4033/5033 Assignment: Low-Rank Matrix Factorization

Sasank Sribhashyam

In this assignment, we will implement the low-rank matrix factorization technique and evaluate it on a real-world user-movie rating matrix. In this matrix, each row corresponds to a user and each column corresponds to a movie. Ratings take value in {1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5}; you should treat '0' entries in the matrix as missing ratings and do not use it for either training or testing. For convenience, we have separated the set of observed ratings that should be used for training from the set of observed ratings that should be used for testing. The former is stored in 'rate_train.csv' and the latter is stored in 'rate_test.csv'.

Task 1. In the lecture, we have derived the update rule for $U_{i:}$. Now, please derive the update rule for $V_{:j}$, which is formula (3) in the lecture note.

Derivation of Update Rule for V_j

Objective Function:

$$\min_{U,V} \sum_{(i,j)\in S} \delta_{ij}(M_{ij} - U_i \cdot V_j)^2 + \lambda_1 ||U||^2 + \lambda_2 ||V||^2$$

Derivation:

1. First Term Derivative:

$$\frac{\partial}{\partial V_j} \sum_{(i,j)\in S} \delta_{ij}(M_{ij} - U_i \cdot V_j)^2 = -2 \sum_{(i,j)\in S} \delta_{ij}(M_{ij} - U_i \cdot V_j)U_i$$

2. Second Term Derivative:

$$\frac{\partial}{\partial V_j}(\lambda_1 ||U||^2) = 2\lambda_1 V_j$$

3. Third Term Derivative:

$$\frac{\partial}{\partial V_j}(\lambda_2 ||V||^2) = 2\lambda_2 V_j$$

4. Setting the Derivative to Zero:

$$-2 \sum_{(i,j)\in S} \delta_{ij}(M_{ij} - U_i \cdot V_j)U_i + 2\lambda_1 V_j + 2\lambda_2 V_j = 0$$

5. Rearranging Terms:

$$\sum_{(i,j)\in S} \delta_{ij}(M_{ij} - U_i \cdot V_j)U_i = (\lambda_1 + \lambda_2)V_j$$

6. Isolating V_j:

$$V_j = \left(\sum_{i=1}^{p} \delta_{ij} U_i^T U_i + \lambda_2 I\right)^{-1} \left(\sum_{i=1}^{p} \delta_{ij} M_{ij} U_i^T + \lambda_1 V_j\right)$$

Update Rule for V_j

The update rule for V_j is given by:

$$V_j = \left(\sum_{i=1}^{p} \delta_{ij} U_i^T U_i + \lambda_2 I \right)^{-1} \left(\sum_{i=1}^{p} \delta_{ij} M_{ij} U_i^T + \lambda_1 V_j \right)$$

This update rule is equivalent to the formula (3) provided in the lecture notes.

Task 2. The matrix factorization technique we studied in class is called Alternate Least Square (ALS). Please implement it from scratch.

Task 3. Learn a rank-k factorization of the rating matrix from the training matrix ('rate_train.csv') and evaluate it on the testing matrix ('rate_test.csv'). Report testing error versus the number updates in Figure 1. Pick the value for k, λ_1, λ_2 and the maximum number of ALS updates yourself. Tip 1: a common stopping criterion for ALS update is when the training error drops to a certain level. Tip 2: after recovering the matrix using UV, you may round the predicted ratings to the nearest value in $\{1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}$ to get a more meaningful prediction and possibly higher prediction accuracy.

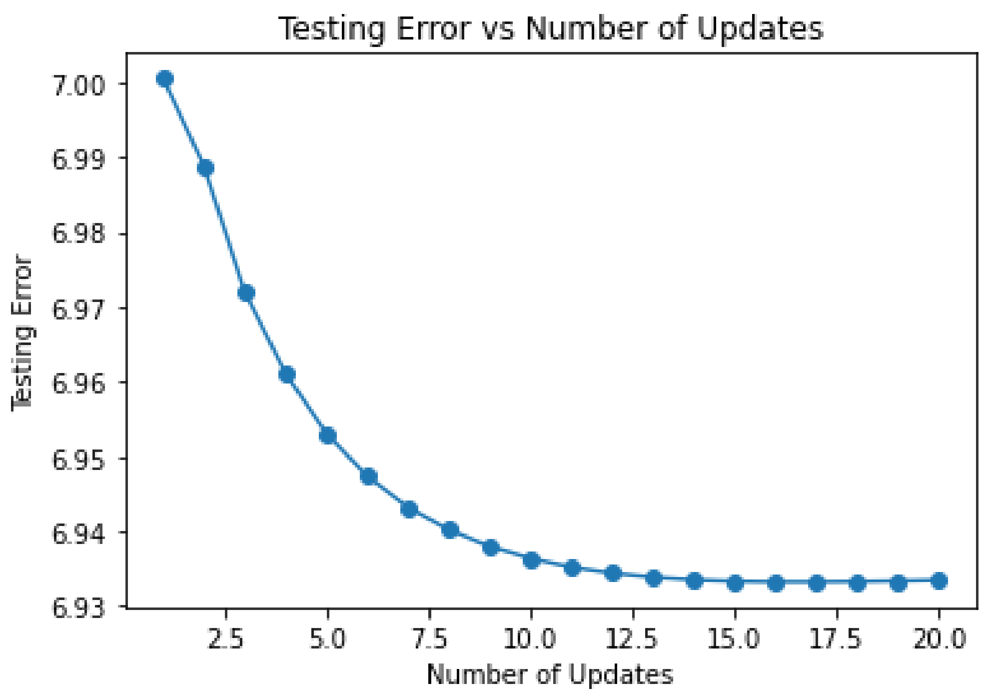

Fig. 1. Testing Error versus ALS Updates with $k = 10$

Task 4. Learn a factorization from the training matrix and evaluate it on the testing matrix. Report the converged testing error versus k in Figure 2. Pick 10 values of k yourself. Tip: a converged testing error is the error that converges in Figure 1 for a given k.

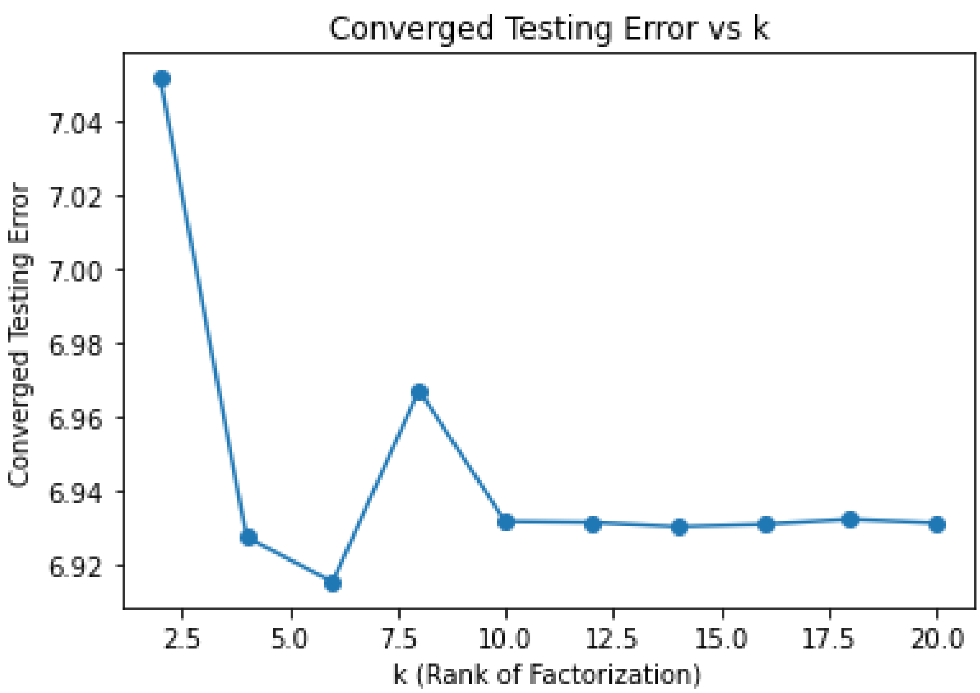

Fig. 2. Testing Error versus k.