

Random Matrix Theory

[Nematrian website page: [RandomMatrixTheory](#), © Nematrian 2015]

The risks expressed by a portfolio of assets or liabilities depend heavily on how these assets and liabilities might ‘co-move’, i.e. move in tandem. One way of analysing these co-movement characteristics is to consider the covariances between the price movements of (or more precisely, the log returns on) different assets and liabilities.

Most usually, investigators encapsulate this information within a single structure, the covariance matrix, often estimating this matrix empirically, using past observed covariances between the different assets and liabilities.

But how do we tell how ‘reliable’ is this estimation process? Suppose we have T time periods that we can use to estimate empirically the characteristics of the covariance matrix and N assets or liabilities (i.e. ‘instruments’). The correlation matrix contains $N(N - 1)/2$ distinct entries. If N is large relative to T (which it usually is in the risk management context) then we should expect the empirical determination of the covariance matrix to be ‘noisy’, i.e. for it to have a material element of randomness derived from measurement ‘noise’. It should therefore be used with caution. Such matrices can be characterised by their eigenvectors and eigenvalues, and it is the smallest (and hence apparently least relevant) of these that are most sensitive to this noise. But these are precisely the ones that determine, in Markowitz theory, the precise structure of optimal portfolios, see [Laloux et al. \(1999\)](#).

It is thus desirable to devise ways of distinguishing between ‘signal’ and ‘noise’, i.e. to distinguish between those eigenvalues and eigenvectors of the covariance matrix that appear to correspond to real characteristics exhibited by these assets and liabilities and those that are mere artefacts of this noise.

One way of doing this is to use random matrix theory. This theory has a long history in physics since Eugene Wigner and Freeman Dyson in the 1950s. It aims to characterise the statistical properties of the eigenvalues and eigenvectors of a given statistical ‘ensemble’ (i.e. the set of random matrices exhibiting pre-chosen symmetries or other sorts of constraints). Amongst other things, we might be interested in the average density of eigenvalues and in the distribution of spacing between consecutively ordered eigenvalues etc.

For example, we might compare the properties of an empirical covariance matrix V with a ‘null hypothesis’ that the assets were, in fact, uncorrelated. Deviations from this null hypothesis that were sufficiently unlikely might then suggest the presence of true information.

In the limit of very large matrices (i.e. $N \rightarrow \infty$) this density is analytically tractable and is as follows, for covariance matrices derived from (independent) series that have a common standard deviation, σ . Correlation matrices derived from the above null hypothesis have this property, as they correspond to covariance matrices where each series has a standard deviation in isolation of $\sigma = 1$.

The density, for a given $Q = T/N$, is $f_{Q,\sigma}(\lambda)$ where:

$$f_{Q,\sigma}(\lambda) = \begin{cases} \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda} & \text{if } \lambda_{min} \leq \lambda \leq \lambda_{max} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{where } \lambda_{max} = \sigma^2 \left(1 + \frac{1}{Q} + 2 \sqrt{\frac{1}{Q}} \right) \text{ and } \lambda_{min} = \sigma^2 \left(1 + \frac{1}{Q} - 2 \sqrt{\frac{1}{Q}} \right)$$

In the limiting case where N is large this has some features that get smoothed out in practice for less extreme values of N , in particular the existence of a hard upper and (if $Q < 1$) lower limit above and below which the density falls to zero.

We may therefore adopt the following prescription for ‘denoising’ an empirically observed correlation matrix, if we can assume that the ‘null’ hypothesis is that the instruments are independent (an assumption that would be inappropriate if, say, we should ‘expect’ two instruments to be correlated, e.g. the two might be listings of the same underlying asset on two different stock exchanges, or two bonds of similar terms issued by the same issuer), see [Scherer \(2007\)](#):

- (a) Work out the empirically observed correlation matrix (which by construction has standardised the return series so that $\sigma = 1$).
- (b) Identify the largest eigenvalue and corresponding eigenvector of this matrix.
- (c) If the eigenvalue is sufficiently large, e.g. materially larger than the cut-off derived from the above (or some more accurate determination of the null hypothesis density applicable to the finite N case), then deem the eigenvalue to represent true information rather than noise and move on to step (d), otherwise stop.
- (d) Record the eigenvalue and corresponding eigenvector. Determine the contribution to each instrument’s return series from this eigenvector. Strip out these contributions from each individual instrument return series, calculate a new correlation matrix for these adjusted return series and loop back to (b).

The reason we in theory need to adjust each instrument return series in step (d) is that otherwise the ‘residual’ return series can no longer be assumed to have a common σ , so we can no longer directly use a formula akin to that above to identify further eigenvectors that appear to encapsulate true information rather than noise. However, if we ignore this nicety (i.e. we adopt the null hypothesis that all residual return series, after stripping out ‘significant eigenvectors’ are independent identically distributed Gaussian random series with equal standard deviations) then the computation of the cut-offs simplifies materially. This is because:

- i. The trace of a symmetric matrix (i.e. the sum of the leading diagonal elements) is the same as the sum of its eigenvalues (and is therefore invariant relative to a change in basis for the relevant vector space).
- ii. So, removing the leading eigenvector as above merely involves removing the leading row and column, if the basis used involves the eigenvectors.
- iii. The variance of the residual series under this null hypothesis is therefore merely the sum of the eigenvalues not yet eliminated iteratively.
- iv. Hence, we can calculate the cutoff $\lambda_{max}(i)$ for the i ’th eigenvalue ($1 < i < N$) as follows, where e_j is the magnitude of the j ’th eigenvalue ($1 < j < N$):

$$\lambda_{max}(i) = \left(\frac{1}{N} \sum_{j=i}^N e_j \right) \left(1 + \frac{N - (i - 1)}{T} + 2 \sqrt{\frac{N - (i - 1)}{T}} \right)$$

Using this prescription, we would exclude any eigenvalues and corresponding eigenvectors beyond the first one for which the eigenvalue is not noticeably above this cutoff. The Nematrian website function [MnEigenvalueSpreadsForRandomMatrices](#) calculates these $\lambda_{max}(i)$. More precise tests of significance could be identified by simulating spreads of results for random matrices.

[Edelman and Rao \(2005\)](#) describe how in many cases it is possible to calculate eigenvalue densities for a wide range of transformations of random matrices, including both deterministic and stochastic transformations. They express the view that the usefulness of random matrix theory will through time follow that of numerical analysis more generally, i.e. most disciplines in science and engineering will in due course find random matrix theory a valuable tool. Its history started in the physics of heavy atoms and multivariate statistics. It has already found its way into wireless communications and combinatorial mathematics and as seen above is potentially also becoming increasingly used in the field of financial analysis and risk management.

Relevant Nematrian Web service tools

[\[RandomMatrixTheoryTools\]](#)

The function [MnEigenvalueSpreadsForRandomMatrices](#) returns an array containing the $\lambda_{max}(i)$ for the limiting case where N is large. It takes as inputs the eigenvalues e_i and a value for T (which is not necessarily integral, to cater for the possibility that the effective number of observations is less than the actual number, e.g. if observations for different time periods are given different weights).

References

[\[RandomMatrixTheoryRefs\]](#)

[Edelman, A. and Rao, N.R. \(2005\)](#). Random matrix theory. *Acta Numerica* **14**, pp. 233-297

[Laloux, L., Cizeau, P., Bouchard, J-P. and Potters, M. \(1999\)](#). Random matrix theory. *Risk*, March 1999

[Scherer, B. \(2007\)](#). *Portfolio Construction and Risk Budgeting*. 3rd ed. RiskBooks