FACTSET 〉 SEE THE ADVANTAGE

ROBUST ESTIMATION OF RISK FACTOR MODEL COVARIANCE MATRIX

Dmitri Mossessian
dmossessian@factset.com

Viviana Vieli
vvieli@factset.com

Robust Estimation of Risk Factor Model Covariance Matrix

Contents

Introduction . . . 2
Covariance matrix of a risk factor model . . . 2
Covariance matrix estimation using random matrix theory (RMT) . . . 3
Simple example . . . 4

Introduction

Every multifactor risk model is based on a factor covariance matrix that is employed to reproduce, model, and analyze the joint distribution of risk factor returns. The moments of that distribution are widely used by market practitioners as measures of risk. The factor covariance matrix is computed using historical time series of risk factor returns, and the limited number of samples in these time series always leads to an estimation errors in covariance matrix itself. If a relatively small number of historical samples (comparable to the number of factors in the model) are used to estimate the matrix, this error may become very large. The presence of this noise in the matrix not only decreases the accuracy of the risk forecasts, but can also cause the matrix estimator to have undesired properties that prevent such forecasts altogether. In particular, the estimator of the covariance matrix may not be positive definite, rendering the matrix useless for Monte Carlo simulation of risk. This note addresses the problem of noise in a factor model covariance matrix and outlines the method of finding the optimal matrix estimator that is based on random matrix theory. The method is similar to a simple regularization method proposed by Ribonato in [1]. The advantage of the proposed method is that it not only makes the resulting estimator positive definite, but also reduces the amount of noise in the estimator and minimizes the differences between the estimator and (generally unknown) true covariance matrix of the model.

Covariance matrix of a risk factor model

Factor models are widely used to describe the return process of financial securities and analyze the risks of portfolios. They allow us to reduce the dimensionality of the problem and at the same time help to define major economic sources of risk of portfolios. A factor model decomposes the total return of a security into a sum of systematic returns that are due to movements of common factors, and an idiosyncratic component. The idiosyncratic returns are generally independent of the systematic return and uncorrelated, so that correlations between securities are defined by the correlations between systematic risk factors. The following equations describe the components of security return in a linear factor model (a nonlinear model might have more complex relationship between factor returns and security returns, but the basic relationship between factor covariances and portfolio risk remains the same). A linear factor model expresses return r_j of a security S_j as a linear combination of factor returns $f_i, i = 1, ...M$, where M is the number of factors in the model:

$$r_j = \sum_i^M l_{ij} f_i + \varepsilon_j \tag{1}$$

Here l_{ij} is the loading (sensitivity) of the security j to the factor f_i, and ε_j is the idiosyncratic component of the security return (the portion of the return not explained by systematic factors). If we introduce the vector of loadings $\mathbf{l}_j^T = (l_{1j}, ...l_{Mj})$ for each security, and define the matrix of loadings $\mathbf{L}$ that consists of columns $l_j j$ ($M \times K$ matrix), the vector of security returns $\mathbf{r}_T = (r_1, ...r_K)$ for all securities in the portfolio will be:

$$\underset{K\times 1}{\mathbf{r}} = \underset{K\times M}{\mathbf{L}^T} \times \underset{M\times 1}{\mathbf{f}} + \varepsilon \tag{2}$$

where ε is the vector of ε_j components. Thus, a portfolio return can be written as:

$$r_p = \mathbf{w}^T (\mathbf{L}\mathbf{f} + \mathbf{E}) = (\mathbf{L} \times \mathbf{w})^T \times \mathbf{f} + \mathbf{w}^T \mathbf{E} \tag{3}$$

where **E** is a diagonal matrix with ε_j on diagonal and the vector $\mathbf{w}^T$ consists of weights of securities in the portfolio. Each factor return f_i is a random variable, and a factor model of returns is basically a model that forecasts the joint distribution of all the factors in the model at certain time horizon T. This forecast allows us to use Eq.(3) to model the distribution of portfolio return at that horizon, which is fully described by the covariance matrix of the factor return $\mathbf{\Sigma}$. For example, volatility of portfolio return can be computed as:

$$\sigma_p^2 = (\mathbf{L} \times \mathbf{w})^T \times \mathbf{\Sigma} \times (\mathbf{L} \times \mathbf{w}) + \mathbf{w}^T \times \mathbf{E} \times \mathbf{w} \tag{4}$$

where **E** is the diagonal matrix of idiosyncratic return volatilities.

Thus, estimating the factor model is equivalent to building an estimator for the factor covariance matrix. The only available information we have for that are the historical time series of factor returns, which effectively represent the samples of random processes $f_i(t)$, and the problem at hand is the problem of building the best estimator for a sample covariance matrix. This problem is best addressed from the point of view of the random matrix theory.

Covariance matrix estimation using random matrix theory (RMT)

Random matrix theory (a good review of the theory can be found in [2]) states that if **X** is $N \times M$ random matrix with i.i.d. standard normal elements, the eigenvalues λ of the correlation matrix $\hat{\mathbf{\Sigma}} = \mathbf{X}^T\mathbf{X}$ are distributed with the following probability distribution function:

$$\rho_\lambda = \frac{c}{2\pi\lambda}\sqrt{(\lambda - \lambda_{min})(\lambda_{max} - \lambda)} \tag{5}$$

where $c = \frac{N}{M}$ is the ratio of the number of samples to the number of variables, and the boundaries for the eigenvalues are

$$\lambda_{max,min} = \left(1 \pm \sqrt{\frac{1}{c}}\right)^2 \tag{6}$$

All eigenvalues of the matrix are concentrated in the region between λ_{min} and λ_{max}—the support region of the eigenvalue probability distribution function. Since the matrix $\mathbf{X}^T\mathbf{X}$ is the sample estimator of the correlation matrix of independently distributed standard normals, it does converge to the identity matrix when number of samples goes to infinity. That corresponds to the case of $c \to \infty$, when all eigenvalues are equal to one and $\hat{\mathbf{\Sigma}}$ is the true estimator of the correlation matrix of M independent normals. But when the number of samples is limited, c becomes smaller, the size of the support region widens, and the matrix $\hat{\mathbf{\Sigma}}$ develops random off-diagonal elements. Thus the width of the support region effectively identifies the range of errors in the matrix estimator due to the limited sample size. Let us look now at an M-factors risk model that have a true correlation matrix $\mathbf{\Sigma}(M \times M)$ (we can always normalize the covariance matrix to the factor volatilities to obtain the correlation matrix). The sample correlation matrix of the factor returns can then be written as (assuming that the factors have zero mean):

$$\hat{\mathbf{\Sigma}} = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{X}^T\mathbf{X}\mathbf{\Sigma}^{\frac{1}{2}} \tag{7}$$

It is clear from this equation that the sample correlation $\hat{\mathbf{\Sigma}}$ will be equal to the actual matrix $\mathbf{\Sigma}$ only if the product $\mathbf{X}^T\mathbf{X}$ equals the identity matrix. In the finite sample size case, the noise in the matrix $\mathbf{X}^T\mathbf{X}$ will determine the errors in the full model covariance matrix estimator $\hat{\mathbf{\Sigma}}$. The distribution of eigenvalues of the sample correlation matrix will be wider than in the case of a pure random matrix as there will be a number of large eigenvalues lying outside of the support region of a random matrix. A simple and intuitive assumption to make is that the components of the correlation matrix that are defined by the small eigenvalues within the random matrix support region (i.e. are orthogonal to the space of the large eigenvalues) are dominated by noise. In other words, only the eigenvalues of the sample matrix $\hat{\mathbf{\Sigma}}$ that lie outside of the support region $[\lambda_{min}, \lambda_{max}]$ contain information relevant to the actual correlation matrix.

The underlying assumption we are going to use to construct the optimal estimator is that the eigenvectors of the optimal estimator are the same as the eigenvectors of the sample estimator $\hat{\mathbf{\Sigma}}$ (for detailed discussion see [3]). This assumption

allows us to construct the optimal matrix estimator by doing spectral decomposition of the original estimator, adjusting the eigenvalues and reconstructing the optimal estimator using the same eigenvector matrix. This, in fact, is similar to the method proposed by Ribonato in [1], with the only difference being the algorithm for eigenvector adjustment. The adjustment we use is similar in nature to principal component analysis when only the principal components that carry information are used, and the ones that carry only the noise are discarded. In the same spirit we construct the optimal estimator of the correlation matrix using only the eigenvalues outside of the RMT support region $\lambda_i \in U \equiv \{\lambda_i > \lambda_{max}\}$:

$$\hat{\mathbf{\Sigma}} = \sum_{\lambda_i \in U} \lambda_i \mathbf{q}_i \mathbf{q}_i^T + diag \left\{ \mathbf{I} - \sum_{\lambda_i \in U} \lambda_i \mathbf{q}_i \mathbf{q}_i^T \right\} \tag{8}$$

where $\mathbf{q}_i$ are the eigenvectors of the unadjusted correlation matrix. The second term in the equation (8) replaces the diagonal elements of the new estimator with ones. This will ensure that the estimator can be used as a correlation matrix - that it is positive definite and its diagonal elements are all 1.

Simple example

In this section we discuss a simple example that illustrates the approach and its results and talk about the implication of the proposed matrix regularization method (RMT method) for portfolio risk measures. To illustrate the effect of finite sampling and subsequent regularization on a correlation matrix let us look at a (100×100) matrix with all correlations equal 0.1. We use this matrix to construct the estimator $\hat{\mathbf{\Sigma}}$ using a (100×1000) set of random samples $\mathbf{X}$ drawn from standard normal distribution using Equation (7). Because of the finite number of samples, the resulting correlation coefficients vary between -0.1 and 0.27.

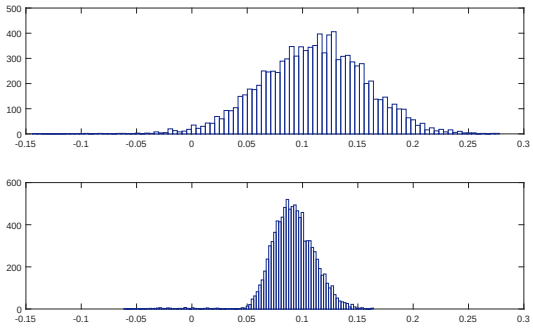

Figure 1: Distributions of correlation values in the sample matrix before (top) and after (bottom) regularization

The top panel of Figure (1) shows the distribution of the elements of the matrix $\hat{\mathbf{\Sigma}}$ (excluding diagonal elements) before regularization is applied. While the distribution of the true correlation matrix should be a single vertical line at 0.1, the estimation noise results in the distribution approximately centered around the true value, but having a final width (standard deviation) of about 0.05. The bottom panel of that figure shows the same distribution after we applied the regularization procedure described in previous section. Clearly, the amount of noise in the correlation matrix (the width of the distribution) is significantly reduced.

The parameters of the distributions of the values of matrix elements before and after regularization are summarized in Table (1). When the correlation matrix is reconstructed using the regularized eigenvalues the scatter of the correlation values around its true value of 0.1 is reduced by more than a factor of two. The method moves the sample correlations much closer to their true value by eliminating random noise in eigenvalues.

	Before	After
μ	0.11	0.091
σ	0.047	0.021

Table 1: Mean and standard deviation of all elements of a constant correlation matrix before and after regularization

Applying a new regularization method to the risk model covariance matrix will inevitably change the characteristics of the resulting joint distribution of the factors and , ultimately, affect the estimated portfolio risk measure. To gauge the effect of the new regularization method on actual portfolio VaR we ran the risk model on several test portfolios. Table (2) shows the results of the tests for Dow Jones Industrial Index (with S&P500 index as a benchmark) and for Barclays Global Aggregate Index ran against USD cash. We show the estimates of 95% daily VaR for three cases - factor covariance matrix estimated without any regularization, covariance method regularized using simple shrinkage (the approach discussed in the FactSet MAC white paper [4]), and covariance matrix regularized using the new method.

The results show that the random noise present in the non-regularized estimator of the covariance matrix in general leads to slightly lower values of VaR. Intuitively that can be easily understood as a diversification effect caused by the presence of uncorrelated noise in model factors, just as the presence of large number of assets with uncorrelated returns leads to reduction of portfolio risk because of the diversification. Application of the RMT regularization method result in partial reduction of the covariance matrix noise, thus reducing the diversification effect and an increasing in the estimated VaR value.

	DJI vs SP500	Barclays Agg
No Regularization	1.92	3.15
Shrinkage	1.91	3.20
RMT Method	2.13	3.42

Table 2: MAC 1 day 95% VaR estimated using covariance matrices with different regularization method applied

References

[1] Ribonato R., Jackel P. (1999). The most simple methodology to create a valid correlation matrix for risk management and option pricing purposes. Quantitative Research Center Papers, http://www.quarchome.org/correlationmatrix.pdf.

[2] Bai Z. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. Statistica Sinica, 9, 611-677

[3] Ledoit O., Wolf M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. The Annals of Statistics, 40(2), 1024–1060

[4] Steven P. Greiner, William McCoy, Nikita S. Imennov, Christopher Carpentier. (2014) FactSet Multi-Asset Class Risk Model. FactSet White Paper.