IEOR 4500 Application Programming for FE

Assignment 1

Anurag Dutt, Shrey Goel, Jatindeep Singh, Vinayak Shinde, Aditya Zalte

September 24, 2018

This file includes te write-up and explanations for any of the problems in the aforementioned assignment

Problem 1

Since the starter code for reading the russell's matrix was given in python 2-7 we converted the text input file into a csv (code added in eigen.py) into a csv file which we used as a input for all the other parts. We had to do this because we used python 3.6 for programming the assignment problems.

Problem 2

For handling the missing values, we arranged the returns matrix based on ascending order of time. The oldest values on top and the most recent values towards the last. Then we applied forward filling of the NA values such that corresponding cells took the returns of the previous date. We could also have done mean imputation where the cells containing NA values could have beenfilled with the mean of the previous and the next day's returns, but we didn't do mean imputation because it wouldn't have made sense to incorporate the future values into imputation of NA values. Hence we chose to stick with forward filling process.

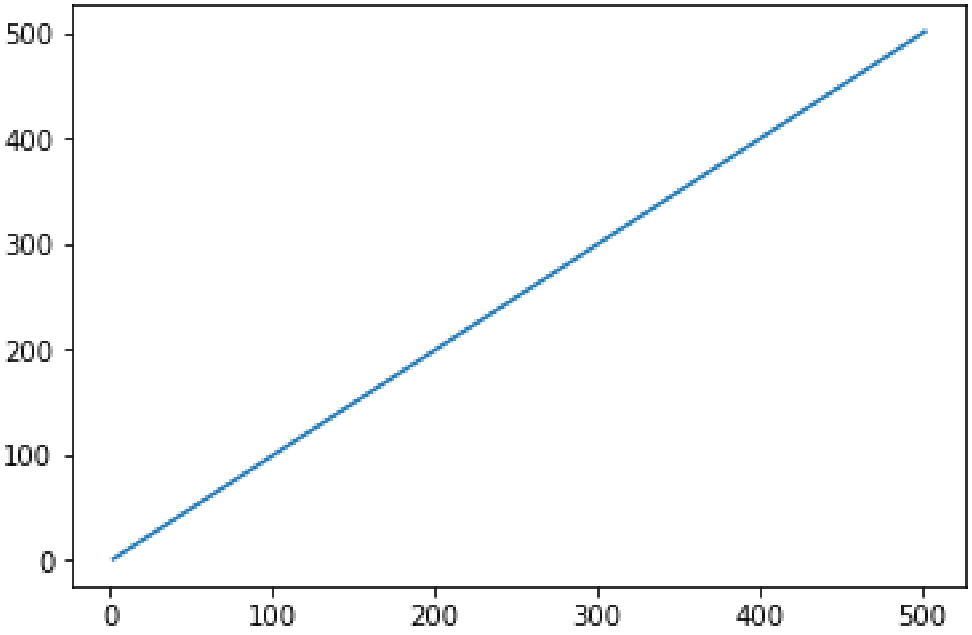
The formula used for variance for jth security was (diagonal elements of the covariance matrix - 947 securities and 504 days):-

$$Var(x_j) = \sigma_x^2 = \sum_{i=1}^{504} (\mu_{ij} - \overline{\mu})^2 \tag{1}$$

The covariance between jth and kth security was calculated as:-

$$Cov(x_j x_k) = \sum_{i=1}^{504} (\mu_{ij} - \overline{\mu_j})(\mu_{ik} - \overline{\mu_k}) \tag{2}$$

Problem 3



The rank of a symmetric matrix is representative of the number of non-zero eigen values of the corresponding vector space. As we increase the length of time-series in consideration for our analysis, the rank of the corresponding matrices increase. Hence we can infer that the number of eigen values increase as we increase the number of days in the sample.

Problem 4

We used A^2 instead of A and it drastically reduced the runtime by almost 5. Thr initial runtime was 5.1 secs which reduced to 0.9 secs.