

# Covariance Estimations<sup>1</sup>

---

**UCLA Anderson MFE MGMT 277H**  
**Quantitative Asset Management**  
**Professor Jason C. Hsu**

## Introduction

The covariance matrix of stock returns is one of the major inputs to the mean-variance optimization process. In particular for the minimum-variance strategy, in which maximizing expected return is not part of the objective, the resulting portfolio is highly sensitive to the covariance matrix estimate. A mis-specification of covariance matrix could lead to undesirable portfolio.

This document covers a brief overview of how to implement various methods of estimating the covariance matrix. You are highly encouraged to study the reference material for theoretical details.

See the SAS IML Introduction document for a guide on linear algebra computation in SAS.

## Sample Covariance Matrix

$$\hat{\Sigma} = \frac{1}{N-1} (\overline{R^e}' \overline{R^e})$$

Where  $\overline{R^e}$  is an N-sample by K-asset matrix of demeaned excess returns.

Estimating covariances directly with recent historical returns without any robust statistical methods is the easiest implementation. The drawback is its sensitive to outliers in data samples, causing the portfolio optimizer to tilt aggressively towards stocks that recently co-move less with the market.

Unlike other methods introduced below, this method does not suffer from any model risk as it makes no assumption regarding the return time series or the covariance matrix structure. It is a sensible method for estimating covariances of diversified portfolios or when number of available data sample (N) is a lot higher than the number of assets (K). However, it should not be used for estimating covariance matrix of a large set of stocks.

---

<sup>1</sup> This material is prepared by Tzee-man Chow, UCLA MFE Class of 2009. Tzee-man is a researcher at Research Affiliates®, LLC.

## Estimation by Diagonal Model (a.k.a. Single Index Model)

Assuming co-movement of all assets is due to one common factor, as in CAPM, excess returns can be decomposed to

$$R^e_i = \alpha_i + \beta_i R^e_{Mkt} + \varepsilon_i$$

With assumptions

$$E[\varepsilon_i] = 0 \quad E[\varepsilon_i R^e_{Mkt}] = 0 \quad E[\varepsilon_i \varepsilon_j] = 0$$

Then,

$$\sigma_i^2 = V(\beta_i R^e_{Mkt} + \varepsilon_i) = \beta_i^2 \sigma_{Mkt}^2 + 2cov(\beta_i R^e_{Mkt}, \varepsilon_i) + V(\varepsilon_i) = \beta_i^2 \sigma_{Mkt}^2 + \sigma_{\varepsilon,i}^2$$

and

$$\sigma_{i,j} = cov(\beta_i R^e_{Mkt} + \varepsilon_i, \beta_j R^e_{Mkt} + \varepsilon_j) = \beta_i \beta_j \sigma_{Mkt}^2$$

Written in matrix form,

$$\hat{\Sigma} = \beta \sigma_{Mkt}^2 \beta' + D_\varepsilon$$

Where  $\beta$  is a column vector of beta,  $D_\varepsilon$  is a diagonal matrix of  $\sigma_{\varepsilon,i}^2$ .

See Chapter 8 of Professor Jorion's "Value At Risk" for more details.

## Estimation by Multi-Factor Model

A better precision than single-index model can be achieved if more than one factor can be identified for describing stock return co-movements.

$$R^e_i = \alpha_i + \beta_{i,1} F_1 + \beta_{i,2} F_2 + \cdots + \beta_{i,L} F_L + \varepsilon_i$$

With assumptions

$$E[\varepsilon_i] = 0 \quad E[\varepsilon_i F_l] = 0 \quad E[\varepsilon_i \varepsilon_j] = 0$$

Then

$$\hat{\Sigma} = \beta \Sigma_F \beta' + D_\varepsilon$$

Where  $\beta$  is a K-asset by L-factor matrix of betas,  $\Sigma_F$  is the L by L covariance matrix estimate of factors,  $D_\varepsilon$  is a diagonal matrix of  $\sigma_{\varepsilon,i}^2$ .

Assuming the selected factors are the Fama-French market, size & value factors, each factor represents returns of diversified traded portfolios, then  $\Sigma_F$  can be estimated as sample covariances.

The model can be further simplified if factors are assumed to be uncorrelated ( $E[F_m F_n] = 0$ ),

$$\hat{\Sigma} = \beta D_F \beta' + D_\varepsilon$$

Where  $D_F$  is a diagonal matrix of  $\sigma_{F,l}^2$ .

## Estimation by Principal Component Analysis

Without relying on any theoretical assumption, principal component analysis (PCA) can be used to determine the underlying drivers of the stock returns. The PCA method transforms the vector space of  $N$  assets into another vector space of  $N$  factors by singular value decomposition (SVD) of the sample covariance matrix. Each factor, an eigenvector from the SVD, represents a linear combination of the original  $N$  assets, and the factors are uncorrelated by definition, with variances equal to the eigenvalues from the SVD.

Asset returns and sample covariance matrix can be written as

$$R^e_i = \beta_{i,1}F_1 + \beta_{i,2}F_2 + \cdots + \beta_{i,N}F_N$$

$$\hat{\Sigma} = \beta D_F \beta'$$

Where  $\beta$  represents  $N$  columns of eigenvectors, and  $D_F$  is the  $N$  by  $N$  diagonal matrix of eigenvalues.

PCA is often employed to reduce dimensionality of the data. If the first  $L$  factors govern most of the variability of the asset returns, i.e. if  $\sum_{l=0}^L \sigma_{F,l}^2 / \sum_{l=0}^N \sigma_{F,l}^2$  is very close to 1, then the last  $N - L$  factors shall be dropped,

$$\hat{\Sigma} = \tilde{\beta} \tilde{D}_F \tilde{\beta}' + D_\varepsilon$$

Where  $\tilde{\beta}$  is the  $N$ -asset by  $L$ -factor matrix of factor loadings (first  $L$  eigenvectors),  $\tilde{D}_F$  is the  $L$  by  $L$  diagonal matrix of the first  $L$  eigenvalues, and  $D_\varepsilon$  is the  $N$ -asset by  $N$ -asset diagonal matrix of variances of idiosyncratic components not explained by the first  $L$  factors.

See Chapter 8 of Professor Jorion's "Value At Risk" for more details.

## Estimation by Bayesian Shrinkage

Instead of assuming an asset pricing model on stock returns, the Bayesian shrinkage assumes the covariance matrix itself follow certain structure:

$$\hat{\Sigma} = \lambda \Sigma_{prior} + (1 - \lambda) \Sigma_{sample}$$

Where  $\Sigma_{sample}$  is the sample covariance matrix,  $\Sigma_{prior}$  is the target covariance matrix, and  $\lambda$ , expected to be within 1 & 0, is the intensity of shrinking sample estimate towards the target.

Specifically, a fairly stringent target covariance may be one that is populated by only two constants – assuming the true variance of all assets equal to the average of all sample variances, and the true covariances of all pairs of assets equal to the average of all sample covariances,

$$\Sigma_{sample} = R^e R^{e'} \quad \Sigma_{prior} = \begin{bmatrix} v & \cdots & c \\ \vdots & \ddots & \vdots \\ c & \cdots & v \end{bmatrix}$$

Where  $R^e$  is the N-sample by K-asset matrix of excess returns,  $v$  is the average of the diagonal elements of  $\Sigma_{sample}$  and  $c$  is the average of the off-diagonal elements of  $\Sigma_{sample}$ . Then, the optimal shrinkage intensity is

$$\lambda = \frac{SUM[SQ(R^e)'SQ(R^e)] - SUM[SQ(\Sigma_{sample})]}{N \times SUM[SQ(\Sigma_{sample} - \Sigma_{prior})]}$$

Where  $SQ(\ )$  is the element-by-element squaring function for a matrix argument, and  $SUM[ \ ]$  is the sum of the argument matrix elements.

See Ledoit & Wolf [2004] for theoretical basis detail analytical derivations.

See Clarke, de Silva & Thorley [2006] for application and empirical results of applying the said Bayesian Shrinkage method to form minimum variance portfolio.

## References

- Clarke, Roger G., Harindra de Silva, and Steven Thorley. 2006. "Minimum-Variance Portfolios in the U.S. Equity Market." *Journal of Portfolio Management*, vol. 33, no. 1 (Fall): 10—24.
- Jorion, Philippe. 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3<sup>rd</sup> ed. ISBN 0-07-146495-6.
- Ledoit, Olivier, and Michael Wolf. 2004. "Honey, I Shrunk the Sample Covariance Matrix." *Journal of Portfolio Management*, vol. 30, no. 4 (Summer): 110—119.