CSE 8803RS: Recommendation Systems

Lecture 4: Hybrid Models

Hongyuan Zha

School of Computational Science & Engineering
College of Computing
Georgia Institute of Technology

Baseline Estimates

- User and item effects
  — systematic tendencies for some users to give higher ratings
- Baseline estimate $b_{ui} = \mu + b_u + b_i$
- Example: a critical user Joe on Titanic
  — average overall rating: $\mu = 3.7$
  — Titanic better than average $b_i = .5$
  — Joe critical: $b_u = -.3$
  — $b_{ui} = 3.7 - .3 - +.5 = 3.9$

Item-Based Methods

Better scalability, accuracy and explainability

- Key component: item-item similarity

$$s_{ij} = \frac{n_{ij}}{n_{ij} + n} \rho_{ij}$$

- $S^k(i; u)$: k most similar items rated by u
- Adjusting for user and item effects

$$\hat{A}_{ui} = b_{ui} + \frac{1}{\sum_{j \in S^k(i;u)} s_{ij}} \sum_{j \in S^k(i;u)} s_{ij}(A_{uj} - b_{uj})$$

Latent Factor Methods

Low-rank matrix factorization

- Key ideas: user and item *latent* profiles F_u and G_i,

$$\hat{A}_{ui} = b_{ui} + F_u G_i^T$$

- Low-rank matrix factorization,

$$\min \sum_{(u,i) \in O} (A_{uj} - \mu - b_u - b_i - F_u G_i^T)^2 + \lambda(\|F_u\|^2 + \|G_i\|^2 + b_u^2 + b_i^2)$$

Item-Based Methods: Extensions

$R(u)$: items rated by u; $S^k(i)$ k items most similar to i measured by s_{ij}; $R^k(i;u) = R(u) \cap S^k(i)$; $N(u)$: items with implicit feedback from u; $N^k(i;u) = N(u) \cap S^k(i)$;

- New formula,

$$\hat{A}_{ui} = b_{ui} + |R^k(i;u)|^{-1/2} \sum_{j \in R^k(i;u)} w_{ij}(A_{uj} - b_{uj}) + |N(u)|^{-1/2} \sum_{j \in N(u)} c_{ij}$$

- Optimization problem,

$$\min \sum_{(u,i) \in O} \Big(A_{uj} - \mu - b_u - b_i -$$

$$-|R^k(i;u)|^{-1/2} \sum_{j \in R^k(i;u)} w_{ij}(A_{uj} - b_{uj}) - |N^k(i;u)|^{-1/2} \sum_{j \in N^k(i;u)} c_{ij}\Big)^2$$

$$\lambda\Big(\sum_{j \in R^k(i;u)} w_{ij}^2 + \sum_{j \in N(u)} c_{ij}^2 + b_u^2 + b_i^2 \Big)$$

Item-Based Methods: Connection with Matrix factorization

- Define the *symmetric* similarity matrix $W = [w_{ij}]$ with $w_{ii} = 0$
- Also define $\tilde{A} = [\tilde{A}_{ui}]$

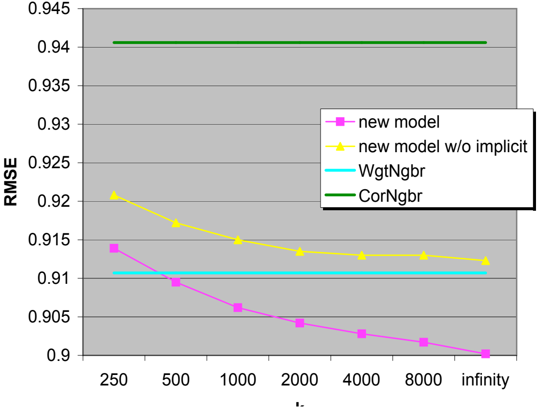
$$\tilde{A}_{uj} = A_{uj} - b_{uj}, \quad j \in R^k(i;u)$$

  and $\tilde{A}_{ui} = 0$ otherwise
- The doc-product $\sum_{j \in R^k(i;u)} w_{ij}(A_{uj} - b_{uj})$gives rise to the matrix product $\tilde{A}W$
- So the model is $A \approx \tilde{A}W$, ignoring other correction

Netflix Data

- Validation/Probe set: 1.4 million recent ratings
- Test set/Quiz set: 1.4 million recent ratings
- RMSE

Experimental Results



Latent Factor Methods: Extensions

Low-rank matrix factorization

- Extending

$$\hat{A}_{ui} = b_{ui} + F_u G_i^T$$

with implicit feedback

$$\hat{A}_{ui} = b_{ui} + (F_u + |N(u)|^{-1/2} \sum_{j \in N(u)} y_j) G_i^T$$

An Integrated Model

- New formula

$$\hat{A}_{ui} = b_{ui} + (F_u + |N(u)|^{-1/2} \sum_{j \in N(u)} y_j) G_i^T$$

$$+ |R^k(i;u)|^{-1/2} \sum_{j \in R^k(i;u)} w_{ij}(A_{uj} - b_{uj}) + |N(u)|^{-1/2} \sum_{j \in N(u)} c_{ij}$$

Experimental Results

Model	50 factors	100 factors	200 factors
SVD	0.9046	0.9025	0.9009
Asymmetric-SVD	0.9037	0.9013	0.9000
SVD++	0.8952	0.8924	0.8911