MF 796 Assignment 4

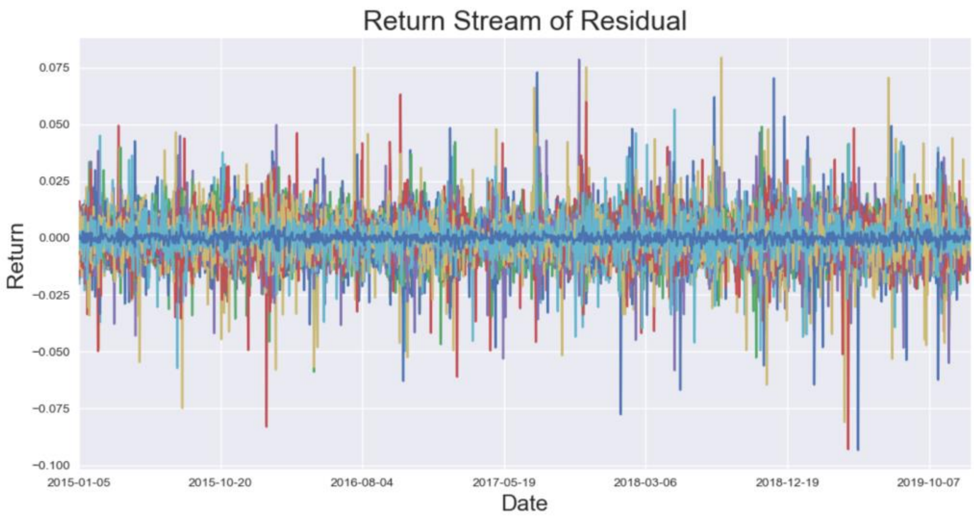
Xinyu Guo
xyguo@bu.edu
U03375769

Problem 1

1. I chose the price data of front 120 stocks from SPY 500 as the sample. There are several days in the data that is Nan, and I filled them using the front-fill method, since the number of missing value is relatively small and disperse.
2. Head of the Stock Return Table:

	A	AAL	AAP	...	SCHW	T	XEC
Date				...			
2015-01-05	-0.018804	-0.000587	-0.013316	...	-0.033962	-0.009475	-0.059677
2015-01-06	-0.015677	-0.015596	-0.000709	...	-0.037419	0.001189	-0.017801
2015-01-07	0.013037	-0.000597	0.021313	...	0.019440	0.001188	-0.008438
2015-01-08	0.029690	0.012071	0.008675	...	0.025586	0.010236	0.024587
2015-01-09	-0.007469	-0.030962	-0.005083	...	-0.034138	-0.003138	0.005257

3. The result of decomposition shows that all of eigenvalues are positive, and none of eigenvalues are negative. Theoretically, the eigenvalues should all be positive since the covariance matrix is a positive definite matrix. If there is any negative eigenvalue, there must be something wrong in data.
4. 8 eigenvalues are required to account for 50% of variance. 71 eigenvalues are required to account for 90% of variance. This makes sense because there is a lot of noise in the return data, which can be reflected by the remaining 10% eigenvalues. And the eigenvalues that account for 90% of variance could be seen as the common factors these stock return have.
5. The figures below show the original stock return and the residual return.



From these figures, we could see that the residual return that remove the impact of top 90% eigenvalues presents a smaller volatility compared with the original return. That is mainly because the residual return mainly consists of the noise and the specific risk of each stock.

Problem 2

In this problem, I use the 'inv' function form numpy.linalg to invert the matrix. And the portfolio allocation (weights of each asset) is presented in the following figure.

From the figure, we can see that some of the asset weights are greater than 1, and some of them even reach 2 to 3. This is not very realistic for most of mutual funds because this allocation demands a very high leverage while the mutual funds are usually not very aggressive and have a relatively high risk aversion. In order to address the problem, I would add an extra constraint to the optimization which requires that each weight of the asset should not exceed some limit (such as 1) to restrict the risk and leverage.