Cracking VAR with kernels

Value-at-risk analysis has become a key measure of portfolio risk in recent years, but how can we calculate the contribution of some portfolio component? ***Eduardo Epperlein*** *and* ***Alan Smillie*** *show how kernel estimators can be used to provide a fast, accurate and robust estimate of component VAR in a simulation framework*

The notion of component value-at-risk (CVAR) originated in the papers of Garman (1996, 1997) and Litterman (1997a, 1997b), and has been used by banks as a practical risk analysis tool since at least Epperlein & Sondhi (1997). The goal is to calculate how much some component of a portfolio contributes to the total VAR of that portfolio. We denote the profit and loss (P&L) of the portfolio as PL and the P&L of the ith component as PL_i, so that:

$$\sum_i PL_i = PL$$

Then the CVAR of the ith component is defined to be the expected value of PL_i given that the portfolio P&L is equal to the VAR, that is:

$$CVAR_i = E\left[PL_i \middle| PL = VAR\right] \tag{1}$$

For notational simplicity, VAR is treated as the appropriate percentile of the P&L distribution, so VAR and CVAR will usually be negative. Component VAR has the desirable property of additivity: if we sum all the CVARs of a given portfolio we recover the portfolio VAR:

$$\sum_i CVAR_i = VAR \tag{2}$$

Also note that CVAR is always defined with respect to a parent portfolio PL and a child portfolio PL_i.

Our definition (1) is different from that often seen in the literature, where CVAR is defined as the incremental change in the portfolio VAR given a small change in the size of the ith exposure w_i, times the size of that exposure, that is:

$$CVAR_i = w_i \frac{\partial VAR}{\partial w_i} \tag{3}$$

Subject to some technical conditions, the two definitions are equivalent (see Gourieroux, Laurent & Scaillet, 2000), but we consider (1) to be more intuitive, at least in the simulation setting we shall employ.

Estimators of component VAR

When all the components of a portfolio are elliptically distributed[1], the CVAR can be calculated analytically, as shown by Carroll *et al* (2001):

$$CVAR_i = E\left[PL_i\right] + \frac{\operatorname{cov}\left[PL, PL_i\right]}{\operatorname{var}\left[PL\right]}\left(VAR - E\left[PL\right]\right) \tag{4}$$

The assumption of elliptically distributed P&L is of course very restrictive, and will be violated by a market risk portfolio containing options, for example.

A more general approach to VAR estimation is Monte Carlo simulation. We generate N random scenarios for the P&L of each component, and sum to find the portfolio P&L. We write the P&L of the ith component in the jth Monte Carlo scenario as $PL_i^{(j)}$, so that the simulated portfolio P&Ls are:

$$PL^{(j)} = \sum_i PL_i^{(j)}$$

Equation (1) then suggests an estimator of the form:

$$CVAR_i^{[S]} = PL_i^{(n)} \tag{5}$$

where n denotes the 'VAR scenario', that is, the scenario such that:

$$PL^{(n)} = VAR$$

In this approach, which we shall refer to as scenario extraction, the estimator will automatically satisfy the additivity property, that is:

$$\sum_i CVAR_i^{[S]} = VAR$$

We shall see later that this estimator gives an unbiased estimate for CVAR but suffers badly from noise.

We can ameliorate (5) by taking an average of the values of PL_i around the nth value in the estimator, and weighting them according to their distance from n. *Ad hoc* smoothers are used by Litterman (1997b) and Hallerbach (2002). We attempt to make the method more rigorous by using a kernel estimator to measure the distance from the nth scenario. A kernel is a function of the form:

$$K(x;h) = K\left(\frac{x}{h}\right)$$

which is symmetric about zero, takes a maximum at $x = 0$ and is

[1] *Note that the multivariate normal distribution is an example of an elliptical distribution*

non-negative for all x. A particularly simple example is the triangle kernel:

$$K(x;h) = \max\left(1 - \left|\frac{x}{h}\right|, 0\right)$$

which we have chosen for ease of implementation.[2] A kernel estimator of CVAR can thus be constructed as:

$$CVAR_i^{[K]} = \frac{\sum_{j=1}^{N} K\left(PL^{(j)} - VAR; h\right) PL_i^{(j)}}{\sum_{j=1}^{N} K\left(PL^{(j)} - VAR; h\right)} \tag{6}$$

The numerator in the above expression can be seen as the weighted average of PL_i, while the denominator is a normalisation factor. A similar approach is hinted at by Gourieroux, Laurent & Scaillet (2000), though these authors never make the use of kernels explicit nor conduct any statistical tests of the accuracy of the estimator in the finite sample.

Using (6) we find that the sum of the CVARs is not in general equal to the VAR, violating the additivity property (2). This is because the kernel estimator is biased low[3] due to asymmetry in the P&L distribution around the point $PL = VAR$ (there are more scenarios with loss 'a little' smaller than VAR than with loss 'a little' greater than VAR). We can correct for the bias by rescaling according to the bias in the kernel estimate of the VAR itself, which leads to a rescaled kernel estimator for CVAR:

$$CVAR_i^{[K,R]} = VAR \frac{\sum_{j=1}^{N} K\left(PL^{(j)} - VAR; h\right) PL_i^{(j)}}{\sum_{j=1}^{N} K\left(PL^{(j)} - VAR; h\right) PL^{(j)}} \tag{7}$$

Using estimator (7) guarantees that the additivity property holds and, as we see in the next section, that the estimate is unbiased.

The performance of the kernel estimator depends on the choice of the smoothing parameter h: in our application, letting $h = 0$ corresponds to the scenario extraction estimator (5), while letting $h \to \infty$ corresponds to taking an unweighted average of PL_i. It can be shown (see Silverman, 1986) that the optimal choice for h (in the sense of minimising the mean square error) for the triangular kernel is:

$$h = 2.575\sigma N^{-\frac{1}{5}}$$

where σ is the standard deviation of PL. We examine the performance of the estimators (5) and (7) in what follows, but first we mention two alternative numerical approaches as a basis for comparison.

The alternative definition of CVAR (3) suggests using numerical differentiation, as proposed by Epperlein (1998). Letting:

$$VAR_i^{+} = VAR\left(PL + \delta PL_i\right)$$

and:

$$VAR_i^{-} = VAR\left(PL - \delta PL_i\right)$$

be the estimates of VAR with the ith exposure perturbed upward and downward respectively, a finite difference estimator for $CVAR_i$ is:

$$CVAR_i^{[D]} = \frac{VAR_i^{+} - VAR_i^{-}}{2\delta}$$

The parameter δ controls the size of the perturbation, and plays a role analogous to h in the kernel estimator. This time it is not clear how to choose δ *a priori*, but numerical experimentation shows that a value of $\delta = 0.1$ achieves a reasonable compromise between bias and variance. Due to Monte Carlo error, the estimated CVARs will again fail to sum to the VAR, but a rescaling method similar to (7) can be used to correct for this, yielding the rescaled finite difference estimator:

$$CVAR_i^{[D,R]} = VAR \frac{CVAR_i^{[D]}}{\sum_i CVAR_i^{[D]}} \tag{8}$$

Finally, we mention a very simple method to calculate CVAR from a simulated sample. If we are willing to assume that the P&Ls are approximately elliptically distributed, equation (4) suggests the semi-parametric estimator proposed by Carroll *et al* (2001):

$$CVAR_i^{[P]} = \frac{\sum_j PL_i^{(j)}}{N} + \frac{\sum_j PL_i^{(j)} PL^{(j)}}{\sum_j PL^{(j)2}} \left(PL^{(n)} - \frac{\sum_j PL^{(j)}}{N} \right)$$

The semi-parametric estimator is additive by construction.

Statistical comparison of the estimators

We shall now compare the estimators (5), (7), (8) and (9) for a selection of sample portfolios. We use Monte Carlo simulation with N = 10,000 and calculate VAR at a confidence level of 99%. The experiment is run 1,000 times, and the mean and standard deviation of the estimated CVAR is recorded. For comparison, we have also included results on the error in the Monte Carlo estimate of the VAR itself.

Example 1: linear portfolio. First, let us consider a very simple example where the portfolio risk factors are normally distributed with zero mean, zero correlation and unit variance, that is:

$$\begin{pmatrix} RF_1 \\ RF_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

We assume linear exposure to w_1 and w_2 units of RF_1 and RF_2 respectively, so we have payout functions:

$$PL_1 = w_1 RF_1$$
$$PL_2 = w_2 RF_2$$

While this may seem like a trivial example, it is of interest since in this case (4) can be used to calculate VAR, $CVAR_1$ and $CVAR_2$ analytically as:

$$VAR = \Phi^{-1}(0.01)\sqrt{w_1^2 + w_2^2}$$

$$CVAR_1 = \Phi^{-1}(0.01)\frac{w_1^2}{\sqrt{w_1^2 + w_2^2}}$$

$$CVAR_2 = \Phi^{-1}(0.01)\frac{w_2^2}{\sqrt{w_1^2 + w_2^2}}$$

[2] *Numerical experiments (not shown) indicate that our conclusions are unaffected by the particular form of kernel function used*

[3] *A referee has pointed out that this can be avoided by estimating the VAR itself using kernels. This will work, but only at the expense of inducing a similar bias in the VAR. Also note that this bias is not the same as the boundary bias observed when kernels are used for non-parametric density estimation*

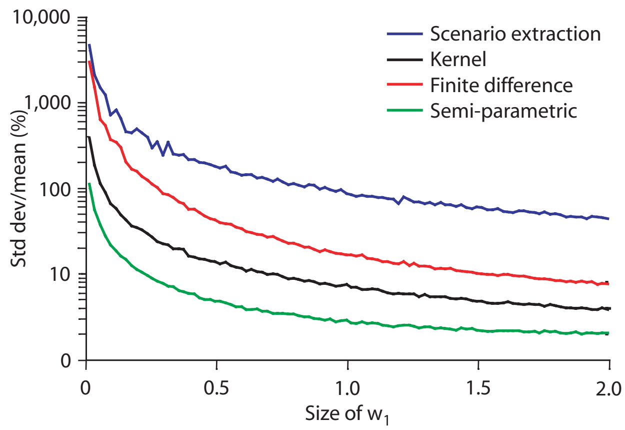
A. Comparison of the estimators for linear, uncorrelated components

		Mean	Std dev	Std dev/mean
Scenario extraction (5)	$CVAR_1$	−1.03	0.928	89.49%
	$CVAR_2$	−4.17	0.928	22.26%
Kernel (7)	$CVAR_1$	−1.04	0.070	6.71%
	$CVAR_2$	−4.16	0.093	2.23%
Finite difference (8)	$CVAR_1$	−1.05	0.170	16.24%
	$CVAR_2$	−4.15	0.178	4.28%
Semi-parametric (9)	$CVAR_1$	−1.04	0.029	2.79%
	$CVAR_2$	−4.16	0.072	1.73%
	VAR	−5.20	0.085	1.63%

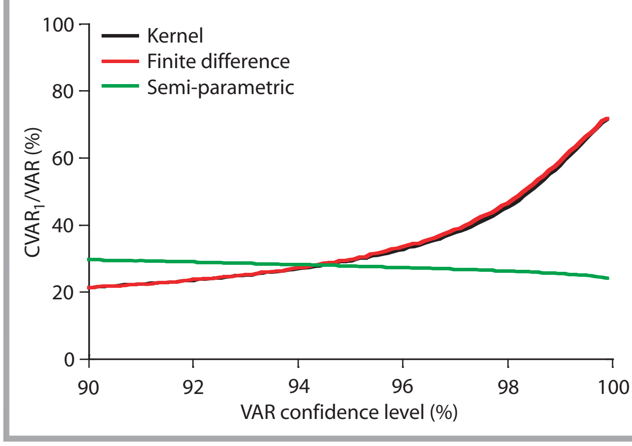
B. Comparison of the estimators for non-linear components

		Mean	Std dev	Std dev/mean
Scenario extraction (5)	$CVAR_1$	−1.89	0.810	42.86%
	$CVAR_2$	−1.35	0.804	59.56%
Kernel (7)	$CVAR_1$	−1.88	0.155	8.24%
	$CVAR_2$	−1.36	0.121	10.00%
Finite difference (8)	$CVAR_1$	−1.92	0.201	10.47%
	$CVAR_2$	−1.32	0.178	13.48%
Semi-parametric (9)	$CVAR_1$	**−0.79**	0.040	2.05%
	$CVAR_2$	**−2.45**	0.046	2.38%
	VAR	−3.24	0.073	2.27%

1 Noise in the CVAR estimators as a function of exposure size



2 CVAR varying with confidence level for a portfolio containing non-linear exposures



where Φ^{-1} denotes the inverse standard normal distribution function. Monte Carlo simulations using (5), (7), (8) and (9) with $w_1 = 1$ and $w_2 = 2$ yield the results in table A.

The mean estimates for $CVAR_1$ and $CVAR_2$ are close to their analytic values, indicating that the estimators are unbiased, and thanks to the rescaling procedure additivity holds exactly for each estimator. The main difference is in the noise of the estimate – the semi-parametric method gives the best results, followed by the kernel. The finite difference method is somewhat worse, and the scenario extraction approach leads to very significant noise even in this simple example.

Notice how the size of the noise relative to the mean is greater for the smaller component – this behaviour is typical of numerical procedures to estimate CVAR. The relative noise can become very large for components that make a minimal contribution to the VAR, as illustrated in figure 1, where we have plotted the relative noise in $CVAR_1$ (that is, standard deviation/mean) using each of the estimators for $w_1 = 0.02, 0.04, \ldots, 2$ while $w_2 = 2$. The very high levels of noise for low values of w_1 may be of less concern than we might initially expect, however, since for such small exposures the position contributes a very small proportion of the total risk.

Example 2: non-linear portfolio. Next, we consider a portfolio containing an instrument with non-linear payout function. The underlying risk factors are again standard normal and independent, but this time we set:

$$PL_1 = -2\max\left(-RF_1 - 1, 0\right)$$
$$PL_2 = RF_2$$

PL_1 represents an option-like payout, specifically a short position in out-of-the-money puts close to expiry. The non-linearity in PL_1 means that in this case we are unable to calculate VAR or CVAR analytically. Running the same test as before gives the results in table B.

The results for the scenario extraction, rescaled kernel and rescaled finite difference estimators are roughly in agreement, and exhibit similar levels of noise to the previous example. The results for the semi-parametric method are very different to the rest, suggesting that in this example estimator (9) does not provide a good approximation to the CVAR. This is unsurprising since PL_1 and PL do not follow an elliptical distribution. Since it is not clear how this error could have been estimated *ex ante* and thus corrected, the presence of such a large bias must be considered a serious flaw in the semi-parametric estimator.

We can use this portfolio to highlight another property of CVAR: its dependence on the VAR confidence level. In figure 2, we plot $CVAR_1$ as a proportion of the total VAR for confidence levels from 90–99.9% using the kernel, finite difference and semi-parametric estimators (we omit scenario extraction since the results are very close to the former two estimators). Notice how the non-linear component contributes a greater proportion of the VAR at higher confidence levels, but the semi-parametric estimator fails to capture this.

Example 3: correlated portfolio. Finally, we examine a special case where all the proposed estimators perform poorly. We return to the example of normally distributed P&L, but this time we have a correlation of ρ and set $w_1 = w_2 = 1$:

$$\begin{pmatrix} RF_1 \\ RF_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$
$$PL_1 = RF_1$$
$$PL_2 = RF_2$$

Using (3) we calculate:

$$VAR = \Phi^{-1}(0.01)\sqrt{2+2\rho}$$
$$CVAR_1 = \Phi^{-1}(0.01)\frac{\sqrt{1+\rho}}{\sqrt{2}}$$
$$CVAR_2 = \Phi^{-1}(0.01)\frac{\sqrt{1+\rho}}{\sqrt{2}}$$

In this example, both of the CVARs are significant, in the sense that they are of the same order of magnitude as the VAR (indeed, they are each 50% of the VAR).

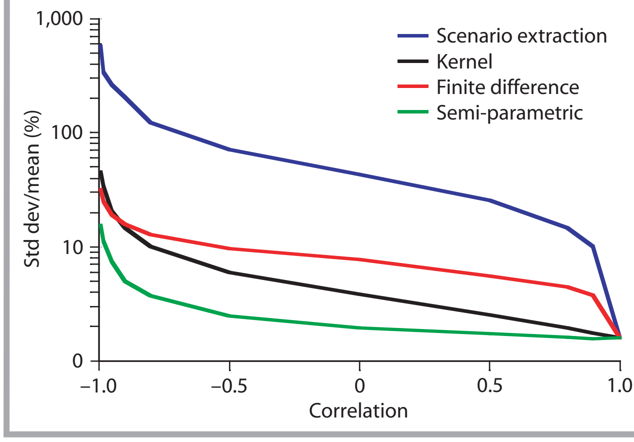
Figure 3 shows the relative noise of each estimator for a range of values of ρ. Overall, the error is of an acceptable size, but for strong negative correlation between the components it increases substantially. Intuitively, this is because the noise in the CVAR estimator is affected by the magnitude of PL_i, while the magnitude of the CVAR is proportional to the magnitude of the VAR. For strong negative correlation, the magnitude of the VAR falls but the magnitude of PL_i does not, causing the relative noise to increase. Such an increase is observed in all the estimators, though the impact on the finite difference estimator seems to be a little smaller than for the others. For very strong negative correlations ($\rho < -0.95$), the finite difference estimator actually gives a better performance than the kernel. It may seem that this is of purely academic interest since market factors will very rarely exhibit such strong negative correlation, but recall that components can comprise any subset of a portfolio. We could have, for example, two components of a trading portfolio that are constructed to hedge each other, where we would expect to see strong negative correlation between the components.

Implementation

We have shown how the kernel method yields the most accurate and robust CVAR estimates for a general portfolio structure, but practioners will also be interested in practical issues such as computational time. Computational times for CVAR using a desktop computer are given in table C, where '20 CVARs' refers to the computation of the CVAR for 20 components of the same parent portfolio. Clearly the question of which method is fastest is of somewhat academic interest, since for all the approaches the time to compute CVAR is likely to be much less than the time to compute the VAR itself.

Scenario extraction is, unsurprisingly, the most efficient approach, since in this case the 'computation' amounts to looking up the *n*th scenario for PL_i. The kernel and semi-parametric methods are also very efficient, particularly where we wish to calculate multiple CVARs on the same parent portfolio. For the kernel method, this is because the weights $K(PL^{(j)} - VAR; h)$ depend only on the PL, so only have to be computed once for each parent portfolio. To speed up the calculation further, we have chosen to use a very simple form of kernel based on a 'triangle function', which means that most (typically 95% for a VAR confidence level of 99%) kernel weights are zero (in contrast to a Gaussian kernel, for example, where each scenario receives some finite positive weight). The finite difference estimator is a little less efficient,

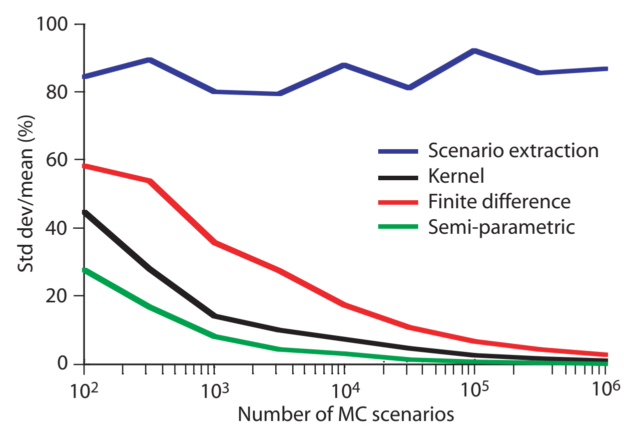
3 Noise in the CVAR estimators as a function of correlation



C. Computing time (in seconds) for the CVAR estimators

	N = 10,000		N = 250,000	
	1 CVAR	20 CVARs	1 CVAR	20 CVARs
Scenario extraction (5)	0.0001	0.0020	0.0001	0.0020
Kernel (7)	0.0069	0.0113	0.2119	0.4505
Finite difference (8)	0.0080	0.1592	0.2413	4.9174
Semi-parametric (9)	0.0058	0.0356	0.1826	1.3305

4 Noise in the CVAR estimators as the number of Monte Carlo scenarios is increased



largely because it requires the computation of two new VARs for every CVAR that is required.

We might expect that the noise in the CVAR estimators will follow the usual Monte Carlo rule and fall as $\sqrt{N}$. We investigate this by computing the standard deviation in $CVAR_1$ from example 1 using increasing values of N. The results are recorded in figure 4. We see that the square-root rule does indeed apply for the kernel, finite difference and semi-parametric estimators, but not for scenario extraction. The reason for this is quite subtle: recall that CVAR is defined as the expectation of the component P&L, given that the parent P&L is equal to the VAR. Using estimator (5), one computes:

$$PL_i | PL = VAR$$

not:

D. CVARs and errors for market risk on the bond portfolio (% of portfolio VAR)

Risk rating	Scenario extraction		Kernel		Finite difference		Semi-parametric	
AAA	4.4	(1.0)	4.7	(0.1)	4.8	(0.5)	4.7	(0.1)
AA	12.0	(1.8)	11.8	(0.3)	12.0	(1.0)	11.8	(0.2)
A	7.8	(1.8)	8.2	(0.3)	7.9	(0.8)	8.2	(0.1)
BBB	44.5	(5.8)	46.2	(1.0)	45.6	(2.4)	46.4	(0.7)
BB	33.7	(5.7)	32.5	(0.6)	32.7	(2.2)	32.6	(0.6)
B	–12.5	(5.7)	–13.1	(0.5)	–12.8	(2.4)	–13.3	(0.2)
CCC	10.1	(3.4)	9.6	(0.5)	9.7	(2.0)	9.7	(0.2)

E. CVARs and errors for credit risk on the CDS portfolio (% of portfolio VAR)

Risk rating	Scenario extraction		Kernel		Finite difference		Semi-parametric	
AAA	0.4	(7.7)	0.1	(0.2)	0.2	(0.2)	0.2	(0.2)
AA	1.2	(13.8)	1.4	(0.9)	1.4	(1.2)	3.1	(0.7)
A	5.4	(27.3)	5.4	(1.7)	5.5	(2.3)	5.8	(0.5)
BBB	94.1	(57.5)	93.5	(4.2)	94.6	(5.3)	40.2	(1.1)
BB	5.2	(19.4)	5.4	(1.2)	5.2	(2.4)	10.3	(0.3)
B	9.5	(32.9)	10.7	(1.9)	10.0	(3.5)	30.5	(0.7)
CCC	–15.8	(32.6)	–16.5	(1.9)	–16.9	(3.5)	9.9	(0.6)

$$E\left[PL_i \middle| PL = VAR\right]$$

Thus, by using scenario extraction, we simply draw a single sample from a (conditional) probability distribution. The distribution will have some variance, which we cannot expect to reduce simply by increasing the number of Monte Carlo runs.

The Monte Carlo error could also be reduced by applying one of the numerous variance reduction techniques available in the literature, but we consider a discussion of these methods to be beyond the scope of this article. Note that the bias in the semi-parametric estimator when a portfolio contains non-linear instruments (such as in example 2) cannot be reduced by increasing the number of Monte Carlo scenarios or by applying a variance reduction technique.

Traded portfolios

We conclude by presenting two examples of CVAR applied to realistic trading portfolios, analysed using internal risk models. We look at market risk on a portfolio of 1,500 US corporate bonds and credit risk on a portfolio of 700 credit default swaps (CDSs). In both cases we aim to assess the VAR contribution by risk rating, and report the mean and standard deviation of CVAR as a percentage of the total portfolio VAR. The composition of the portfolios and the details of the risk models used are proprietary, but the key point is that since the bond portfolio contains only linear exposures the P&L will be approximately elliptical, while the credit loss distribution on the CDS portfolio will exhibit fat tails and non-Gaussian dependence. The market risk model uses 10,000 simulations, while the credit loss model uses 250,000.

CVARs (standard deviation in parenthesis) for the bond portfolio are given in table D. We can see that most of the risk in the bond portfolio comes from BBB and BB rated issuers, and that in this case the semi-parametric estimator yields the best estimate of CVAR.

Results for the CDS portfolio are shown in table E. This portfolio is mainly comprised of exposures to AA, A and BBB grade issuers, but almost all the default risk comes from the latter due to the small probability of default on the AA and A rated issuers. The loss distribution of this portfolio is fat-tailed, which means the CVAR estimates suffer more from noise than for an elliptical portfolio, even given the far higher number of Monte Carlo simulations used. The violation of the elliptical distribution assumption means that the semi-parametric estimator mis-estimates the CVARs considerably, failing, for example, to detect the negative CVAR that indicates a hedge position in CCC issuers.

Summary

The rescaled kernel exhibits the best performance overall, since it is robust in the presence of non-linearity, is less noisy than the finite difference method for all portfolios without strong negative correlation, and is highly efficient, particularly where we wish to compute many CVARs from a single portfolio. Where we can be sure that all the components follow an elliptical distribution (for example, when we have linear exposures to multivariate normal risk factors), the semi-parametric estimator is more effective, but in this case both VAR and CVAR can be computed analytically so there is little reason to use Monte Carlo simulation at all.

We remark that although we have focused on parametric Monte Carlo simulation, the proposed methods could also be used to decompose the risks in a historical simulation-based risk engine, though here the relatively low number of scenarios may be problematic. Kernel-based estimators can also be applied to compute component expected shortfall (see Scaillet, 2004). ■

Eduardo Epperlein is managing director and head of market, model and counterparty risk analytics at Citigroup. Alan Smillie is a quantitative analyst in Citigroup's risk architecture group. Email: eduardo.epperlein@citigroup.com, alan.smillie@citigroup.com

References

Carroll R, T Perry, H Yang and A Ho, 2001
A new approach to component VAR
Journal of Risk 3(3), pages 57–67

Epperlein E, 1998
Component VAR in large scale Monte Carlo simulations and its relationship to scenario extraction
ICBI Risk Management Conference, December

Epperlein E and P Sondhi, 1997
Are factor sensitivity limits consistent with value-at-risk limits?
Working paper, Citibank

Garman M, 1996
Improving on VAR
Risk May, pages 61–63

Garman M, 1997
Taking VAR to pieces
Risk October, pages 70–71

Gourieroux C, J-P Laurent and O Scaillet, 2000
Sensitivity analysis of value-at-risk
Journal of Empirical Finance 7(3), pages 225–245

Hallerbach W, 2002
Decomposing portfolio value-at-risk: a general analysis
Journal of Risk 5(2), pages 1–18

Litterman R, 1997a
Hot spots and hedges (I)
Risk March, pages 42–45

Litterman R, 1997b
Hot spots and hedges (II)
Risk May, pages 38–42

Scaillet O, 2004
Nonparametric estimation and sensitivity analysis of expected shortfall
Mathematical Finance 14, pages 115–129

Silverman B, 1986
Density estimation for statistics and data analysis
Chapman & Hall, London