Probabilistic Matrix Factorization

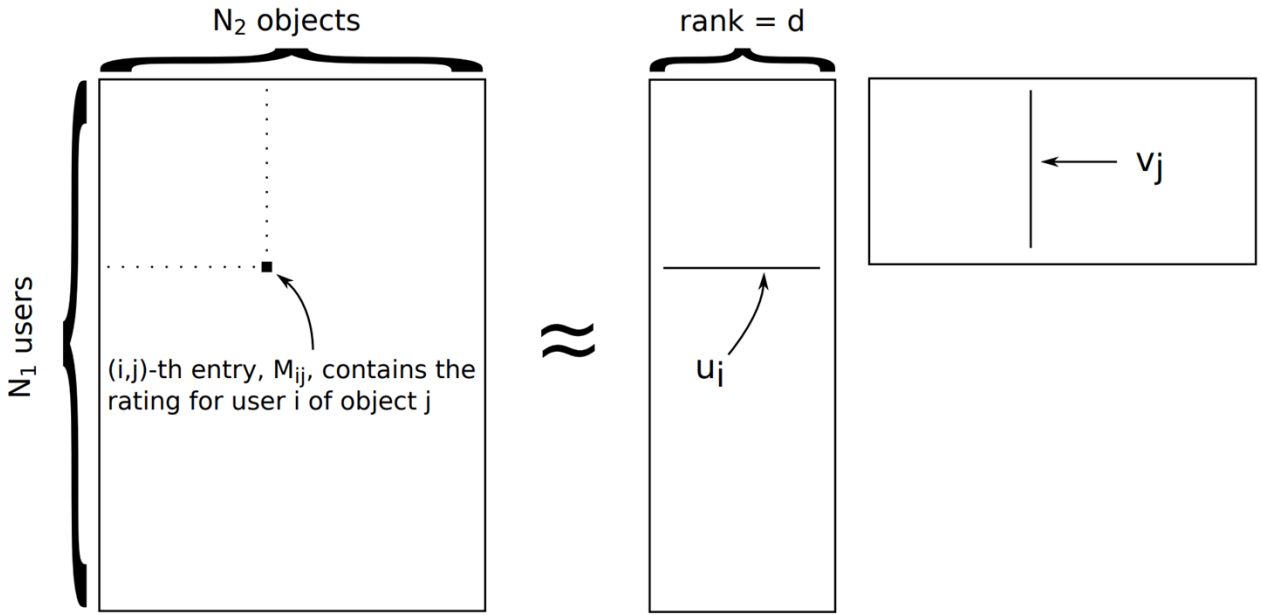
Group 7
Sim, Young js5134@columbia.edu
Sohn, Jongyoon js5342@columbia.edu
Gao, Xin xg2298@columbia.edu
Yang, Siyu sy2796@columbia.edu
Meng, Yang ym2696@columbia.edu

Collaborative Filtering

Use previous user inputs to make future recommendations

- Model-based approach (matrix factorization)

Matrix Factorization



Probabilistic Matrix Factorization

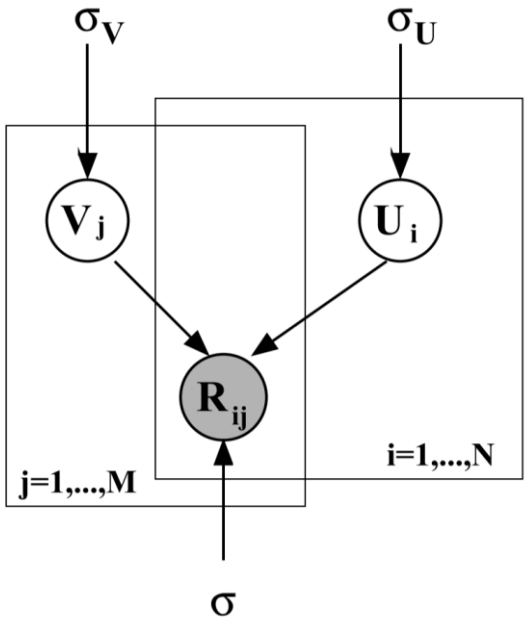
- Assumptions

1. Conditional distribution over the observed ratings:

$$r_{ij}|U_i, V_j, \sigma^2 \sim N(\widehat{r_{ij}}|U_i, V_j\ , \sigma^2)$$

2. $V_j \sim N(0, \sigma_V^2)$
3. $U_i \sim N(0, \sigma_U^2)$

Probabilistic Matrix Factorization Continued



Probabilistic Matrix Factorization Continued

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(R_{ij} - U_i^T V_j\right)^2 + \frac{\lambda_U}{2}\sum_{i=1}^{N} \parallel U_i \parallel_{Fro}^2 + \frac{\lambda_V}{2}\sum_{j=1}^{M} \parallel V_j \parallel_{Fro}^2$$

Post-Processing(KNN)

Define similarity as cosine similarity obtained from

$$s(v_j, v_{j_2}) = \frac{v_j^T v_{j_2}}{||v_j||||v_{j_2}||}$$

Post-Processing(Kernel Ridge)

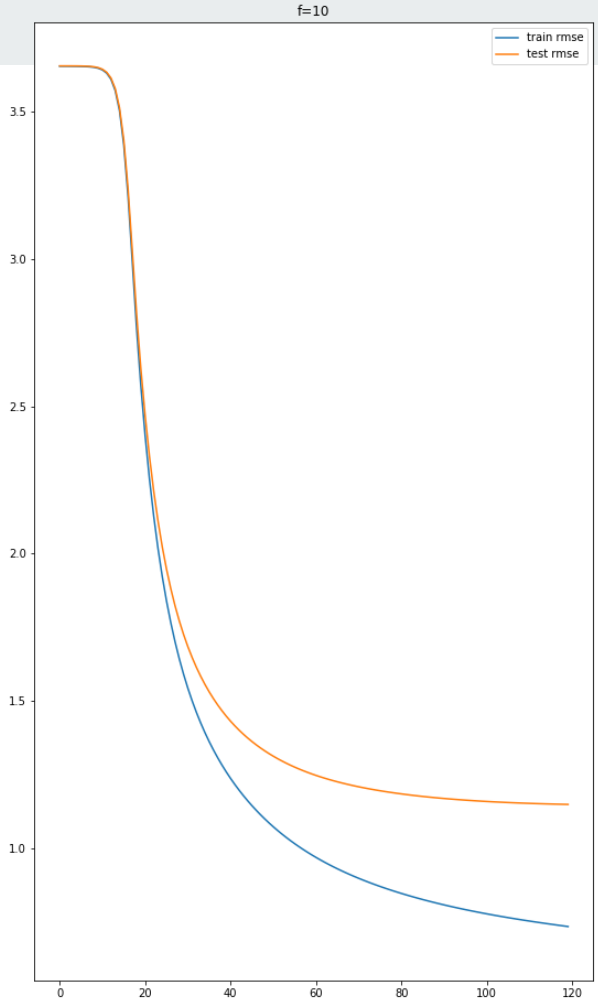
1. Discard all weights from U
2. Define y as vector of user specific ratings
3. X consists of normalized vector of factors for movies rated by the user in each row.
4. Solve Kernel Ridge Regression

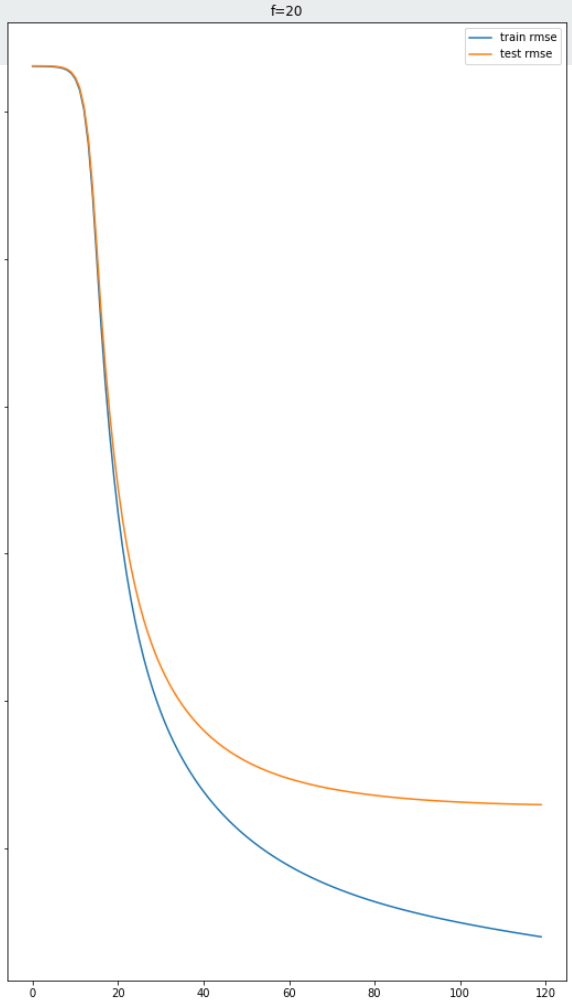
Post-Processing(Kernel Ridge) Continued

Prediction can be obtained by the equation

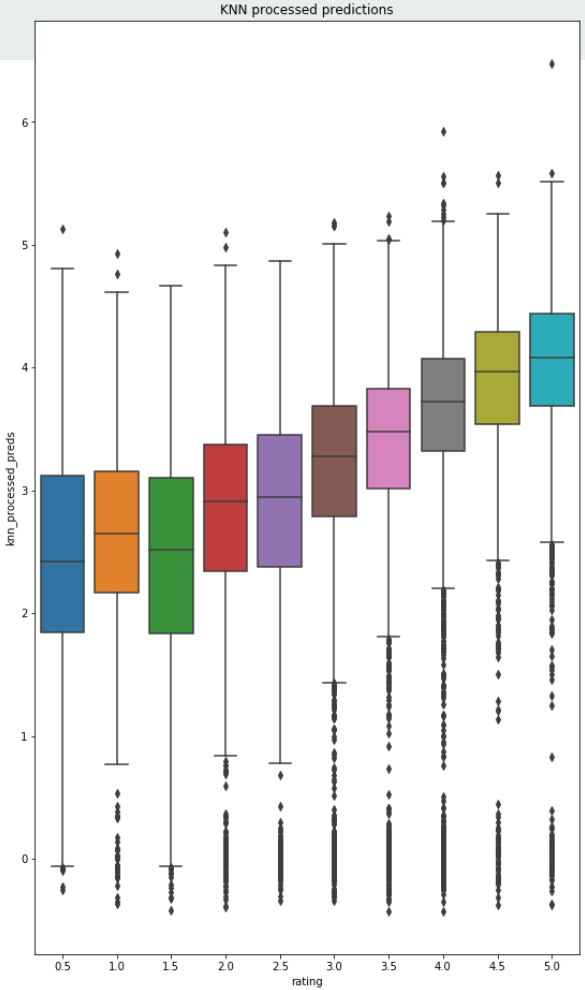
$$\hat{y}_i = K(x_i^{\mathrm{T}}, X)(K(X, X) + \lambda I)^{-1} y$$

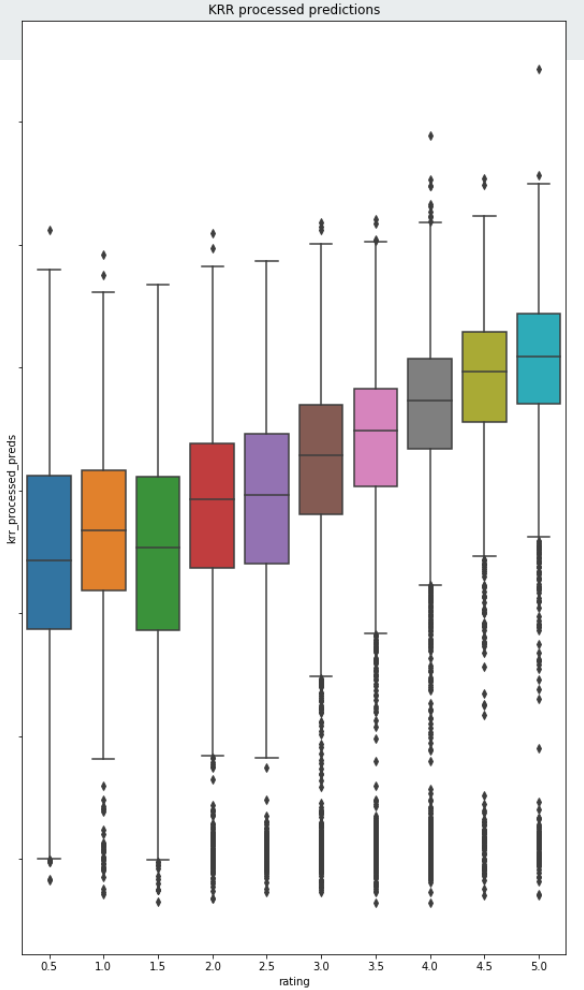
Result-PMF


f=10
train rmse
test rmse
3.5
3.0
2.5
2.0
1.5
1.0
0
20
40
60
80
100
120


f=20
train rmse
test rmse
0
20
40
60
80
100
120

Result-
Post-Processing





Result continued

Post-processing Method	RMSE
PMF	1.1536
PMF with KNN	1.1541
PMF with KRR	1.1433

Questions?

Thank you!