Risk Measures, Risk Aggregation and Capital Allocation

We consider risk measures, risk aggregation and capital allocation in these lecture notes and build on our earlier introduction to Value-at-Risk (VaR) and Expected Shortfall (ES). We will follow Chapter 8 of the 2^{nd} edition of *Quantitative Risk Management* by *MFE* quite closely. This chapter, however, contains considerably more material than we will cover and it should be consulted if further details are required.

1 Coherent Measures of Risk

In 1999 Artzner et al. proposed a list of properties that any good risk measure should have and this list gave rise to the concept of coherent and incoherent measures of risk. Since then a substantial body of research has developed on the theoretical properties of risk measures and we describe some of these results here.

Let $\mathcal{M}$ denote the space of random variables representing portfolio losses over some fixed time interval, Δ. We assume that $\mathcal{M}$ is a *convex cone* so that if $L_1 \in \mathcal{M}$ and $L_2 \in \mathcal{M}$ then $L_1 + L_2 \in \mathcal{M}$ and $\lambda L_1 \in \mathcal{M}$ for every $\lambda > 0$. A *risk measure* is then a real-valued function, $\varrho : \mathcal{M} \rightarrow \mathbb{R}$, that satisfies certain desirable properties. $\varrho(L)$ may be interpreted as the riskiness of a portfolio or the amount of capital that should be added to a portfolio with a loss given by L, so that the portfolio can then be deemed acceptable from a risk point of view. Note that under this latter interpretation, portfolios with $\varrho(L) < 0$ are already acceptable and do not require capital injections. In fact, if $\varrho(L) < 0$ then capital could even be *withdrawn* while the portfolio would still remain acceptable. The following properties of a risk measure merit special attention:

Axiom 1 : (**Translation Invariance**) For all $L \in \mathcal{M}$ and every constant $a \in \mathbb{R}$, we have $\varrho(L + a) = \varrho(L) + a$.

This property is necessary if the risk-capital interpretation we stated above is to make sense.

Axiom 2 : (**Subadditivity**) For all $L_1, L_2 \in \mathcal{M}$, we have $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$.

This axiom reflects the idea that pooling risks helps to diversify a portfolio. While this has been the most debated of the risk axioms, it allows for the decentralization of risk management. For example, if a risk manager has a total risk budget of B, he can divide B into B_1 and B_2 where $B_1 + B_2 = B$. He can then allocate risk budgets of B_1 and B_2 to different trading desks or operating units in the organization, safe in the knowledge that the firm-wide risk will not exceed B.

Axiom 3 : (**Positive Homogeneity**) For all $L \in \mathcal{M}$ and every $\lambda > 0$ we have $\varrho(\lambda L) = \lambda \varrho(L)$.

This axiom is also somewhat controversial and has been criticized for not penalizing concentration of risk and any associated liquidity problems. In particular, if $\lambda > 0$ is very large, then some people claim that we should require $\varrho(\lambda L) > \lambda \varrho(L)$. However, such a result would be inconsistent with the subadditivity axiom. This is easily seen if we write

$$\varrho(nL) \;=\; \varrho(L + \cdots + L) \;\leq n\varrho(L) \tag{1}$$

where $n \in \mathbb{N}$ and the inequality follows from subadditivity. The positive homogeneity assumption states that we must have equality in (1). This reflects the fact that there are no diversification benefits when we hold multiples of the same portfolio, L.

Axiom 4 : (**Monotonicity**) For $L_1, L_2 \in \mathcal{M}$ such that $L_1 \leq L_2$ almost surely, we have $\varrho(L_1) \leq \varrho(L_2)$.

It is clear that any risk measure should satisfy this axiom.

Definition 1 *A risk measure, ϱ, acting on the convex cone $\mathcal{M}$ is called* **coherent** *if it satisfies the translation invariance, subadditivity, positive homogeneity and monotonicity axioms.*

Remark 1 *The criticisms of the subadditivity and positive homogeneity axioms have led to the study of* **convex risk measures**. *A convex risk measure satisfies the same axioms as a coherent risk measure except that the subadditivity and positive homogeneity axioms are replaced by the convexity axiom:*

*Axiom 5 : (***Convexity***) For $L_1, L_2 \in \mathcal{M}$ and $\lambda \in [0,1]$,*

$$\varrho(\lambda L_1 \; + \; (1-\lambda)L_2) \; \leq \; \lambda\varrho(L_1) \; + \; (1-\lambda)\varrho(L_2).$$

It is possible within the convex class to find risk measures that satisfy $\varrho(\lambda L) \geq \lambda\varrho(L)$ for $\lambda > 1$.

1.1 Value-at-Risk

Value-at-risk is **not** a coherent risk measure because it fails to be subadditive. This is perhaps the principal[1] criticism that is made of VaR when it is compared to other risk measures. We will see two examples below that demonstrate this. We first recall the definition of VaR.

Definition 2 *Let $\alpha \in (0,1)$ be some fixed confidence level. Then the VaR of the portfolio loss, L, at the confidence interval, α, is given by*

$$\mathit{VaR}_\alpha := q_\alpha(L) \; = \; \inf\{x \in \mathbb{R} \; : \; F_L(x) \geq \alpha\}.$$

where $F_L(\cdot)$ is the CDF of the random variable, L.

Example 1 Consider[2] two assets, X and Y, that are usually normally distributed but are subject to occasional shocks. In particular, assume that X and Y are independent and identically distributed with

$$X \; = \; \epsilon \; + \; \eta \quad \text{where } \epsilon \sim \mathsf{N}(0,1) \quad \text{and } \eta \; = \; \begin{cases} 0, & \text{with prob .991} \\ -10, & \text{with prob .009.} \end{cases}$$

Consider a portfolio consisting of X and Y. Then

$$\mathsf{VaR}_{.99}(X+Y) \; = \; 9.8 \; > \; \mathsf{VaR}_{.99}(X) + \mathsf{VaR}_{.99}(Y) \; = \; 3.1 + 3.1 = 6.2$$

thereby demonstrating the non-subadditivity of VaR. ∎

Exercise 1 *Confirm that the VaR values of 3.1 and 9.8 in the previous example are correct.*

We now give a more meaningful and disturbing example of how VaR fails to be sub-additive.

Example 2 (**VaR for a Portfolio of Defaultable Bonds (E.G. 6.7 in 1^{st} ed. of *MFE*)**)
Consider a portfolio of $n = 100$ defaultable corporate bonds where the probability of a default over the next year is identical for all bonds and is equal to 2%. We assume that defaults of different bonds are independent from one another. The current price of each bond is 100 and if there is no default, a bond will pay 105 one year from now. If the bond defaults then there is no repayment. This means we can define L_i, the loss on the i^{th} bond, as

$$L_i \; := \; 105Y_i \; - \; 5$$

where $Y_i = 1$ if the bond defaults over the next year and $Y_i = 0$ otherwise. By assumption we also see that $P(L_i = -5) = .98$ and $P(L_i = 100) = .02$. Consider now the following two portfolios:

[1] Of course, the general criticism that summarizing an entire loss distribution with just a single number can be applied to all risk measures, coherent or not. Furthermore, there is the implicit assumption that we know the loss distribution when determining the value of a risk measure. This assumption is often unjustifiable: it can and indeed has often led to financial catastrophe!

[2] This example is taken from "Subadditivity Re-Examined: the Case for Value-at-Risk" by Daníelsson et al.

A: A fully concentrated portfolio consisting of 100 units of bond 1.

B: A completely diversified portfolio consisting of 1 unit of each of the 100 bonds.

We can compute the 95% VaR for each portfolio as follows:

Portfolio A: The loss on portfolio A is given by $L_A = 100L_1$ so that $\mathrm{VaR}_{.95}(L_A) = 100\mathrm{VaR}_{.95}(L_1)$. Note that $P(L_1 \leq -5) = .98 > .95$ and $P(L_1 \leq l) = 0 < .95$ for $l < -.5$. We therefore obtain $\mathrm{VaR}_{.95}(L_1) = -5$ and so $\mathrm{VaR}_{.95}(L_A) = -500$. So the 95% VaR for portfolio A corresponds to a gain(!) of 500.

Portfolio B: The loss on portfolio B is given by

$$L_B \;=\; \sum_{i=1}^{100} L_i \;=\; 105 \sum_{i=1}^{100} Y_i \;-\; 500$$

and so $\mathrm{VaR}_{.95}(L_B) = 105\ \mathrm{VaR}_{.95}(\sum_{i=1}^{100} Y_i) - 500$. Note that $M := \sum_{i=1}^{100} Y_i \;\sim\; \mathrm{Bin}(100, .02)$ and by inspection we see that $P(M \leq 5) \approx .984 > .95$ and $P(M \leq 4) \approx .949 < .95$. Therefore $\mathrm{VaR}_{.95}(M) = 5$ and so $\mathrm{VaR}_{.95}(L_B) = 525 - 500 = 25$.

So according to $\mathrm{VaR}_{.95}$, portfolio B is riskier than portfolio A. This is clearly nonsensical. Note that we have shown that

$$\mathrm{VaR}_{.95}\left(\sum_{i=1}^{100} L_i\right) \;\geq\; 100\ \mathrm{VaR}_{.95}(L_1) \;=\; \sum_{i=1}^{100} \mathrm{VaR}_{.95}(L_i)$$

demonstrating again that VaR is not subadditive. ∎

Remark 2 *Let ϱ be any coherent risk measure that depends only on the distribution of L. Then we obtain*

$$\varrho\left(\sum_{i=1}^{100} L_i\right) \;\leq\; \sum_{i=1}^{100} \varrho(L_i) \;=\; 100\varrho(L_1)$$

and so in the previous example, ϱ would correctly classify portfolio A as being riskier than portfolio B.

We now describe a situation where VaR is always subadditive.

Theorem 1 (Subadditivity of VaR for Elliptical Risk Factors (Theorem 6.8 in *MFE*)) *Suppose that $\mathbf{X} \sim E_n(\mu, \mathbf{\Sigma}, \psi)$ and let $\mathcal{M}$ be the set of linearized portfolio losses of the form*

$$\mathcal{M} \;:=\; \{L : L \;=\; \lambda_0 \;+\; \sum_{i=1}^{n} \lambda_i X_i,\ \lambda_i \in \mathbb{R}\}.$$

Then for any two losses $L_1, L_2 \in \mathcal{M}$, and $0.5 \leq \alpha < 1$,

$$\mathit{VaR}_\alpha(L_1 + L_2) \;\leq\; \mathit{VaR}_\alpha(L_1) \;+\; \mathit{VaR}_\alpha(L_2).$$

Proof: Without loss of generality we may assume that $\lambda_0 = 0$. Recall also that if $\mathbf{X} \sim \mathrm{E}_n(\mu, \mathbf{\Sigma}, \psi)$ then $\mathbf{X} = \mathbf{AY} + \mu$ where $\mathbf{A} \in \mathbb{R}^{n \times k}$, $\mu \in \mathbb{R}^n$ and $\mathbf{Y} \sim S_k(\psi)$ is a spherical random vector. Any element $L \in \mathcal{M}$ can therefore be represented as

$$\begin{aligned} L \;=\; \lambda^T \mathbf{X} \;&=\; \lambda^T \mathbf{AY} \;+\; \lambda^T \mu \\ &\sim\; ||\lambda^T \mathbf{A}||\, Y_1 \;+\; \lambda^T \mu \end{aligned} \tag{2}$$

where (2) follows from part 3 of Theorem 2 in the *Multivariate Distributions* lecture notes. Now the translation invariance and positive homogeneity of VaR imply

$$\mathrm{VaR}_\alpha(L) \;=\; ||\lambda^T \mathbf{A}||\ \mathrm{VaR}_\alpha(Y_1) \;+\; \lambda^T \mu.$$

Suppose now that $L_1 := \lambda_1^T \mathbf{X}$ and $L_2 := \lambda_2^T \mathbf{X}$. The triangle inequality implies

$$||(\lambda_1 + \lambda_2)^T \mathbf{A}|| \;\leq\; ||\lambda_1^T \mathbf{A}|| \;+\; ||\lambda_2^T \mathbf{A}||$$

and since $\mathsf{VaR}_\alpha(Y_1) \geq 0$ for $\alpha \geq .5$ (why?), the result follows from (2). $\Box$

Remark 3 *It is a widely held belief that if the individual loss distributions under consideration are continuous and symmetric then VaR is subadditive. This is not true and a counterexample may be found in Section* 6.2 *of* MFE. *The loss distributions in the counterexample are smooth and symmetric but the copula is highly* asymmetric. *VaR can also fail to be subadditive when the individual loss distributions have heavy tails.*

1.2 Expected Shortfall

We now show that expected shortfall (ES) or CVaR is a coherent measure of risk. We first recall the definition of ES.

Definition 3 *For a portfolio loss, L, satisfying $\mathsf{E}[|L|] < \infty$ the expected shortfall at confidence level $\alpha \in (0,1)$ is given by*

$$ES_\alpha(L) \;:=\; \frac{1}{1-\alpha} \int_\alpha^1 q_u(F_L)\, du.$$

The relationship between ES_α and VaR_α is therefore given by

$$\mathsf{ES}_\alpha(L) \;:=\; \frac{1}{1-\alpha} \int_\alpha^1 \mathsf{VaR}_u(L)\, du \tag{3}$$

from which it is clear that $\mathsf{ES}_\alpha(L) \geq \mathsf{VaR}_\alpha(L)$. When the CDF, F_L, is continuous then a more well known representation of $\mathsf{ES}_\alpha(L)$ is given by

$$\mathsf{ES}_\alpha(L) \;:=\; \frac{\mathsf{E}\left[L;\; L \geq q_\alpha(L)\right]}{1-\alpha} \;=\; \mathsf{E}\left[L \mid L \geq \mathsf{VaR}_\alpha\right]. \tag{4}$$

The following result demonstrates that expected shortfall is a coherent risk measure. We again follow the proof in *MFE*.

Theorem 2 *Expected shortfall is a coherent risk measure.*

Proof: The translation invariance, positive homogeneity and monotonicity properties all follow from the representation of ES in (3) and the same properties for quantiles. We therefore only need to demonstrate subadditivity.

Let $L_1, \ldots, L_n$ be a sequence of random variables and let $L_{1,n} \geq \cdots \geq L_{n,n}$ be the associated sequence of order statistics. Note that

$$\sum_{i=1}^m L_{i,n} \;=\; \sup\{L_{i_1} + \cdots + L_{i_m} \;:\; 1 \leq i_1 < \cdots < i_m \leq n\} \tag{5}$$

where $m \in \mathbb{N}$ satisfying $1 \leq m \leq n$ is arbitrary. Now let $(L, \tilde{L})$ be a pair of random variables with joint CDF, F, and let $(L_1, \tilde{L}_1), \ldots, (L_n, \tilde{L}_n)$ be an IID sequence of bivariate random vectors with this same CDF. Then

$$\begin{aligned}
\sum_{i=1}^m (L + \tilde{L})_{i,n} \;&=\; \sup\{(L+\tilde{L})_{i_1} + \cdots + (L+\tilde{L})_{i_m} \;:\; 1 \leq i_1 < \cdots < i_m \leq n\} \\
&\leq\; \sup\{L_{i_1} + \cdots + L_{i_m} \;:\; 1 \leq i_1 < \cdots < i_m \leq n\} \;+\; \sup\{\tilde{L}_{i_1} + \cdots + \tilde{L}_{i_m} \;:\; 1 \leq i_1 < \cdots < i_m \leq n\} \\
&=\; \sum_{i=1}^m L_{i,n} \;+\; \sum_{i=1}^m \tilde{L}_{i,n}. \qquad (6)
\end{aligned}$$

Now set[3] $m = \lfloor n(1-\alpha) \rfloor$ and let $n \to \infty$. It may be shown[4] that $\frac{1}{m}\sum_{i=1}^{m} L_{i,n} \to \mathrm{ES}_\alpha(L)$, $\frac{1}{m}\sum_{i=1}^{m} \tilde{L}_{i,n} \to \mathrm{ES}_\alpha(\tilde{L})$ and $\frac{1}{m}\sum_{i=1}^{m} (L+\tilde{L})_{i,n} \to \mathrm{ES}_\alpha(L+\tilde{L})$. The subadditivity of ES then follows immediately from (6). □

There are many other examples of risk measures that are coherent. They include, for example, risk measures based on *generalized scenarios* and *spectral* risk measures of which expected shortfall is an example.

2 Bounds for Aggregate Risk

Let $\mathbf{L} = (L_1, \ldots, L_n)$ denote a vector of random variables, each one representing a loss on a particular trading desk, portfolio or operating unit within a firm. Sometimes we wish to *aggregate* these losses into a single random variable, $\psi(\mathbf{L})$, say. Common examples of the aggregating function, $\psi(\cdot)$, include:

- The total loss so that $\psi(\mathbf{L}) = \sum_{i=1}^{n} L_i$.
- The maximum loss where $\psi(\mathbf{L}) = \max\{L_1, \ldots, L_n\}$.
- The *excess-of-loss treaty* so that $\psi(\mathbf{L}) = \sum_{i=1}^{n} (L_i - k_i)^+$.
- The *stop-loss treaty* in which case $\psi(\mathbf{L}) = \left(\sum_{i=1}^{n} L_i \ - \ k\right)^+$.

We wish to understand the risk of the aggregate loss function, $\varrho(\psi(\mathbf{L}))$, but to do so we need to know the distribution of $\psi(\mathbf{L})$. In practice, however, we often know only the distributions of the L_i's and have little or no information about the dependency or copula of the L_i's. In this case we can try to compute lower and upper bounds on $\varrho(\psi(\mathbf{L}))$. In particular we can formulate the two problems

$$\begin{aligned} \varrho_{min} &:= \inf\{\varrho(\psi(\mathbf{L})) \ : \ L_i \sim F_i, \ i = 1, \ldots, n\} \\ \varrho_{max} &:= \sup\{\varrho(\psi(\mathbf{L})) \ : \ L_i \sim F_i, \ i = 1, \ldots, n\} \end{aligned}$$

where F_i is the CDF of the loss, L_i. Problems of this type are referred to as *Frechet* problems and solutions are available in some circumstances. Indeed, when we study copulas we will see an example of such a problem when we address the question of attainable correlations given known marginal distributions. In a risk management context, these problems have been studied in some detail when $\psi(\mathbf{L}) = \sum_{i=1}^{n} L_i$ and $\varrho(\cdot)$ is the VaR function. Results related to this problem are generally of more theoretical than practical interest and so we will not discuss them any further. Results and references, however, can be found in Section 6.2 of *MFE*.

3 Capital Allocation

Consider again a total loss given by $L = \sum_{i=1}^{n} L_i$ and suppose we have determined the risk, $\varrho(L)$, of this loss. The *capital allocation* problem seeks a decomposition, $AC_1, \ldots, AC_n$, such that

$$\varrho(L) \ = \ \sum_{i=1}^{n} AC_i \tag{7}$$

and where AC_i is interpreted as the risk capital that has been allocated to the i^{th} loss, L_i. This problem is important in the setting of performance evaluation where we want to compute a risk-adjusted return on capital (RAROC). This return might be estimated, for example, by *Expected Profit / Risk Capital* and in order to

[3] $\lfloor x \rfloor$ is defined to be the largest integer less than or equal to x, i.e. the *floor* of x.
[4] See, for example, Lemma 2.20 in *MFE*.

compute this we must determine the risk capital of each of the L_i's. Obviously, we would require the corresponding risk capitals to sum to the total risk capital so that (7) is satisfied.

More formally, let $L(\lambda) := \sum_{i=1}^n \lambda_i L_i$ be the loss associated with the portfolio consisting of λ_i units of the loss, L_i, for $i = 1, \ldots, n$. The loss on the actual portfolio under consideration is then given by $L(\mathbf{1})$. Let $\varrho(\cdot)$ be a risk measure on a space $\mathcal{M}$ that contains $L(\lambda)$ for all $\lambda \in \Lambda$, an open set containing $\mathbf{1}$. Then the associated *risk measure function*, $r_\varrho : \Lambda \rightarrow \mathbb{R}$, is defined by $r_\varrho(\lambda) = \varrho(L(\lambda))$. We have the following definition.

Definition 4 *Let r_ϱ be a risk measure function on some set $\Lambda \subset \mathbb{R}^n \setminus \mathbf{0}$ such that $\mathbf{1} \in \Lambda$. Then a mapping, $f^{r_\varrho} : \Lambda \rightarrow \mathbb{R}^n$, is called a per-unit* capital allocation principle *associated with r_ϱ if, for all $\lambda \in \Lambda$, we have*

$$\sum_{i=1}^n \lambda_i f_i^{r_\varrho}(\lambda) \;=\; r_\varrho(\lambda). \tag{8}$$

We then interpret $f_i^{r_\varrho}$ as the amount of capital allocated to one unit of L_i when the overall portfolio loss is $L(\lambda)$. The amount of capital allocated to a position of $\lambda_i L_i$ is therefore $\lambda_i f_i^{r_\varrho}$ and so by (8), the total risk capital is fully allocated.

Definition 5 (Euler Capital Allocation Principle) *If r_ϱ is a positive-homogeneous risk-measure function which is differentiable on the set Λ, then the per-unit Euler capital allocation principle associated with r_ϱ is the mapping*

$$f^{r_\varrho} : \Lambda \rightarrow \mathbb{R}^n \;:\; f_i^{r_\varrho}(\lambda) \;=\; \frac{\partial r_\varrho}{\partial \lambda_i}(\lambda).$$

The Euler allocation principle is seen to be a full allocation principle since a well-known property of any positive homogeneous and differentiable function, $r(\cdot)$ is that it satisfies $r(\lambda) = \sum_{i=1}^n \lambda_i \frac{\partial r}{\partial \lambda_i}(\lambda)$. The Euler allocation principle therefore gives us different risk allocations for different positive homogeneous risk measures. It should also be mentioned that there are good economic reasons[5] for employing the Euler principle when computing capital allocations. We will not discuss those reasons here, however.

We now describe the Euler allocation for some specific risk measure below. Section 6.3 of *MFE* should be consulted for proofs and further details if necessary.

Standard Deviation and the Covariance Principle

Let $r_{sd}(\lambda) = \text{std}(L(\lambda))$ be our risk measure function and write $\mathbf{\Sigma}$ for the variance-covariance matrix of $L_1, \ldots, L_n$. Then $r_{sd}(\lambda) \;=\; \left(\lambda^T \mathbf{\Sigma} \lambda\right)^{1/2}$ and using the Euler allocation principle it follows that

$$f_i^{r_{sd}}(\lambda) \;=\; \frac{\partial r_{sd}}{\partial \lambda_i}(\lambda) \;=\; \frac{(\mathbf{\Sigma}\lambda)_i}{r_{sd}(\lambda)} \;=\; \frac{\sum_{j=1}^n \mathrm{Cov}(L_i, L_j)\lambda_j}{r_{sd}(\lambda)} \;=\; \frac{\mathrm{Cov}(L_i, L(\lambda))}{\sqrt{\mathrm{Var}(L(\lambda))}} \tag{9}$$

and the actual capital allocation, AC_i, for L_i is obtained by setting $\lambda = \mathbf{1}$ in (9). This is then known as the *covariance principle*.

Value-at-Risk and Value-at-Risk Contributions

If $r_{VaR}^\alpha(\lambda) = \text{VaR}_\alpha(L(\lambda))$ is our risk measure function, then subject to technical conditions it can be shown that

$$f_i^{r_{VaR}^\alpha}(\lambda) \;=\; \frac{\partial r_{VaR}^\alpha}{\partial \lambda_i}(\lambda) \;=\; \mathrm{E}\left[L_i \mid L(\lambda) = \text{VaR}_\alpha(L(\lambda))\right], \quad \text{for } \; i = 1, \ldots, n. \tag{10}$$

[5] See Section 6.3.3 of *MFE* for these reasons.

Once again, the actual capital allocation, AC_i, for L_i is then obtained by setting $\lambda = \mathbf{1}$ in (10).

Expected Shortfall and Shortfall Contributions

If $r_{ES}^{\alpha}(\lambda) = \mathrm{E}\left[L(\lambda) \mid L(\lambda) \geq \mathsf{VaR}_{\alpha} L(\lambda)\right]$ is our risk measure function, then subject again to technical conditions it can be shown that

$$f_i^{r_{ES}^{\alpha}}(\lambda) \;=\; \frac{\partial r_{ES}^{\alpha}}{\partial \lambda_i}(\lambda) \;=\; \frac{1}{1-\alpha}\, \mathrm{E}\left[L_i \mid L(\lambda) \geq \mathsf{VaR}_{\alpha}(L(\lambda))\right], \quad \text{for } i = 1, \ldots, n. \tag{11}$$

We therefore have the capital allocation $AC_i = \mathrm{E}\left[L_i \mid L \geq \mathsf{VaR}_{\alpha}(L)\right]$ for the risk, L_i, where $L := L(\mathbf{1})$.

3.1 An Application: Estimating Value-at-Risk Contributions

We now consider an application[6] where we will use (10) to estimate the VaR contributions from each security in a portfolio. We will do this via Monte-Carlo simulation, a general approach that can be used for complex portfolios where (10) cannot be calculated analytically. Recall that the total portfolio loss is given by $L = \sum_{i=1}^n L_i$. According to (10) with $\lambda = 1$ we know that

$$AC_i \;=\; \mathrm{E}\left[L_i \mid L = \mathsf{VaR}_{\alpha}(L)\right] \tag{12}$$

$$\;=\; \left.\frac{\partial\, \mathsf{VaR}_{\alpha}(\lambda)}{\partial \lambda_i}\right|_{\lambda=1}$$

$$\;=\; w_i \frac{\partial\, \mathsf{VaR}_{\alpha}}{\partial w_i} \tag{13}$$

for $i = 1, \ldots, n$ and where w_i is the number of units of the i^{th} security held in the portfolio.

Question: How might we use Monte-Carlo to estimate the VaR contribution, AC_i, of the i^{th} asset?

Solution: There are three approaches we might take:

1. As AC_i is a (mathematical) derivative we could estimate it numerically. In particular a finite-difference estimator based on (13) would take the form

$$\widehat{AC}_i := \frac{\mathsf{VaR}_{\alpha}^{i,+} - \mathsf{VaR}_{\alpha}^{i,-}}{2\delta_i} \tag{14}$$

   where $\mathsf{VaR}_{\alpha}^{i,+}$ ($\mathsf{VaR}_{\alpha}^{i,-}$) is the portfolio VaR when the number of units of the i^{th} security is increased (decreased) by $\delta_i w_i$ units. Each term in the numerator of (14) can be estimated via Monte-Carlo. For variance reduction purposes, the same set of random returns should be used to estimate each term. It remains to choose an appropriate value of δ_i. There is a bias-variance tradeoff to be made in this choice and a value of $\delta_i = .05$ or $.1$ seems to lead to reasonable results in practice.

   Note that this estimator will not satisfy the additivity property so that $\sum_i^n AC_i \neq \mathsf{VaR}_{\alpha}$. It easy to perform a re-scaling of the estimated $\widehat{AC}_i$'s so that the property will be satisfied.

2. Another approach is to estimate (12) directly. We could do this by simulating N portfolio losses $L^{(1)}, \ldots, L^{(N)}$ with $L^{(j)} = \sum_{i=1}^n L_i^{(j)}$ where $L_i^{(j)}$ is the loss on the i^{th} security in the j^{th} simulation. We could then set (why?) $AC_i = L_i^{(m)}$ where m denotes the VaR_{α} scenario, i.e. $L^{(m)}$ is the $\lceil N(1-\alpha) \rceil^{th}$ largest of the N simulated portfolio losses.

   Question: Will this estimator satisfy the additivity property, i.e. will $\sum_i^n AC_i = \mathsf{VaR}_{\alpha}$?

   Question: What is the problem with this approach? Will this problem disappear if we let $N \to \infty$?

---

[6]See *"Cracking VAR with kernels"* (RISK, 2006) by E. Epperlein and A. Smillie for a more complete application and discussion. See also *"Simulations with Exact Means and Covariances"* (2009) by A. Meucci for an application where the VaR contributions of a equity options portfolio are estimated.

3. An alternative approach that resolves the problem alluded to in the previous question is to take a weighted average of the losses in the i^{th} security around the VaR_α scenario. One convenient way to do this is via a **kernel** function. In particular, we say $K(x;h) := K\left(\frac{x}{h}\right)$ is a kernel function if it is symmetric about zero, takes a maximum at $x = 0$ and is non-negative for all x. A simple choice is to take the triangle kernel so that
$$K(x;h) := \max\left(1 - \left|\frac{x}{h}\right|, 0\right).$$
The kernel estimate of AC_i is then given by
$$\widehat{AC}_i^{ker} := \frac{\sum_{j=1}^N K\left(L^{(j)} - \widehat{\mathsf{VaR}}_\alpha; h\right) L_i^{(j)}}{\sum_{j=1}^N K\left(L^{(j)} - \widehat{\mathsf{VaR}}_\alpha; h\right)} \tag{15}$$
where $\widehat{\mathsf{VaR}}_\alpha := L^{(m)}$ with m as defined above. One minor problem with (15) is that the additivity property doesn't hold. We can easily correct this by instead setting
$$\widehat{AC}_i^{ker} := \widehat{\mathsf{VaR}}_\alpha \frac{\sum_{j=1}^N K\left(L^{(j)} - \widehat{\mathsf{VaR}}_\alpha; h\right) L_i^{(j)}}{\sum_{j=1}^N K\left(L^{(j)} - \widehat{\mathsf{VaR}}_\alpha; h\right) L^{(j)}}. \tag{16}$$
It remains to choose an appropriate value of the smoothing parameter, h. It can be shown that an optimal choice (in the sense of minimizing mean-squared error) is to set
$$h = 2.575\sigma N^{-1/5}$$
where σ is the standard deviation of L, a quantity that we can easily estimate.

Exercise 2 *How would you use Monte-Carlo to estimate AC_i when we use expected shortfall as our risk measure?*

When Losses Are Elliptically Distributed

If $L_1, \ldots, L_N$ have an elliptical distribution then it may be shown that
$$AC_i \;=\; \mathsf{E}[L_i] + \frac{\mathrm{Cov}(L, L_i)}{\mathrm{Var}(L)} \left(\mathsf{VaR}_\alpha(L) - \mathsf{E}[L]\right). \tag{17}$$
In our numerical example below, we will assume that the portfolio losses are elliptically distributed so that the VaR contributions can also be computed analytically via (17) and then compared to the Monte-Carlo estimates obtained from (16).

Example 3 (Elliptically Distributed Losses)
In Figure 3 we have plotted the estimated $VaR_{\alpha=.99}$ contributions of a portfolio consisting of $n = 10$ securities. We assumed losses were multivariate normally distributed so that $(L_1, \ldots, L_n) \sim \mathsf{MN}_n(\mathbf{0}, \mathbf{\Sigma})$. The first eight securities were all positively correlated with one another, the second-to-last security was uncorrelated and the last security had a correlation of -.2 with the remaining securities. As a result we see that the last two securities have a *negative* contribution to the total portfolio VaR. The "naive" Monte-Carlo estimator refers to the estimator outlined in approach #2 above.

∎

3.2 Estimating Factor Contributions to Value-at-Risk

We have focussed on computing the VaR contribution from individual securities. We might prefer, however, to compute the VaR contribution from a collection of risk factors and it is relatively straightforward to do this. In fact, our earlier calculations can be viewed as a specific case of this where we have one (change in risk) factor

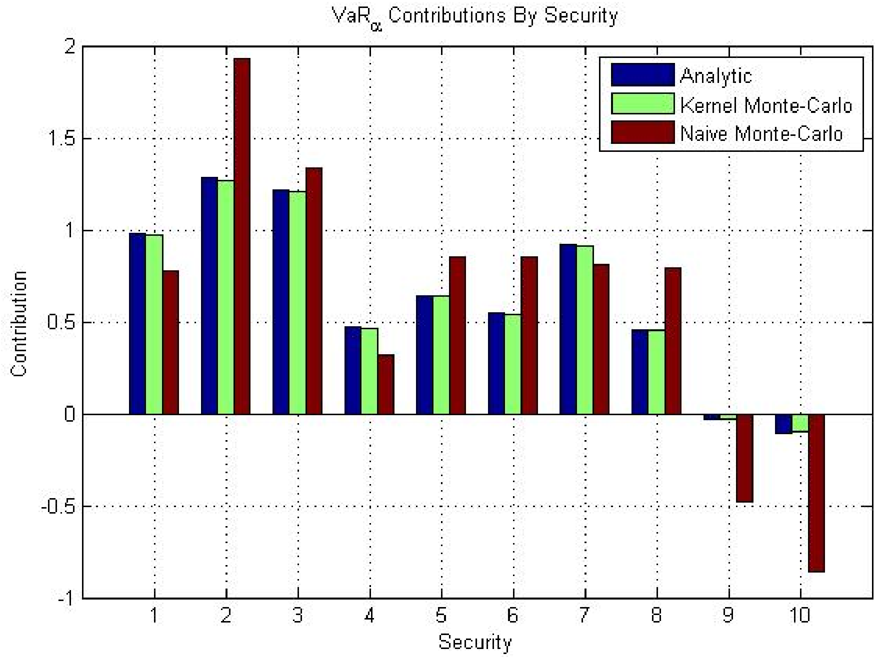

Figure 1: $\mathrm{VaR}_{.99}(L)$ by Security. The analytic numbers are correct as the underlying losses were elliptically – in this case multivariate normally – distributed. The Monte-Carlo estimates were based on 100k simulations. The kernel-smoothed estimate is very accurate whereas the naive approach is (not surprisingly) very poor. The true $\mathrm{VaR}_{.99}(L)$ was \$6.37m and the estimated $\mathrm{VaR}_{.99}(L)$ was \$6.32m.

for each security in the portfolio. In that case the i^{th} factor, F_i, is simply the return on the i^{th} security and we can write $L = \sum_{i=1}^{n} w_i F_i$. The contribution of F_i to the portfolio VaR is then AC_i as computed earlier.

More generally we might be interested in risk factors, $\tilde{\mathbf{F}}$, defined according to $\tilde{\mathbf{F}} := \mathbf{P}\mathbf{F}$ for some constant matrix $\mathbf{P}$. In the case that $\mathbf{P}$ is invertible we can write

$$L = \mathbf{w}^\top \mathbf{F} = \mathbf{w}^\top \mathbf{P}^{-1}\mathbf{P}\mathbf{F} = \tilde{\mathbf{w}}^\top \tilde{\mathbf{F}}$$

where

$$\tilde{\mathbf{w}}^\top := \mathbf{w}^\top \mathbf{P}^{-1}. \tag{18}$$

We can therefore view the portfolio as a portfolio with individual losses $\tilde{F}_1, \ldots, \tilde{F}_n$ with positions $\tilde{w}_1, \ldots, \tilde{w}_n$ and compute the VaR contributions as above. If however, we have already estimated the VaR contributions from the original securities (or risk factors $\mathbf{F}$), then we can use these estimates by noting that

$$\frac{\partial\, \mathsf{VaR}_\alpha}{\partial \tilde{\mathbf{w}}} = \mathbf{P}\, \frac{\partial\, \mathsf{VaR}_\alpha}{\partial \mathbf{w}}. \tag{19}$$

As in (13) the VaR contribution of $\tilde{F}_i$ is then given by

$$\tilde{AC}_i = \tilde{w}_i \frac{\partial\, \mathsf{VaR}_\alpha}{\partial \tilde{w}_i}$$

which is easily calculated from the previously calculated $AC_1, \ldots, AC_n$ using (18) and (19).

In the case where $\mathbf{P}$ is not invertible then a little more work is required. But see A. Meucci's *"Risk Contributions from Generic User-Defined Factors "* for how to handle this case as well as additional examples.