On multi-task collaborative rating problem

Jimmy Ba,

University of Toronto

Dec. 1, 2012

Base line algorithm

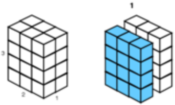
Formal definition

Define: U users and V rating options in the system
Given: A tensor $\mathbf{R} \in \Re^{UxUxV}$, each of its face $R_v \in \Re^{UxU}$ is the pairwise rating matrix for v^{th} rating option.

- There are matrices $P \in \Re^{Uxk}$ and $W_v \in \Re^{kxU}$, such that,

$$R_v = PW_v$$

- P stays fixed for $v \in \{1, 2, ..., V\}$



Base line algorithm

Objective Function

Define: mask matrix $M_v \in R^{UxU}$, has ones for all the rating entry and zeros else where. **Goal:** Minimize the Frobenius norm of the matrix completion problem

$$\min_{P,W_v} \sum_{v=1}^{V} ||(PW_v - R_v) \cdot M_v||_F^2$$

Base line algorithm

Training model

Updates: Take the gradient of objective function with respect to the P and W_v

$$\frac{\partial J}{\partial P} = \sum_{v=1}^{V} [(PW_v - R_v) \cdot M_v] W_v^T$$
$$\frac{\partial J}{\partial W_v} = P^T [(PW_v - R_v) \cdot M_v]$$

Use any non-convex optimization optimizer to estimate P and W_v

Base line algorithm

Ranking

Given: W_v for each task
Score vector:

$$S = Diag[W_v^T Q W_v]$$

where, Q is a weighting matrix

Experimental results - toy data

- We start with a 5-by-5 distance matrix that describe the similarity among users:

$$\begin{pmatrix} 0 & 2 & 6 & 8 & 9 \\ 2 & 0 & 4 & 6 & 7 \\ 6 & 4 & 0 & 2 & 3 \\ 8 & 6 & 2 & 0 & 1 \\ 9 & 7 & 3 & 1 & 0 \end{pmatrix}$$

- We also create a absolute skill level matrix for 3 tasks

$$\begin{pmatrix} 1. & 0.5 & 0.7 \\ 0.5 & 0.3 & 0.8 \\ 0.7 & 0.2 & 2 \\ 2. & 1. & 0.5 \\ 1. & 2. & 0.1 \end{pmatrix}$$

Experimental results - toy data

- We generate a rating tensor from the ground truth.
- We can recover the relative distance matrix.
  - k = 2, and L_2 weight decay is applied to both P and W
  - The Pearson correlation of the ground truth and estimated distance matrix is 0.5, meaning the two are strongly correlated.

Experimental results - simulated large data

- We simulated a generative process for rating tensor for 159 users and 18 tasks. The final rating tensor is 95% sparse
- The algorithm can still recover the relative distance matrix with 0.5 Pearson correlation.
  - k = 50, and L_2 weight decay is applied to both P and W