Portfolio Management - Applications
Exercise 1: Minimum Variance Portfolios

Date due: November 28, 2018

Instructions: Work in groups of max 2 students.
Upload your solution to Learn@WU. Name the uploaded files
`Assignment1_NameFirstname1_NameFirstname2.pdf`,
....Rnw, ...R, etc. It is sufficient if one group members uploads the zip file.
The pdf file is your report. It has to include results, text answers, references, charts etc. It is very important that you interpret your results. Relate your results to the paper assigned as a reading and perhaps other literature. Results that are too cumbersome (e.g. a 500×500 matrix) should be saved in a separate data file along with other data. In the text you have to state clearly where those results can be found.

Each group member has to indicate individually whether he/she would like to present the solution in class. Both group members can (and are encouraged to) indicate their willingness to present. This counts towards the grade!

Goal: In this exercise you will construct and analyze minimum variance portfolios, using historical data from S&P 500 index constituents.

Data: I provide a data set `data_ex1_covariance_20181122.Rdata` that contains

- `SP`. This dataframe gives for companies information on membership in the S&P 500 index (gvkeyx 000003). If the variable `thru` is missing, the company is still in the index. Note that sometimes companies has several issues traded, see issue id `iid`.
- `DT.members`. This is almost the same information presented differently. The variable `ym` is a `zoo yearmon` variable. If a `gvkey/iid` was in the index for at least one day in a given calendar month, there is an appropriate row in this dataframe. The index information here starts in 1970.
- `gvkeys`. A vector of the gvkeys in character format.
- `cnam`. A dataframe that gives the most recent company name `conm` and industry classification `sich` for the `gvkey/iid`.

- `returns`. A dataframe that gives the return in percent `trt1m` for a `gvkey/iid` in a calendar month `ym`.
- `spx`. A dataframe that gives the monthly return `ret` of the S&P 500 as a decimal number, and a total return index `tri` that starts end of 1969 with a value of 1.

Literature: The main reference for this exercise is Dangl and Kashofer (2013). In section 4 of their paper a number of methods are described that allow to obtain estimated covariance matrices of full rank. In this exercise we will use two of them - Shrinkage Towards the Average-Correlation Estimator (SAC) and Principal Component Estimator (PC).

Shrinkage Towards the Average-Correlation Estimator (SAC): The idea behind the Average-Correlation Estimator (AC) is to avoid errors in the covariance matrix by assuming that the pairwise correlation of asset returns is identical over all pairs of stocks. It imposes following structure on the covariance

$$\hat{\Sigma}_{AC} = \Delta C \Delta, \quad \text{with } \Delta = diag(\hat{\sigma}_i),\ C = \begin{pmatrix} 1 & \hat{\rho} & \cdots & \hat{\rho} \\ \hat{\rho} & 1 & \cdots & \hat{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho} & \hat{\rho} & \cdots & 1 \end{pmatrix},$$

where $\hat{\sigma}_i$ is the estimated standard deviation of stock i's return and $\hat{\rho}$ is the sample estimate of the mean of the pair-wise correlations.

In the SAC we combine both the sample covariance matrix $\hat{\Sigma}_S$ and the restricted estimate $\hat{\Sigma}_{AC}$ in following way

$$\hat{\Sigma}_{SAC} = \delta_{SAC}\hat{\Sigma}_{AC} + (1 - \delta_{SAC})\hat{\Sigma}_S$$

The Principal Component Estimator (PC): This approach uses the eigenvalues decomposition of the sample covariance matrix. The idea behind this method is that usually only a small number of eigenvectors explain almost the entire sample variance. So PC selects the n largest eigenvalues and their corresponding eigenvectors. The estimate has following form

$$\hat{\Sigma}_{PCn} = U'\Lambda_n U + \Sigma_{I,PC,n}$$

where U is the matrix that contains the orthogonal eigenvectors of $\hat{\Sigma}_S$ in its rows. Λ_n is diagonal matrix containing the n largest eigenvalues, other diagonal elements are replaced by zero. The matrix $\Sigma_{I,PC,n}$ is a diagonal matrix containing the residual variance from projecting the individual returns onto the subspace spanned by the first n eigenvectors.

Specific Tasks: First use only a recent index composition (as of June 2018) and monthly return data over the last 5 years (60 months ending in June 2018).

1. Describe the data. Are there missing values? Decide on a strategy to deal with missing values and discuss possible consequences.
2. Compute three estimates of the covariance matrix:
   (a) sample covariance matrix $\hat{\Sigma}_S$
   (b) SAC estimate $\hat{\Sigma}_{SAC}$ using $\delta_{SAC} = 0.5$
   (c) PC estimate $\hat{\Sigma}_{PC}$ using $n = 3$
3. A covariance matrix has to be symmetric and positive semi-definite. For further computation we also need nonsingularity (full rank). Check if $\hat{\Sigma}_S, \hat{\Sigma}_{SAC}$ and $\hat{\Sigma}_{PC}$ have these properties. If not give an explanation.
4. Compute the weights of the stocks in the minimum variance portfolio using $\hat{\Sigma}_{SAC}$ and $\hat{\Sigma}_{PC}$
   (a) with no restriction
   (b) with restriction on short sales $w \geq 0$

   Round the weights to 4 digits.
5. Describe the minimum variance portfolios
   (a) How many of the stocks have positive, zero and negative weights?
   (b) How similar/different are the MVP using different methods? What are the most important stocks and their weights?
   (c) Calculate the average monthly return and standard deviation of each portfolio and compare it with the index.
6. Is it a good choice to use of 5 years of monthly data for the calculation? What are the pro and cons for the use of shorter / longer periods, and higher/lower frequency?

For the next questions, use the whole data set.

- Calculate the MVP each month, starting at the point in time when you have 5 years of historic data for the stocks that are in the index at your starting point. Each subsequent month, stick to 5 years of monthly data for calculation of covariances (so-called rolling windows).
- Calculate the performance of each portfolio that you have formed over the subsequent month. This way you get a time series of returns that an investor could have obtained in real-time.
- Compare the portfolio returns and standard deviations to the S&P 500 index and an equally weighted (1/N) portfolio. Plot diagrams.
- Interpret and discuss.

Literature

- Dangl, Kashofer (2013), Minimum-Variance Stock Picking - A Shift in Preferences for Minimum-Variance Portfolio Constituents, Working paper
- DeMiguel, Garlappi, and Uppal (2009), Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? The Review of Financial Studies, 22(5), 1915-1953.
- Connor, Goldberg, Korajczyk (2010), Portfolio Risk Analysis, Princeton University Press.