Building a Radio Recommendation Engine

John Min

May 1, 2014

Approaches to recommendation

1. Content-based
   - Mel-frequency cepstral coefficients (MFCC) feature extraction
2. Dyadic data analysis
   - Latent factor models: (Bayesian) PMF
   - Mixed membership models: Latent Dirichlet Allocation (LDA)
   - $\text{M}^3\text{F} = \text{BPMF} + \text{LDA}$

Mixed Membership Matrix Factorization (M^{3}F)

- K^U: the number of user topics
- K^M: the number of item topics
- $\Lambda^U \sim \text{Wishart}(W_0, \nu_0), \Lambda^M \sim \text{Wishart}(W_0, \nu_0)$
- $\mu^U \sim N(\mu_0, (\lambda_0 \Lambda^U)^{-1}), \mu^M \sim N(\mu_0, (\lambda_0 \Lambda^U)^{-1})$
- For each user $i \in \{1, \ldots, N\}$:
  - $u_i \sim N(\mu^U, (\lambda^U)^{-1})$
  - $\theta_i^U \sim \text{Dir}(\alpha / K^U)$
- For each item $j \in \{1, \ldots, M\}$:
  - $v_i \sim N(\mu^V, (\lambda^V)^{-1}$
  - $\theta_i^V \sim \text{Dir}(\alpha / K^V)$
- For each rating $r_i j$:
  - $z_{ij}^U \sim \text{Multi}(1, \theta_i^U), z_{ij}^V \sim \text{Multi}(1, \theta_j^V)$
  - $r_i j \sim N(\beta_{ij}^{kl} + u_i \cdot v_j, \sigma^2)$

M^{3}F-TIB (Topic-Indexed Bias)

- $\beta_{ij}^{kl} = \chi_0 + c_i^k + d_j^l$
  - χ_0 is a fixed global bias
  - c_i^k's and d_j^l's are drawn from Gaussian priors
- "Napoleon Dynamite" effect
- each user and item can choose a different topic and thus, a different bias for each rating (such as in the case that multiple users share a single account)

Gibbs sampler (simplified)

- sample hyperparameters, $\{(\mu_U, \Lambda_U), (\mu_M, \Lambda_M)\}$
- sample topics, (z_{ij}^U, z_{ij}^M)
- sample user parameters $(\theta_i^U, u_i, \{c_i^k\}_{k=1}^{K^U})$
- sample item parameters $(\theta_j^V, v_i, \{d_j^l\}_{k=1}^{K^V})$

Coming to theaters (or online radios) near you

- get results!
- compare content-based approach with matrix factorization
- integrate the MFCC feature vector with MF algorithms