Report laboratorio 1

Luca Baruffini, Matteo Casartelli, Denis Suozzi, Giuseppe Tarantino

Aprile 2019

1 Esercizio 2

Per modellare questo problema, vogliamo utilizzare la funzione dei least squares, in modo da minimizzare l'errore dovuto dalle rilevazioni nella formula

$$y = Ax + v \tag{1}$$

dove v è un disturbo piccolo e centrato in 0.
I termini noti per questo scenario sono l'insieme dei vettori di input

$$x(1), ..., x(N) \tag{2}$$

e l'insieme corrispettivo dei vettori di output

$$y(1), ..., y(N) \tag{3}$$

Si vuole minimizzare la funzione

$$J = \sum_{k=1}^{N} ||Ax(k) - y(k)||^2 \tag{4}$$

Notando che il termine non noto del quale produrre una versione least squares è la matrice

$$A \in \mathbb{R}^{mxn} \tag{5}$$

In questo caso, però, il termine è dal lato sbagliato del prodotto matriciale per poter essere minimizzato utilizzando il metodo della matrice pseudoinversa.
Introduciamo quindi le due matrici

$$X = \begin{bmatrix} x(1) & ... & x(N) \end{bmatrix} \tag{6}$$

$$Y = \begin{bmatrix} y(1) & ... & y(N) \end{bmatrix} \tag{7}$$

In modo da poter riscrivere la funzione J come

$$J = ||AX - Y||_2^2 \tag{8}$$

Ricordiamo che se A è una matrice, la sua LUNGHEZZA sarà

$$||A||_2 = \sqrt{\sum_{i,j} \widetilde{a}_{ij}^2} \Rightarrow ||A||_2^2 = \sum_{i,j} \widetilde{a}_{ij}^2 = ||A^T||_2^2 \tag{9}$$

Possiamo quindi scrivere:

$$AX = \begin{bmatrix} Ax(1) & ... & Ax(N) \end{bmatrix} \tag{10}$$

$$AX - Y = \begin{bmatrix} Ax(1) - y(1) & ... & Ax(N) - y(N) \end{bmatrix} \tag{11}$$

Dove Ax(k) - y(k) è la k-esima colonna di AX - Y, riconducendoci alla forma iniziale del problema.
Da qui:

$$J = ||AX - Y||_2^2 = ||(AX - Y)^T||_2^2 = ||X^T A^T - Y^T||_2^2 = \sum_i^m ||X^T \widetilde{a}_i - \widetilde{y}_i||^2 \tag{12}$$

Abbiamo portato la matrice dalla parte giusta del prodotto e notiamo che la somma di diversi minimi indipendenti produrrà sempre una somma minima.
Possiamo scrivere:

$$\widetilde{a}_{ils} = (X^T)^\dagger \widetilde{y}_i \Rightarrow \begin{pmatrix} \widetilde{a}_{ls1} & ... & \widetilde{a}_{lsm} \end{pmatrix} = (X^T)^\dagger \begin{pmatrix} \widetilde{y}_1 & ... & \widetilde{y}_m \end{pmatrix} \tag{13}$$

E quindi

$$A^T = (X^T)^\dagger Y^T \Rightarrow A = Y((X^T)^\dagger)^T = YX^\dagger \tag{14}$$

Che è la soluzione desiderata.

Si nota che per far si che questo modello funzioni vanno rispettate le seguenti condizioni

X^T deve essere INTO
(Full rank skinny)
.
$N \geq n$
(Devo avere più campioni che numero di variabili in input)
.
Devo avere almeno N $x(i)$ INDIPENDENTI
.