Finding The Most Important Principal Component

[Nematrian website page: ERMMTFindingTheMostImportantPrincipalComponent, © Nematrian 2015]

Suppose we have a set of n series of returns (or losses, …). A principal component is a set of exposures (and a principal component series is a series of returns) corresponding to an eigenvector of the relevant $n \times n$ covariance matrix, V. Eigenvectors satisfy the vector equation $Vx = \lambda x$ for some scalar λ.

Typically principal components are identified in practice using suitable software packages designed to identify eigenvectors and eigenvalues, applied to the relevant covariance matrix, V, e.g. using using Nematrian web services functions that target principal components, i.e. MnPrincipalComponents, MnPrincipalComponentsSizes and MnPrincipalComponentsWeights.

However, for the first, i.e. most important, principal component there is a conceptually simpler approach as follows.

We note that any vector can be written as a combination of the eigenvectors of a matrix, and that these eigenvectors can be chosen to be orthonormal (if suitably chosen if some eigenvalues take the same value) so we can write any vector, a, of active positions as the sum of positions, a_i, in the relevant eigenvectors, q_n, i.e. as:

$$a = a_1 q_1 + \cdots + a_n q_n$$

Then, $Y = a^T V a = a_1^2 \lambda_1 + \cdots + a_n^2 \lambda_n$. If we order the eigenvectors (principal components) so that the most important ones are first, i.e. $\lambda_1 \geq \cdots \geq \lambda_n$ then Y is maximised, subject to $a^T a = 1$, if $a_1 = \pm 1$ and $a_2 = a_3 = \cdots = 0$. Thus, we can identify the most important principal component by reference to the set of positions of unit magnitude that exhibit the largest risk (here equated with ex-ante tracking error/variance/VaR).