Matrix Factorization Recommender Systems and Cold Start

Aaron Tuor and Brian Hutchinson, Computer Science Department, Western Washington University

Overview

Recommender Systems

Created to direct users to items of interest, given a dataset, e.g., user ratings or product descriptions.

Cold Start Problem

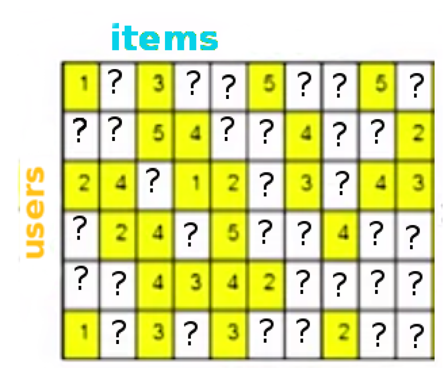
How to give ratings for new users and new items for which no ratings are recorded in the dataset?

Outline

- Survey state of the art Matrix Factorization solutions to the Cold Start problem.
- Suggest avenues of further research.

User-Item Utility Matrix

Ratings for m users and n items in $m \times n$ matrix, R.
$R_{u,i}$ = user u's rating of item i.

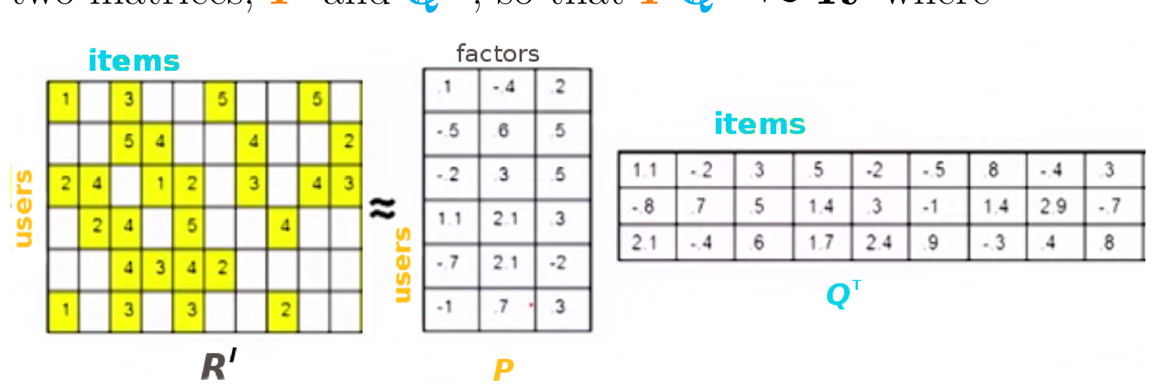
$R =$ 

Modeling Global Behavior [1]

- $R'_{u,i} = R_{u,i} - b_{u,i}$
- $b_{u,i} = \mu + \sigma_u + \sigma_i$ where $\mu \approx$ mean rating, $\sigma_u, \sigma_i \approx$ observed deviations of user u's and item i's ratings from μ.
- $\min_{b_*} \sum_{(u,i)} (R_{u,i} - \mu - \sigma_u - \sigma_i)^2 + \gamma(\sum_u \sigma_u^2 + \sum_i \sigma_i^2)$

Matrix Factorization Model [1]

Assume low rank to complete R'. R' may be factored into two matrices, P and Q^T, so that $PQ^T \approx R'$ where



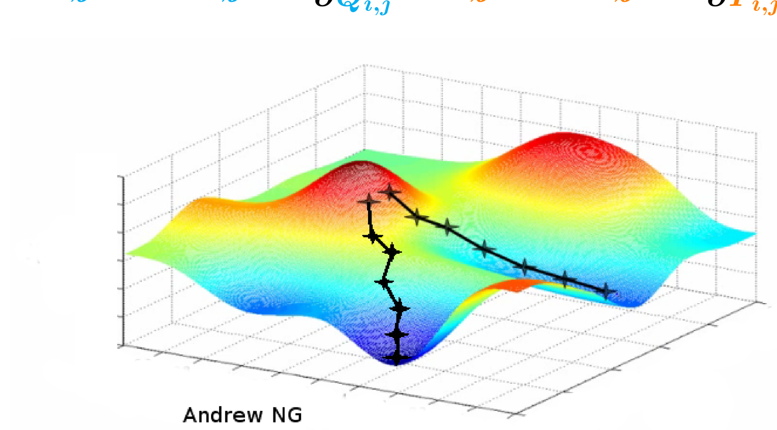
$P_{u,x}$ = user u's association with latent factor x
$Q_{i,x}$ = item i's association with latent factor x.
$q_i^T p_u$ = User-Item interaction.

- **Objective Function**

$$L(Q,P) = \sum_{(u,i)} (R'_{u,i} - q_i^T p_u)^2 + \lambda(\|q_i\|^2 + \|p_u\|^2)$$

- **Gradient Descent Step**

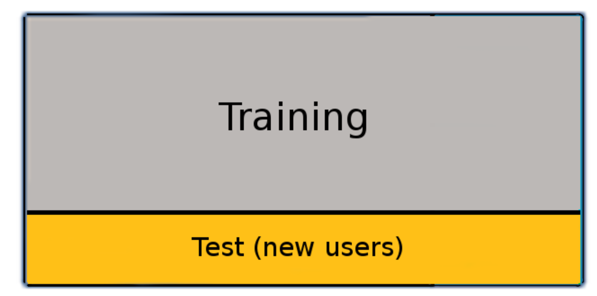
$$\forall(i,j)\ Q_{i,j} = Q_{i,j} - \frac{\partial L}{\partial Q_{i,j}},\ P_{i,j} = P_{i,j} - \frac{\partial L}{\partial P_{i,j}}$$



User Cold Start

User Cold Start Problem

A basic MF model can't predict ratings for new users.



Functional Matrix Factorization [3]

- Hybrid MF-Decision Tree Recommender System
- Uses MF to find best queries to estimate user profiles

- **Recommender Decision Trees**
  - New users are asked about their preference for certain items. Possible responses: **Like**, **Dislike**, **Unknown**.
  - a_u = user u's set of responses to queries.
  - $T(a_u) \approx p_u$ (function for approx. user profiles)
- **Objective Function**

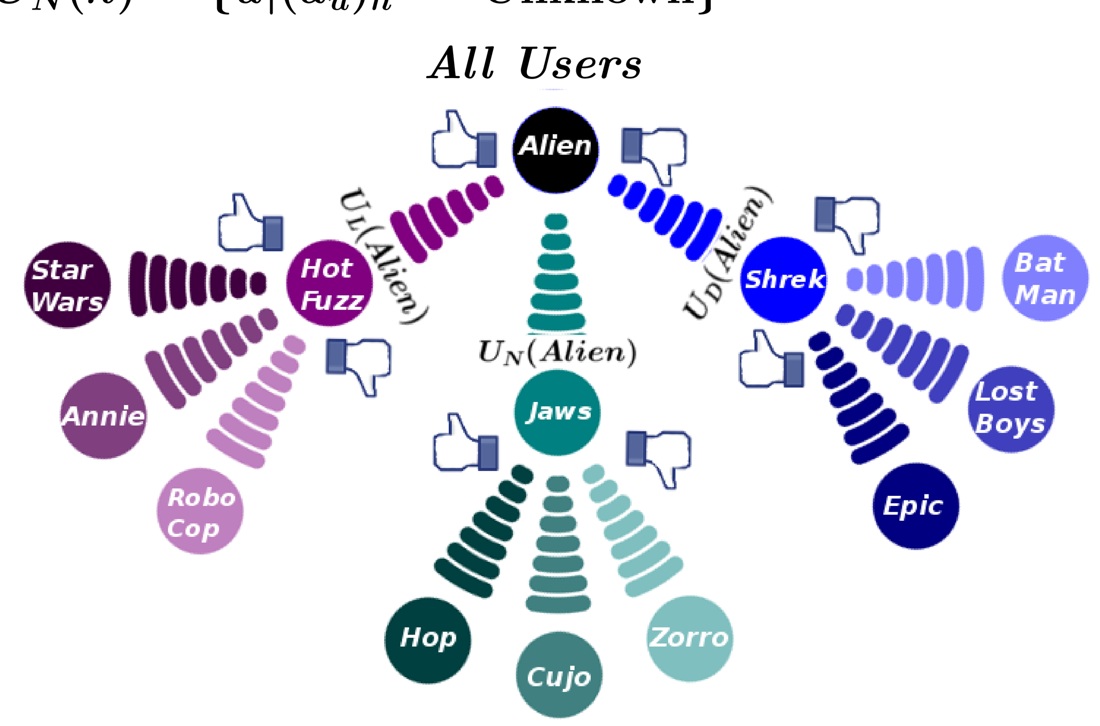
$$\min_{Q,T} \left(L(Q,T) = \sum_{(u,i)} (R'_{u,i} - q_i^T T(a_u))^2 + \lambda \|q_i\|^2\right)$$

- **Fix T and Optimize $w.r.t.$ Q**

$$\forall i,\ q_i = \left(\sum_{(u,i)} T(a_u)T(a_u)^T + \lambda I\right)^{-1}\left(\sum_{(u,i)} R'_{u,i} T(a_u)\right)$$

- **Fix Q and Optimize $w.r.t.$ T**
  - Starting at root, partition users into sets by query h:

$$U_L(h) = \{u | (a_u)_h = \text{ Like}\}$$
$$U_D(h) = \{u | (a_u)_h = \text{ Dislike}\}$$
$$U_N(h) = \{u | (a_u)_h = \text{ Unknown}\}$$



  - **Find Optimal Profiles for users in Child Nodes**

Let γ be a variable ranging over L, D, N.

$$h_\gamma = \operatorname{argmin}_p \sum_{u \in U_\gamma(h)} \sum_{(i,j)} (R'_{i,j} - q_j^T p)$$

  - **Find Queries to Optimize Profiles**

$$\min_h \sum_{u \in U_L(h)} \sum_{(i,j)} (R'_{i,j} - q_i^T h_L) + \sum_{u \in U_N(h)} \sum_{(i,j)} (R'_{i,j} - q_i^T h_D)$$
$$+ \sum_{u \in U_D(h)} \sum_{(i,j)} (R'_{i,j} - q_i^T h_N)$$

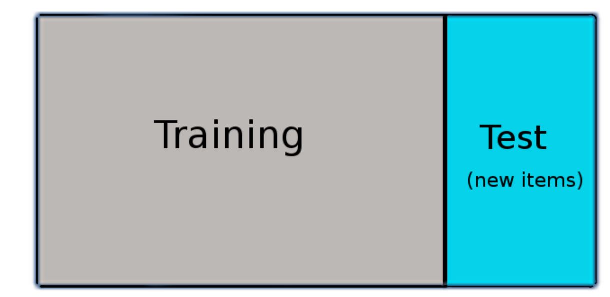
User Cold Start Summary

For new user z, use decision tree and queries to derive user profile p_z. User z's estimated rating for i is then $q_i^T p_z + \mu$.

Item Cold Start

Item Cold Start Problem

How to recommend new unrated items in MF framework?



Local Collective Embeddings [4]

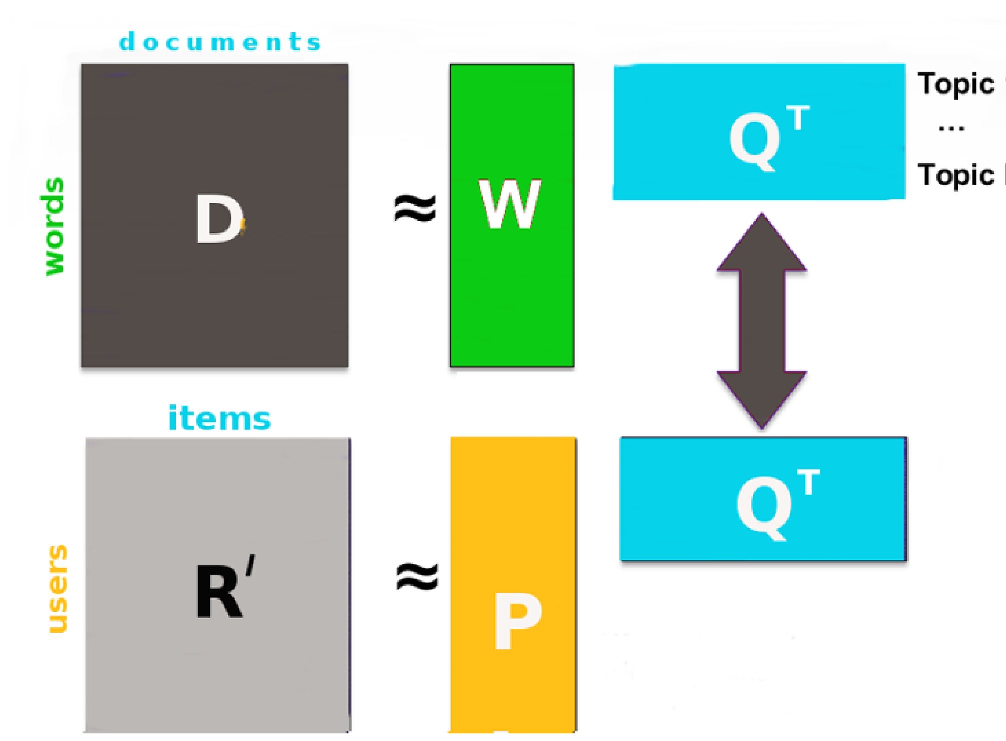
- Uses collective factorization to link user ratings to item text.
- **TF-IDF** [5]

t_k = a corpus term, d_j = a corpus document
$f_{k,j}$ = # times t_k occurs in d_j .
$T_j = \{t : t$ occurs in document $d_j\}$
N = # documents in corpus
$n_k = |\{d : t_k$ in $d\}|$

$$\text{TF-IDF}(t_k, d_j) = \frac{f_{k,j}}{\max_{z \in T_j} f_{z,j}} \log \frac{N}{n_k}$$

- **Collective Factorization**

$$D_{j,i} = \text{TF-IDF}(t_j, d_i)$$



- **Exploiting Locality**

Measurement of local smoothness for factor Q:

$$S = \frac{1}{2} \sum_{(i,j)} \|q_i - q_j\|_F^2 \frac{\mathrm{d_i} \cdot \mathrm{d_j}}{\|\mathrm{d_i}\| \|\mathrm{d_j}\|}$$

- **Objective Function**

$$L = \min_{Q,P,W} \frac{1}{2}[\alpha \sum_{(u,i)} (R'_{u,i} - q_i^T p_u)^2]$$
$$+ \frac{1}{2}[(1-\alpha) \sum_{(j,i)} (D_{j,i} - q_i^T w_j)^2]$$
$$+ \beta S + \lambda(\|q_i\|^2 + \|p_u\|^2 + \|w_j\|^2)$$
$$s.t.\ W \geq 0,\ Q \geq 0$$

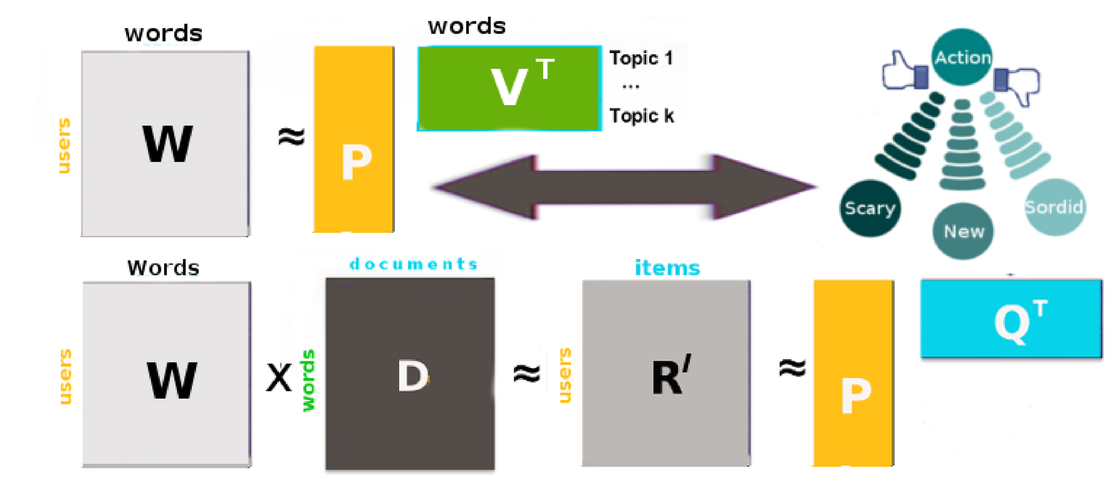
Item Cold Start Summary

For new item z and TF-IDF vector d_z, derive item profile q_z using common factor Q by solving $Wq_z = d_z$. User u's estimated rating for i is then $q_z^T p_u + \mu$.

General Cold Start

Further Research

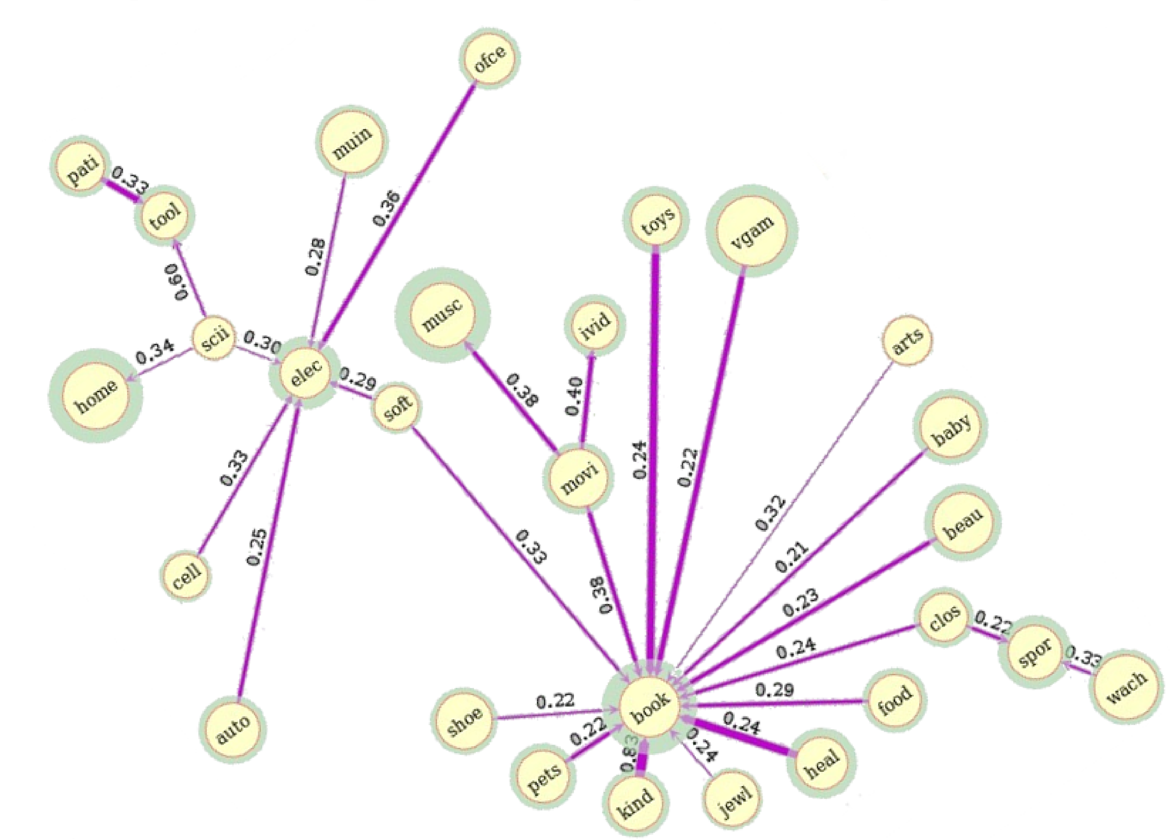
- Benchmark state of the art methods using Amazon dataset
- Explore cold start metric which leverages performance of both user and item cold start scenarios
- Develop method which addresses user and item cold start
- Use product descriptions to generate topic factorization and topic keywords to generate decision tree.



$$\alpha \sum_{(u,i)} (R'_{u,i} - q_i^T p_u)^2 + \beta \sum_{(u,w)} (W_{u,w} - v_w^T p_u)^2$$
$$+ \gamma \sum_{(u,i)*} (d_i^T w_u - q_i^T p_u)^2$$

Amazon Dataset [Leskovec, McAuley. 2013. SNAP]

- 34,686,770 reviews (text & ratings) & 6,643,669 users
- 28 product categories & descriptions for all products



References

[1] Koren, Yehuda, and Robert Bell. "Advances in collaborative filtering." Recommender systems handbook. Springer US, 2011. 145-186.
[2] Sarwar, Badrul, et al. "Item-based collaborative filtering recommendation algorithms." Proceedings of the 10th international conference on World Wide Web. ACM, 2001.
[3] Zhou, Ke, Shuang-Hong Yang, and Hongyuan Zha. "Functional matrix factorizations for cold-start recommendation." Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ACM, 2011.
[4] Saveski, Martin, and Amin Mantrach. "Item cold-start recommendations: learning local collective embeddings." Proceedings of the 8th ACM Conference on Recommender systems. ACM, 2014.
[5] Amatriain, Xavier, et al. "Data mining methods for recommender systems." Recommender Systems Handbook. Springer US, 2011. 39-71.