9 Recommender System

This assignment is dedicated to helping you understand SGD and recommender systems. You need to download the **ml-latest-small.zip** at https://grouplens.org/datasets/movielens/

9.1 Data Set

Note that the zip file contains side information (e.g. tag applications) that will not be used in the project: we consider only the ratings from the users. Therefore, the first step is to pre-process the data, and organize all the users' ratings as a matrix. Suppose there are n users and p movies. Then the size of the rating matrix M is $n \times p$. Let us denote the index set of observed entries by Ω.

The second step is to divide Ω into two sets Ω_1 and Ω_2: Ω_1 for training and Ω_2 for testing. To this end, we randomly 90 percent of entries in Ω to form Ω_1, and Ω_2 consists of the remaining.

9.2 Learning

Then you will have to solve the following non-convex program to learn the prediction matrix:

$$\min_{U,V} \; F(U,V) := \frac{1}{2} \sum_{(i,j)\in\Omega_1} (M_{ij} - \boldsymbol{u}_i \boldsymbol{v}_j^\top)^2 + \frac{\lambda}{2}(\|U\|_F^2 + \|V\|_F^2) \tag{9.1}$$

where M_{ij} is the (i,j)th entry of M, $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ are the ith and jth row of U and V respectively.

1. For a given index (i,j), derive the stochastic gradient $\frac{\partial F(U,V)}{\partial \boldsymbol{u}_i}$ and $\frac{\partial F(U,V)}{\partial \boldsymbol{v}_i}$.
2. Suppose $\lambda = 1$. Describe the update rule of SGD and implement it with Python. You can randomly initialize all $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$. Note that you need to carefully choose the learning rate.
3. Plot the objective value against the number of iterations, and summarize your findings.

9.3 Evaluation

After we terminate SGD, we will obtain the solution U, V. Our prediction matrix X is then given by $X = UV^\top$. We evaluate the performance of our prediction matrix X by root-mean-square error (RMSE):

$$\text{RMSE} := \sqrt{\frac{1}{|\Omega_2|} \sum_{(i,j)\in\Omega_2} (M_{ij} - X_{ij})^2}.$$

1. Record the RMSE for the choice $\lambda = 1$.
2. Now pick λ from $\{10^{-6}, 10^{-3}, 0.1, 0.5, 2, 5, 10, 20, 50, 100, 500, 1000\}$. For each value, learn and evaluate the your model. Plot RMSE against λ and summarize your findings.