

Laboratorio Sperimentale di Matematica Computazionale

Lezione 5

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

4 Aprile 2014

1 HITS

Scopo dell'esercitazione è studiare la relazione tra l'ordinamento dato da Degree, con quello di HITS e di Pagerank. Dapprima si implementa HITS secondo lo schema iterativo o con il calcolo dell'autovettore dominante con il comando `eigs`. Nella seconda parte si analizza un altro approccio nel quale il grado di un nodo (indegree e outdegree) viene usato per filtrare i nodi che sono poco connessi e sul grafo più piccolo ottenuto mantenendo solo questi nodi si calcola Pagerank e HITS. La motivazione si basa sull'osservazione che in generale i nodi poco connessi hanno bassi punteggi di Pagerank e authority.

1. Si consideri un grafo diretto con n nodi. Si costruisca la matrice di adiacenza G . Si scriva una funzione `function [h, a]=Iterate_HITS(G, k)` che partendo da due vettori $\mathbf{h}_0, \mathbf{a}_0$ costanti calcoli i punteggi di Authority e Hub dopo k iterazioni. Alla fine di ogni iterazione si normalizzino i due punteggi ottenuti.
2. Si confrontino i risultati ottenuti con quelli ottenuti calcolando l'autovettore relativo all'autovalore di modulo massimo di $G^T G$ e di $G G^T$ con il comando `eigs` di Octave.
3. Si scriva una funzione `function [top, pr, a]=ranking(G)`. G è la matrice di adiacenza di un grafo web. La funzione deve implementare il seguente algoritmo
 - Si calcoli l'in-degree e out-degree di ogni nodo
 - Si ordinino i nodi in base alla somma di outdegree e indegree
 - Si prendano i primi 1000 nodi (restituiti in `top`)
 - Si crei un nuovo grafo dove solo questi nodi sono inclusi. Il nuovo grafo dovrà mantenere solo i link presenti tra questi 1000 nodi.
 - Si calcoli pagerank (`pr`) e HITS (`a`) su questo nuovo grafo.