

# Esperienze di programmazione

## Lezione 6

Gianna Del Corso <delcorso@di.unipi.it>

25 Marzo 2015

### 1 PageRank

Scopo dell'esercitazione è l'implementazione dell'algoritmo PageRank usando il metodo delle potenze o il metodo di Lanczos (cioè utilizzando la funzione `eigs` di MatLab/OCTAVE). Si faranno in particolare esperimenti per valutare l'algoritmo per differenti valori del parametro  $\alpha$ .

1. Si consideri un grafo diretto con  $n$  nodi. Si costruisca la matrice di adiacenza  $G$ , e la matrice  $P$  ottenuta dividendo gli elementi di ogni riga per l'outdegree del nodo corrispondente. Sia  $\bar{P} = P + \mathbf{d}\mathbf{v}'$  con  $d_i = 1$  se il nodo  $i$  non ha nessun arco uscente e  $d_i = 0$  altrimenti e  $v_i = 1/n$ .

Sia  $\hat{P} = \alpha\bar{P} + (1 - \alpha)\mathbf{e}\mathbf{v}'$ .

Utilizzando il comando `eigs` si calcoli e si faccia un plot del vettore PageRank, cioè dell'autovettore sinistro di  $\hat{P}$ , per i seguenti valori di  $\alpha = 0.85, 0.9, 0.99, 1 - 10^{-8}, 1 - 10^{-12}$ . Quali conclusioni si possono trarre a proposito della dipendenza di PageRank da  $\alpha$ ?

2. Si implementi il metodo delle potenze per il calcolo del pagerank tenendo conto di come possono essere effettuate le moltiplicazioni matrice-vettore continuando a lavorare con matrici sparse. Si usi la norma del residuo come criterio d'arresto. Si testi sulla matrice di Stanford e si faccia un plot del residuo.
3. Si implementi il calcolo di PageRank utilizzando il comando `eigs`. Per non dover memorizzare la matrice piena  $\hat{P}$ , si chiami la funzione `eigs` con una handle alla moltiplicazione matrice-vettore, ovvero si realizzi una funzione `function y=Px(x, P, alpha)` che restituisce  $\mathbf{y} = \hat{P}^T \mathbf{x}$ . Si noti che poichè in `eigs` non si può controllare la norma dei vettori nella moltiplicazione matrice-vettore, non si può utilizzare il "trucco" del punto precedente.

4. Si testino i due algoritmi usando i valori 0.85 e 0.99 per il parametro  $\alpha$ . Si confronti la velocità di convergenza dei due algoritmi e il tempo impiegato (usando di tic e toc di OCTAVE). Si confrontino i vettori ottenuti di pagerank per i due valori del parametro di teletrasporto.

## 2 DATI

A scopo di test si può utilizzare la matrice di adiacenza stanford-web.dat.

La matrice può essere letta con `load stanford-web.dat` e poi convertendo la lista con il comando `A=spconvert(stanford_web)`. Si modifichi *A* affinché la matrice risulti quadrata.