Cracking the covariation code

David Rowe suggests a relative rank-based representation of covariance is more useful than one based on actual values

Combining exposure to diverse uncertain events commonly leads to greatly reduced levels of risk. Peter Bernstein has referred to this phenomenon of diversification as the only free lunch allowed by the laws of economics (see Bernstein's recent article in *Risk* January 2001, page 54, for example). It is the essential reality that makes insurance and various forms of portfolio management possible. As a result, rigorous analysis of covariation and diversification is an essential component of effective risk management. Unfortunately, it is also fraught with difficulties.

Virtually all statistical analysis is based on an assumption of 'stable random distributions'. The Black-Scholes option pricing formula, for example, is explicitly predicated on a stable random distribution for changes in the price of the underlying asset, although if this assumption were accurate there would be no need for option markets!

Stable random distributions characterise the realm of first-order uncertainty. It can reasonably be said that statistical and option theory have tamed this form of risk. In fact, of course, second-order uncertainty is also a pervasive phenomenon in the social environment. Indeed, I contend that a significantly higher degree of secondary uncertainty is what distinguishes the social realm from the physical. It is also what limits the full applicability of many tools and techniques that are very successful when applied in the physical sciences.

The essential question revolves around just how stable the underlying random process actually is. In some cases, attempts have been made to push back the frontiers of our knowledge to include certain aspects of this second-order uncertainty. This involves building predictive models of how the parameters of traditionally estimated distributions change over time. To the extent that such efforts are successful, they effectively bring such changes into the realm of first-order uncertainty. Nevertheless, such models have their own statistical error properties whose residual instability remains in the realm of second-order uncertainty. Social scientific modelling requires frequent review and revision of the estimates of these 'stable' parameters.

David Rowe is group executive vice-president for risk management at SunGard Trading and Risk Systems
e-mail: david.rowe@risk.sungard.com

Estimating covariation

Covariation is plagued with an especially high degree of secondary uncertainty. First, it often behaves differently for different magnitudes of change in the underlying variables. As a result, estimates based on a full sample, including observations from all periods, often break down in times of crisis. One way to model the common phenomenon of fat tails is to assume that observations represent draws from two different distributions. In this context, one also can assume distinctive covariation behaviour for the different distributions. The empirical problem is deciding which historical observations to attribute to each distribution.

Second, covariation estimates tend to be less stable over time than other statistical parameters. This requires more frequent re-calibration, which leads to yet another difficulty. Correlation coefficients are the most commonly used measures of covariation. Unfortunately, these are hard to combine with the well-known departures from normal distribution behaviour that characterise observed market data. The standard way to impose an observed set of correlation coefficients on a Monte Carlo sample is to apply the corresponding Cholesky transform. This involves taking a weighted average of uncorrelated variables. Such an averaging process, however, tends to drive the resulting correlated variables toward normality. This disturbs the desired imposition of skewness (asymmetry) and leptokurtosis (fat tails) on the individual distributions.

Ordinal versus cardinal covariation

One attractive technique for resolving this dilemma is the use of a statistical concept known as 'copulas'. Complete characterisation of a multivariate stochastic system requires the joint distribution function of all the variables. But, in practice, it is often desirable to separate the covariation characteristics of the system from the marginal behaviour of each random variable individually. The copula of a joint distribution function can be thought of as capturing the dependence structure of the variables without reference to their marginal characteristics.

Given a fixed copula, the corresponding quantiles of the component variables are deemed to display consistent covariation. One advantage of this definition of covariation is that it is invariant to strictly increasing transformations of the marginal distributions. (This is not true of standard correlation coefficients.) As a result, using copulas allows revision of the marginal distribution parameters without triggering the need to make corresponding adjustments in the covariation structure. The opposite is also true. The covariation structure, as represented by the copula, can be revised more frequently without requiring a wholesale re-estimation of the marginal distribution parameters. Perhaps most important, a rank-based specification of the covariation structure is perfectly consistent with non-normal behaviour of the marginal distributions.

Summary

A common problem in simulating the behaviour of multiple market variables is how to impose both observed covariation characteristics and non-normal behaviour such as skewness and fat tails. A covariation structure based on relative rank rather than actual values helps to accomplish both these objectives. This approach also decouples the estimation process for the covariation structure from that for the marginal distributions. It allows parameters for one to be revised without necessarily affecting those for the other. ■