

Recommender System Project: Approccio statistico per la tecniche SLIM

Alessio Baccelli, Flavio Di Palo

December 2017

1 Introduzione

Lo scopo del progetto è valutare se l'introduzione di un vincolo sulla matrice di similarità in fase di apprendimento dell'algoritmo SLIM[1] possa portare ad un significativo miglioramento rispetto alle tecniche SLIM tradizionali. In particolare in questo documento proponiamo l'introduzione di un vincolo che consenta di mantenere la somma delle righe della matrice di similarità pari ad 1 in fase di apprendimento.

2 Formalizzazione

Sia U l'insieme di tutti gli utenti ed I l'insieme di tutti gli item: vogliamo imporre un vincolo sulla matrice di similarità W , di dimensioni $I \times I$, che definisce quanto un'item sia simile ad un altro item.

3 Slim

Data una User Rating Matrix A della dimensione $m \times n$, l'obiettivo della tecnica SLIM è imparare una matrice sparsa W di dimensione $n \times n$ in grado di risolvere il seguente problema di ottimizzazione:

$$\begin{aligned} & \text{minimize}_W \quad \frac{1}{2} \|A - AW\|_F^2 + \frac{\beta}{2} \|W\|_F^2 + \lambda \|W\|_1 \\ & \text{subject to} \quad W > 0, \text{diag}(W) = 0 \end{aligned}$$

Ai vincoli classicamente espressi nel problema aggiungiamo un'ulteriore vincolo che impone la somma delle righe della matrice di similarità pari ad 1 in fase di apprendimento.

Introduciamo la matrice B i cui elementi vengono definiti come:

$$b_{ij} = \frac{w_{ij}}{\sum_{k=1}^n w_{ik}}$$

che quindi per costruzione rispetta il vincolo $\sum_i b_{ij} = 1$

Il problema di ottimizzazione sopra proposto può quindi essere rappresentato, tenendo conto dei vincoli introdotti, come:

$$\begin{aligned} & \underset{W}{\text{minimize}} && \frac{1}{2} \|A - AB\|_F^2 + \frac{\beta}{2} \|B\|_F^2 + \lambda \|B\|_1 \\ & \text{subject to} && B > 0, \text{diag}(B) = 0 \end{aligned}$$

Detta f la funzione da minimizzare è possibile calcolare che:

$$\frac{\partial f}{\partial w_{ij}} = ((a_{ij} - \sum_{k=1}^n a_{ik} * b_{kj})(-a_{ii}) + \beta b_{ij} + \lambda \frac{|b_{ij}|}{b_{ij}}) (\frac{\sum_{k=1}^n (w_{ik}) - w_{ij}}{(\sum_{k=1}^n w_{ik})^2})$$

dove a_{ij} è l'elemento della matrice A presente sulla i -esima riga e la j -esima colonna.

Dopo aver calcolato il gradiente rispetto a b_{ij} procediamo con la fase di update in modo da avere:

$$w_{ij}^{t+1} = w_{ij}^t - \alpha \frac{\partial f}{\partial w_{ij}}(w_{ij}^t)$$

Dopo questo step procederemo a valutare se tale approccio apporta dei miglioramenti rispetto alla tecnica classica esposta nella fonte[1]

References

- [1] Xia Ning and George Karypis. *SLIM: Sparse Linear Methods for Top-N Recommender Systems*. Data Mining (ICDM), 2011 IEEE 11th International Conference on