Teoria dell'Approssimazione e Applicazioni, A.A. 2014/15

Esercitazione di laboratorio del 21 maggio 2015

Prof. Stefano De Marchi

Approssimazione ai minimi quadrati quando i centri differiscono dai "data sites"

Supponiamo che la funzione f sia nota su data sites $X = \{x_1, ..., x_N\}$ mentre le funzioni di base siano centrate in $\Xi = \{\xi_1, \dots, \xi_M\}$ con $M \le N$. L'approssimante RBF sarà

$$Q_f(x) = \sum_{i=1}^{M} c_j \Phi(x, \xi_j), \quad x \in \mathbb{R}^s. \tag{1}$$

I coefficienti c_j si potranno determinare, come noto, come soluzione ai minimi quadrati del sistema $A\mathbf{c} = \mathbf{f}$ (con A matrice rettangolare $N \times M$ con componenti $A_{j,k} = \Phi(x_j, \xi_k)$, $\mathbf{c}$, $\mathbf{f}$, vettori di lunghezze M ed N, rispettivamente) minimizzando il funzionale quadratico $\|Q_f - f\|_2^2$.

1. Usare lo script `RBFApproximation2D.m`, che utilizza due insiemi di punti, i "data sites" e i centri, per approssimare la funzione `sinc` in $[0,1]^2$ usando come RBF la funzione gaussiana con $\epsilon = 1$. Aumentare o diminuire ϵ e vedere come cambia l'errore e il rango della matrice A. Fare anche esperimenti con diversi insiemi di punti (sia data sites che centri): equispaziati e di Chebyshev.

2. Usare poi lo script `RBFApproximation2Dlinear.m` (che consente di usare funzioni di base *Thin Plate Splines (TPS)*, definite nella funzione `tps.m`, con l'aggiunta di un termine polinomiale di primo grado, lineare appunto) per approssimare *dati affetti da errore*. La simulazione proposta consiste nel campionare la *funzione di Franke* sull'insieme $X \subset [0,1]^2$ dei data sites, quindi aggiungendo un rumore random uniformemente distribuito di lunghezza variabile (la riga dello script che fa questo è `rhs=rhs+0.03*randn(size(rhs))`, ovvero un'aggiunta del 3% di rumore).

   (a) La prima proposta è di determinare l'approssimazione ai minimi quadrati di tanti punti con poche funzioni di base (come fatto coi minimi quadrati al punto precedente). Questo approccio in statistica si chiama *regressione con splines*. In questo esperimento i centri coincidono con i data sites. Si noterà dai plot delle approssimanti, che questa strategia funziona meglio che non l'interpolazione.

   (b) La seconda strategia per "lisciare" dati affetti da rumore è il cosidetto *ridge regression method*, cioè il metodo dei *minimi quadrati regolarizzati* che fa uso di un parametro ω di smoothing, risolvendo non il sistema sovradeterminato $A\mathbf{c} = \mathbf{f}$ ma

   $$\left(A + \frac{1}{2\omega} I\right) \mathbf{c} = \mathbf{f}\,.$$

   Per questo, usare lo script `TPS_RidgeRegression2D.m`, che usa le sempre le TPS per approssimare ancora la funzione di Franke, facendo variare ω. Mediante il metodo di *cross-validation* trovare il parametro ω^* che rende lo smoothing migliore.