Project Proposal for 2012 Mathematical Problems in Industry Workshop
University of Delaware
June, 2012

Company: Standard & Poor's

Approximating Correlation Matrices

Assessing the potential for gains and losses on portfolios of financial assets as well as understanding the drivers of realized returns are important parts of performance tracking and risk management operations for banks, asset managers, insurance companies, etc. Understanding systematic exposure to macro-economic, regional and sector factors across the portfolio helps explain concentration risks and predict the potential for large losses. Banks often use the modeling of portfolio loss distributions to help establish capital requirements to survive periods of extreme credit stress, while portfolio managers employ similar models to track the performance of their portfolios against benchmark portfolios.

One common approach to portfolio return modeling is to combine a process for changes in individual exposure returns with a dependence structure across exposures that captures the joint evolution. In some models, dependence structure is represented by a normalized Gaussian distribution (or other multivariate distribution) used to model joint changes in the underlying value of the exposures; such distributions are often characterized by a correlation matrix - a symmetric positive definite matrix with ones on the diagonal.

A general correlation matrix is often derived empirically from historical asset returns, and working directly with pair-wise correlations provides a framework for stress testing views in a natural way – e.g., assessing the impact of in an increase in correlation among U.S. banks. However, for large portfolios with tens or hundreds of thousands of exposures, working directly with an explicit correlation matrix of extremely large dimension is impractical, particularly as a Monte Carlo simulation is often required to compute the portfolio return distribution. Hence approximating a correlation matrix with a much lower dimensional factor model structure is of interest.

The solution to the question of how to optimally approximate a given correlation matrix with a fixed number of factors is known. However, the solution is expressed in terms of principal components of the matrix and does not necessarily have an intuitive economic interpretation. If the original correlation matrix has a block structure (with constant correlation within a block), it is also known how to exactly represent the matrix in terms of a factorization into much lower dimensional components; again, however, the interpretation of the decomposition is not intuitive.

We would like to study the estimation of the optimal factor loadings when a specified factor structure is imposed. Specifically, we consider 'localized one or two factor models'. This means that each exposure in the portfolio can weight on a shared global factor and on either one or two additional factors that would typically represent the sector and/or region of the exposure. For example, there may be ten possible sectors, leading to an eleven factor model (one global and one for each sector), but any specific exposure would only have weight on the global factor and one of the sector factors.

Assuming we start with either a general correlation matrix or a block correlation matrix, questions of interest are:

1) How can the optimal localized one or two factor model be computed?
2) Can bounds for the impact of the approximation error on portfolio risk measures be estimated?

Another related question arises when we are given a sector correlation matrix and a regional correlation matrix. Here we would like a method to estimate a localized two factor model that is consistent with the given sector and regional correlation matrices.