

Laboratorio Sperimentale di Matematica Computazionale

Lezione 6

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

8 Aprile 2014

1 Reputation System

1. Si scriva una funzione `function [r, t]=Reputation(E, c)` che presa la matrice delle valutazioni E , e il parametro c , restituisce il vettore delle reputazioni degli item r e il vettore di fiducia dei raters t . Inizialmente ogni votazione dei raters risulta essere fidata e quindi gli elementi la matrice di fiducia T vengono inizializzati ad 1. Si implementi l'algoritmo descritto a lezione.
2. Sui dati messi a disposizione riguardo alla valutazione di film, si confrontino i risultati ottenuti con questo algoritmo con quelli ottenuti prendendo come reputazione degli item il voto medio dato dai raters. Si aggiunga poi il 25% di spammers che assegnano sempre un punteggio uguale a 1 tranne che al loro film preferito al quale assegnano un punteggio di 5 e si faccia di nuovo il confronto tra la reputazione dei film rispetto alla reputazione data dalla media.
3. (fac) Vedendo il problema come un problema di ottimizzazione non vincolata, la soluzione \mathbf{r}^* restituita dal metodo precedente risulta essere il punto di massimo della funzione scalare $\psi(r) = \sum_{i=1}^n \sum_{j:i \rightarrow j} T_{ij}^2$. Si scriva una funzione `function r=NR_reputation(E, c, nsteps)` che applica il metodo di Newton-Raphson al sistema di equazioni $\nabla(\psi)(r) = 0$, si vedano i lucidi per la descrizione della funzione ψ e della matrice Hessiana.

2 DATI

Si utilizzi il database [MovieLens](#) che contiene 100.000 valutazioni di 1682 film da parte di 943 raters. Il file `u.data` contiene la matrice sparsa delle valutazioni. Lo si legga con `A=load('u.data')`, si rimuova la 4^o colonna e lo si trasformi in matrice sparsa con il comando `E= spconvert(A)`. Il file `u.item` contiene i titoli dei film considerati. Ogni rater valuta almeno 20 film.