7-5: Probabilistic Matrix Factorization

Introduction to Recommender Systems

Introduction

- Linear algebra isn't the only matrix factorization basis
- This lecture: matrix factorization approaches with a probabilstic basis

Probabilistic Modeling

- Many interesting algorithms are based on probabilistic models
- Basic idea:
  - Assume that data is generated by random process (with known structure)
  - Learn parameters that would generate data that looks like what you have

Popularity Model

- Non-personalized probabilistic model
- P(I) is # of times i was bought, scaled to probability
  - Divide by total # of purchases

Personalized

- Goal: compute P(i|u)
  - Probability that user u will select item i
- Problem: many, many parameters
- Also, we don't know for items the user hasn't bought!

Probabilistic Latent Semantic Analysis

- Goal: estimate P(i|u)
  - Probability that user u will select item i
- Decompose with latent factors

$$P(i|u) = \sum_z P(i|z)P(z|u)$$

  - P(z|u) – user picks a random feature
  - P(i|z) – user picks movie for feature
  - User preference broken down to feature preference

PLSI and LSI (SVD)

- Probabilistic has same form as SVD
  - $M = U\Sigma V^T$
- Left and right vectors are *stochastic*, not orthogonal
  - That is, they encode probability distributions
- Learned with *expectation maximization*

PLSI with Ratings

- Basic idea: model ratings as a distribution (e.g. gaussian)
- Distribution parameters determined by item, user, and/or latent feature

PMF

- Alternate formulation of probabilistic factorization of ratings matrix
- Models ratings as drawn from normal distributions
  - Mean determined by user and item via features
  - Not simple probability matrices like PLSI
- Faster to train than PLSI

Further Reading

- Matrix factorization, both probabilistic and not, is core to many current algorithms
- Other latent feature models
  - Restricted Boltzman machines
  - Neural nets generally

7-5: Probabilistic Matrix Factorization

$$P(i) = \frac{\text{\# of times } i \text{ was played}}{\text{\# of total plays}} \qquad \frac{1k}{1M} = 0.001$$

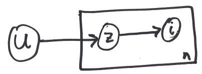
$$\sum_i P(i) = 1$$

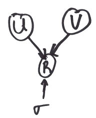
SVD:

$M = U \Sigma V^T$

$P(i|u) = ?$ k factors Z

$$= \sum_{z \in Z} p(z|u)\, p(i|z) \qquad M = U\ V^T$$





$$P(r_{ai} = r) = N(\vec{u}_a \cdot \vec{v}_i, \sigma)$$