Please provide complete and well-written solutions to the following exercises.

Due October 24, 4PM PST, to be uploaded as a single PDF or Jupyter Notebook document to brightspace.

Homework 6

Exercise 1. In a previous exercise, we performed Alternating Least Squares to check how well we could fill in the unknown entries of `rating_matrix`, which was obtained from the MovieLens 1M dataset. In this exercise, we will compare the performance of ALS to an elementary singular value truncation (SVT).

Here is some code to perform such a truncation on the original rating matrix.

```
U, D_vector, V = np.linalg.svd(rating_matrix)
mean_rating = np.mean(rating_matrix[rating_matrix != 0])
cutoff = 1000
q = np.sum(D_vector > cutoff)
D = np.zeros(rating_matrix.shape)
D[:q, :q] = np.diag(D_vector[:q])
pca_data = U @ D @ V
mean_new_data = np.mean(pca_data[rating_matrix == 0])
output_data = pca_data * (mean_rating / mean_new_data)
output_data = np.round(np.maximum(np.minimum(output_data, 5), 1)).astype('uint8')
```

Compare the performance of ALS to SVT, as in our previous ALS exercise. Which method performs better? Which method is faster? (Try a few different values of `cutoff`.)

(Optional: Repeat the above for a larger MovieLens dataset such as the 10M or 20M datasets.)

Exercise 2. Modify the code from the notes to fill in the data values that appear to be missing from the Google Financial Data plots. Then, re-compute the correlation matrix of the stock prices. Are these tech stocks highly correlated?

Exercise 3. I would like to know the trend of the sales prices of a USC Trojans Jersey on eBay. Do you notice any spikes in the quantity of sales or the price of jerseys that are sold? Can you come up with an explanation for any of your observations?

I would also like to compare these results to the prices of listed (but unsold) items on eBay.

Exercise 4. Do some work on your independent study problem.