Université Paris Sud — Département de Mathématiques d'Orsay
Agrégation de mathématiques
Modélisation et calcul scientifique — 14 novembre 2008

Approximation par moindres carrés

1) Droite de régression
On se donne un entier m supérieur ou égal à 1 et des valeurs ponctuelles "issues de données" (x_j, y_j) pour $1 \leq j \leq m$. On cherche la "meilleure" droite d'équation

(1) $\quad y \,=\, a + b\,x \;\equiv\; f(x)$

qui passe "au plus près" des points au sens des moindres carrés. On cherche pour cela à minimiser la fonctionnelle

(2) $\quad J(f) \,\equiv\, \dfrac{1}{m} \displaystyle\sum_{j=1}^{m} \big(y_j - f(x_j)\big)^2 \,.$

• En paramétrant la droite d'équation (1) par ses coefficients a et b, expliciter le système linéaire qui doit être résolu par le couple (a, b). A quelle condition sur les données (x_j, y_j) $(1 \leq j \leq m)$ la solution (a, b) est-elle unique ? Mettre en œuvre sous "Matlab" l'algorithme obtenu en choisissant l'ensemble des données *via* une perturlation aléatoire de la fonction $[0, 1] \ni x \longmapsto \sin(2\pi x)$.

2) Cubique de régression
Reprendre l'ensemble du problème précédent en remplaçant la droite d'équation (1) par une cubique de la forme

(3) $\quad f(x) \,=\, a + b\,x\, + c\,x^2 + d\,x^3 \,.$

3) Introduction d'une contrainte
On se place dans le cadre proposé à la seconde question. On se donne deux réels $x_0 \in]0, 1[$ et $y_0 \in \mathbb{R}$ arbitraire. On ajoute au problème de minimisation de la fonctionnelle J donnée à la relation (2) la contrainte suivante :

(4) $\quad f(x_0) \,=\, y_0$

pour exprimer que la valeur de la fonction f au point particulier x_0 est connue avec certitude. Pour résoudre le problème de minimisation qui consiste à chercher f de la forme (3) qui minimise la fonctionnelle J donnée en (2) sous la contrainte (3), on introduit le Lagrangien

$$(5) \quad L(f,\, \lambda) \,=\, J(f) + \lambda \big(f(x_0) - y_0 \big)\,.$$

Montrer que les équations d'Euler-Lagrange de ce problème se ramènent ici à un système linéaire que l'on précisera pour les cinq inconnues $(a,\, b,\, c,\, d,\, \lambda)$. Construire ce système linéaire et le résoudre numériquement à l'aide de "Matlab".

4) Une seconde contrainte

On remplace la contrainte (4) par une condition de moyenne. On se donne $\mu \in \mathbb{R}$ et on impose à la fonction f la condition

$$(6) \quad \int_0^1 f(x)\, \mathrm{d}x \,=\, \mu\,.$$

Reprendre la troisième question avec cette nouvelle contrainte.

FD, 12 novembre 2008.