Fixing a broken correlation matrix

Phil Joubert and Stephen Langdell

The correlation matrix is central to risk calculations, whether it is used as part of the definition of a copula or other dependency structure used to generate risk factors, or more directly to combine capital requirements using the "VCV" technique (e.g. in the Standard Formula method of Solvency 2). In this article we discuss a common problem: a matrix which superficially "looks like" a correlation matrix may not be one mathematically. First we review some theory of correlation matrices, before discussing what can go wrong. We then present some ways of constructing valid correlation matrices or "fixing" broken ones.

Correlation Matrix Theory

The correlation between two random variables describes the strength of any linear relationship between those variables on a scale which ranges from -1 to 1. We exclude the extremes of the scale as they would imply that the variables are deterministically related. For jointly normal random variables, this single number summarises all there is to know about the relationship between the two. For other distributions the situation can be more complicated, but we will not consider this. Correlations and correlation matrices are fundamentally a multivariate normal concept.

When more than two random variables (e.g. risk factors) are involved it is useful to draw up the correlations between each pair of values into a table. Appropriately set out this table looks like a matrix. We take the convention that the correlation of a variable with itself is 1 (the upper limit of the scale), and we note that the relationship is symmetric: the correlation between variables X and Y is the same as between Y and X.

From this we deduce that a correlation matrix is any matrix which:

1. Is symmetric
2. Has elements in (-1,1) off the main diagonal
3. Has elements = 1 on the main diagonal.

But there is a fourth requirement. If we know the relationship between variables X and Y, and between Y and Z, then we should have some idea of the relationship between X and Z. All three have to be consistent.

For normal random variables we can try to force consistency by construction. If Z_1 and Z_2 are independent standard normal variables (i.e. their correlation and means are zero and their variances are one), then we can construct two dependent standard normal variables X_1 and X_2 by setting $X_1 = Z_1$ and $X_2 = \rho Z_1 + \sqrt{1-\rho^2}Z_2$, where ρ is the desired correlation. The first term builds the dependence, and the second is a "balancing item" to make sure the new variable (X_2) has variance one. We can continue by adding a third variable X_3 such that X_3 is a linear combination of Z_1, Z_2 and Z_3. This builds up a set of linear simultaneous equations which we can write in matrix form as

$$X = LZ$$

where L is a lower triangular matrix (i.e. it has zeroes above the main diagonal). Then we set the correlations of these variables to our desired correlation matrix C, giving us the equation

$$Corr[X] = E[XX^T] = E[(LZ)(LZ)^T] = LE[ZZ^T]L^T = LL^T = C.$$

The last identity is key here – a matrix C with properties 1, 2 and 3 above is only a correlation matrix if we can find a lower triangular matrix L such that $LL^T = C$. If it exists, this is called the Cholesky decomposition of C and is analogous to the square root of a real number. Matrices which have Cholesky decompositions also have several other characteristics which can be useful, and are called *positive definite* matrices. So we add a fourth characteristic:

4. Is a positive definite matrix.

Why a correlation matrix might be broken

Correlation matrices in some applications (e.g. portfolio risk) are calculated from historic data, but rarely in a consistent way. Data might be missing because a particular stock didn't trade on a given day, or a particular market was closed, or because the company didn't exist until five years ago. Correlations are calculated pairwise, and then put into the form of a symmetric matrix. There is no guarantee that this matrix satisfies requirement 4.

In insurance, lack of data means that a firm's correlation matrix is frequently set, at least partly, by "expert judgment". Managers might be asked to set correlations between risk factors as {high, medium, low, none} = {0.75, 0.5, 0.25, 0}. Even the technically inclined will struggle to keep the resulting matrix consistent (in the sense of requirement 4.) for a matrix of more than a couple of elements. For a typical case with dozens or even hundreds of risk factors this is almost impossible.

Whatever the origin of the problem, we are often presented with a broken correlation matrix and asked to fix it. Next we present a couple of repair techniques. We'll use a simple example to illustrate: the offending matrix is

1	0.95	0
	1	0.95
		1

Patch it up

A quick and dirty method of patching up a broken matrix exploits one of the characteristics of positive definite matrices: a positive-definite matrix always has positive eigenvalues (and vice-versa). Recall that a symmetric matrix C can be decomposed as

$$C = Q\Lambda Q^T$$

where Q is the matrix whose columns are the eigenvectors of C and Λ is a matrix with the corresponding eigenvalues on the diagonal and zero elsewhere. We don't need any more theory of eigenvalues than this, a decent statistical or numerical library will be able to calculate them for us[1].

We can attempt to fix the matrix by calculating the eigen-decomposition of the matrix, setting any negative eigenvalues to zero, and then reconstructing the resulting matrix. So we form $Q\Lambda'Q^T$, where Λ' is the same as Λ, but with negative entries replaced by zero and scale the result to give a matrix C' with ones on the diagonal.

For our matrix the eigenvalues are {-0.3435, 1.0000, 2.3435}. Setting the first (negative) entry to zero and calculating C' as above results in

1.0000	0.73454	0.07908
	1.0000	0.73454
		1.0000

Find the closest correlation matrix

While the above method works, it is a bit arbitrary – we don't really understand its relationship to the original matrix. On the other hand, the closest correlation matrix method (Higham, 2002) explicitly minimises the distance between the original matrix and the "fixed" matrix. In this case we define the distance between two matrices as the square root of the sum of the squared differences of their elements:

$$\|X, Y\| = \sqrt{\sum_i \sum_j \left(x_{i,j} - y_{i,j}\right)^2}.$$

This is called the Frobenius distance. Computing the nearest correlation matrix to a given matrix is done numerically by iteratively and alternately projecting onto the spaces of positive definite and unit diagonal symmetric matrices, eventually converging to the closest matrix in the intersection of those spaces (see figure).
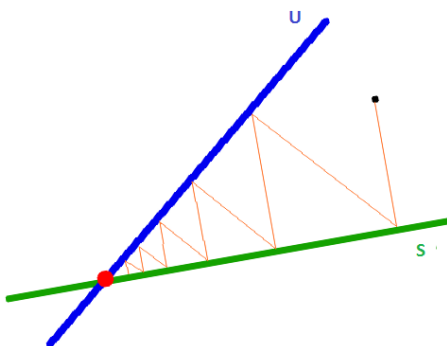
We find the matrix below[2]:

---

[1] The NAG library was used for this article

1.00000	0.75217	0.13151
	1.00000	0.75217
		1.00000

which looks similar to the matrix we found in the previous section. A quick computation of the Frobenius distance using the formula above would show that this matrix is indeed "closer" to the original.

Having an analytic framework to hang our method on means we can introduce more flexibility. Suppose the manager setting the correlations miscommunicated his intentions, and that the 0 in his matrix should have been interpreted as "I don't care". In this case we can add a weighting to the algorithm, so that the important correlations are disturbed as little as possible. Assigning appropriate weights we now get[3]:

1.00000	0.94044	0.82438
	1.00000	0.94044
		1.00000

Complete the matrix

The last scenario occurs often in practice, especially when dealing with large matrices. Matrices might be set in "blocks" by different business units, and need to be combined, or managers may be agnostic about the values of certain correlations, but certain about others. These are known as matrix completion problems.

Given the matrix

1.00	0.95	?
	1.00	0.95
		1.00

we can set the missing entry to the product of the known correlations, giving

1.00	0.95	0.9025
	1.00	0.95
		1.00

A proof that this does result in a positive definite matrix is given in Kahl and Günther (2005), which confirms the intuition that it should work. This method can be used to combine block matrices. It also suggests an alternative method of correlation matrix

Numeric Software: Build or Buy?

As actuarial models get more sophisticated there is a growing requirement for more numerical and statistical methods, which need to be accurate, reliable and high-performance. Actuaries are often quite keen on building models in-house, but there are clear advantages to using proven algorithms from an established numerical library:

- Development time is reduced.
- Actuarial resource can be devoted to actuarial problems, rather than software problems.
- Code from a high-quality software house is exhaustively tested.
- The best algorithm for a given problem may not be the most obvious. For example, although the alternating projections algorithm for the nearest correlation matrix is guaranteed to converge, it is not the most efficient. The NAG Library codes use a Newton algorithm that is more complicated but converges much more quickly - especially for large matrices.
- Bugs can be subtle, especially since their effect can be masked by random noise in data.

---

[2] The NAG library algorithm (G02AAF) was used
[3] The NAG library algorithm (G02AJF) was used

construction, whereby business units are asked to provide pairwise correlations with one central random variable (say GDP growth or something similar), with all other correlations being inferred.

Conclusions

Correlation matrices are a key input to risk systems, and a major source of headaches. Technical actuaries and quants don't understand why business actuaries can't check that the matrices they provide are positive definite, and the business actuaries don't understand why the quants refuse to use a perfectly reasonable looking matrix. We've demonstrated a couple of ways out of this bind: either by finding a matrix which resembles the given matrix but which is positive definite, or by changing the method of specifying correlation matrices in the first place. These techniques are robustly implemented and we have successfully used them in large scale solutions for several insurers and other financial institutions.

Phil Joubert is an independent actuarial consultant specialising in implementation of risk models and blogs at www.not-normal-consulting.co.uk .
Stephen Langdell is a Senior Technical Consultant at the Numerical Algorithms Group
Note: A version of this article also appears in The Actuary magazine (September 2013).

References

Nicholas J. Higham, Computing the Nearest Correlation Matrix---A Problem from Finance, IMA Journal of Numerical Analysis, 22, (2002)

Christian Kahl and Michael Günther, Complete the Correlation Matrix (2005)

The NAG Library, The Numerical Algorithms Group (NAG), www.nag.com (2013)