}\tilde{z}_1 - (Q^T b)_1\|^2,$$

d'où

$$\|Az - b\|^2 = \|\bar{V}\tilde{z}_1 - (Q^T b)_1\|^2 + \|(Q^T b)_2\|^2.$$

La matrice $\bar{V}$ est inversible, donc $\|Az - b\|$ est minimal pour $\bar{V}\tilde{y}_1 - (Q^T b)_1 = 0$ et le minimum est de nouveau $\|(Q^T b)_2\|$. Choisissons

$$\tilde{y} = \left(\begin{array}{c} \tilde{y}_1 \\ \hline 0 \end{array} \right), \quad y = P^T \tilde{y}$$

Alors y est solution du problème de minimisation. y est de norme minimale : toutes les autres solutions s'écrivent sous la forme

$$z = P^T \tilde{z}, \quad \tilde{z} = \left(\begin{array}{c} \tilde{y}_1 \\ \hline \tilde{z}_2 \end{array} \right)$$

si bien que (puisque la matrice P^T est orthogonale),

$$\begin{array}{rcl} \|z\|^2 & = & \|\tilde{z}\|^2 \\ & = & \|\tilde{y}_1\|^2 + \|\tilde{z}_2\|^2 \\ & = & \|\tilde{y}\|^2 + \|\tilde{z}_2\|^2 \\ & = & \|y\|^2 + \|\tilde{z}_2\|^2 \\ & \geq & \|y\|^2 \end{array}$$

6.5.3 Décomposition QR par matrices de Householder

Définition 6.3 *On appelle matrice de Householder associée au vecteur $u \in \mathbb{R}^p$ de norme $\sqrt{2}$ la matrice $p \times p$ donnée par*

$$H_u = I - u\,u^T$$

Propriétés 6.1 *Pour tout u dans $\mathbb{R}^p$ de norme $\sqrt{2}$, la matrice H_u est symétrique, orthogonale. C'est la matrice de la réflexion sur l'hyperplan orthogonal à u.*

Pour effectuer la décomposition QR de la matrice A, on va procéder comme dans la méthode de Gauss : on va multiplier successivement la matrice A à gauche par des matrices élémentaires de façon à mettre successivement des zéros sous la diagonale de A.

Lemme 6.1 *Soit x un vecteur non colinéaire à e_1 (premier vecteur de base). Alors il existe un nombre σ réel, et une matrice H_u telle que $H_u x = \sigma e_1$.*

σ est donné par $|\sigma| = \|x\|$, et son signe est l'opposé de celui de x_1. Le vecteur u est alors

$$u = \sqrt{2}\frac{x - \sigma e_1}{\|x - \sigma e_1\|}$$

Ecrivons la matrice A à l'aide de ses vecteurs colonne :

$$A = (a^1 \dots a^n), \quad H_u A = (H_u a^1 \dots H_u a^n),$$

Supposons que a^1 n'est pas colinéaire à e_1 (sinon on ne fait rien). Utilisons le lemme pour trouver $(\sigma_1, u_1) \in \mathbb{R} \times \mathbb{R}^N$ tel que $H_{u_1} a^1 = \sigma_1 a^1$, et définissons $A_1 = H_{u_1} A$. La première colonne de la matrice A_1 est le vecteur

$$\begin{pmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

et la matrice A_1 se décompose par blocs :

$$A_1 = \left(\begin{array}{c|ccc} \sigma_1 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & \bar{A}_1 & \\ 0 & & & \end{array}\right)$$

Nous renouvelons la construction sur $\bar{A}_1$: construisons $(\sigma_2, \bar{u}_2) \in \mathbb{R} \times \mathbb{R}^{N-1}$ tel que

$$H_{\bar{u}_2}\bar{A}_1 = \left(\begin{array}{c|ccc} \sigma_2 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & \bar{A}_2 & \\ 0 & & & \end{array}\right)$$

Attention la matrice $\bar{A}_2$ est de taille $(N-2)\times(N-2)$. Notons que si nous définissons le vecteur

$$u_2 = \left(\begin{array}{c} 0 \\ \hline \bar{u}_2 \end{array}\right)$$

alors on a

$$\left(\begin{array}{c|ccc} 1 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & H_{\bar{u}_2} & \\ 0 & & & \end{array}\right) = H_{u_2}$$

et

$$A_2 = H_{u_2}A_1 = \left(\begin{array}{c|ccc} \sigma_1 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & H_{\bar{u}_2}\bar{A}_1 & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|c|ccc} \sigma_1 & & \times & \dots & \times \\ \hline 0 & \sigma_2 & \times & \dots & \times \\ \hline \vdots & 0 & & & \\ \vdots & \vdots & & \bar{A}_2 & \\ 0 & 0 & & & \end{array}\right)$$

On itère jusqu'à A_{n-1} qui est triangulaire supérieure, avec la construction d'une famille $u_1, \dots, u_{n-1}$ et

$$A_{n-1} = H_{u_{n-1}} \dots H_{u_1} A$$

et

$$A = (H_{u_{n-1}} \dots H_{u_1})^{-1} A_{n-1} = (H_{u_1} \dots H_{u_{n-1}}) A_{n-1}$$

puisque les matrices H_u sont orthogonales et symétriques. La matrice $P = H_{u_1} \dots H_{u_{n-1}}$ est orthogonale et nous avons fini la construction.

6.5.4 Lien avec l'orthogonalisation de Gram-Schmidt

Notons a^j les vecteurs colonne de A et q^j les vecteurs colonne de Q. Alors

$$A = QR \iff \forall j, 1 \leq j \leq n, a^j = \sum_{k=1}^{j} R_{kj} q^k$$

La décomposition QR correspond donc à l'orthogonalisation des colonnes de A.