British Journal of Mathematics & Computer Science
4(13): 1797-1814, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Market Value-At-Risk: ROM Simulation, Cornish-Fisher Var and Chebyshev-Markov Var Bound

Werner Hürlimann[1*]

[1]*Wolters Kluwer Financial Services Switzerland AG, Seefeldstrasse 69, CH-8008 Zürich, Switzerland.*

Original Research Article

Received: 24 March 2014
Accepted: 18 April 2014
Published: 05 May 2014

Abstract

We apply the recently developed sampling algorithm, called random orthogonal matrix (ROM) simulation by Ledermann et al. [3], to compute VaR of a market risk portfolio. Typically, the covariance matrix has a large influence on ROM VaR. But VaR, being a lower quantile of the portfolio return distribution, is also much impacted by the skewness and kurtosis of the risk factor returns. With ROM VaR it is possible to stress test risk factors under adverse market conditions by targeting other sample moments that are consistent with periods of financial crisis. In particular, the important effects of skewness or kurtosis in the tail of the portfolio returns can be incorporated in ROM VaR. In a simulation study, we integrate ROM VaR into other methods that take into account skewness and kurtosis, namely the Cornish-Fisher VaR approximation and a robust approximation to the Chebyshev-Markov VaR upper bound in Hürlimann [7].

Keywords: MC simulation, orthogonal matrix, cornish-fisher expansion, chebyshev-markov inequalities, skewness, kurtosis, value-at-risk.

1 Introduction

According to [1] there are three basic types of VaR models, namely the *normal linear* VaR model, also called *parametric* VaR or *variance-covariance* VaR, the *historical* VaR *simulation* model, and the *Monte Carlo* VaR (MC VaR) model. The Monte Carlo framework is the most flexible of all, and may be used with a great diversity of market risk factor return distributions. However, a main disadvantage of the MC VaR model is the lack of fast computation due to the large number of simulation steps required to reach a given level of accuracy. However, with the increasing computer power this drawback becomes less relevant. Two equally important design aspects of MC VaR are the *sampling algorithm* ([1], IV.4.2) and the different *statistical models for risk factor returns* to which the algorithm is applied ([1], IV.4.3 and IV.4.4). Besides these technical tools it is very important to control two sources of *model risk* in MC VaR, namely the simulation errors through an appropriate choice of the *sampling method*, and the errors due to inappropriate

**Corresponding author: whurlimann@bluewin.ch;*

behavioural models for risk factor returns. Since *variance reduction techniques* used to reduce the simulation errors in MC methods are well-known (e.g. [2], Chap. 4), we concentrate here on the recent and important sampling algorithm called *random orthogonal matrix* (ROM) *simulation*, which is introduced in [3] (see also [4,5,6]). The authors describe this novel Monte Carlo algorithm as follows:

"ROM simulation eliminates sampling error in the sample mean vector, covariance matrix and the Mardia multivariate skewness and kurtosis measures, so that in each simulation they match exactly their target values."

This attractive property leads to the following advantages. It implies that within ROM simulation it is possible to specify in advance the mean vector and the covariance matrix of the risk factor returns. In industry practice, these risk characteristics are often estimated using the so-called "Risk Metrics VaR Methodology" ([1], IV.2.10.3). Typically, the covariance matrix has a large influence on ROM VaR. But VaR, being a lower quantile of the portfolio return distribution, is also much impacted by the skewness and kurtosis of the risk factor returns. Like historical VaR the new method is non-parametric. However, the limitation of historical VaR to past observations implies that history will repeat itself in the sense that the risk factor returns over the risk horizon are identical to their distributions in the historical sample (inappropriate behavioural pattern). With ROM VaR it is possible to simulate a very large number of realized risk factor returns that are all consistent with given observed historical sample moments. Moreover, we can stress test risk factors under adverse market conditions by targeting other sample moments that are consistent with periods of financial crisis. In particular, the important effects of skewness and kurtosis in the tail of the portfolio returns can be incorporated in ROM VaR. In a simulation study, we compare ROM VaR with other methods that take into account skewness and kurtosis, namely the Cornish-Fisher VaR approximation and the Chebyshev-Markov VaR upper bound and its robust approximation introduced in [7].

A brief account of the content follows. Section 2 describes in equation (2.5) the fundamental *ROM sampling algorithm* that generates random samples with exact mean and covariance matrix using a random permutation matrix, a random orthogonal matrix, and a deterministic so-called L matrix (in honor of W. Ledermann). An introduction to the multivariate Mardia skewness and kurtosis of the L matrices is given in Section 2.1 and the required properties are described in Section 2.3. A short review on the random square orthogonal matrices, which pre- or post-multiply a given L matrix during a ROM simulation is contained in Section 2.2. Section 3 considers the application of ROM simulation to Market VaR (Section 3.1) in combination with two semi-parametric models that take into account skewness and kurtosis, namely the Cornish-Fisher VaR approximation (Section 3.2) and two robust approximations derived from the Chebyshev-Markov VaR upper bound (Section 3.3). Section 4 illustrates with a numerical case study and provides some comments and conclusions. Finally, the importance of the present topic for the revised Basel III project as well as its potential application to other sources of risk like credit risk and liquidity risk should be emphasized.

2 ROM Simulation

Monte Carlo simulation consists to generate a random sample $X_{m,n}$ of size m on $n<m$ random variables $X_1,...,X_n$. Consider the multivariate normal (MVN) model, where the sample mean vector and sample covariance matrix of $X_{m,n}$ match the mean (column) vector μ_n and covariance matrix C_n such that (without bias adjustments for ease of notation)

$$m^{-1}\cdot(X_{m,n}-1_m\cdot\mu_n^T)^T\cdot(X_{m,n}-1_m\cdot\mu_n^T)=C_n. \tag{2.1}$$

Setting $X_{m,n}=Z_{m,n}+1_m\cdot(\mu_n^T-\bar{z}_n^T)$ with $Z_{m,n}$ a MVN simulation with sample mean $\bar{z}_n$, yields a simulation with mean μ_n. A priori, it is however not obvious that a random matrix $X_{m,n}$ satisfying (2.1) will exist whatever the choice of the covariance matrix. Since C_n is a positive semi-definite matrix, it is always possible to find a decomposition of the form $C_n=B_n^T\cdot B_n$, for example the Cholesky decomposition, the spectral decomposition (e.g. [8]) or the hyper-sphere decomposition (e.g. [9]). Then, applying the transformation

$$L_{m,n}=m^{-1/2}\cdot(X_{m,n}-1_m\cdot\mu_n^T)\cdot B_n^{-1}, \tag{2.2}$$

one sees that solving (2.1) is equivalent to finding a matrix $L_{m,n}$ satisfying the following conditions (*orthonormal relation with hyper-plane constraint*)

$$L_{m,n}^T\cdot L_{m,n}=E_n,\quad 1_m^T\cdot L_{m,n}=0_n^T, \tag{2.3}$$

where E_n is the identity matrix and 0_n is the null vector. Now, solving (2.3) and inverting the transformation (2.2) enables the generation of *exact MVN samples* for any prescribed sample mean μ_n and covariance matrix C_n. Any rectangular orthonormal matrix $L_{m,n}$ satisfying the hyper-plane constraint $1_m^T\cdot L_{m,n}=0_n^T$ is called an *L matrix* (in honor of W. Ledermann), which is fundamental to ROM simulation. The set of all *L* matrices has been classified into deterministic, parametric, data-specific and hybrid *L* matrices (see [3], Section 1). In particular, *deterministic ROM simulation* includes the so-called *Ledermann matrix* $L_{m,n}^*=\left(\ell_{m-n},...,\ell_{m-1}\right)$, which consists of the last $n\in\{2,...,m-1\}$ orthonormal columns of the matrix $L_{m,m-1}=\left(\ell_1,...,\ell_{m-1}\right)$ defined by

$$\ell_i=(1/\sqrt{i(i+1)})\cdot(1,...,1,-i,0,...,0)^T,\quad i=1,...,m-1, \tag{2.4}$$

where in ℓ_i the single entry $-i$ is preceded by a number of i one entries and followed by $n-1-i$ zero entries. In fact, adding the last column $\ell_m = \left(1/\sqrt{m},...,1/\sqrt{m}\right)^T$ to $L_{m,m-1}$, one obtains the orthogonal matrix $L_m = (\ell_1,...,\ell_m) \in O(m)$ (the *orthogonal group* of order m) that satisfies the hyper-plane constraint $1_m^T \cdot L_{m,m-1} = 0_{m-1}^T$. This matrix corresponds to the transpose of the "Helmert orthogonal matrix" introduced by [10]. Besides ROM simulation it has found many other applications (e.g. [11], example (1.5), [12], [13], p.1).

Now, given an arbitrary *L* matrix $L_{m,n}$, a random permutation matrix P_m and a random orthogonal matrix $R_n \in O(n)$, one sees that the specification,

$$X_{m,n} = 1_m \cdot \mu_n^T + \sqrt{m} \cdot P_m \cdot L_{m,n} \cdot R_n \cdot B_n, \tag{2.5}$$

which defines the *ROM sampling* algorithm, generates a random sample with exact mean μ_n and covariance matrix C_n. Indeed, it is clear that $L_{m,n} \cdot R_n$ is an *L* matrix, and the fact that $P_m \cdot L_{m,n}$ is also an *L* matrix follows from the validity of the property $1_m^T \cdot P_m = 1_m^T$, which implies that the columns of the product $P_m \cdot L_{m,n}$ sum to zero, i.e. the hyper-plane constraint in (2.3) is fulfilled. If P_m is not a permutation matrix, then the latter property does not necessarily hold. This is why the *L* matrices appearing in (2.5) can be pre-multiplied by permutations, but general orthogonal matrices can only be post-multiplied. The equation (2.5) is the *foundation of ROM simulation* as a means to simulate infinitely many random samples that have identical sample mean vectors and covariance matrices.

What about the multivariate *skewness* and *kurtosis* of ROM samples (2.5)? How are these characteristics related to the choice of a given *L* matrix? These questions are discussed in the next Subsection.

2.1 Multivariate Skewness and Kurtosis

As in the univariate case, there are many different ways to measure the skewness and kurtosis of a multivariate sample (e.g. [14,15,16]). To fix ideas we focus on the Mardia multivariate measure of skewness and kurtosis, which has also been used in [3].

For a mxn random sample $X_{m,n} = \left(x_1^T,...,x_m^T\right)^T$ in row vector notation with $x_i = (x_{i1},...,x_{in}),\ i = 1,...,m$, the Mardia measures of skewness and kurtosis are defined by the formulas

$$\tau_M(X_{m,n}) = m^{-2} \cdot \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{(x_i - \bar{x}) \cdot S_X^{-1} \cdot (x_j - \bar{x})^T\right\}^3,$$
$$\kappa_M(X_{m,n}) = m^{-1} \cdot \sum_{i=1}^{m}\left\{(x_i - \bar{x}) \cdot S_X^{-1} \cdot (x_i - \bar{x})^T\right\}^2, \tag{2.6}$$

where S_X is the mxm sample covariance matrix of $X_{m,n}$ and $\bar{x}$ is the row vector of sample means. These measures are known to be invariant under non-singular affine transformation of the type $Y_{m,n} = X_{m,n} \cdot A_{n,n} + 1_m \cdot b_n$, where $A_{n,n}$ is any invertible matrix and b_n is any row vector, i.e. $\tau_M(Y_{m,n}) = \tau_M(X_{m,n}),\ \kappa_M(Y_{m,n}) = \kappa_M(X_{m,n})$. The invariance property is the *cornerstone of ROM simulation*. Indeed, besides preserving the mean vector and the covariance matrix, the multivariate skewness and kurtosis sampling properties of (2.5) are encrypted in the matrix $L_{m,n}$ in the sense that $\tau_M(X_{m,n}) = \tau_M(L_{m,n}),\ \kappa_M(X_{m,n}) = \kappa_M(L_{m,n})$ for all ROM simulated samples $X_{m,n}$. This is due to the fact that (2.5) is a non-singular affine transformation of $L_{m,n}$, and therefore, the multivariate skewness and kurtosis measures are preserved under such transformations.

It is therefore of great importance to study the skewness and kurtosis of deterministic *L* matrices. Without loss of generality it suffices to consider *L* matrices of the type $L_{m,m-1}$. Recall that $L_{m,m-1}$ denotes the rectangular matrix obtained from an orthogonal matrix $L_m \in O(m)$ satisfying the constraint $1_m^T \cdot L_{m,m-1} = 0_{m-1}^T$ by deleting the last column. For each $n = 2,...,m-1$, let $L_{m,n}$ (respectively $L_{m,n}^*$) be the rectangular orthonormal matrices built up by the first (respectively last) n orthonormal columns of $L_{m,m-1}$. Clearly, if $m \geq 3$ a number of $\frac{1}{2}(m-1)m$ such rectangular orthonormal matrices could be build up from $L_{m,m-1}$. For ease of notation, only the simplest specified cases are used. In order to describe the skewness and kurtosis of such rectangular matrices through simple algebraic formulas, it is appropriate to use various (partial) *inner products* over subspaces of the Euclidean space R^{m-1} defined and denoted by

$$< x, y >_n = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j, \quad < x, y >_n^* = \sum_{i=m-n}^{m-1}\sum_{j=m-n}^{m-1} x_i y_j, \tag{2.7}$$

for each pair of $1x(m-1)$ row vectors $x = (x_1,...,x_{m-1}),\quad y = (y_1,...,y_{m-1}) \in R^{m-1}$. The corresponding natural (partial) *norms* are defined and denoted by

$$\|x\|_n = \sqrt{< x, x >_n}, \quad \|x\|_n^* = \sqrt{< x, x >_n^*}. \tag{2.8}$$

For the full space R^{m-1} the inner products and norms for $n = m-1$ coincide. In this situation the lower indices are omitted and one just writes $< x, y > = < x, y >_{m-1} = < x, y >^*_{m-1}$ and $\|x\| = \|x\|_{m-1} = \|x\|^*_{m-1}$.

Lemma 2.1. (*Skewness and kurtosis of deterministic L matrices*) Suppose that $L_{m,n} = \left(\ell_1^T, ..., \ell_m^T\right)^T$, $n = 2,...,m-1$, is an arbitrary mxn orthonormal matrix with $\ell_i = (\ell_{i1}, ..., \ell_{in})$, $i = 1,...,m$. Then one has the formulas

$$\tau_M(L_{m,n}) = m \cdot \sum_{i=1}^{m} \sum_{j=1}^{m} < \ell_i, \ell_j >_n^3, \quad \kappa_M(L_{m,n}) = m \cdot \sum_{i=1}^{m} \|\ell_i\|_n^4. \qquad (2.9)$$

Proof. The sample mean vector of $L_{m,n}$ is $\mu_n = 0_n$, and its sample covariance matrix is $C_n = m^{-1} \cdot E_n$, with inverse $C_n^{-1} = m \cdot E_n$. Insert into (2.6) to get the expressions (2.9). ◊

For the Helmert-Ledermann matrix (2.4) the Mardia skewness and kurtosis are respectively given by (Proposition 2.1 in [3]):

$$\tau_M(L^*_{m,n}) = n \cdot [(m-3) + (m-n)^{-1}], \quad \kappa_M(L^*_{m,n}) = n \cdot [(m-2) + (m-n)^{-1}]. \qquad (2.10)$$

Unfortunately, these statistical measures are linked by the relationship $\kappa_M(L^*_{m,n}) - \tau_M(L^*_{m,n}) = n$, and it is therefore not possible to target both skewness and kurtosis using (2.5) in ROM simulation (cf. [3], Section 2.1). To analyze whether it is possible to get rid of this disadvantage by allowing a broader range of variation for skewness and kurtosis, the author [17] considers a larger set of deterministic *generalized Helmert-Ledermann* (GHL) orthogonal matrices $L_m \in O(m)$ that have fixed last column $\ell_m = \left(1/\sqrt{m}, ..., 1/\sqrt{m}\right)^T$ and satisfy the hyper-plane constraint $1_m^T \cdot L_{m,m-1} = 0_{m-1}^T$.

More generally, to study the maximum range of variation of skewness and kurtosis over the space of all deterministic *L* matrices, we are interested in the constrained optimization on the Stiefel manifold (introduced in [18]) of the objective functions (2.9) subject to the constraints $L_{m,n}^T \cdot L_{m,n} = E_n$, $1_m^T \cdot L_{m,n} = 0_n^T$. To tackle these problems several algorithms are available. Besides manifold versions of the Newton and the conjugate gradient method (e.g. [19], Sections 3.2 and 3.4, [20], Sections 6 and 8), there exist (curve)linear search algorithms (e.g. [20], Section 4). Within the last class of algorithms [21] have developed a feasible (constraints preserving) retraction method. We hope that these complex topics will be tackled in the future.

2.2 Random Orthogonal Matrices and Sample Characteristics in ROM Simulation

The focus is now on the random square orthogonal matrices, which pre- or post-multiply a given L matrix during a ROM simulation in (2.5). They fall into three categories: permutations, reflections and rotations. There exist many different ways to generate random orthogonal matrices. They include techniques based on Givens rotations matrices, skew-symmetric matrices and Cayley transforms, or the matrix exponential function (e.g. [3], Section 2.2, and [4]). According to [22] (see also [23,24]) it is well-known that every orthogonal matrix $R_n \in O(n)$ can be written as a product of $\frac{1}{2}n(n-1)$ rotation matrices and reflections:

$$R_n = (G_{12}G_{13}...G_{1n}) \cdot (G_{23}...G_{2n}) \cdot ... \cdot (G_{n-1n}) \cdot D_\varepsilon, \tag{2.11}$$

where the matrix $D_\varepsilon = diag(\varepsilon_1,...,\varepsilon_n),\ \varepsilon_i = \pm 1,\ i = 1,...,n$, represents reflections, and

$$G_{ij} = \begin{pmatrix} E & 0 & 0 & 0 & 0 \\ 0 & \cos\theta_{ij} & 0 & \sin\theta_{ij} & 0 \\ 0 & 0 & E & 0 & 0 \\ 0 & -\sin\theta_{ij} & 0 & \cos\theta_{ij} & 0 \\ 0 & 0 & 0 & 0 & E \end{pmatrix} \tag{2.12}$$

are Givens rotation matrices with angles of rotation $\theta_{ij} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The generation of such random orthogonal matrices is described in [22] using Bernoulli random variables for D_ε and a set of $\frac{1}{2}n(n-1)$ mutually independent beta random variables. We note that alternative algorithms for this have been developed in [24]. For simplicity, and like [3], we restrict the attention to random *upper Hessenberg orthogonal matrices* $H_n \in O(n)$, which can be written as a product of $n-1$ Givens rotation matrices of the form (e.g. [25], [26])

$$H_n = G_n(\theta_1) \cdot G_n(\theta_2) \cdot ... \cdot G_n(\theta_{n-1}), \tag{2.13}$$

where $G_n(\theta_i)$ is an nxn identity matrix except for the 2x2 principal sub-matrix with entries

$$G_n(\theta_i)[i, i+1; i, i+1] = \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix}, \tag{2.14}$$

where θ_i is chosen at random in the interval $[0, 2\pi),\ i = 1,...,n-1$.

An investigation of how random permutations, reflections and other random orthogonal matrices alter the sample characteristics of multivariate time series ROM simulations is found in [3], Section 2.2. A detailed study of these effects for different rotational matrices is found in [4].

2.3 Properties of Multivariate Moments under Sample Concatenation

First of all, one observes that the value of $n < m$, which achieves a desired skewness or kurtosis level along the line of Section 2.1, will be much smaller than the number of observations required for a standard simulation. To overcome this disadvantage one obviously repeats simulations of size m until enough observations have been generated, a technique called *sample concatenation*. Once the desired mean, covariance matrix and multivariate skewness or kurtosis have been matched for a given small $n < m$, what about the first four multivariate moments of the concatenated sample ? The result is summarized in Proposition 2.1. For the skewness this depends upon the notion of co-skewness.

Definition 2.1. Given two different samples $X_{m_X,n} = \left(x_1^T,...,x_{m_X}^T\right)^T$ and $Y_{m_Y,n} = \left(x_1^T,...,x_{m_Y}^T\right)^T$ on the same n random variables, the multivariate *co-skewness* is defined by

$$\tau_C(X_{m_X,n}, Y_{m_Y,n}) = (m_X + m_Y)^{-2} \cdot \sum_{i=1}^{m_X} \sum_{j=1}^{m_Y} \left\{2(x_i - \bar{x}) \cdot (S_X + S_Y)^{-1} \cdot (y_j - \bar{y})^T\right\}^3, \quad (2.15)$$

where $\bar{x}, \bar{y}$ and S_X, S_Y are the means and covariance matrices of $X_{m_X,n}, Y_{m_Y,n}$ respectively.

The co-skewness is invariant under non-singular affine transformations of the form

$$\tilde{X}_{m_X,n} = X_{m_X,n} B_n + 1_{m_X} \cdot b_n^T, \quad \tilde{Y}_{m_Y,n} = Y_{m_Y,n} B_n + 1_{m_X} \cdot b_n^T, \quad (2.16)$$

with any invertible matrix B_n and column vector b_n. That is, under (2.16), one has $\tau_C(\tilde{X}_{m_X,n}, \tilde{Y}_{m_Y,n}) = \tau_C(X_{m_X,n}, Y_{m_Y,n})$ (see [3], Appendix A.2).

Proposition 2.1. Consider r random samples $X_{m_1,n},..., X_{m_r,n}$, each with sample mean μ_n and sample covariance matrix C_n. Set $m = \sum_{k=1}^{r} m_k$ and define $X_{m,n} = \left(X_{m_1,n}^T,...,X_{m_r,n}^T\right)^T$. Then

$$m^{-1} \cdot 1_m^T \cdot X_{m,n} = \mu_n^T, \quad m^{-1} \cdot X_{m,n}^T \cdot X_{m,n} = C_n, \quad (2.17)$$

$$\begin{aligned} \tau_M(X_{m,n}) &= m^{-2} \cdot \sum_{k=1}^{r} m_k^2 \tau_M(X_{m_k,n}) + 2m^{-2} \cdot \sum_{k<\ell}^{r} (m_k + m_\ell)^2 \tau_C(X_{m_k,n}, Y_{m_\ell,n}), \\ \kappa_M(X_{m,n}) &= m^{-1} \cdot \sum_{k=1}^{r} m_k \kappa_M(X_{m_k,n}). \end{aligned} \tag{2.18}$$

Proof. See [3], Appendix A.3.

The equations (2.17) state that sample means and covariance matrices are preserved under sample concatenation of smaller samples with identical sample means and covariance. In particular, a large exact moment simulation, which targets a given covariance matrix C_n, can be constructed by concatenating many smaller exact moment simulations, each with the same target covariance matrix C_n.

However, the equations (2.18) show that skewness and kurtosis will be preserved under sample concatenation only under much more restrictive conditions. Firstly, the kurtosis will be preserved by concatenating smaller samples with equal size and identical kurtosis. This implies that the same *L* matrix must be used for each smaller simulation, so that

$$X_{rm,n} = 1_{rm} \cdot \mu_n^T + \sqrt{m} \cdot ((L_{m,n} \cdot R_n^{(1)})^T, \ldots, (L_{m,n} \cdot R_n^{(r)})^T)^T \cdot B_n. \tag{2.19}$$

The behaviour of skewness is more complex due to the co-skewness terms in (2.18). Since $\tau_C(X_{m,n}, X_{m,n}) = \tau_M(X_{m,n})/4$ one sees that $\tau_M(X_{rm,n}) = \tau_M(L_{m,n})$ provided

$$X_{rm,n} = 1_{rm} \cdot \mu_n^T + \sqrt{m} \cdot (L_{m,n}^T, \ldots, L_{m,n}^T)^T \cdot R_n \cdot B_n. \tag{2.20}$$

3 Application to Market Value-at-risk (Market VaR)

ROM simulation may be potentially applied to any problem that can be resolved with Monte Carlo simulation. In general, these kinds of problems require the forecast of future multivariate distributions using historical or scenario sample data. It is shown how ROM simulation applies to VaR estimation, which is a main industry benchmark to assess financial risk, and how to get some useful semi-parametric analytical sample approximations to VaR.

3.1 ROM VaR Methodology

Suppose a portfolio with n risk factors is given. Its target mean vector of returns is μ_n and its covariance matrix is C_n. A sample matrix $X_{m,n}$ with $m > n$ is generated using a ROM simulation of the form (2.5). It represents m observations for the returns $X_1, \ldots, X_n$ on the portfolio's n risk factors and generates m observations of the overall portfolio return

$S = \sum_{j=1}^{n} X_j$ through $S_i = \sum_{j=1}^{n} X_{i,j}, i = 1,...,m$. ROM VaR considers the mean, variance, skewness and kurtosis sample characteristics $(\hat{\mu}_S, \hat{\sigma}_S^2, \hat{\gamma}_S, \hat{\gamma}_{2,S})$ of the portfolio return calculated as

$$\begin{aligned}
&\hat{\mu}_S = \frac{1}{m}\sum_{i=1}^{m} S_i, \quad \hat{\sigma}_S^2 = \frac{1}{m-1}\sum_{i=1}^{m}(S_i - \hat{\mu}_S)^2, \\
&\hat{\gamma}_S = \frac{\hat{\kappa}_{3,S}}{\hat{\sigma}_S^3}, \quad \hat{\kappa}_{3,S} = \frac{m}{(m-1)(m-2)}\sum_{i=1}^{m}(S_i - \hat{\mu}_S)^3 \\
&\hat{\gamma}_{2,S} = \frac{\hat{\kappa}_{4,S}}{\hat{\sigma}_S^4}, \quad \hat{\kappa}_{4,S} = \frac{m(m+1)}{(m-1)(m-2)(m-3)}\sum_{i=1}^{m}(S_i - \hat{\mu}_S)^4 - 3\cdot\frac{1}{(m-2)(m-3)}\left(\sum_{i=1}^{m}(S_i - \hat{\mu}_S)^2\right)^2.
\end{aligned} \tag{3.1}$$

Under the usual MVN assumption the portfolio's ROM VaR is calculated as

$$VaR_\varepsilon[S] = -\hat{\mu}_S + \hat{\sigma}_s \cdot z_\varepsilon, \tag{3.2}$$

with $z_\varepsilon = \Phi^{-1}(1-\varepsilon)$ the ε-percentile of the standard normal and ε the probability of loss. Typically, the covariance matrix C_n has a big impact on (3.2) because $\hat{\sigma}_S^2 \approx 1_n^T \cdot C_n \cdot 1_n$. Since VaR is a lower quantile of the return distribution it is also much dependent upon the variation of skewness and kurtosis. To take this effect into account it is possible to specify either parametric analytical distributions of return (see [7], Section 2.1, for some useful choices) or adopt a semi-parametric approach that does not assume a specific distribution but accounts for non-trivial higher moments. Since parametric models suffer from *model risk* we consider only a semi-parametric approach based on the skewness and kurtosis, namely the Cornish-Fisher VaR approximation (Section 3.2) and an approximation based on the Chebyshev-Markov VaR upper bound (Section 3.3).

We note that the (non-parametric) historical VaR could also be used because it obviously takes into account the historical skewness and kurtosis However, as already mentioned in the introduction, the limitation to historical observations implies that history will repeat itself. With ROM VaR it is possible to simulate a very large number of realized risk factor returns that are all consistent with the observed historical sample moments. Moreover, we can stress test risk factors under adverse market conditions by targeting other sample moments that are consistent with periods of financial crisis.

For the above reasons we choose a ROM VaR simulation, where the deterministic *L* matrix is chosen to reflect the risk manager's point of view on Mardia kurtosis while still preserving its historical observed value as well as other properties of the historical data. As seen in Section 2.2, equation (2.19), to control the Mardia kurtosis with a deterministic *L* matrix, say the Helmert-Ledermann matrix (2.4), we choose a concatenated ROM sample of the form

$$X_{rp,n} = 1_{rp} \cdot \mu_n^T + \sqrt{p} \cdot ((L_{p,n}^* \cdot R_n^{(1)})^T, ..., (L_{p,n}^* \cdot R_n^{(r)})^T)^T \cdot B_n, \tag{3.3}$$

where $R_n^{(1)},...,R_n^{(r)}$ are different random orthogonal matrices. For this, we know by (2.10) that $\kappa_M(X_{rp,n}) = \kappa_M(L^*_{p,n}) \approx n \cdot (p-2)$. Therefore, to target a kurtosis of $(1+\beta) \cdot n \cdot (n+2), \beta \geq 0$, where $\beta = 0$ corresponds to the Mardia kurtosis of a MVN model (e.g. [3], Proposition 2.3)), the parameter p is set equal to the nearest integer matching the equation

$$p \approx 2 + (1+\beta) \cdot (n+2). \tag{3.4}$$

3.2 Cornish-fisher VaR Approximation

This semi-parametric approach makes use of the Cornish-Fischer [27] expansion. For a random variable S with mean μ_S, variance σ_S^2, skewness γ_S and kurtosis $\gamma_{2,S}$, which represents here a profit, one has the following *Cornish-Fischer* VaR *approximation* (CF VaR) (e.g. [28] or [1], IV.5.3.3, for the context of delta-gamma-normal approximation):

$$\begin{aligned} &VaR_\varepsilon^{CF}[S] = -\mu_S + \sigma_S \cdot z_\varepsilon^{CF}, \\ &z_\varepsilon^{CF} = z_\varepsilon + \frac{1}{6}\left(z_\varepsilon^2 - 1\right) \cdot \gamma_S + \frac{1}{24}\left(z_\varepsilon^3 - 3z_\varepsilon\right) \cdot \gamma_{2,S} - \frac{1}{36}\left(2z_\varepsilon^3 - 5z_\varepsilon\right) \cdot \gamma_S^2, \end{aligned} \tag{3.5}$$

with $z_\varepsilon = \Phi^{-1}(1-\varepsilon)$ the ε-percentile of the standard normal distribution. The Cornish-Fisher approximation consists to transform the quantile of a normal law into the realization of a random variable S with non-vanishing skewness and kurtosis such that $F_S(S) = \Phi(z_\varepsilon)$. To be well-defined such a transformation must be one-to-one. A necessary and sufficient condition for this is the non-vanishing of the derivative dS/dz_ε, which holds provided the following inequality is satisfied:

$$4 \cdot \left(\tfrac{\gamma_{2,S}}{8} - \tfrac{\gamma_S^2}{6}\right) \cdot \left(1 - \tfrac{\gamma_{2,S}}{8} + \tfrac{5 \cdot \gamma_S^2}{36}\right) - \tfrac{\gamma_S^2}{9} \geq 0. \tag{3.6}$$

In practice γ_S and $\gamma_{2,S}$ are small and $\gamma_{2,S}$ is positive, hence the condition is often fulfilled.

3.3 Chebyshev-Markov VaR Upper Bound and Robust Approximation

Since Chebyshev and Markov it is possible to construct universal semi-parametric bounds for the evaluation of VaR (and expected shortfall) based on the first few moments of higher order. Rather simple and practical analytical bounds, which are based on the mean, variance, skewness and kurtosis of the portfolio loss distribution, have been derived in [29], Theorems 4.1, 4.2 and Corollary 4.1 (see also [30,31], Section 2.4). In general, one has the following implicitly defined *Chebyshev-Markov* VaR (CM VaR) *upper bound*:

$$VaR_{\varepsilon}^{CM}[S] = -\mu_S + \sigma_S \cdot z_{\varepsilon}^{CM}, \quad p\left(z_{\varepsilon}^{CM}\right) = \varepsilon, \quad \varepsilon \le \frac{1}{2}\left(1 - \frac{\gamma_S}{\sqrt{4+\gamma_S^2}}\right),$$
$$p(u) = \frac{\Delta_S}{q(u)^2 + \Delta_S(1+u^2)}, \quad q(u) = 1 + \gamma_S u - u^2, \quad \Delta_S = 2 + \gamma_{2,S} - \gamma_S^2. \tag{3.7}$$

Since this upper bound takes into account the extreme effects that skewness and kurtosis may have on VaR, it is clearly only a crude upper bound. In the limiting case of a normal distribution with $\gamma_S = \gamma_{2,S} = 0$ one gets as special case $z_{\varepsilon}^{CM} = \left(\frac{2-3\varepsilon}{\varepsilon}\right)^{\frac{1}{4}}$, which considerably overestimates the true normal VaR coefficient $z_{\varepsilon} = \Phi^{-1}(1-\varepsilon)$. It is possible to transform this upper bound into a more robust formula. Roughly speaking a model is *robust* when a small change in the assumptions does not produce big changes in the results. A simple device to adjust the upper bound (in order to get a robust version of it in the normal case) is through a multiplication factor, chosen here as the ratio of these two coefficients. We get the *robust Chebyshev-Markov* VaR *approximation* (robust CM VaR):

$$VaR_{\varepsilon}^{RCM}[S] = -\mu_S + \sigma_S \cdot \frac{z_{\varepsilon}}{\left(\frac{2-3\varepsilon}{\varepsilon}\right)^{\frac{1}{4}}} \cdot z_{\varepsilon}^{CM}. \tag{3.8}$$

In the special case of a symmetric distribution with vanishing skewness $\gamma_S = 0$, one has the explicit *symmetric Chebyshev-Markov* VaR *upper bound* (sym CM VaR) ([29], Example 4.1):

$$VaR_{\varepsilon}^{SCM}[S] = -\mu_S + \sigma_S \cdot z_{\varepsilon}^{SCM}, \; z_{\varepsilon}^{SCM} = \frac{\sqrt{2}}{2} \cdot \left[\sqrt{\gamma_{2,S}^2 + \frac{4(1-\varepsilon)}{\varepsilon} \cdot (\gamma_{2,S} + 3) - \frac{4}{\varepsilon}} - \gamma_{2,S}\right]^{\frac{1}{2}}. \tag{3.9}$$

Similarly to the preceding situation, one gets the *robust symmetric Chebyshev-Markov* VaR *approximation* (robust sym CM VaR):

$$VaR_{\varepsilon}^{RSCM}[S] = -\mu_S + \sigma_S \cdot \frac{z_{\varepsilon}}{\left(\frac{2-3\varepsilon}{\varepsilon}\right)^{\frac{1}{4}}} \cdot z_{\varepsilon}^{SCM}. \tag{3.10}$$

4 A Numerical Case Study

The present Section illustrates the ROM VaR simulation method. Given the mean vector and the covariance matrix of the risk factor returns, the Mardia kurtosis of portfolio returns is targeted as explained in Section 3.1. Based on the simulated mean, volatility, skewness and kurtosis of the portfolio return, VaR is estimated using the Cornish-Fisher VaR approximation, the Chebyshev-Markov VaR upper bound and its two robust approximations.

For a number of risk factors $n \in \{3,4,5,6,10\}$ we generate $m = r \cdot p = 10'000$ concatenated ROM samples of the type (3.3) with $p \in \{8,10,16,20\}$, $p > n$, and random upper Hessenberg orthogonal matrices $R_n^{(1)},...,R_n^{(r)}$ as specified in (2.13)-(2.14). For simplicity, we assume a zero-mean vector of returns $\mu_n = 0_n$. This assumption is made in the Risk Metrics VaR methodology (e.g. [1], IV.2.10.3)), which should be compared in practice with the present approach. For illustration only, we use the following full-rank column-homogeneous angles parameterization of the correlation matrix $\rho_n = (\rho_{ij}^{(n)})$ by [32]:

$$\rho_{ij}^{(n)} = \sqrt{(1-\alpha_i^2)(1-\alpha_j^2)}\frac{1-(\alpha_i\alpha_j)^{n-1}}{1-\alpha_i\alpha_j} + (\alpha_i\alpha_j)^{n-1}, \quad \alpha_k \in (-1,1), \quad 1 \le i, j \le n. \quad (4.1)$$

In this situation, the spectral decomposition $\rho_n = b_n^T \cdot b_n$ with $b_n = (b_{ji}^{(n)})$ reads

$$b_{ji}^{(n)} = (1-\alpha_i^2)^{\frac{1\{j<d\}}{2}} \alpha_i^{j-1}. \quad (4.2)$$

Further, we assume a constant vector $\sigma_n = (0.1,...,0.1)^T$ of standard deviations of returns, so that the spectral decomposition of the associated covariance matrix $C_n = \sigma_n^T \cdot \rho_n \cdot \sigma_n = B_n^T \cdot B_n$, $B_n = (B_{ji}^{(n)})$, is given by $B_{ji}^{(n)} = b_{ji}^{(n)}\sigma_i$. The vector $\alpha^{(n)} = (\alpha_1,...,\alpha_n)$ of constants in (4.1)-(4.2) is taken from the following specification:

$$\alpha^{(10)} = (\alpha_1,...,\alpha_{10}) = (-0.5,0,0.5,0,-0.5,0,0.5,0,-0.5,0). \quad (4.3)$$

Once a ROM simulated sample of returns $X_{m,n}$ has been generated, the overall sample portfolio returns $S_i = \sum_{j=1}^{n} X_{i,j}$, $i = 1,...,m$, as well as the sample characteristics (3.1), are calculated. Table 4.1 summarizes the results and displays corresponding VaR approximations for the small probability of loss $\varepsilon = 0.005$ (Basel III and Solvency II compatible).

The following observations can be made. While the normal linear VaR (3.2) remains for fixed n (up to MC errors) rather stable across choices of $p > n$, the CF VaR (3.5) and the CM VaR (3.7)-(3.8) vary much with p. Since the kurtosis is increasing with p one expects that VaR is also increasing with p in agreement with increased risk by increased kurtosis. This is true for the CM VaR approximations but not for the CF VaR approximation (a counterexample is the pair (6, 8)-(6, 10)). Since for small kurtosis the CF VaR can be below the normal VaR, one also observes that the CF VaR might discriminate too much with respect to normal VaR. Since the CM VaR approximations are above or at least close to the normal VaR for the robust CM VaR, these approximations do not share this disadvantage. On the other hand, the CM VaR upper bound (3.7) might be too conservative to be useful in practice. The robust CM VaR lies often on the safe side with respect to the CF VaR and appears to be a reasonable compromise. Whether the above observations remain true in more general settings is open for further investigation.

<u>Example 4.1</u>: VaR approximations using sample concatenated ROM simulation, $\varepsilon = 0.005$

(n,p)	sample characteristics of portfolio return $\boxed{\hat{\mu}_S}$	$\boxed{\hat{\sigma}_s}$	$\boxed{\hat{\gamma}_S}$	$\boxed{\hat{\gamma}_{2,S}}$	VaR sample approximations Normal appr. (3.2)	CF VaR appr. (3.5)	CM VaR appr. (3.7)	robust CM VaR (3.8)
(3,8)	-0.00090	0.27943	-0.72004	1.10760	0.72067	0.56673	1.23146	0.71101
(3,10)	-0.00197	0.28019	-0.92752	2.50553	0.72369	0.61086	1.31745	0.76108
(3,16)	-0.00068	0.28020	-1.41438	6.87933	0.72243	0.77055	1.47935	0.85396
(3,20)	-0.00075	0.27377	-1.81896	10.15159	0.70594	0.78652	1.49431	0.86262
(4,8)	0.00249	0.37493	-0.71406	1.14419	0.96327	0.76608	1.65659	0.95489
(4,10)	-0.00043	0.37501	-0.95445	2.60568	0.96638	0.80929	1.76262	1.01731
(4,16)	0.00293	0.37249	-1.42248	7.10152	0.95655	1.04491	1.97560	1.13879
(4,20)	-0.00040	0.37571	-1.75623	10.05203	0.96818	1.13616	2.06661	1.19272
(5,8)	0.00222	0.46503	-0.50198	0.79422	1.19561	1.05113	2.08150	1.20021
(5,10)	0.00169	0.46496	-0.80554	2.32418	1.19598	1.08728	2.21010	1.27464
(5,16)	0.00119	0.47206	-1.23465	6.60437	1.21475	1.45785	2.54250	1.46667
(5,20)	-0.00216	0.46816	-1.53516	9.49184	1.20806	1.61382	2.62973	1.51842
(6,8)	0.00060	0.56451	-0.34874	0.63112	1.45349	1.36697	2.56020	1.47713
(6,10)	0.00012	0.55771	-0.73850	2.19073	1.43644	1.34630	2.66379	1.53711
(6,16)	0.00037	0.55892	-1.27634	6.66214	1.43931	1.68324	2.99690	1.72923
(6,20)	0.00217	0.55993	-1.54007	9.81751	1.44011	1.98895	3.16545	1.82573
(10,16)	0.00194	0.94655	-1.11161	5.85605	2.43622	2.91843	5.06326	2.92097
(10,20)	0.00942	0.92664	-1.45863	9.54017	2.37744	3.39026	5.25837	3.03040

As an important remark, one can state that the considered combined ROM VaR *semi-parametric approach* clearly demonstrates the important variability of VaR due to higher than normal kurtosis in portfolio returns. Unlike parametric models like Student t, normal mixture and Student t mixture models, this approach does not suffer from model risk and is an appropriate complementary substitute of the normal linear VaR model, which is not able to capture non-normality effects. Moreover, the simplicity and reasonable performance of concatenated ROM simulation competes well with a general MC VaR method that includes non-linear (via delta-gamma approach) and non-normality effects.

In practice, one should compare the above proposals with the Risk Metrics VaR methodology (as already stated above) and other methods like the Gram-Charlier approximation and the delta-gamma approach. An example for the latter is found in [1], IV.5.3.3, who comments also on comparing it with the Cornish-Fisher VaR. In particular, one should take care of highly leptokurtic data by using appropriate parametric statistical distributions that enable the use of skewness and kurtosis. A very tractable possibility is Johnson's distribution (e.g. [1], Example IV.5.2, [33],

[34]). Similarly popular are the normal inverse Gaussian and the skew Student t that are applied in [7] among many other papers. Further valuable alternatives include the variance-gamma, the normal variance-gamma, the truncated Lévy flight and the normal tempered stable distribution (see [35,36,37]). The approximation by Silitto [38] has been recommended by [39]. Unfortunately, detailed and comprehensive calculations are not within the scope of the present paper. One cannot conclude without mentioning the relevance of the present combined ROM VaR semi-parametric approach to other risk management topics like credit risk and liquidity risk.

5 Conclusion

In the present paper, the relatively recent and fundamental ROM sampling algorithm has been integrated into two semi-parametric methods that take into account skewness and kurtosis in order to compute VAR, namely the Cornish-Fisher VaR (CF VaR) approximation and a robust approximation to the Chebyshev-Markov VaR upper bound (robust CM VaR). The important variability of VaR due to higher than normal kurtosis in portfolio returns is demonstrated by the combined ROM VaR semi-parametric approach at a numerical case study. In this study, the robust CM VaR variant is often on the safe side with respect to the CF VaR and appears to be a reasonable compromise.

Competing Interests

Author has declared that no competing interests exist.

References

[1] Alexander C. Market Risk Analysis. Volume IV: Value-at-Risk Models. Chichester, West Sussex. J Wiley & Sons Ltd.; 2008.

[2] Glasserman P. Monte Carlo Methods in Financial Engineering. Springer, NY; 2004.

[3] Ledermann W, Alexander C, Ledermann D. Random orthogonal matrix simulation. Linear Algebra and its Applications. 2011;434:1444-67.

[4] Ledermann D, Alexander C. ROM simulation with rotation matrices. ICMA Center Discussion Papers in Finance. 2011;DP2011-06.

[5] Ledermann D, Alexander C. Further properties of random orthogonal matrix simulation. Mathematics and Computers in Simulation. 2012;83:56-79.

[6] Alexander C, Ledermann D. ROM simulation. Applications to stress testing and VaR; 2012. Accessed 28 April 2014. Available: http://ssrn.com/abstract=2049239

[7] Hürlimann W. Robust variants of the Cornish-Fischer approximation and Chebyshev-Markov bounds: Application to value-at-risk and expected shortfall. Advances and Applications in Mathematical Sciences. 2009;1(2):239-60.

[8] Rebonato R, Jäckel P. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. Journal of Risk. 2000;2(2):17-27.

[9] Rebonato R. On the simultaneous calibration of multifactor lognormal interest rate models to Black volatilities and to the correlation matrix. J Comp Finance. 1999;2:5-27.

[10] Helmert FR. Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers direkter Beobachtungen gleicher Genauigkeit. Astronom Nachr. 1876;88:115-132.

[11] Ogawa J. On the sampling distributions of classical statistics in multivariate analysis. Osaka Mathematical Journal. 1953;5(1):13-52.

[12] Lancaster HO. The Helmert marices. Amer Math Monthly. 1965;72(1):4-12.

[13] Tsagris MT, Preston S, Wood ATA. A data-based power transformation for compositional data. Preprint; 2011. Accessed 28 April 2014.
Available: http://arxiv.org/pdf/1106.1451.pdf

[14] Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika. 1970;57(3):519.

[15] Srivastava MS. A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. Statistics & Probability Letters. 1984;2:263-67.

[16] Kollo T. Multivariate skewness and kurtosis measures with an application in ICA. Journal of Multivariate Analysis. 2008;99:2328-38.

[17] Hürlimann W. Generalized Helmert-Ledermann orthogonal matrices and ROM simulation. Linear Algebra and its Applications. 2013;439(7):1716-29.

[18] Stiefel E. Richtungsfelder und Fernparallelismus in Mannigfaltigkeiten. Comment Math Helv. 1936;8:3-51.

[19] Edelman, A., Arias, T.A., Smith, S.T. The geometry of algorithms with orthogonality constraints. SIAM J. Matrix Anal. Appl. 1998; 20(2): 303-53.

[20] Absil PA, Mahony R, Sepulchre R. Optimization Algorithms on Matrix Manifolds. Princeton and Oxford: Princeton University Press; 2008.

[21] Wen Z, Yin W. A feasible method for optimization with orthogonality constraints. Mathematical Programming. 2013;142(1-2):397-434.

[22] Anderson TW, Olkin I, Underhill LG. Generation of random orthogonal matrices. SIAM J Sci Stat Comput. 1987;8:625-629.

[23] Golub G, Van Loan C. Matrix Computations. Baltimore: Johns Hopkins University Press; 1996.

[24] Genz A. Methods for generating random orthogonal matrices. Proceedings of the MCQMC98 Meeting; 1998.

[25] Berry M, Dongarra J, Kim Y. A highly parallel algorithm for the reduction of a non-symmetric matrix to block upper Hessenberg form. Parallel Computing. 1995;21:1189.

[26] Gemignani L. A unitary Hessenberg QR-based algorithm via semiseparable matrices. J Computational Appl Math. 2005;184:505-517.

[27] Cornish EA, Fisher RA. Moments and Cumulants in the Specification of Distributions. Review of the International Statistical Institute. 1937;5:307-20.

[28] Jaschke SR. The Cornish-Fisher-Expansion in the Context of Delta-Gamma-Normal Approximations. Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät 2001, Discussion Paper 2001-54.

[29] Hürlimann W. Analytical bounds for two value-at-risk functionals. ASTIN Bulletin. 2002;32:235-65.

[30] Hürlimann W. Extremal Moment Methods and Stochastic Orders – Application in Actuarial Science. Bulletin of the Venezuelan Mathematical Association (BAMV) 2008;15 :1-2.

[31] Nadarajah S, Zhang B, Chan S. Estimation methods for expected shortfall. Quantitative Finance. 2014;14(2):271-91.

[32] Brigo D. A note on correlation and rank reduction; 2002. Accessed 28 April 2014. Available: http://www.damianobrigo.it/correl.pdf

[33] Simonato JG. The performance of Johnson's distributions for computing value-at-risk and expected shortfall. The Journal of Derivatives. 2011;19(1):7-24.

[34] Lien D, Stroud Ch, Ye K. Comparing VaR approximation methods which use the first four moments as inputs. Working Paper, WP # 0050MSS-202-2013, University of Texas at San Antonio; 2013.

[35] Hürlimann W. Portfolio ranking efficiency (I) Normal variance gamma returns. Int J Math Archive. 2013;4(5):192-218.

[36] Hürlimann W. Portfolio ranking efficiency (II) Truncated Lévy flight returns. Int J Math Archive. 2013;4(8):39-48.

[37] Hürlimann W. Portfolio ranking efficiency (III) Normal tempered stable returns. Int J Math Archive. 2013;4(10):20-29.

[38] Sillitto GP. Derivation of approximants to the inverse distribution function of a continuous univariate population from the order statistics of a sample. Biometrika. 1969;56:641-650.

[39] Lien D, Yang X, Ye K. Alternative approximations to value-at-risk: A comparison. Working Paper, WP # 0037ECO-202-2012. University of Texas at San Antonio; 2012.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
www.sciencedomain.org/review-history.php?iid=509&id=6&aid=4456