

Esercizi

ACS2_09 – Analisi delle Componenti Principali.

1. Scrivere una funzione MATLAB per implementare l’algoritmo “*PCA incrementale*” (dove il massimo autovalore ed il corrispondente autovettore sono calcolati da una funzione che implementa il “*power method*”). Confrontare i suoi risultati con quelli restituiti dalla funzione MATLAB `pca()`. Calcolare anche il tempo di esecuzione della funzione per l’algoritmo “*PCA incrementale*” mediante: `tic` <codice da valutare> `T=toc`, e confrontarlo col tempo richiesto dalla funzione `pca()`. Il codice seguente mostra come generare i dati random (dal punto di vista statistico) e come usare le funzioni `tic ... T=toc` per l’elapsed time.

```
N=50; % o più (numero dei campioni)
% media mu e matrice di covarianza Sigma della popolazione
mu=[3 1 2 0]; Sigma=[1 .2 .7 .3; .2 1 0 .6; .7 0 1 .5; .3 .6 .5 1];
X=mvnrnd(mu,Sigma,N); % campione da distribuzione normale multivariata N(mu,Sigma)
tic
[basis,comp,lambda]=pca(X); % PCA della matrice dei campioni
T=toc;
fprintf("Tempo richiesto da pca() su %d campioni: %g\n", N,T)
```

2. Per i campioni del precedente esercizio implementare l’algoritmo *PCA*, calcolando gli autovalori e gli autovettori mediante le funzioni MATLAB:
 - `pca()`: *PCA* applicata alla matrice dei dati, alla matrice dei dati centrati ed alla matrice dei dati standardizzati;
 - `svd()`: fattorizzazione SVD della matrice centrata e della matrice standardizzata;
 - `eig()`: autovalori/autovettori della matrice di covarianza e della matrice di correlazione; e confrontarne i risultati.

Confrontare anche i loro tempi di esecuzione mediante: `tic` <codice da valutare> `T=toc`.
3. Implementare l’algoritmo “*Eigenfaces*” per il riconoscimento di una faccia mediante la funzione MATLAB `pca()`. Per i dati dell’esercizio (il database delle immagini delle facce) eseguire il download del file `db_400faces_112x92_col_uint8.mat` dalla pagina del corso sulla piattaforma di e-Learning.
4. Implementare l’algoritmo “*Eigenfaces*”, calcolando gli autovalori/autovettori mediante le funzioni MATLAB:
 - `pca()`: *PCA* applicata alla matrice dei dati, alla matrice dei dati centrati ed alla matrice dei dati standardizzati.
 - `svd()`: fattorizzazione SVD della matrice centrata e della matrice standardizzata.
 - `eig()`: autovalori/autovettori della matrice di covarianza e della matrice di correlazione. e confrontarne i risultati.

Confrontare anche i loro tempi di esecuzione mediante: `tic` <codice da valutare> `T=toc`.
5. Calcolare e visualizzare la retta di regressione di y su x e quella di x su y per i dati X che seguono. Calcolare e visualizzare anche il valore del funzionale $J_{LS}^{(1)}(a,b)$.

```
N=20; % o più (numero dei campioni)
% media mu e matrice di covarianza Sigma della popolazione
mu=[1 2]; Sigma=[1 .2; .2 .7];
X=mvnrnd(mu,Sigma,N); % campione da distribuzione normale multivariata
```

6. Calcolare e visualizzare il piano di regressione di z su x,y , quello di x su y,z e quello di y su x,z per i dati \mathbf{X} che seguono. Calcolare e visualizzare anche i corrispondenti valori dei tre funzionali $J_{LS}(a,b,c)$.

```
N=20; % o più (numero dei campioni)
% media mu e matrice di covarianza Sigma della popolazione
mu=[3 1 2]; Sigma=[1 .2 .8; .2 1 0; .8 0 1];
X=mvnrnd(mu,Sigma,N); % campione da distribuzione normale multivariata
```

7. Per gli stessi dati dell'esercizio precedente, calcolare e visualizzare il piano *PCA* (generato dalle prime due direzioni principali) ed il piano *Least Squares* di z su x,y . Quale condizione ottimale soddisfa ciascun piano rispetto ai campioni? Come verificarlo?