Mathematical Statistics

Stockholm University

Reverse Stress Testing in Skew-Elliptical Models

Jonathan von Schroeder

Taras Bodnar

Thorsten Dickhaus

Research Report 2019:04

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2019:04**,
http://www.math.su.se

Reverse Stress Testing in Skew-Elliptical Models

April 2019

Jonathan von Schroeder[a], Taras Bodnar[b] and Thorsten Dickhaus[a,1]

[a] *Institute for Statistics, University of Bremen, Bibliothekstraße 1, D-28359 Bremen, Germany*

[b] *Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden*

Abstract

We extend previous results on reverse stress testing under elliptical models to the broader class of skew-elliptical models. In particular, under the assumption of a linear Profit and Loss function, we are concerned with finding the most likely scenarios given that the loss exceeds a given threshold. In the elliptical case, an explicit formula for the solution is provided. In the skew-elliptical case, we characterize the solution in terms of an easy-to-implement numerical optimization problem. As a specific example, we investigate the class of skew-normal models in some detail.

Keywords: Bank regulation, constrained optimization, density generator, Profit and Loss function, risk management

[1]Corresponding author. E-mail address: dickhaus@uni-bremen.de, Phone: +49-421-218-63651

1 Introduction

Recently the Basel Committee on Banking Supervision has recognized the importance of including extreme scenarios as part of an overall stress testing programme:

> A stress testing programme should also determine what scenarios could challenge the viability of the bank (for example by reverse stress testing) and thereby uncover hidden risks and interactions among risks. (cf. [1, p. 18 ff.])

Thus one of the methods relevant to stress testing is the scenario analysis. The utilized "[s]cenarios usually involve some kind of coherent, logical narrative or 'story' as to why certain events and circumstances can occur and in which combination and order" [2, p. 9]. Stress scenarios should reflect an "organizations unique vulnerabilities to factors that affect its exposures, activities, and risks" [2, p. 9]. This however means that an analyst has to select plausible scenarios based on knowledge about the organization which will invariably introduce a bias. Furthermore plausibility is a constraint that, based on historical data, can exclude certain "break the bank" scenarios that can occur during extreme events like the financial crisis of 2008. One of the main goals of reverse stress testing is to overcome this limitation by "assum[ing] a known adverse outcome ... and then deduc[ing] the types of events that could lead to such an outcome" [2, p. 9]. To that end the question of what happens to a given collection of assets (the portfolio under investigation), if a market factor changes in a certain way, is reversed by asking instead what could cause a certain portfolio event (e.g. a loss exceeding a certain threshold, cf. [3]). Reverse stress testing is usually based on the density of the Profit and Loss (P&L) distribution, which has to be estimated for real-world portfolios. Addressing the arising confounding issue requires "that the distribution that serves as the foundation for reverse stress tests is consistent with stylized facts of actual tail behavior" [3]. As a matter of fact, many loss distributions are highly skewed (cf. [4, p. 44]) and thus normality can certainly not be assumed. The high skewness furthermore means, that even the family of elliptical distributions is not rich enough. Therefore the normality assumption of [3] as well as the more general assumption of ellipticity of [5] are not always practical. Therefore we explore some aspects of the representation of the solution given in [5, Proposition 1] and give conditions for global optimality. We then extend our results to the more general class of skew-elliptical distributions.

The rest of the material is structured as follows. In Section 2, we introduce basic notation and assumptions. Section 3 contains our main results. The specific case of skew-normal models is investigated in Section 4, and we conclude with a discussion in Section 5.

2 Preliminaries

2.1 The Profit and Loss function

To formalize the notion of the "distribution of the return" consider the case of a portfolio whose value linearly depends on the change of value in the individual assets (cf. [4, p. 4]):

Definition 1 (P&L function)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *with* $\mathbb{P} : \mathcal{A} \to [0, 1]$ *be a probability space where* $\mathcal{A}$ *is a* $\sigma-$*algebra over* Ω *and* $\mathbb{P}$ *is a probability measure. Consider an* $\mathcal{A}$*-measurable random variable* $X : \Omega \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ *defined on* $(\Omega, \mathcal{A}, \mathbb{P})$*. If*

- X_i *denotes the change in value of i-th asset of the portfolio over a given holding period*
- *the change in the value of the portfolio is linear in* X *and the dependency is given by the (non-stochastic) portfolio weights* $c \in \mathbb{R}^n$, $c \neq 0$,

then the P&L function is given by

$$v(\omega) = c^T X(\omega) = \sum_{i=1}^{n} c_i X(\omega)_i, \tag{1}$$

and it is a random variable $v : \Omega \to \mathbb{R}$ *on* $(\Omega, \mathcal{A}, \mathbb{P})$.

In the following it will be assumed that the P&L function is of the form (1). Obviously a more general, non-linear dependence could be considered. However for analytical and / or numerical tractability a restriction to the linear or quadratic case is often necessary (see e.g. [6]).

Remark 1. *Definition 1 is closely related to [6, Definition 1.3]: Note that the* ω_i *in [6, Definition 1.3] are related to the* X_i *in Definition 1 by* $\omega_i = X_i - \mathbb{E}[X_i]$*. Consider the special case where u in [6, Definition 1.3] is a linear and homogenous function (i.e.* $u(\omega) = c^T\omega$ *for some* $c \in \mathbb{R}^n$*):*

$$u\left(\overline{\psi} + \omega\right) - u\left(\overline{\psi}\right) = u(\omega) = c^T X - c^T \mathbb{E}[X]$$

The remaining difference stems mostly from the fact, that [6] considers "market rates" whereas Definition 1 considers the change in value of the assets. Which treatment is more appropriate depends on the use case. For the mathematical treatment one can simply replace X *by* $X - \mathbb{E}[X]$ *and vice versa.*

2.2 The family of (skew-)elliptical distributions

This section aims to give a short introduction to the family of skew-elliptical distributions as well as relevant special properties. This exhibition closely follows [7, Chapter 6].

Definition 2 (Elliptical Distribution). *Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ where $\mathbb{R}_+ := \{r \in \mathbb{R} \,|\, r \geq 0\}$ and let*

$$k_n := \int_0^\infty r^{n-1} g\left(r^2\right) \mathrm{d}r$$

If $k_n < \infty$ then g is called a density generator and

$$f\left(x; \mu, \Sigma\right) := \frac{\Gamma\left(\frac{n}{2}\right)}{2k_n\sqrt{\pi^n \det\left(\Sigma\right)}} g\left((x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

is the density of an elliptical distribution $EC_n\left(\mu, \Sigma, g\right)$ where $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n\times n}$ is a positive definite symmetric matrix.

It is interesting to note that Σ is (in general) not the covariance matrix. It is however always a scalar multiple of it:

Corollary 1 (to [8, Theorem 2.17]). *If X is elliptically distributed according to the preceding definition and possesses second moments, then the covariance matrix of X is a real multiple of Σ, i.e. there exists $\kappa > 0$ such that $\mathrm{Cov}\left(X\right) = \kappa\Sigma$ (cf. [8, Theorem 2.17]).*

The generators for common elliptical probability distributions are (see e.g. [9, p. 62 ff.]):

Normal distribution $g(u) := \exp\left(-\frac{u}{2}\right)$

Student-t distribution $g(u) := \left(1 + \frac{u}{c_p}\right)^{-p}$ for $c_p \in \mathbb{R}$, $p > \frac{1}{2}$

Logistic Distribution $g(u) := \frac{e^{-u}}{(1+e^{-u})^2}$

Exponential distribution $g(u) := \exp\left(-ru^s\right)$ for $r, s > 0$

Since it is of great importance for the following section we note that the following holds:

Lemma 1. *For the Normal, Student-t, Logistic and Exponential distribution the generator is strictly decreasing on $\mathbb{R}_+$.*

The family of elliptical distributions can be extended to allow for skewness by multiplying elliptical densities with a scaling function (cf. [7, Equation (6.11)]):

Definition 3 (Skew-elliptical distribution). *Let $\tilde{f}$ be the density of an elliptical distribution (cf. Definition 2) and let F be the cumulative distribution function of a univariate elliptical distribution. Then*

$$f(x; \mu, \Sigma, \lambda) := 2 \cdot \tilde{f}(x; \mu, \Sigma) \cdot F\left(\lambda^T(x-\mu)\right)$$

is the density of a skew-elliptical distribution.

3 Reverse Stress Testing in (Skew-)Elliptical Models

Assume that the density of the n-dimensional random vector X is given by f, which is the density of an elliptical distribution (see Definition 2), and that X is supported on whole $\mathbb{R}^n$. Consider the P&L function (see Definition 1) given by $c^T X$, where large positive values represent large losses. For a given loss threshold the goal of a reverse stress test is then, following [5], to find the most likely loss scenarios $x^*(\ell)$ given that the loss exceeds ℓ[2]

$$x^*(\ell) = \arg\max_{x \in \mathbb{R}^n} f\left(x \,\middle|\, c^T X \geq \ell\right)$$

The first result derived by [5] is that there is a connection to the conditional expectation (which can be estimated using the empirical likelihood method):

Theorem 1 ([5, Proposition 1])**.** *Assuming appropriate tail behaviour there exists a sequence* $\{\kappa_\ell\}_\ell$ *(which depends on the tail behaviour) such that* $\kappa_\ell \to \kappa \in \mathbb{R}$ *as* $\ell \to \infty$ *and*

$$x^*(\ell) = \kappa_\ell \cdot \mathbb{E}\left[X | c^T X \geq \ell\right]. \tag{2}$$

Since the conditional expectation depends on the tail behaviour and so does κ_ℓ, the solution given by (2) could depend on the tail behaviour of the distribution of X. It is therefore interesting to derive an explicit equation for $x^*(\ell)$ to gain more insight into its properties. To this end note that for any (Lebesgue-)measurable A it holds that

$$\begin{aligned}\int_A f\left(x \,\middle|\, c^T X \geq \ell\right) \mathrm{d}x &= \mathbb{P}\left(X \in A \,\middle|\, c^T X \geq \ell\right) \\ &= \frac{\mathbb{P}\left(X \in A \cap c^T X \geq \ell\right)}{\mathbb{P}\left(c^T X \geq \ell\right)} \\ &= \frac{1}{\mathbb{P}\left(c^T X \geq \ell\right)} \cdot \int_{A \cap B} f(x)\, \mathrm{d}x\end{aligned}$$

where $B := \left\{x \in \mathbb{R}^n \,\middle|\, c^T x \geq \ell\right\}$. Hence, it follows that

$$f\left(x \,\middle|\, c^T X \geq \ell\right) = \frac{1}{\mathbb{P}\left(c^T X \geq \ell\right)} \begin{cases} 0 & \text{if } c^T x < \ell \\ f(x) & \text{otherwise} \end{cases}$$

i.e. the conditional density is zero on the complement of B and proportional to f on B. Thus finding the maximizer of the conditional density is equivalent to finding a maximizer of the unconditional density over B. The following theorem shows that the reverse stress testing problem has a global optimum and gives an explicit expression in terms of the parameters of the density:

[2]Note that in the notation of [5] $X_n = L$ is the loss and therefore $c_i = \delta_{i,n} = \mathbf{1}\{i = n\}$.

Theorem 2. *A global optimum of the constrained optimization problem*

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & f(x) \\ \text{subject to} \quad & c^T x \geq \ell, \end{aligned} \tag{3}$$

where $f(x) = g\left((x-\mu)^T\Sigma^{-1}(x-\mu)\right)$ *is an elliptical density with a decreasing function* g *is given by*

$$x^*(\ell) := \begin{cases} \mu & \text{if } \ell \leq c^T\mu, \\ \mu + (\ell - c^T\mu)\frac{\Sigma c}{c^T\Sigma c} & \text{otherwise.} \end{cases} \tag{4}$$

Proof of Theorem 2. Define $a = x - \mu$. Then (3) can be written by

$$\begin{aligned} \arg\max_{a} \quad & g\left(a^T\Sigma^{-1}a\right) \\ \text{subject to} \quad & c^T a \geq \ell - c^T\mu \end{aligned} \tag{5}$$

Let

$$\Omega_v = \{a : \quad c^T a = v\} \quad \text{for} \quad v \in [\ell - c^T\mu, \infty).$$

Since g is a decreasing function, the solution of (5) is obtained by the following optimization problem:

$$\min_{v \geq \ell - c^T\mu} \min_{a \in \Omega_v} a^T\Sigma^{-1}a$$

If $\ell \leq c^T\mu$, then $v = 0$ is admissible which implies $0 = a \in \Omega_0$ is admissible and therefore the minimum of $a^T\Sigma^{-1}a$ is attained at $a = 0$ independently of v and consequently $x^*(\ell) = \mu$ is the only solution of (3). Otherwise, using that finding the minimum of $a^T\Sigma^{-1}a$ over Ω_v is a quadratic optimization problem under a single linear equality constraint we straightforwardly get its solution as

$$a_v = v\frac{\Sigma c}{c^T\Sigma c}$$

or

$$x^*(\ell) = \mu + v\frac{\Sigma c}{c^T\Sigma c} \tag{6}$$

and

$$\min_{a \in \Omega_v} a^T\Sigma^{-1}a = \frac{v^2}{c^T\Sigma c}$$

Hence,

$$\min_{v \geq \ell - c^T\mu} \min_{a \in \Omega_v} a^T\Sigma^{-1}a = \min_{v \geq \ell - c^T\mu} \frac{v^2}{c^T\Sigma c} = \frac{(\ell - c^T\mu)^2}{c^T\Sigma c}$$

for $v = \ell - c^T\mu$ which together with (6) leads to the statement of the theorem. □

Equation (4) shows that, for fixed ℓ, the solution does not depend on (the tail behaviour of the generator) g, but only on the parameters of the (elliptical) distribution of X. In practically relevant applications, the latter parameters will typically not be known. Therefore, the following remark is insightful.

Remark 2. *The solution*

$$x^*(\ell, \Sigma) := \mu + \left(\frac{\ell - c^T\mu}{c^T\Sigma c}\right)\Sigma c$$

does not depend on the scaling of Σ, *i.e. for* $\alpha \in \mathbb{R}_+, \alpha \neq 0$ *it holds that*

$$\begin{aligned} x^*(\ell, \alpha\Sigma) &= \mu + \left(\frac{\ell - c^T\mu}{c^T\alpha\Sigma c}\right)\alpha\Sigma c \\ &= \mu + \left(\frac{\ell - c^T\mu}{c^T\Sigma c}\right)\Sigma c = x^*(\ell, \Sigma). \end{aligned}$$

Therefore, assuming existence of second moments of X, *the dispersion matrix* Σ *can be replaced by the covariance matrix* $\mathrm{Cov}\,(X)$. *The latter matrix can be estimated from data by standard techniques.*

3.1 Reverse Stress Testing in Skew-Elliptical Models

Next, we derive the solution of (3) in the case of a skew-elliptical model. We start with two lemmas needed to prove the main result of this section.

Lemma 2. *Let* f *be the density of a skew-elliptical distribution (in* $\mathbb{R}^n$ *for* $n \geq 2$*), i.e.*

$$f(x; \mu, \Sigma, \lambda) = 2 \cdot \tilde{f}(x; \mu, \Sigma) \cdot F\left(\lambda^T(x - \mu)\right) \tag{7}$$

where $\tilde{f}$ *is the density of an elliptical distribution and* F *is a cumulative distribution function of a univariate distribution. Let* M *be a non-singular matrix. If* X *is a random vector with density given by* (7), *then the density of* $Y := M(X - \mu)$ *is given by*

$$f(y; \mu, \Sigma, \lambda) = 2 \cdot \tilde{f}(y; 0, M\Sigma M^T) \cdot F\left(\lambda^T M^{-1} y\right),$$

that is Y, *is also skew-elliptically distributed with parameters* 0, $M\Sigma M^T$, *and* $(M^{-1})^T\lambda$.

Proof. The statement of lemma follows directly from the application of the change-of-variables formula which is applied first to the density $f(x; \mu, \Sigma, \lambda)$ and then to $\tilde{f}(x; \mu, \Sigma)$. □

Lemma 3. *Write* $X = (X_1, X_2^T)^T$, *where* X_1 *takes values in* $\mathbb{R}$. *Write analogously* $x = (x_1, x_2^T)^T$ *for a realization of* X. *Assume that the density of* X *is given by*

$$f(x; 0, \Omega, \lambda) = 2 \cdot \tilde{f}(x; 0, \Omega) \cdot F(x_1)$$

where $\tilde{f}$ is the density of an elliptical distribution with density generator $\tilde{g}$ and F is a cumulative distribution function of a univariate distribution. Then, conditionally on X_1, X_2 is elliptically contoured distributed with location parameter $\Omega_{21}\frac{x_1}{\omega_{11}}$, dispersion matrix $\Omega_{22}-\frac{\Omega_{21}\Omega_{12}}{\omega_{11}}$, and density generator given by

$$\breve{g}(\cdot|x_1) \propto \tilde{g}(\omega_{11}^{-1}x_1^2 + \cdot),$$

where $\breve{g}(\cdot|x_1)$ is a decreasing function as soon as $\tilde{g}$ is decreasing.

Proof. Let Ω and $B = \Omega^{-1}$ be partitioned as

$$\Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} b_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \tag{8}$$

where $B_{12} = B_{21}^T$ and $b_{11} > 0$. Moreover, using the formula for the inverse of the partitioned matrix we get

$$B_{22} = \left(\Omega_{22} - \frac{\Omega_{21}\Omega_{12}}{\omega_{11}}\right)^{-1}, \quad B_{22}^{-1}B_{21} = -\frac{\Omega_{21}}{\omega_{11}}, \quad b_{11} - B_{12}B_{22}^{-1}B_{21} = \omega_{11}^{-1}.$$

Then, the application of

$$\begin{aligned} x^T\Omega^{-1}x &= x^TBx = b_{11}x_1^2 + 2B_{21}^Tx_2x_1 + x_2^TB_{22}x_2 \\ &= (x_2 + B_{22}^{-1}B_{21}x_1)^TB_{22}(x_2 + B_{22}^{-1}B_{21}x_1) + (b_{11} - B_{12}B_{22}^{-1}B_{21})x_1^2 \\ &= \omega_{11}^{-1}x_1^2 + (x_2 - \omega_{11}^{-1}\Omega_{21}x_1)^T\left(\Omega_{22} - \frac{\Omega_{21}\Omega_{12}}{\omega_{11}}\right)^{-1}(x_2 - \omega_{11}^{-1}\Omega_{21}x_1) \end{aligned}$$

leads to

$$\begin{aligned} \tilde{f}(x_1, x_2; 0, \Omega, \lambda) &= \frac{\Gamma\left(\frac{n}{2}\right)}{k_n\sqrt{\pi^n \det(\Omega)}}F(x_1) \\ \times\quad \tilde{g}&\left(\omega_{11}^{-1}x_1^2 + (x_2 - \omega_{11}^{-1}\Omega_{21}x_1)^T\left(\Omega_{22} - \frac{\Omega_{21}\Omega_{12}}{\omega_{11}}\right)^{-1}(x_2 - \omega_{11}^{-1}\Omega_{21}x_1)\right) \end{aligned}$$

from which the conditional distribution of X_2 given X_1 follows as provided in the statement of the lemma. □

In the following we assume that $\lambda \neq 0$, since otherwise the results of Theorem 2 can be applied. Let $\lambda = \left(\lambda_1, \lambda_2^T\right)^T$ with $\lambda_1 \neq 0$ which can be assumed without loss of generality, since otherwise the components of the vector $x - \mu$ can be rearranged. We define

$$M = \begin{pmatrix} \lambda_1 & \lambda_2^T \\ 0_{n-1,1} & I_{n-1} \end{pmatrix} \tag{9}$$

and applying [10, Proposition 2.31, p.45] to M yields

$$M^{-1} = \begin{pmatrix} \lambda_1^{-1} & -\lambda_1^{-1} \cdot \lambda_2^T \\ 0_{n-1,1} & I_{n-1} \end{pmatrix}$$

from which it follows that, for $c = \left(c_1, c_2^T\right)^T$,

$$(m_1, m_2^T)^T = (M^T)^{-1} c = \frac{c_1}{\lambda_1} \cdot \begin{pmatrix} 1 \\ -\lambda_2 \end{pmatrix} + \begin{pmatrix} 0 \\ c_2 \end{pmatrix},$$

i.e. $m_1 = c_1/\lambda_1$ and $m_2 = -c_1/\lambda_1 \cdot \lambda_2 + c_2$. Define $\Omega = M \Sigma M^T$ and consider the representation

$$\Sigma = \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

of the dispersion matrix. Furthermore, let

$$\Sigma_1 = \begin{pmatrix} \sigma_{11} \\ \Sigma_{21} \end{pmatrix} \qquad \text{and} \qquad \Sigma_2 = \begin{pmatrix} \Sigma_{12} \\ \Sigma_{22} \end{pmatrix}.$$

For symmetric Σ, it follows that

$$\Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

where $\omega_{11} = \lambda^T \Sigma \lambda$, $\Omega_{22} = \Sigma_{22}$ and $\Omega_{21}^T = \Omega_{12} = \lambda^T \Sigma_2$. In preparation of Theorem 3 we let

$$\begin{aligned} k &= m_1 + \omega_{11}^{-1} m_2^T \Omega_{21} = \frac{c_1}{\lambda_1} \cdot \left[1 - \frac{\lambda_2^T \Sigma_2^T \lambda}{\lambda^T \Sigma \lambda}\right] + \frac{c_2^T \Sigma_2^T \lambda}{\lambda^T \Sigma \lambda} = \frac{c^T \Sigma^T \lambda}{\lambda^T \Sigma \lambda} \\ A &= \Omega_{22} - \omega_{11}^{-1} \Omega_{21} \Omega_{12} \end{aligned}$$

Theorem 3. *Let f be the density of a skew-elliptical distribution (in $\mathbb{R}^n$ for $n \geq 2$) in the sense of Definition 3, with decreasing density generator $\tilde{g}$. Let $\omega_{11}, \Omega_{21}, k$ be defined as previously.*

Then, a global maximum of the constrained optimization problem

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & f(x; \mu, \Sigma, \lambda) \\ \text{subject to} \quad & c^T x \geq \ell \end{aligned} \tag{10}$$

is given by

$$x^*(\ell) = \mu + M^{-1} \begin{pmatrix} z_{1;max}(\ell) \\ z_{2;max}(\ell) \end{pmatrix}$$

where

$$z_{1;max}(\ell) \in \begin{cases} \arg \max_{k z_1 \geq \ell - c^T \mu} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2\right) & \text{if } c = \eta\lambda \\ \arg \max_{z_1} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + k_1 \cdot (\max\left\{0, \ell - c^T\mu - k z_1\right\})^2\right) & \text{otherwise} \end{cases} \tag{11}$$

$$z_{2;max}(\ell) \;=\; \omega_{11}^{-1}\Omega_{21} z_{1;max}(\ell) + k_1 \cdot \max\{0, \ell - c^T\mu - k z_{1;max}(\ell)\} \cdot A m_2$$

and $k_1 = \begin{cases} 0 & \textit{if } c = \eta\lambda \\ (m_2^T A m_2)^{-1} & \textit{otherwise} \end{cases}.$

Proof. We consider the following transformation given by

$$z = M(x - \mu),$$

where M is given in (9). Then, from Lemma 2, the density of $z = (z_1, z_2^T)^T$ is given by

$$f(z; 0, \Omega, \lambda) = 2 \cdot \tilde{f}(z; 0, \Omega) \cdot F(z_1) = \frac{\Gamma\left(\frac{n}{2}\right)}{k_n \sqrt{\pi^n \det(\Omega)}} \cdot F(z_1)$$
$$\times\, \tilde{g}\left(\omega_{11}^{-1} z_1^2 + (z_2 - \omega_{11}^{-1}\Omega_{21} z_1)^T A^{-1} (z_2 - \omega_{11}^{-1}\Omega_{21} z_1)\right),$$

where Ω is partitioned as in (8).

Furthermore, the condition $c^T x \ge \ell$ is equivalent to $((M^T)^{-1}c)^T z \ge \ell - c^T\mu$ or $m_2^T z_2 \ge \ell - c^T\mu - m_1 z_1$. It follows that

$$\max_{x: c^T x \ge \ell} f(x; \mu, \Sigma, \lambda)$$

is equivalent to

$$\max_{z_1} \max_{z_2 : m_2^T z_2 \ge \ell - c^T\mu - m_1 z_1} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + (z_2 - \omega_{11}^{-1}\Omega_{21} z_1)^T A^{-1} (z_2 - \omega_{11}^{-1}\Omega_{21} z_1)\right),$$

and by Lemma 3 the objective function is, for every fixed z_1, the density of an elliptical distribution with location parameter $\tilde{\mu} = \omega_{11}^{-1}\Omega_{21} z_1$ and dispersion matrix $\tilde{\Sigma} = \Omega_{22} - \omega_{11}^{-1}\Omega_{21}\Omega_{12}$. If $c = \eta\lambda$ then $m_2 = 0$ and the optimizer is $z_2 = \omega_{11}^{-1}\Omega_{21} z_1$. Otherwise let $\tilde{c} = m_2^T$ and $\tilde{\ell} = \ell - c^T\mu - m_1 z_1$. Then, by Theorem 2, it follows that, for $v = (z_2 - \omega_{11}^{-1}\Omega_{21} z_1)$,

$$\begin{aligned} z_{1;max} &= \arg\max_{z_1} \max_{z_2 : m_2^T z_2 \ge \tilde{\ell}} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + v^T A^{-1} v\right) \\ &= \arg\max_{z_1} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + \frac{(\max\left\{0, \ell - c^T\mu - k z_1\right\})^2}{m_2^T A m_2}\right) \end{aligned}$$

where $k = m_1 + \omega_{11}^{-1} m_2^T \Omega_{21}$ which is attained at

$$z_{2;max} = \omega_{11}^{-1}\Omega_{21} z_{1;max} + \max\{0, \ell - c^T\mu - k z_{1;max}\} \cdot \frac{A m_2}{m_2^T A m_2}.$$

Finally, using the inverse transformation we get x_{max} as stated in the theorem. □

To derive a more explicit representation of $x^*(\ell)$ one can verify the following three identities:

$$M^{-1} \begin{pmatrix} 1 \\ \omega_{11}^{-1}\Omega_{21} \end{pmatrix} = \frac{\Sigma\lambda}{\lambda^T \Sigma \lambda}$$

$$M^{-1}\begin{pmatrix} 0 \\ Am_2 \end{pmatrix} = \Sigma c - \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda}\Sigma\lambda$$

$$m_2^T A m_2 = c^T\left(\Sigma c - \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda}\Sigma\lambda\right)$$

These imply the following corollary.

Corollary 2. *Under the assumptions of Theorem 3 it holds that*

$$x^*(\ell) = \mu + \begin{cases} \eta^{-1} \cdot z_{1;max}(\ell) \cdot \frac{\Sigma c}{c^T \Sigma c} & \text{if } c = \eta\lambda \\ z_{1;max}(\ell) \cdot \frac{\Sigma\lambda}{\lambda^T \Sigma \lambda} + \max\left\{0, \ell - c^T\mu - k \cdot z_{1;max}(\ell)\right\} \frac{v}{c^T v} & \text{otherwise} \end{cases}$$

where $v = \Sigma c - \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda}\Sigma\lambda$ *and* $k = \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda}$.

Thus $x^*(\ell)$ depends on $\|\lambda\|$ (for $\lambda \neq 0$) only through $z_{1;max}(\ell) \cdot \|\lambda\|$. The behaviour of this term is obviously governed by the properties of the density generator $\tilde{g}$ and the cdf F. Furthermore a quick calculation shows that, for $c \neq \eta\lambda$, $c^T x^*(\ell) = \ell$ if $\ell - c^T\mu - k \cdot z_{1;max}(\ell) \geq 0$.

A simple consequence of equation (11) is the following corollary:

Corollary 3. *The number of solutions to* (10) *is bounded from above by the number of optimizers of the two one-dimensional optimization problems that can be derived from* (11).

4 Example: Reverse Stress Testing in (Skew-)Normal Models

In light of Theorem 3 it is interesting to consider, when a unique solution $z_{1;max}(\ell)$ of (11) exists. In this section, we will answer this question for the skew-normal distribution, i.e. for the case where f is a normal density and F is the cumulative distribution function of the univariate normal distribution, which we will denote by Φ.

The following lemma is the first step.

Lemma 4. *Consider the optimization problem*

$$\arg\max_{q} \quad g\left(\eta_1 q^2 + \eta_2 q + \eta_3\right) \cdot \Phi\left(q\right), \tag{12}$$

where $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$ *with* $\eta_1 > 0$. *The objective function is unimodal and therefore the optimization problem always has a unique solution* q^* *which satisfies* $\phi\left(-\frac{\eta_2}{2\eta_1}\right)/\Phi\left(-\frac{\eta_2}{2\eta_1}\right) \geq q^* + \frac{\eta_2}{2\eta_1} \geq 0$.

An immediate consequence of Corollary 3 and the preceding Lemma is the following corollary:

Corollary 4. *In the skew-normal case there are at most two solutions to* (10).

The next step is to characterize the solution for large losses ℓ:

Corollary 5. *Let f be a skew-normal density. Then there exists L such that, for all $\ell \geq L$ and c, λ not collinear, the solution $z_{1;max}(\ell)$ of* (11) *is unique.*

Proof. Since $\tilde{g}$ is monotonically decreasing and $k_1 > 0$ it is sufficient to show that there exists $L \in \mathbb{R}$ such that, for all $\ell \geq L$, the solution q^* of

$$\begin{aligned} q^*(\ell) &= \arg\max_q F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} q^2 + k_1 \cdot (\ell - c^T \mu - kq)^2\right) \\ &= \arg\max_q F(z_1) \cdot \tilde{g}\left((\omega_{11}^{-1} + k_1 k^2) q^2 - 2k_1 k(\ell - c^T \mu) q + (\ell - c^T \mu)^2\right) \end{aligned}$$

(which is unique by Lemma 4) satisfies $\ell - c^T\mu - kq^*(\ell) \geq 0$. Consider the case $k > 0$ (the case $k < 0$ is analogous). Then we need to show that $q^*(\ell) \leq (\ell - c^T\mu)/k$ which is implied (cf. Lemma 4) by $\phi\left(-\frac{\eta_2}{2\eta_1}\right) / \Phi\left(-\frac{\eta_2}{2\eta_1}\right) - \frac{\eta_2}{2\eta_1} \leq (\ell - c^T\mu)/k$ where $-\frac{\eta_2}{2\eta_1} = \frac{k_1 k(\ell - c^T\mu)}{\omega_{1,1}^{-1} + k_1 k^2} \underset{\ell \to \infty}{\to} \infty$. Since $\phi(x)/\Phi(x) \underset{x \to \infty}{\to} 0$ it is sufficient to show that $\frac{k_1 k}{\omega_{1,1}^{-1} + k_1 k^2} < \frac{1}{k}$. But this follows immediately from the fact that $\omega_{11} > 0$ and $k_1 > 0$:

$$\frac{k_1 k}{\omega_{1,1}^{-1} + k_1 k^2} = \frac{1}{k} \cdot \frac{k_1 k^2}{\omega_{1,1}^{-1} + k_1 k^2} < \frac{1}{k} \cdot \frac{k_1 k^2}{k_1 k^2} = \frac{1}{k}$$

□

From a numerical point of view the following approach is well-suited for solving (12):

Remark 3. *Noting that the logarithm is monotonically increasing on $\mathbb{R}_+$ and since*

$$\log\left(g\left(\eta_1 q^2 + \eta_2 q + \eta_3\right) \cdot \Phi(q)\right) = -\frac{1}{2}\left(\eta_1 q^2 + \eta_2 q + \eta_3\right) + \log(\Phi(q))$$

the optimization problem (12) *can be rewritten as*

$$\arg\max_q \quad -\frac{1}{2}\left(\eta_1 q^2 + \eta_2 q + \eta_3\right) + \log(\Phi(q)),$$

which is numerically more suitable since exp *grows very quickly. Because the objective function is differentiable it is sufficient to find the unique root of the derivative, which is given by*

$$\begin{aligned} &\frac{\partial}{\partial q}\left(-\frac{1}{2}\left(\eta_1 q^2 + \eta_2 q + \eta_3\right) + \log(\Phi(q))\right) \\ &= -\eta_1 q - \frac{\eta_2}{2} + \frac{\phi(q)}{\Phi(q)}. \end{aligned}$$

In order to apply Newton's method, also the second derivative is of interest. To this end, note that the derivative of $\phi(q)/\Phi(q)$ is given in [7, Equation (2.20)]; see the quantity ζ_2 there.

It is now possible to illustrate the behaviour of the solution to the reverse stress testing problem in the skew-elliptical model. Only the case that ℓ is large enough, such that Corollary 5 holds, is depicted:

Example 1. *To illustrate how the selected scenario in the skew-normal model (cf. Theorem 3) differs from that without skewness (cf. Theorem 2) consider the points*

$$\lambda_\varphi = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

for $\varphi \in (0, 2\pi) \setminus \{\pi/2\}$ *which lie on the unit circle in* $\mathbb{R}^n$*. Denote by* x_φ *the scenario for* $\lambda = \lambda_\varphi$ *and* $c = \lambda_{\pi/2}$*. Furthermore denote by* x *the scenario for* $\lambda = 0$ *(i.e. the setting of Theorem 2) for the same portfolio weights* c*.*

Figure 1 considers four different cases for the covariance structure:

(a) Negative correlation and equal variances $\Sigma = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$

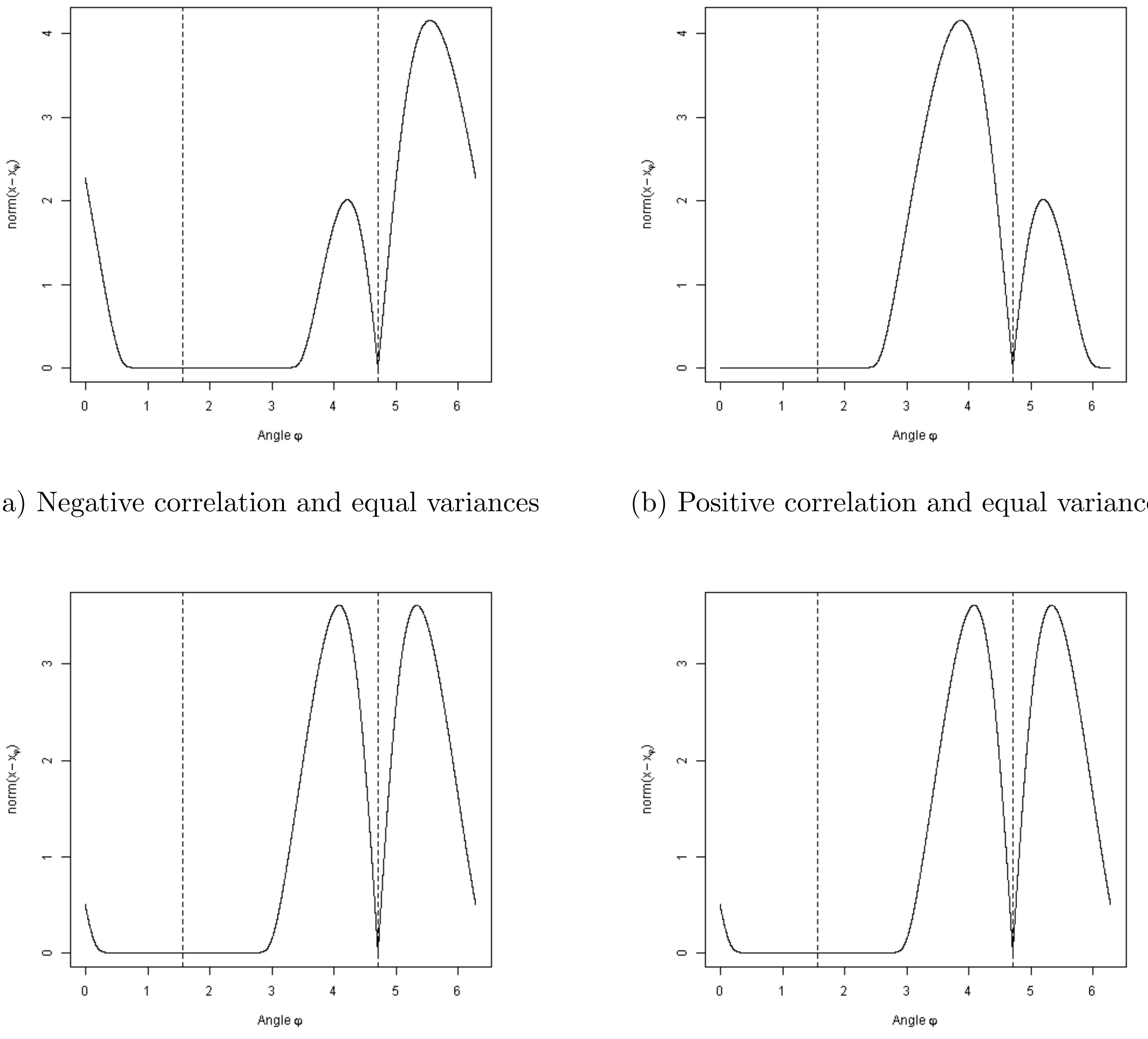
(b) Positive correlation and equal variances $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$

(c) Uncorrelated and equal variances $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(d) Uncorrelated and unequal variances $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$

As can be seen in Figure 1 the covariance structure does not significantly alter the qualitative behaviour: If λ_φ *points approximately in the direction of* c*, then the difference is comparatively small (since both vectors have norm 1).*

A numerical visualization can be done in the case that X *takes values in* $\mathbb{R}^3$*, by considering points on the unit sphere and plotting the distance against the scalar product. The results for two random covariance matrices are shown in Figure 2. Noting that the scalar product is zero if the vectors are orthogonal it is clear that the qualitative results are very similar to those in* $\mathbb{R}^2$*. For comparison Figure 3 contains the plots against the scalar product in the* $\mathbb{R}^2$ *case. Finally, Figure 4 compares the run times and the accuracy of our proposed methods with those of the all-purpose constrained optimization routine* `constrOptim` *in* `R`*.*



(a) Negative correlation and equal variances

(b) Positive correlation and equal variances

(c) Equal variances and uncorrelated

(d) Unequal variances and uncorrelated

Figure 1: Distance between the selected scenario in the skewed normal and non-skewed model in $\mathbb{R}^2$ where $\text{norm}(x) := \sqrt{x^T x}$ is the usual Euclidean norm. The dashed vertical lines mark the 'direction' (as well as its opposite) in which c points. For further details see Example 1.

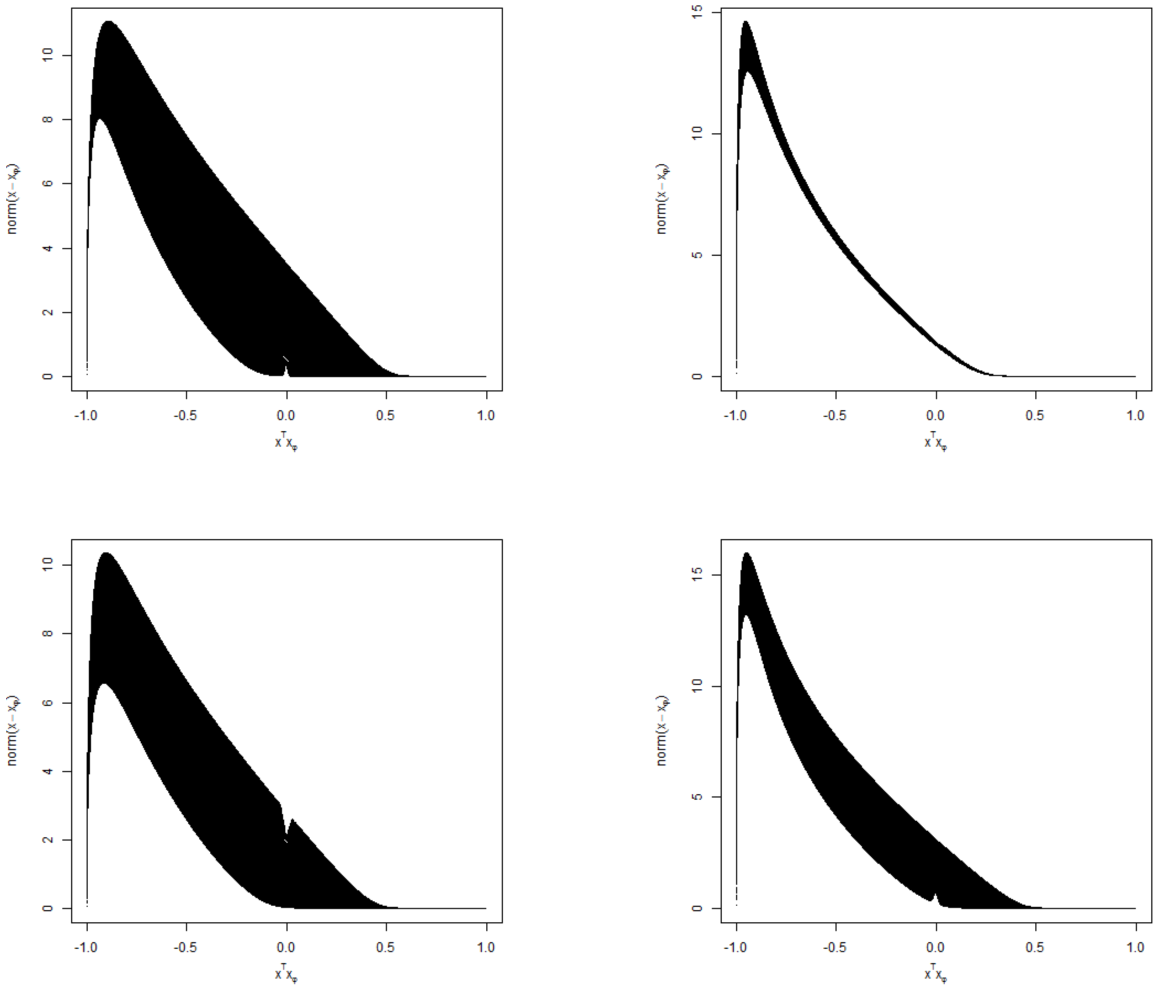
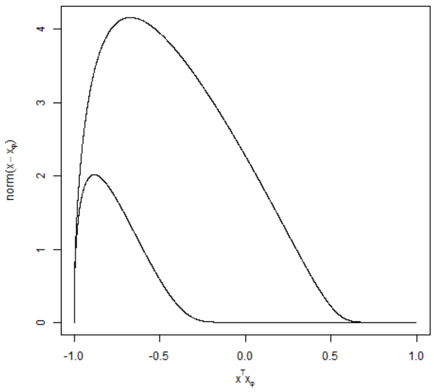
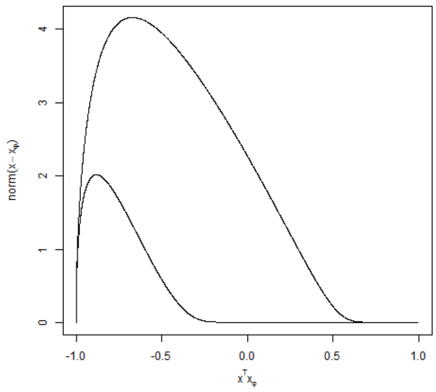


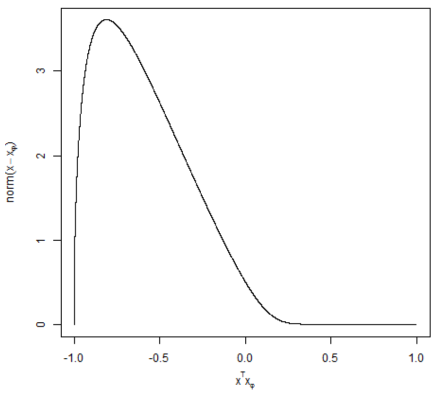
Figure 2: Distance between the selected scenario in the skewed normal and non-skewed model in $\mathbb{R}^3$ where $\mathrm{norm}(x) := \sqrt{x^T x}$ is the usual Euclidean norm. In each of the four sub-figures, the covariance matrix was chosen randomly. For further details see Example 1.



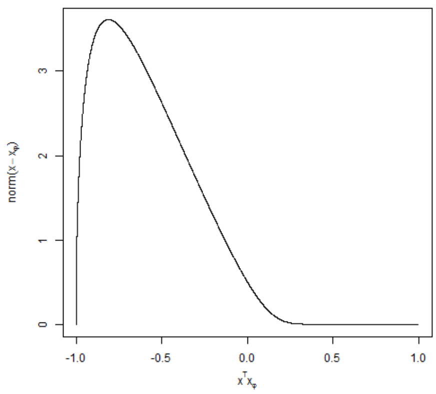
(a) Negative correlation and equal variances



(b) Positive correlation and equal variances



(c) Equal variances and uncorrelated



(d) Unequal variances and uncorrelated

Figure 3: Distance between the selected scenario in the skewed normal and non-skewed model in $\mathbb{R}^2$ where $\text{norm}(x) := \sqrt{x^T x}$ is the usual Euclidean norm. For the details see Example 1. The difference to figure 1 is, that the horizontal axis refers to the scalar-product and not the angle.

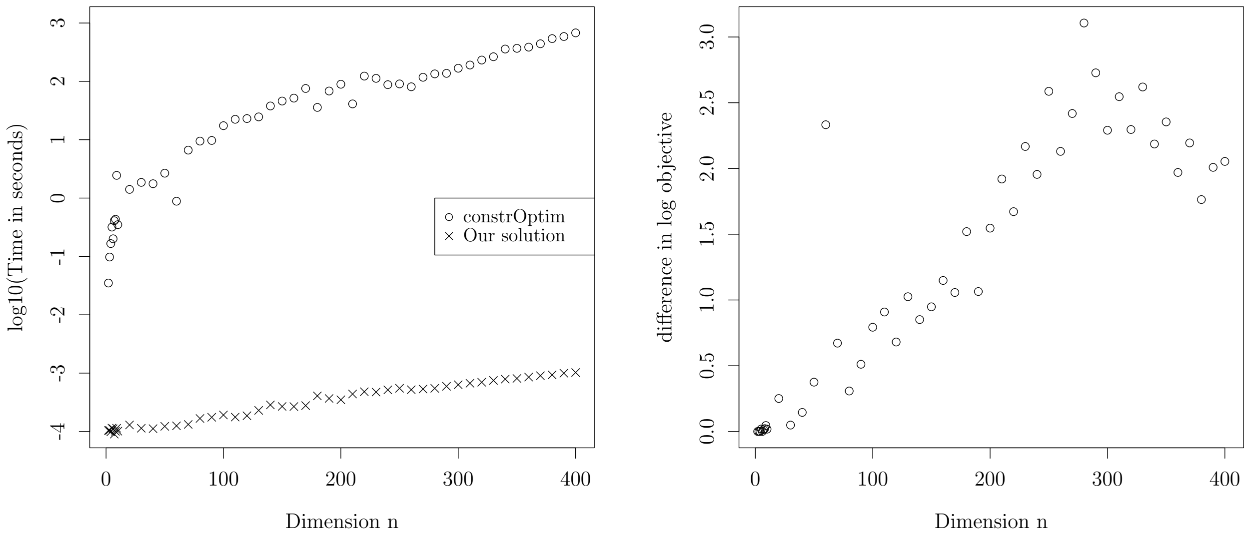


Figure 4: Comparison of the run time and achieved objective value when the dimensionality of n of the optimization problem increases. We compared our solution to a naive implementation using R's constrOptim method. In the right sub-figure, we calculated the difference of the value of the logarithmic objective function obtained with our method minus the corresponding value obtained by constrOptim. Positive values are in favour of our method.

5 Conclusion

Reverse stress testing is a highly relevant task in the context of bank regulation. Therefore, it is essential that reliable and numerically stable methods are available. With the present work, we have contributed (i) an explicit solution for the most likely scenario $x^*(\ell)$ given that the loss exceeds ℓ under the scope of elliptical models, and (ii) a characterization of $x^*(\ell)$ in terms of a numerically stable and easy-to-implement optimization problem under the broader scope of skew-elliptical models.

Potential extensions of our work would be to consider more general (non-linear) P & L functions as well as different distributional models, for instance by utilizing copula theory. A statistical task, which will be pursued in future research, is to estimate the loss distribution from data and to quantify the uncertainty which propagates from the latter estimation to the obtained value of $x^*(\ell)$.

Computer programs, with which all results of the present paper can be reproduced, are available from the first author upon request.

Acknowledgments

J. von Schroeder is supported by the Deutsche Forschungsgemeinschaft (DFG) within the framework of GRK 2224 / π^3: Parameter Identification - Analysis, Algorithms, Applications.

Appendix

Proof of Lemma 1. For the normal distribution it holds that $g(u) := \exp\left(-\frac{u}{2}\right)$ and since

$$g'(u) = -\frac{1}{2}\exp\left(-\frac{u}{2}\right) < 0$$

the generator is strictly decreasing.

For the Student-t distribution it holds that $g(u) := \left(1 + \frac{u}{c_p}\right)^{-p}$ for $c_p \in \mathbb{R}$, $p > \frac{1}{2}$. Since it holds that

$$g'(u) = -\frac{p\left(1 + \frac{u}{c_p}\right)^{-1-p}}{c_p}$$

and, noting that the integrability condition of definition 2 is only fulfilled for $c_p > 0$, it is immediate thate the generator is strictly decreasing since $g' < 0$.

For the Logistic distribution it holds that $g(u) := \frac{e^{-u}}{(1+e^{-u})^2}$. By the quotient rule

$$\begin{aligned} g'(u) &= \frac{-e^{-u}\left(1+e^{-u}\right)^2 + 2e^{-u}\left(1+e^{-u}\right)}{\left(1+e^{-u}\right)^4} \\ &= \frac{e^{-u}\left(1+e^{-u}\right)\cdot\left(-1-e^{-u}+2\right)}{\left(1+e^{-u}\right)^4} \\ &= \frac{e^{-u}\left(1+e^{-u}\right)\cdot\left(1-e^{-u}\right)}{\left(1+e^{-u}\right)^4} \end{aligned}$$

Since $1 < e^{-u}$ for $u > 0$ it follows that $g'(u) < 0$ on $\mathbb{R}_+$ and therefore g is strictly decreasing on $\mathbb{R}_+$.

For the Exponential distribution it holds that $g(u) := \exp\left(-ru^s\right)$ for $r, s > 0$. Furthermore

$$g'(u) = -\exp\left(-ru^s\right) rsu^{s-1}$$

and thus for $u > 0$ it holds that $g'(u) < 0$ and consequently g is strictly decreasing on $\mathbb{R}_+$. □

Proof of Lemma 4. First note that the objective function $\varphi(q) := g\left(\eta_1 q^2 + \eta_2 q + \eta_3\right)\cdot\Phi\left(q\right)$ looks very similar (cf. [7, Equation (2.1)]) to the density of an univariate skew-normal distribution. It is indeed possible to generalize [7, Proposition 2.6] to functions like φ: Showing that the second derivative of $\log\varphi(q)$ is still strictly negative will (as in the proof of [7, Proposition 2.6]) imply log-concavity which in turn will imply that φ has a unique mode which is the desired result. Let ζ_1, ζ_2 be given as in [7, Equation (2.20)]. Then it holds that

$$\begin{aligned}
\frac{\partial^2}{\partial q^2} \log \varphi(q) &= \frac{\partial^2}{\partial q^2} \left[-\frac{1}{2} \left(\eta_1 q^2 + \eta_2 q + \eta_3 \right) + \log \left(\Phi \left(q \right) \right) \right] \\
&= \frac{\partial}{\partial q} \left[-\eta_1 q - \frac{\eta_2}{2} \right] + \zeta_2(q) \\
&= -\eta_1 - \zeta_1(q) \left[q + \zeta_1(q) \right]
\end{aligned}$$

The first summand is negative by the assumption $\eta_1 > 0$ and by [7, Equation (2.21)] the second summand is negative, too. Thus $\frac{\partial^2}{\partial q^2} \log \varphi(q) < 0$. Furthermore from the first order conditions it follows that

$$\begin{aligned}
0 = \frac{\partial}{\partial q} \log \varphi(q^*) = -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi(q^*)}{\Phi(q^*)} \\
\geq -\eta_1 q^* - \frac{\eta_2}{2}
\end{aligned}$$

from which the lower bound follows directly. Furthermore this implies (since $x \mapsto \phi(x)/\Phi(x)$ is strictly decreasing)

$$0 = -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi(q^*)}{\Phi(q^*)} \leq -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi\left(-\frac{\eta_2}{2\eta_1}\right)}{\Phi\left(-\frac{\eta_2}{2\eta_1}\right)}$$

which implies $q^* \leq -\frac{\eta_2}{2\eta_1} + \frac{\phi\left(-\frac{\eta_2}{2\eta_1}\right)}{\Phi\left(-\frac{\eta_2}{2\eta_1}\right)}$. □

References

[1] Basel Committee on Banking Supervision, Principles for sound stress testing practices and supervision (2009).
URL `http://www.bis.org/publ/bcbs155.pdf`

[2] Board of Governors of the Federal Reserve System, Guidance on Stress Testing for Banking Organizations with Total Consolidated Assets of More Than $10 Billion (2012).

[3] Y. Kopeliovich, A. Novosyolov, D. Satchkov, B. Schachter, Robust Risk Estimation and Hedging: A Reverse Stress Testing Approach, The Journal of Derivatives 22 (4) (2015) 10–25. doi:10.3905/jod.2015.22.4.010.

[4] A. J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management, Princeton University Press, 2005.

[5] P. Glasserman, C. Kang, W. Kang, Stress scenario selection by empirical likelihood, Quantitative Finance 15 (1) (2015) 25–41. doi:10.1080/14697688.2014.926019.

[6] G. Studer, Maximum Loss for Measurement of Market Risk, Ph.D. thesis, ETH Zurich (1997).
URL http://www2.risklab.ch/ftp/papers/ThesisGeroldStuder.pdf

[7] A. Azzalini, The Skew-Normal and Related families, Cambridge University Press, Cambridge, 2014. doi:10.1017/CBO9781139248891.

[8] K.-T. Fang, S. Kotz, K.-W. Ng, W.-Y. Loh, Symmetric Multivariate and Related Distributions, Springer, 1992.

[9] Z. M. Landsman, E. A. Valdez, Tail Conditional Expectations for Elliptical Distributions, North American Actuarial Journal 7 (4) (2003) 55–71. doi:10.1080/10920277.2003.10596118.

[10] P. J.Dhrymes, Mathematics for Econometrics, Fourth edition, Springer, 2013.