1. a) In problem set 1 you plotted one of possible trajectories of the process \( \eta(t) \) approximating “white noise”. Now, we plot approximate trajectory of Brownian motion. Compute numerically \( w(t) = \int_0^t \eta(s)ds \approx \sum \eta(t_i)dt \) and plot one of the trajectories. Use \( N = 125 \) and time step \( dt = 0.001 \).

b) Approximate the quadratic variation on the interval \([0,1]\) by \( \sum [w(t_{i+1}) - w(t_i)]^2 \). Compute this number for the plotted trajectory for the case \( t_{i+1} - t_i = 0.01 \). What number do you get and what number should you get theoretically? Repeat the same calculation for \( t_{i+1} - t_i = 0.001 \). Explain your result.

c) For a theoretical Brownian motion compute \( E[|x(t)|] \). Let the interval \([0,1]\) be partitioned into intervals \([t_i, t_{i+1}]\). Set \( t_{i+1} - t_i = h \). What can you say about \( \lim_{h \to 0} \sum |x(t_{i+1}) - x(t_i)| \). Can you support your result by numerical simulation?

d) Generate two independent trajectories \( w_1(t) \) and \( w_2(t) \). Plot a trajectory of two-dimensional Brownian motion in \((x,y)\) plane.

2. Let \( Y, X_1, \ldots, X_k \) be a multivariate, centered, Gaussian random vector. We wish to write \( Y \) in the form
\[
Y = X \hat{\beta} + \alpha + \varepsilon
\]
(i.e. \( Y = \sum_{j=1}^k \beta_j X_j + \alpha + \varepsilon \)), where \( \varepsilon \) is a random variable independent of \( X \) (residual). Show that \( \hat{\beta} \) minimizes the variance of the residual:
\[
E[|Y - X \beta - \alpha|^2] \to \min
\]

a) Show that if the matrix \( C = \text{cov}(X, X) \) is non-degenerate the above minimization problem has a unique solution with \( \hat{\beta} = [\text{cov}(X, X)]^{-1} \text{cov}(X, Y) \).
Give a simple numerical example showing that the solution is not unique if the matrix is degenerate.
b) If the matrix $C$ is degenerate we introduce a penalty function $\frac{1}{2} \beta' \beta$ and minimize

$$E \left[ Y - X \tilde{\beta} - a \right]^2 + \gamma \frac{1}{2} \beta' \beta \rightarrow \min.$$ 

Show that this problem has a unique solution $\tilde{\beta}$. Do you think that $\tilde{\beta}$ approaches a limit as $\gamma \rightarrow 0$? If so, characterize the limit.

c) Repeat (b) with a penalty function of the form $\frac{\gamma}{2} \beta' \Omega \beta$, where $\Omega$ is a positive definite symmetric matrix. Using the degenerate example you gave in a), describe how the penalty function method remedies the degeneracy problem.

3. Download historical daily closing prices for the Dow Jones Industrials Average (^DJX) for the last two years, as well as daily closing prices for the 30 component stocks over the same period. [The data can be obtained from finance.yahoo.com.]

a) Using the data, compute the empirical covariance matrix of returns for the 30 Dow components.

b) Use SVD to find eigenvalues of $\hat{C}$. How many components do you need to explain 90% of the variance? 95%? 99%?

c) Rank the stocks by their market capitalization (again, you can use Yahoo!) and do a linear regression of the DJX index returns of the returns of 5 stocks with the largest market capitalization. Use any statistical package or plug-in to do the linear regression analysis. [You may want to use the formula from problem 2, however in this case you are responsible for handling possible degeneracy.] Give the variance of residuals for your regression.

d) Repeat c) for 10, 20 and 30 stocks, computing each time the weights and the variance of the residual.

e) Compare the regression weights obtained with the 30 stocks with the actual percentage weights of the components in the index. What do you observe?