

---

**Problem set # 4**

Due: Wednesday, February 19, by 8am

**Problem 1: Covariance Matrix Decomposition:** Download historical data from your favorite source for 5 years and at least 100 companies or ETFs. In this problem we will look at the covariance matrix for these assets and its properties.

1. Clean the data so that your input pricing matrix is as full as possible. Fill in any gaps using a reasonable method of your choice. Explain why you chose that particular method.
2. Generate a sequence of daily returns for each asset for each date.
3. Calculate the covariance matrix of daily returns and perform an eigenvalue decomposition on the matrix. How many positive eigenvalues are there? How many were negative? If any are negative, what happened?
4. How many eigenvalues are required to account for 50% of the variance? How about 90%? Does this make sense to you?
5. Using the number of eigenvalues in the 90% threshold above, create a return stream that represents the residual return after the principal components that correspond to these eigenvalues have been removed. Plot this return stream and comment on its properties.

**Problem 2: Portfolio Construction:** In Lecture 7, we defined a Lagrangian for portfolio with constraints in matrix form by

$$L(w, \lambda) = \langle R, w \rangle - a \langle w, Cw \rangle - \langle \lambda, Gw - c \rangle \quad (1)$$

1. Form the matrix  $G$  by imposing the budget constraint, which is  $\langle 1, w \rangle = 1$ , and another constraint that allocates 10% of the portfolio to the first 17 securities (to simulate sector allocation). Using  $C$  from Problem 1, use your favorite method and the software package of your choice to invert  $GC^{-1}G^T$  in a nice, stable way. (Hint: consider my favorite method).
2. What does the resulting portfolio look like? Would it be acceptable to most mutual funds? If not, what would you do to fix that?