Assignment 6

Vivek Kumar

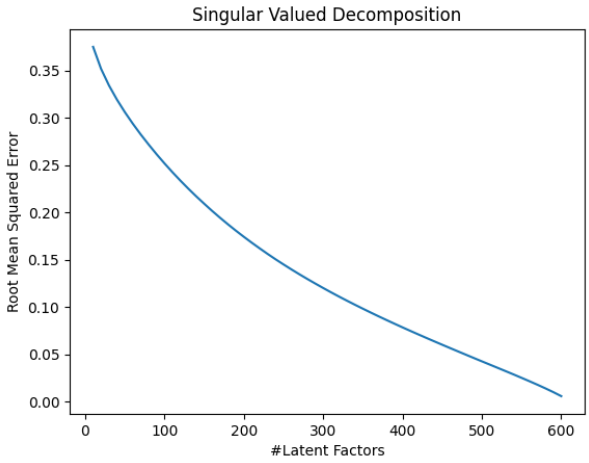
(vivekkumar13@iisc.ac.in)

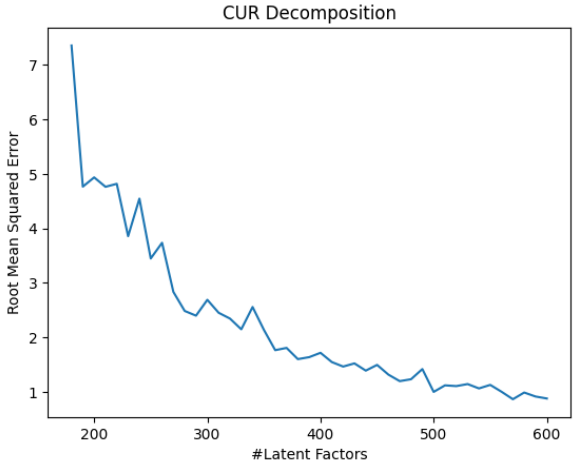
Question 1-5:

I choose 240 for the number of latent factors because when I computed variance captured for different number of latent factors, 240 is capturing more than 90% of variance of original data.

RMSE vs Number of latent factor plots

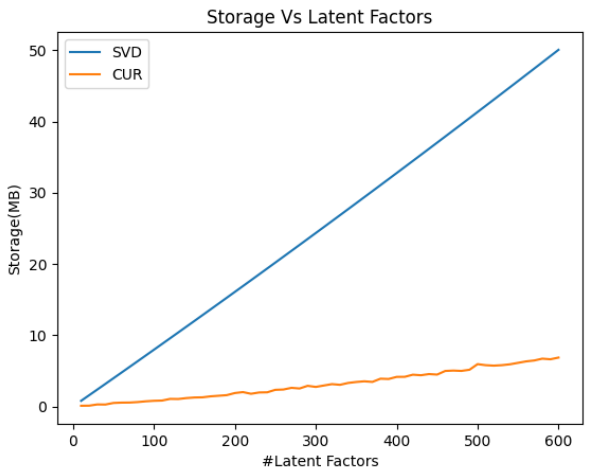
Since CUR is probabilistic algorithm, RMSE can be unexpected in one trial. As CUR is computationally fast, I am running three trials of CUR for given *k* and picking the one with minimum RMSE.

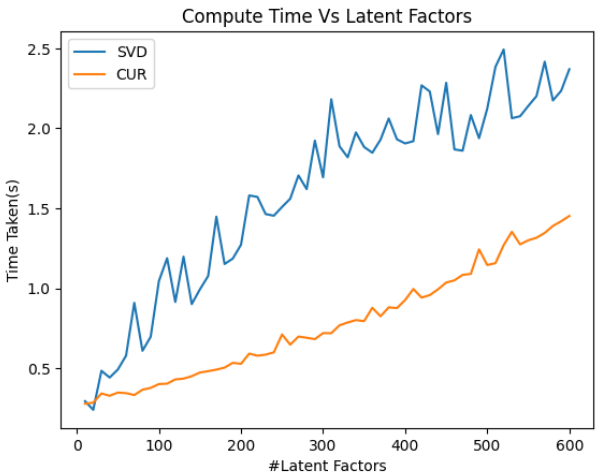




Storage And Compute Time Plots

For computing the space consumed by CUR, I first converted C and R to their sparse representations using *scipy.sparse* as C and R are subset of rows and columns of original matrix respectively, so they will also be sparse. Then I calculated the space consumed by data and corresponding meta data to get the total space required.

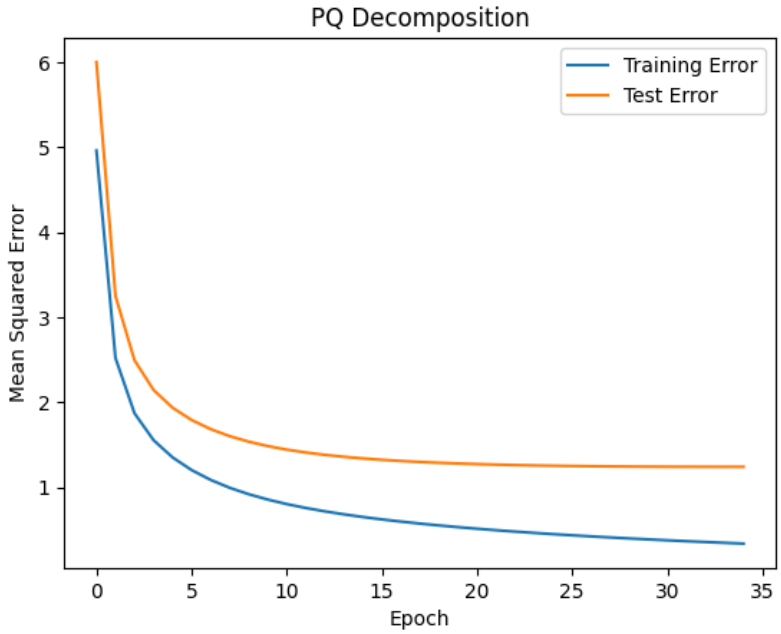




Question 6:

I Tried Different values for the hyperparameters i.e., regularization parameter and learning rate. I got good results at 0.005 for regularization parameter, 0.001 for learning rate and 100 for the number of latent factors.

Here is the Training Error and Test Error I got on different epochs,



Question 7:

I used two-layer multi-layer perceptron as my NCF architecture. First and Second both NCF layers has same number of perceptron which is equal to the 50 and used 0.01 as learning rate.

Dimension for neural embedding of user is M×50 and dimension for neural embedding of item is N×50.

In PQ decomposition I used M×100 and N×100 for P and Q respectively.

While preparing the training and test set, for the negative instances Y^-(i.e., the unobserved interactions) I uniformly sampled them from unobserved interactions in each iteration and control the sampling ratio w.r.t. the number of observed instances Y^+.

At the end of first Epoch,

Training Loss= 0.7093945036288591, Test Loss = 0.7093932232905483