LABORATORIO DI ANALISI NUMERICA

Laurea Magistrale in Statistica e Informatica

Esercitazione su minimi quadrati e SVD

Prof. Stefano De Marchi

`Padova, 19 novembre 2009`

1 Alcuni utili comandi Matlab

La *decomposizione ai valori singolari*, SVD *(Singular Value Decomposition)* di una matrice, fattorizza una matrice $A \in \mathbb{R}^{n\times m}$ (anche rettangolare) nel prodotto

$$A = USV^T,\ \ U \in \mathbb{R}^{n\times n},\ \ S \in \mathbb{R}^{n\times m},\ \ V \in \mathbb{R}^{m\times m}$$

Gli elementi nella "diagonale" di S sono chiamati *valori singolari* di A. Tutti gli altri elementi di S sono nulli. Le matrici U e V sono *ortogonali*, cioè $U^TU = I_n$ e $V^TV = I_m$, ove I_n e I_m indicano le matrici identità di ordine n e m, rispettivamente.

- Il comando `svd` calcola appunto la decomposizione SVD. Vediamo un esempio.

```
>> A=[1,2,1;1,3,1;0,1,1;1,1,1];

>> [U,S,V]=svd(A)

U =

   -0.5346    0.1091   -0.1889   -0.8165
   -0.7186   -0.5603   -0.0547    0.4082
   -0.2738    0.2608    0.9258    0.0000
   -0.3505    0.7786   -0.3230    0.4082


S =

    4.5732         0         0
         0    0.7952         0
         0         0    0.6736
         0         0         0


V =
```

```
   -0.3507    0.4117   -0.8411
   -0.8417   -0.5323    0.0903
   -0.4105    0.7397    0.5332
```

A partire dai fattori U, S e V è possibile calcolare, ancora, la soluzione ai minimi quadrati di un sistema sovradeterminato. Dapprima scelgo un termine noto `b` e quindi procediamo come segue

```
>> b=[1;2;0;1];

>> d=U'*b

d =

   -2.3223
   -0.2329
   -0.6213
    0.4082

>> s=diag(S)

s =

    4.5732
    0.7952
    0.6736
```

`length`

```
>> y=d(1:length(s))./s

y =

   -0.5078
   -0.2929
   -0.9224

>> x=V*y

x =

    0.8333
    0.5000
   -0.5000
```

- In maniera analoga, è possibile calcolare la soluzione ai minimi quadrati di un sistema *quadrato* singolare.

```
>> A=[1,2,1;1,2,1;0,1,0];

>> b=[1;2;0];

>> [U,S,V]=svd(A)

U =

   -0.6873   -0.1664   -0.7071
   -0.6873   -0.1664    0.7071
   -0.2353    0.9719         0


S =

    3.5616         0         0
         0    0.5616         0
         0         0         0


V =

   -0.3859   -0.5925   -0.7071
   -0.8379    0.5458         0
   -0.3859   -0.5925    0.7071
```

Il terzo valore singolare vale 0 e dunque non è possibile calcolare $y_3 = d_3/s_3$. Basta però porre $y_3 = 0$ o cercare con il comando `find` gli elementi non nulli

```
>> d=U'*b;

>> s=diag(S);

>> y=zeros(size(d));

>> index=find(s~=0);

>> y(index)=d(index)./s(index)

y =
```

```
   -0.5789
   -0.8888
         0

>> x=V*y

x =

    0.7500
   -0.0000
    0.7500
```

1.1 Esercizi proposti

1. Si considerino i valori di tabella

x_i	1	2.5	3	5	6.5	8	9.3
y_i	4	2	3	3.5	3.9	7	5.3

   (a) Determinare il polinomio P_m, di grado $m = 3$ approssimante le coppie di valori (x_i, y_i) nel senso dei minimi quadrati discreti. *NB*: Si usi la decomposizione `svd` per risolvere il corrispondente sistema lineare sovradeterminato. Provare anche ad usare \.

   (b) Si giustifichi il fatto che per $m = 6$ il polinomio è interpolante.

   (c) Si consideri il punto $\bar{x} = 4$ e come valore corrispondente $\bar{y}$, quello dell'interpolante lineare sull'intervallo [3, 5]. Sia ora $|P_m(\bar{x}) - \bar{y}|$ l'errore assoluto in $\bar{x}$. Far vedere che per $m = 2$ l'errore è minimo.

2. *Approssimazione non polinomiale ai minimi quadrati.* Supponiamo di voler approssimare un insieme di dati $\{(x_i, y_i)\}_{i=1}^n$ per mezzo di una funzione ottenuta come combinazione delle funzioni dell'insieme $U_1 = \{1, \sin(x), \cos(x)\}$.

$$t_1(x) = a_1 \sin(x) + a_2 \cos(x) + a_3$$

   (i) Calcolarne la soluzione ai minimi quadrati, usando la funzione $t_1(x)$, nel caso in cui i dati provengano da 50 punti random in $[0, \pi]$ di $f(x) = \sin(\pi \cos(x))$.

   (ii) Si consideri ora l'insieme $U_2 = \{1, \sin(x), \cos(x), \sin(2x), \cos(2x)\}$. Costruire l'approssimante ai minimi quadrati usando come approssimante la funzione

$$t_2(x) = a_1 \sin(2x) + a_2 \cos(2x) + a_3 \sin(x) + a_4 \cos(x) + a_5\,.$$

   (iii) Rispetto all'errore assoluto, quale delle due approssimazioni, (i) o (ii), è preferibile?

   (iv) Sviluppando con la serie di MacLaurin la funzione $\sin(y(x))$, con $y(x) = \pi \cos x$, si trovi un'approssimazione che da un errore in norma infinito dell'ordine di 10^{-3}.