Report laboratorio 1

Luca Baruffini, Matteo Casartelli, Denis Suozzi, Giuseppe Tarantino

Aprile 2019

1 Esercizio 1

Vedere il file Exercise1.m nel package con associato il file di definizione della funzione genQR.m

2 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

Vedere il file Exercise2.m per la risoluzione del problema proposto

3 Esercizio 3

3.1 Obiettivo

Trovare le costanti che minimizzano la funzione J:

$$J = \frac{1}{N}\sum_{i}^{N} log(\frac{\hat{T}_i}{T_i})^2 \tag{15}$$

Considerando che con l e $T \in R^{nx1}$:

$$\hat{T}(l) = Cl^B l^{gl} \tag{16}$$

3.2 Soluzione

Per minimizzare la funzione J modifichiamo la funzione logaritmica al suo interno:

$$log(\frac{\hat{T}_i}{T_i}) = log(\hat{T}_i) - log(T_i) = \tag{17}$$

$$log(Cl_i^B l_i^{gl_i}) - log(T_i) = log(C) + log(l_i^B) + log(l_i^{gl_i}) - log(T_i) = \tag{18}$$

$$log(C) + Blog(l_i) + gl_i log(l_i) - log(T_i) \tag{19}$$

Considero:

$$x = \begin{bmatrix} log(C) \\ B \\ g \end{bmatrix} \in R^{px1} \tag{20}$$

dove p = numero di costanti (in questo caso 3) e

$$b = \begin{bmatrix} log(T_1) \\ \ldots\ldots\ldots \\ log(T_i) \end{bmatrix} \in R^{mx1} \tag{21}$$

$$A = \begin{bmatrix} 1 & log(l_1) & l_1 log(l_1) \\ \ldots & \ldots & \ldots \\ 1 & log(l_i) & l_i log(l_i) \end{bmatrix} \in R^{mxp} \tag{22}$$

Di conseguenza:

$$log(\frac{\hat{T}_i}{T_i}) = Ax - b \tag{23}$$

Abbiamo ricondotto la funzione non lineare ad una funzione affine.

La funzione J può essere riscritta come:

$$J = \frac{1}{N} \sum_{i}^{N} (Ax_i - b_i)^2 \tag{24}$$

riconducendo il nostro problema ad un problema di least squares.
Per calcolare x utilizziamo la seguente formula (A deve essere INTO):

$$x = (A^T A)^{-1} A^T y = A^\dagger y \tag{25}$$

dove A è la nostra matrice A, y è la nostra matrice b, $(A^T A)^{-1} A^T$ è la matrice pseudoinversa di A, ovvero $A^\dagger$.
Ricordando che otterremo:

$$x = \begin{bmatrix} log(C) \\ B \\ g \end{bmatrix} \tag{26}$$

dovremo modificare la soluzione per ottenere il valore di C come

$$C = e^{x_1} \tag{27}$$

Dato che abbiamo ricondotto il nostro problema ad un problema di least squares, possiamo dichiarare che questa soluzione riconduce sempre al minimo valore globale di J. Se avessimo dovuto usare il metodo della linearizzazione, il minimo valore ottenuto di J sarebbe stato locale.
Per poter utilizare la formula:

$$x = A^\dagger y \tag{28}$$

la matrice A deve essere into, ovvero deve essere full rank skinny.
Quindi dato che $A \in R^{mxp}$ deve valere: $rank(A) = min(m, p) \Rightarrow m \geq p \Rightarrow rank(A) = p$

Vedere il file Exercise3.m nel package per la soluzione del problema proposto.

4 Esercizio 4

Vedere il file Exercise4.m nel package per la spiegazione e i risultati del problema.
Il problema della massa ha come condizioni iniziali posizione 0 e velocità 1, cioè

$$\begin{cases} p(0) = 0 \\ \dot{p}(0) = 1 \end{cases} \tag{29}$$

E vogliamo che l'output del nostro modello sia la posizione e velocità a fronte di un input di forze in 10 instanti e anche la posizione intemedia all'istante $n = 5$.
Poniamo il nostro vettore di output cioè

$$y = \begin{bmatrix} p(10) \\ \dot{p(10)} \\ p(5) \end{bmatrix} \tag{30}$$

Prima di poter riscrivere il problema dobbiamo costruire la matrice A

$$A = \begin{bmatrix} 10 - j + 0.5 \\ 1 \\ 5 - j + 0.5 \end{bmatrix} \tag{31}$$

Quindi la matrice sarà

$$A = \begin{bmatrix} 9.5 & 8.5 & 7.5 & 6.5 & 5.5 & 4.5 & 3.5 & 2.5 & 1.5 & 0.5 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4.5 & 3.5 & 2.5 & 1.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{32}$$

Avendo le condizioni inziali diverse da zero, il modello non è propriamente lineare ma affine. Quindi avrà la forma

$$y = Ax + b \tag{33}$$

Per trovare la b possiamo ragionare sul fatto che, in assenza di forze, il moto della massa è rettilineo uniforme. Possiamo quindi trovare facilmente b come

$$b = \begin{bmatrix} 10v_0 + p_0 \\ v_0 \\ 5v_0 + p_0 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ 5 \end{bmatrix} \tag{34}$$

Il modello di interesse per il nostro problema è quindi

$$y = Ax + \begin{bmatrix} 10 \\ 1 \\ 5 \end{bmatrix} \tag{35}$$

Ora vogliamo trovare il vettore di forze x che porta la massa in $y = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

In più viene richiesta la condizione che

$$\int_{t=0}^{10} f(t)^2 dt \tag{36}$$

sia minima.

Abbiamo visto come questo sia equivalente a chiedere che la norma del vettore delle forze

$$||x||^2 \tag{37}$$

sia anch'essa minima.

In definitiva dobbiamo trovare la soluzione a norma minima:

$$x_{ls} = -A^{\dagger} \begin{bmatrix} 10 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} -0.622727 \\ -0.437121 \\ -0.251515 \\ -0.065909 \\ 0.119697 \\ 0.180303 \\ 0.115909 \\ 0.051515 \\ -0.012879 \\ -0.077273 \end{bmatrix} \tag{38}$$

E troviamo quindi il vettore delle forze migliore che controlla la massa nella posizione richiesta.

Il secondo punto dell'esercizio richiede invece una soluzione di un problema di minimi quadrati multi obiettivo.
In questo caso particolare è un problema di minimi quadrati "regolarizzati"

$$J_1 = ||Ax - y||^2 \qquad J_2 = ||x||^2 \tag{39}$$

Dobbiamo trovare l'ottimo di Pareto: la curva che minimizza la somma pesata delle due funzioni di costo $J_1 + \mu J_2$ con il parametro μ da settare.

Abbiamo visto a lezione che il problema ha soluzione

$$x(\mu) = (A^T A + \mu I)^{-1} A^T y \tag{40}$$

Come da richiesta tracciamo il grafico di $J_1 + \mu J_2$ per individuare la curva di trade-off. La curva presenta un certo grado di scelta (nessun trade-off per la curva y = -x)

Considerazioni sul parametro μ μ rappresenta il peso che vogliamo dare ai due problemi da minimizzare: se preferiamo cioe' in questo caso particolare minimizzare l'errore o la norma del vettore delle forze. Infatti agli estremi abbiamo:
Se $\rightarrow 0^+$ ci riconduciamo al problema di x_{ln} a norma minima; mentre se facciamo tendere $\mu \rightarrow \infty$ otteniamo $J_2 \rightarrow 0$ quindi $||x|| = 0$.
A questo punto otteniamo l'errore massimo

$$J = ||A0 - \begin{bmatrix} 10 \\ 1 \end{bmatrix} ||^2 = 10^2 + 1^2 = 101 \tag{41}$$

cioè lasciando ferma la massa.