Risk Measures and Their Applications in Asset Management

Ş. İlker Birbil
Faculty of Engineering and Natural Sciences, Sabancı University, 34956 Istanbul, Turkey.
sibirbil@sabanciuniv.edu

J.B.G. Frenk
Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands.
frenk@few.eur.nl

Bahar Kaynar
VU University Amsterdam, De Boelelaan 1081 HV, Amsterdam, The Netherlands.
bkaynar@feweb.vu.nl

Nilay Noyan
Faculty of Engineering and Natural Sciences, Sabancı University, 34956 Istanbul, Turkey.
nnoyan@sabanciuniv.edu

report EI 2008-14

ABSTRACT Several approaches exist to model decision making under risk, where risk can be broadly defined as the effect of variability of random outcomes. One of the main approaches in the practice of decision making under risk uses mean-risk models; one such well-known is the classical Markowitz model, where variance is used as risk measure. Along this line, we consider a portfolio selection problem, where the asset returns have an elliptical distribution. We mainly focus on portfolio optimization models constructing portfolios with minimal risk, provided that a prescribed expected return level is attained. In particular, we model the risk by using Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). After reviewing the main properties of VaR and CVaR, we present short proofs to some of the well-known results. Finally, we describe a computationally efficient solution algorithm and present numerical results.

Keywords: Elliptical distributions; mean-risk; value-at-risk; conditional value-at-risk; portfolio optimization.

1 Introduction

In recent years, off-balance sheet activities, such as trading financial instruments and generating income from loan sales, have begun to be profitable for banks in the competitive environment of the financial world. One of the main goals of such banks trading in these markets is to reduce the risk associated with their activities; however, with the taken positions trading may even be riskier. In particular, after the insolvency of some banks, the collapse of Barings in February 1995, risk management became quite significant in terms of internal control measures. One of these internal controls is recognition of the maximum loss that a portfolio can attain over a given time interval, termed Value-at-Risk (VaR). With VaR methodology, not only is exposed risk identified but VaR can also be used as a decision tool to take positions in the market so as to reduce the risk and, if possible, minimize it. The importance of VaR also stems from its status as universal risk measure employed in banking regulations, like Basel II, to evaluate capital requirements for banks' trading activities. Technically speaking, VaR at the confidence level α of a portfolio is the α-quantile of the distribution function of total random loss associated with the portfolio at a specified probability level. A closely related recent risk measure, Conditional Value-at-Risk (CVaR), on the other hand, is the expectation of loss values exceeding the VaR value with the corresponding probability level (Rockafellar and Uryasev, 2000).

Comparing random outcomes is one of the main interests of decision theory in the presence of uncertainty. Several decision models have been developed to formulate optimization problems in which uncertain quantities are represented by random variables. One method of comparing random variables is via the expected values. For the basic limitations of optimization models considering the expected value see, e.g., Shapiro and Ruszczyński (2006). In cases where the same decisions under similar conditions are repeatedly made, one can justify the optimization of the expected value by the Law of Large Numbers. However, the average of a few results may be misleading due to the variability of the outcomes. Therefore, sound decision models in the presence of uncertainty should take into account the effect of inherent variability, which in turn leads to the concept of risk. The preference relations among random variables can be specified using a risk measure. There are two main approaches for quantifying risk; it can be identified as a function of the deviation from an expected value or as a function of the absolute loss. The former approach is the main idea of the Markowitz mean-variance model. The latter approach involves the two recent risk measures mentioned above, Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR).

The challenge of managing a portfolio that includes finitely many assets has been a mainstay in finance literature. The simplest, most widely used approach for modeling changes in the portfolio value is the variance-covariance method popularized by RiskMetrics (1997). Two main assumptions of this model are as follows (Glasserman et al., 2002): (i) The risk factors are conditionally multivariate normal over a specified short horizon. (ii) The change in the portfolio value, mainly profit-loss function, is a linear function of the changes in the risk factors. In this setting, the term conditionally means that conditioned on the information available at the beginning of the short horizon--such as the price of the instruments or the value of the portfolio--the changes in the risk factors become multivariate normal. The central problem is to estimate the profit-loss function and its relation with the underlying risk factors of a portfolio over a specified horizon. Since VaR deals with the extreme losses, estimating the tail of the loss distribution is crucial for portfolio management. For example, although two different possible distributions for the price changes have the same mean, the probability of facing very large changes maybe much greater for one than it is for the other. Starting from this model, we may relax either the linearity assumption or the normality assumption. Methods such as delta-gamma, interpolation, or low variance Monte Carlo simulation relax the linearity assumption. Monte Carlo simulation is universally adaptable; however, since it is common in risk management to deal with rare-events, Monte Carlo simulation works much slower (Glasserman et al., 2002; Kamdem, 2005). The other option is to relax the normality assumption and use another family of distributions to model the returns of the underlying risk factors. The latter option, in fact, is the main focus of our work.

In this paper, we analyze a general risk management model applied to portfolio problems with VaR and CVaR risk measures. We assume that our portfolio is linear and the risk factor changes have an elliptical distribution. A similar approach was initiated by Rockafellar and Uryasev (2000) for the special case of multivariate normally distributed returns. The class of elliptical distributions is a general class of distributions, which contains the normal and the student t-distributions. Contrary to Rockafellar and Uryasev (2000); we do not explicitly talk about applications of financial concepts (such as hedging), which actually lead to similar models. In the literature, it is observed that market returns have heavier tails compared to normal distributions; many studies moreover discuss the comparisons of portfolios among families of distributions on returns (Fama, 1965; Praetz, 1972; Blattberg and Gonedes, 1974; Embrechts et al., 2002). In particular Blattberg and Gonedes (1974) illustrate that the student model has greater descriptive validity than do the other models. Although most of the works are restricted to t-distribution, Kamdem generalizes VaR and Expected Shortfall to the family of elliptical distributions. Embrechts et al. (2002) also analyze the class of elliptical distributions within the context of risk management. On one hand, these papers concentrate on measuring Value-at-Risk and Conditional Value-at-Risk measures. On the other hand, these do not include a discussion on portfolio optimization. In this paper, we explicitly focus on constructing optimal portfolios.

We first concentrate on general risk measures and then concentrate on VaR and CVaR. We also briefly review the theory of coherent risk measures, thoroughly studied in Arztner et. al. (1999). We then discuss the behavior of VaR and CVaR in terms of coherency. By converting the ideas used for rewards (Ogryczak and Ruszczyński, 2002), a different definition of CVaR is given. Following risk measures, we define general portfolio optimization problems. Additionally, we draw attention to the effect of adding a riskless asset. After a condensed introduction on elliptical distributions, we give short proofs on properties of VaR, CVaR, and linear loss functions. We note that the important risk measure CVaR is also discussed by Embrechts et. al. (2002) under the term expected shortfall or mean excess loss together with properties of elliptical distributions. Using the well-known equivalence between the mean-risk approach employing VaR and CVaR as risk measures and the famous mean-variance approach of Markowitz, we adapt an algorithm for special quadratic programming problems originally proposed by Michelot (1986). In this algorithm, the number of steps to find the optimal allocation of the assets is finite and equals at most the number of the considered assets. Our computational results suggest that the adapted algorithm is a faster alternative to the standard solver used in the financial toolbox of **MATLAB**. We also present some numerical results to emphasize the fact that we can construct optimal portfolios for returns having distributions different than normal; in particular we provide results for multivariate t-distributions.

2 Risk Measures

Consider an optimization problem in which the decision vector $\mathbf{x}$ affects an uncertain outcome represented by a random variable $Z(\mathbf{x})$. Thus, for a decision vector $\mathbf{x}$ belonging to a certain feasible set $\mathcal{X} \subseteq \mathbb{R}^n$, we obtain a realization of the real-valued random variable $Z(\mathbf{x})$, which may be interpreted as some reward or loss of the decision $\mathbf{x}$. In our work $Z(\mathbf{x})$ and $-Z(\mathbf{x})$ represent the loss and the reward of the decision $\mathbf{x}$, respectively. Therefore, smaller values of $Z(\mathbf{x})$ are preferred to larger ones. To find the 'best' values of the decision vector $\mathbf{x}$, we need to compare the random variables $Z(\mathbf{x})$ according to a preference relation. While comparing random variables, sounds decision models should take into account the effect of inherent variability, which leads to the concept of risk. The preference relations among random variables can be specified using a risk measure. One of the main approaches in practice uses mean-risk models. In these models one uses a specified risk measure $\rho : B \to [-\infty,+\infty]$, where ρ is a functional and B is a linear space of F-measurable functions on a probability space (Ω, F, P). Notice that for a given vector $\mathbf{x}$, the argument of the function ρ is a real-valued random variable denoted here by $Z(\mathbf{x})$ with the cumulative distribution function (cdf)

$$F_{Z(\mathbf{x})}(a) := P\{Z(\mathbf{x}) \le a\}. \tag{1}$$

Clearly $\rho(Z(\mathbf{x}_1)) = \rho(Z(\mathbf{x}_2))$ for $Z(\mathbf{x}_1)$ and $Z(\mathbf{x}_2)$ having the same cdf (denoted by $Z(\mathbf{x}_1) \overset{d}{=} Z(\mathbf{x}_2)$).

In the mean-risk approach for a given risk measure ρ one solves the problem

$$\max_{\mathbf{x} \in X} \{E[-Z(\mathbf{x})] - \lambda \rho(Z(\mathbf{x}))\}, \tag{2}$$

where $\lambda \ge 0$ is the trade-off coefficient representing our desirable exchange rate of mean reward for risk. We say that the decision vector $\mathbf{x}$ is efficient (in the mean-risk sense) if and only if for a given level of minimum expected reward, $\mathbf{x}$ has the lowest possible risk, and, for a given level of risk, $\mathbf{x}$ has the highest possible expected reward. In many applications, especially in portfolio selection problems, the mean risk efficient frontier is identified by finding the efficient solutions for different trade-off coefficients.

The classical Markowitz (1952) model discussed in Steinbach (2001) uses variance as a risk measure. One of the problems associated with the Markowitz's mean-variance formulation, however, is that it penalizes over-performance (positive deviation from the mean) and under-performance (negative deviations from the mean) equally. When typical dispersion statistics such as variance are used as risk measures, the mean-risk models may lead to inferior solutions. That is, there may exist other feasible solutions which would be preferred by any risk-averse decision maker to the efficient solution obtained by the mean-risk model.

Example 2.1 *Consider two decision vectors x_1 and x_2 for which the probability mass functions of the random outcomes (losses) are given as follows:*

$$P(Z(\mathbf{x}_1) = a_1) = \begin{cases} 1/2 & a_1 = 2 \\ 1/2 & a_1 = 4 \\ 0 & otherwise \end{cases} \quad and \quad P(Z(\mathbf{x}_2) = a_2) = \begin{cases} 1 & a_2 = 10 \\ 0 & otherwise. \end{cases}$$

Any rational decision maker would prefer the decision vector x_1 with the random loss $Z(\mathbf{x}_1)$. However, when a dispersion type risk measure $\rho(Z(\mathbf{x}))$ is used, then both decision vectors lie *in the efficient frontier of the corresponding mean-risk model, since for each such a risk measure* $\rho(Z(\mathbf{x}_1)) > 0$ *and* $\rho(Z(\mathbf{x}_2)) = 0$.

To overcome the preceding disadvantage, alternative asymmetric risk measures, such as downside risk, have been proposed and significant effort has been devoted to the development of downside risk models (see e.g., Ogryczak and Ruszczyński, 2002). We refer to Ruszczyński and Shapiro (2006) as well as Rockafellar et al. (2006), and the references therein for other stochastic optimization models involving general risk functionals. VaR and CVaR are also among the popular downside risk measures.

Definition 2.1 *The first quantile function* $F_X^{(-1)} : (0,1] \to \mathbb{R}$ *corresponding to a random variable* X *is the left-continuous inverse of the cumulative distribution function* F_X:

$$F_X^{(-1)}(\alpha) = \inf\{\eta \in \mathbb{R} : F_X(\eta) \ge \alpha\}.$$

In the finance literature, the α-quantile $F_X^{(-1)}(\alpha)$ is called the Value at Risk (VaR) at the confidence level α and denoted by $\mathrm{VaR}_\alpha(X)$. Using Definition 2.1 and (1), we can state that the realizations of the random variable X larger than $\mathrm{VaR}_\alpha(X)$ occur with probability less than $1-\alpha$.

A closely related and recently popular risk measure is the Conditional Value-at-Risk (CVaR), also called Mean Excess Loss or Tail VaR. CVaR at level α is defined as follows (Rockafellar and Uryasev, 2000, 2002; Pflug, 2000):

$$\mathrm{CVaR}_\alpha(X) = \inf_{\eta \in \mathrm{R}} \left\{ \eta + \frac{1}{1-\alpha} E\left[\max\{X-\eta, 0\}\right] \right\}. \tag{3}$$

The correspondence between the concepts of CVaR and VaR and the fact that CVaR is also based on a quantile approach can be seen from the following result. Employing a similar argument, a closely related statement has been proven in Ogryczak and Ruszczyński (2002), where the random variable X represents rewards (returns) instead of losses. As in Rockafellar (1972), the set $\partial f(y)$ denotes the subgradient set of a finite convex function $f : \mathbb{R} \to \mathbb{R}$ at y.

Lemma 2.1 *For any real valued random variable* X *having a finite first absolute moment*

$$\mathrm{CVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(X)\, du$$

for all $0 < \alpha < 1$.

Proof. Introducing the convex and continuous function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(y) := E\left[\max\{X - y, 0\}\right] \tag{4}$$

it follows by relation (3) that

$$\mathrm{CVaR}_{\alpha}(X) = (1-\alpha)^{-1} v(Q), \tag{5}$$

where $v(Q) := \inf_{y\in\mathrm{R}} \{(1-\alpha)y + f(y)\}$. Since

$$f(y) = \int_y^{\infty} (x-y) dF_X(x) = \int_y^{\infty} (1 - F_X(x)) dx,$$

we obtain that y^* is an optimal solution of the convex optimization problem

$$\inf_{y\in\mathrm{R}} \{(1-\alpha)y + f(y)\}$$

if and only if $1-\alpha$ belongs to $-\partial f(y^*)$. It is now easy to verify by the definition of the subgradient set and relation (4) that $1-\alpha$ belongs to $-\partial f(F_X^{-1}(\alpha))$. This result shows that $F_X^{-1}(\alpha)$ is an optimal solution of the optimization problem $\inf_{y\in\mathrm{R}} \{(1-\alpha)y + f(y)\}$ and so

$$v(Q) = (1-\alpha)F_X^{-1}(\alpha) + E\left[\max\{X - F_X^{-1}(\alpha), 0\}\right] \tag{6}$$

Since $X \overset{d}{=} F_X^{-1}(U)$, where U is a uniform distributed random variable on $(0,1)$, we obtain

$$E\left[\max\{X - F_X^{-1}(\alpha), 0\}\right] = E\left[\max\{F_X^{-1}(U) - F_X^{-1}(\alpha), 0\}\right] = \int_\alpha^1 \mathrm{VaR}_u(X) du - (1-\alpha)F_X^{-1}(\alpha)$$

and by relations (5) and (6) the assertion holds true. □

It is easy to see that the function $\alpha \mapsto \mathrm{VaR}_{\alpha}(X)$ is increasing; therefore, an immediate consequence of Lemma 2.1 is given by

$$\mathrm{CVaR}_{\alpha}(X) \geq \mathrm{VaR}_{\alpha}(X).$$

Moreover, when F_X is continuous on $(-\infty, \infty)$, we know that $\mathrm{VaR}_{\alpha}(X)$ is not an atom of the distribution of X; therefore, $1-\alpha = P\{X \geq \mathrm{VaR}_{\alpha}(X)\}$. Then we have

$$\int_\alpha^1 \mathrm{VaR}_u(X) du = E\left[F_X^{-1}(U) 1_{\{U\geq\alpha\}}\right] = E\left[F_X^{-1}(U) 1_{\{F_X^{-1}(U)\geq F_X^{-1}(\alpha)\}}\right] = E\left[X 1_{\{X\geq \mathrm{VaR}_{\alpha}(X)\}}\right]$$

It follows from the last equation that

$$\mathrm{CVaR}_{\alpha}(X) = E\left[X \mid X \geq \mathrm{VaR}_{\alpha}(X)\right], \tag{7}$$

which has been also shown in Rockafellar and Uryasev (2000) by using another approach.

This definition provides a clear understanding of the concept of CVaR, i.e., the conditional expectation of values above the Value-at-Risk at the confidence level α. For example, in the portfolio optimization context, VaR_{α} is the α-quantile (a high quantile) of the distribution of losses (negative returns), which provides an upper bound for a loss that is exceeded only with a small probability $1-\alpha$. On the other hand, CVaR_{α} is a measure of severity of loss if we lose more than VaR_{α}.

An axiomatic approach to construct risk measures has been proposed by Artzner et al. (1999). It is now widely accepted that risk measures should satisfy the following set of axiomatic properties:

1. Monotonicity: $\rho(X_1) \geq \rho(X_2)$ for any $X_1, X_2 \in B$ such that $X_1 \geq X_2$ (where the inequality $X_1 \geq X_2$ is assumed to hold almost surely).
2. Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ for $X_1, X_2 \in B$.
3. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$ for any $\lambda > 0$ and $X \in B$.
4. Translation invariant: $\rho(X + a) = \rho(X) + a$ for any $a \in \mathbb{R}$ and $X \in B$.

A risk measure satisfying the above properties is called a coherent risk measure. It is well-known that CVaR is a coherent risk measure, but due to the lack of subadditivity, VaR fails to be a coherent risk measure in general (Pflug, 2000).

Optimization models involving VaR are technically and methodologically difficult; for details see Rockafellar and Uryasev (2002). As also observed by Crouhy et al. (2001), using VaR as a risk measure has been criticized mainly for not being subadditive; hence, not being convex. In some applications, nonconvexity is a significant objection since it does not reward diversification. For example, in the portfolio selection theory, lack of conversity implies that portfolio diversification may increase the risk and considering the advantages of a portfolio diversification strategy, this objection cannot be ignored. However, as discussed by Embrechts et al. (2002), VaR is convex in the elliptical world (see Section 4), and so, within this framework VaR is a coherent risk measure. Therefore, we can use VaR in our portfolio selection problems. In the next section, we present single period portfolio optimization models.

3 A Single Period Portfolio Optimization Problem

In this section, we consider a single period (short term) portfolio selection problem with a set of n risky assets. At the beginning of the period, the length of which is specified, the investor decides on the amount of capital to be allocated on each available asset. At the end of the investment period, each asset generates a return, which is uncertain at the beginning of the period since the future price of an asset is unknown. We represent these uncertain returns with random variables and denote the vector of random returns of assets $1, 2, \ldots, n$ by $\mathbf{Y}^{\mathrm{T}} = (Y_1, Y_2, \ldots, Y_n)$. In finance, the ratio of money gained or lost on an investment relative to the money invested is called the rate of return or percentage return, which throughout the paper we just refer to as simply "return".

We denote the fractions of the initial capital and the amounts of the initial capital invested in assets $j = 1,\ldots,n$ by $\mathbf{x}^{\mathrm{T}} = (x_1,\ldots,x_n)$ and $\tilde{\mathbf{x}}^{\mathrm{T}} = (\tilde{x}_1,\ldots,\tilde{x}_n)$, respectively. Thus if $\tilde{x}_j$ is the amount of the capital invested in asset j and C is the total amount of capital to be invested, we have $x_j = \tilde{x}_j / C$ for $j = 1,\ldots,n$. The constructed portfolio may be represented by either of these two decision vectors. We assume that short-selling is not allowed, which means that investors can not sell assets they do not own presently in the hope of repurchasing them later at a lower price. Therefore, the portfolio decision variables are nonnegative. If short-selling is allowed, however, the decision variables would be unrestricted. As mentioned in Steinbach (2001), the classical Markowitz model to be introduced next has in this case an analytical solution. Clearly, the set of possible asset allocations is:

$$\tilde{\mathcal{X}} = \{\tilde{\mathbf{x}} \in \mathbb{R}^n : \tilde{x}_1 + \cdots + \tilde{x}_n = C, x_j \geq 0,\ j = 1,\ldots,n\},$$

or equivalently,

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : x_1 + \cdots + x_n = 1, x_j \geq 0,\ j = 1,\ldots,n\}.$$

Then, at the end of the investment period, the total value of the portfolio is $C + \tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{Y}$; therefore, the loss of the portfolio for the period under consideration is

$$Z(\tilde{\mathbf{x}}) = C - (C + \tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{Y}) = -\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{Y}.$$

Let $\boldsymbol{\mu}^{\mathrm{T}} = (\mu_1,\ldots,\mu_n)$, where μ_i denotes the expected return of asset i, i.e., $\mu_i = E[Y_i]$ for $i = 1,\ldots,n$. Then the expected total return (reward) of the portfolio $\tilde{\mathbf{x}}$ is $\tilde{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\mu}$.

The problem of choosing between portfolios becomes the problem of choosing between random losses according to a preference relation, which is specified using a risk measure. The mean-risk models have been widely used for portfolio optimization under risk. In these models one uses two criteria: the mean representing the expected total return or loss of a portfolio, and the risk which is a scalar measure of the variability of the random total return or loss of the portfolio. Markowitz's mean-variance model (1952, 1959) which uses variance of return as the risk measure, has been one of the most widely used mean-risk model for the portfolio selection problem. However, as mentioned in Section 2, the model has several disadvantages such as equally treating over-performance as under-performance. Markowitz (1959) also recommends using semivariance rather than variance as risk measure, but even in this case significant deficiencies remain as mentioned in e.g. Ogryczak and Ruszczyński (2002). In particular, here we use VaR and CVaR as risk measures.

There are alternative approaches to implement a mean-risk model. For example, one approach is based on the model constructing a portfolio with minimum risk, provided that a desired level of expected return of the portfolio is attained (by enforcing a lower bound on the expected total return of the portfolio). Another approach is based on the problem formulated in the form of (2), in which the preference relation is defined using a trade-off between the mean (reward) and risk, where a larger value of mean (reward) and a smaller value of risk are preferable. In many applications, the trade-off coefficient does not provide a clear understanding of the decision makers' preferences. The commonly accepted approach to implement a mean-risk model is to minimize the risk of the random outcome while enforcing a specified lower bound on the total expected return (see, e.g. Mansini et al., 2003). We also prefer to use this widely accepted bounding approach. Thus, among alternative formulations of the mean-variance Markowitz model, we consider the formulation constructing a portfolio with minimal risk provided that a prescribed expected return level w is attained:

$$\min\left\{\rho(-\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{Y}) : \mathbf{e}^{\mathrm{T}}\tilde{\mathbf{x}} = C, \tilde{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\mu} = w, \tilde{\mathbf{x}} \geq \mathbf{0}\right\}$$

Notice that $\rho(-\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{Y})$ is the risk of the portfolio represented by $\tilde{\mathbf{x}}$.

With the use of the decision vector $\mathbf{x}$ representing the fractions of the capital invested in each asset, we obtain an equivalent optimization problem:

$$\min\{\rho(-C\mathbf{x}^{\mathrm{T}}\mathbf{Y}) : \mathbf{e}^{\mathrm{T}}\mathbf{x} = 1, \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu} = wC^{-1}, \mathbf{x} \geq \mathbf{0}\}. \tag{8}$$

Suppose that $C + w = (1+r)C$ where r is the desired rate of the return of the portfolio. When the specified risk measure is positive homogeneous, problem (8) takes the form of

$$\min\{\rho(-\mathbf{x}^{\mathrm{T}}\mathbf{Y}) : \mathbf{e}^{\mathrm{T}}\mathbf{x} = 1, \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu} = r, \mathbf{x} \geq \mathbf{0}\}. \tag{Q}$$

If we also consider a non-risky asset characterized by a known return r_0 that usually reflects the interest rate on the money market, this asset would generate a return of $r_0 x_0$ at the end of the period, where x_0 denotes the fraction of the capital invested in the non-risky asset and C the total capital available. In this case, we obtain the following optimization problem

$$\min\{\rho(-\mathbf{x}^{\mathrm{T}}\mathbf{Y}) : \mathbf{e}^{\mathrm{T}}\mathbf{x} + x_0 = 1, \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu} + r_0 x_0 = r, \mathbf{x} \geq \mathbf{0}, x_0 \geq 0\}.$$

In the above portfolio selection problems, no transaction costs are involved. Nonetheless, if the transaction costs are linear functions in terms of the decision vectors, we have similar formulations and our subsequent discussion still applies.

In our work we use the class of multivariate elliptical distributions to model the random returns. This general class of multivariate distributions contains both (multivariate) normal and

t-distributions. The most popular approach for modeling (short term) changes in portfolio value is the analytical variance-covariance approach popularized by RiskMetrics (1997). This method assumes that the vector of rate of returns is multivariate normal. However, there is a considerable amount of evidence that empirical rate of returns over a short horizon have heavier tails than given by the multivariate normal distribution. For example, Fama (1965) and Mandelbrot (1963) show through empirical studies on real stock portfolios that the distribution of returns can be distinguished from the normal distribution. Recent studies by Embrechts et al. (2002), Glasserman et al. (2002) and Huismann et al. (1998) also support this theory. Heavy tails imply that extreme losses are more likely to occur. Thus, if a risk measure based on the tail of the loss distribution, such as VaR, is used to optimize the portfolio, we underestimate the actual risk under a normality assumption. To overcome the problem of heavy tails, several alternative distributions for rates of returns are offered. One of the strongest is the multivariate t-distribution which belongs to the class of elliptical distributions. Empirical support on modeling univariate rate of returns with a t-distribution can be found in Huisman et al. (1998), Praetz (1972), Glasserman et al. (2002), and Blattberg and Gonedes (1974). The multivariate t-distribution is fully characterized by the mean μ, the covariance matrix Σ and an additional parameter called the degree of freedom ν to control the heaviness of the tail. As ν goes to infinity, the multivariate t-distribution approaches the multivariate normal distribution. According to Crouhy et al. (2001) and Glasserman et al. (2002), the values of parameter ν for most of the rate of returns are between 3 and 8--in fact, usually around 4. However, a shortcoming of the multivariate t-distribution is that all the risk factors in the portfolio have the same degrees of freedom. As suggested by Glasserman et al. (2002) to overcome this shortcoming, copulas can be used with different ν values for each rate of return. The other candidate for a multivariate distribution of the rates of returns is the family of multivariate stable distributions (see Feller (1971) for a discussion of univariate stable distributions). The comparison of stable distributions with a t-distribution and the normal distribution can be found in Blattberg and Gonedes (1974) and Praetz (1972).

The class of elliptical distributions within the context of risk management has been studied by Embrechts et al. (2002) and Kamdem (2005). On one hand, both papers concentrate on measuring Value-at-Risk and Conditional Value-at-Risk measures. On the other hand, they do not include a discussion on portfolio optimization. In this paper, we explicitly focus on constructing the optimal portfolios. We next briefly discuss the properties of elliptical distributions.

4 Elliptical World

To analyze our general risk model for portfolio management, we first introduce the following class of multivariate distributions (see also Embrechts et al., 2002 and Fang, 1990). Recall a linear mapping U is called orthogonal if $U^{\mathrm{T}}U = UU^{\mathrm{T}} = I$. We also use the notation $\mathbf{X} : F$, meaning the random vector $\mathbf{X}$ has the joint distribution function F.

Definition 4.1 *A random vector* $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$ *has a spherical distribution if for any orthogonal mapping* $U : \mathbb{R}^n \to \mathbb{R}^n$*, it holds that,*

$$U\mathbf{X} \overset{d}{=} \mathbf{X}.$$

It is well known that $\mathbf{X} : N(\mathbf{0}, I)$ has a spherical distribution, where $N(\mu, \Sigma)$ denotes the multivariate normal distribution with mean μ and covariance matrix Σ. Since $U = -I$ is an orthogonal mapping we obtain for $\mathbf{X}$ having a spherical distribution that

$$-\mathbf{X} \overset{d}{=} \mathbf{X} \tag{9}$$

Hence if a spherically distributed random vector $\mathbf{X}$ has a finite expectation, its expected value equals $\mathbf{0}$. It can be easily shown using the above definition (see Fang, 1990) that the random vector $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$ has a spherical distribution if and only if there exists some real-valued function $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that the characteristic function $\psi(\mathbf{t}) := E\left[\exp(i\mathbf{t}^{\mathrm{T}}\mathbf{X})\right]$ is given by

$$\psi(\mathbf{t}) = \phi(\|\mathbf{t}\|^2). \tag{10}$$

This representation based on the characteristic function provides us with an alternative definition of a spherical distributed random vector. It is easy to show (see Fang, 1990) for any spherically distributed random vector $\mathbf{X}$ that there exists some nonnegative random variable R such that

$$\mathbf{X} \overset{d}{=} R\mathbf{U}^{(n)}, \tag{11}$$

where R is independent of the random vector $\mathbf{U}^{(n)}$ that is uniformly distributed on the unit sphere surface $S_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^{\mathrm{T}}\mathbf{x} = 1\}$. The alternative representation (11) will be useful for the computation of the covariance matrix for an elliptically distributed random variable. As mentioned before, an important member of the class of spherical distributions is the standard multivariate normal distribution $N(\mathbf{0},\mathbf{I})$ with mean $\mathbf{0}$ and covariance matrix I. For this distribution, the generating random variable R in (11) has a chi-distribution χ_n with n degrees of freedom. Another important member is the standard multivariate t-distribution with ν degrees of freedom. In this case $R^2 n^{-1}$ has a $F(n,\nu)$-distribution with n and ν degrees of freedom. As stated by Fang (1990) an important proper subclass of the elliptical distributions is the so-called class of scale mixtures of multivariate normal distributions. The random vector associated with such a scale mixture is given by $\mathbf{X} = S\mathbf{V}$, where $\mathbf{V} : N(\mathbf{0},\mathbf{I})$ and S is a real-valued random variable, which is independent of $\mathbf{V}$. The already introduced standard multivariate t-distribution with ν degrees of freedom belongs to this class. For this distribution, the random variable $\nu^{1/2}S$ has a chi-distribution χ_ν with ν degrees of freedom. This representation will be useful in our computational section. From the representation of the characteristic function we immediately obtain for all $t \in \mathbb{R}$ and $1 \le j \le n$ that

$$E\left[\exp(itX_j)\right] = \phi(t^2). \tag{12}$$

This confirms (see also relation (9)) that

$$-X_1 \overset{d}{=} X_1.$$

Using the characteristic function representation of a spherical distribution another useful description can also be derived. For completeness, here a short proof is presented (see also Fang, 1990).

Lemma 4.1 *The random vector* $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$ *has a spherical distribution if and only if* $\mathbf{a}^{\mathrm{T}}\mathbf{X} \overset{d}{=} \|\mathbf{a}\| X_1$ *for all* $\mathbf{a} \in \mathbb{R}^n$.

Proof. If the random vector $\mathbf{X}$ has a spherical distribution, then by (10) there exists some function $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that $E\left[\exp(i\mathbf{a}^{\mathrm{T}}\mathbf{X})\right] = \phi(\|\mathbf{a}\|^2)$ for all $\mathbf{a} \in \mathbb{R}^n$. Hence for all $t \in \mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^n$ it follows by (12) that

$$E\left[\exp(it\mathbf{a}^{\mathrm{T}}\mathbf{X})\right] = \phi(\|t\mathbf{a}\|^2) = E\left[\exp(it\|\mathbf{a}\|X_1)\right].$$

By using the one to one correspondence between the characteristic function and the cumulative distribution function of the associated random variable (see Feller, 1971), we obtain $\mathbf{a}^{\mathrm{T}}\mathbf{X} \overset{d}{=} \|\mathbf{a}\| X_1$. To prove the reverse implication we observe that

$$E\left[\exp(it\mathbf{a}^{\mathrm{T}}\mathbf{X})\right] = E\left[\exp(i\|\mathbf{a}\|X_1)\right]. \tag{13}$$

This implies for all $\mathbf{a} \in \mathbb{R}^n$ that

$$E\left[\exp(-i\mathbf{a}^{\mathrm{T}}\mathbf{X})\right] = E\left[\exp(i\mathbf{a}^{\mathrm{T}}\mathbf{X})\right],$$

and so the function $\mathbf{a} \mapsto E\left[\exp(i\mathbf{a}^{\mathrm{T}}\mathbf{X})\right]$ is real-valued. Hence by (13), the function $\mathbf{a} \mapsto E\left[\exp(i\|\mathbf{a}\|X_1)\right]$ is also real-valued. Introducing $\phi : \mathbb{R}_+ \to \mathbb{R}$ given by $\phi(t) := E\left[\exp(i\sqrt{t}X_1)\right]$ it follows again by (13) that

$$E\left[\exp(i\mathbf{a}^{\mathrm{T}}\mathbf{X})\right] = \phi(\|\mathbf{a}\|^2).$$

It follows from the representation (10) that $\mathbf{X}$ has a spherical distribution. □

A class of distributions related to spherical distributions is given by the following definition (Embrechts et al., 2002; Fang, 1990).

Definition 4.2 *A random vector* $\mathbf{Y} = (Y_1, \cdots, Y_n)^{\mathrm{T}}$ *has an elliptical distribution if there exists an affine mapping* $\mathbf{x} \mapsto A\mathbf{x} + \mu$ *and a random vector* $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$ *having a spherical distribution such that* $\mathbf{Y} = A\mathbf{X} + \mu$.

For convenience, an elliptical distributed n-dimensional random vector $\mathbf{Y}$ is denoted by $(A, \mu, \mathbf{X})$, where $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$. It is now possible to show the following result.

Lemma 4.2 *If the random vector* $\mathbf{Y}$ *has an elliptical distribution with representation* $(A, \mu, \mathbf{X})$ *then*

$$\mathbf{x}^{\mathrm{T}}\mathbf{Y} \stackrel{d}{=} \|A^{\mathrm{T}}\mathbf{x}\| X_1 + \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}$$

for all portfolio vectors $\mathbf{x} \in \mathbb{R}^n$. *Moreover, the parameters of the spherical (marginal) distribution of the random variable* X_1 *are independent of* $\mathbf{x}$.

Proof. Since the elliptical distributed random vector $\mathbf{Y}$ has representation $(A, \mu, \mathbf{X})$ and $\mathbf{X}$ has a spherical distribution, it follows that

$$\mathbf{x}^{\mathrm{T}}\mathbf{Y} = \mathbf{x}^{\mathrm{T}}A\mathbf{X} + \mathbf{x}^{\mathrm{T}}\boldsymbol{\mu}.$$

Applying Lemma 4.1 with $\mathbf{a} = A^{\mathrm{T}}\mathbf{x}$ yields the desired result. □

To compute the covariance matrix Σ of the random vector $\mathbf{Y}$ having an elliptical distribution with representation $(A, \mu, \mathbf{X})$ we first observe that

$$\Sigma = Cov(\mathbf{Y}, \mathbf{Y}) = Cov(A\mathbf{X} + \mu, A\mathbf{X} + \mu) = ACov(\mathbf{X}, \mathbf{X})A^{\mathrm{T}}.$$

Moreover, since $\mathbf{X}$ has a spherical distribution it follows by (11) that there exists some nonnegative random variable R satisfying $\mathbf{X} \stackrel{d}{=} R\mathbf{U}^{(n)}$ and independent of $\mathbf{U}^{(n)}$. This implies that $Cov(\mathbf{X}, \mathbf{X}) = n^{-1}E\left[R^2\right]\mathbf{I}$ and hence with $c = \frac{E\left[\mathbf{R}^2\right]}{n} > 0$ we obtain that

$$\Sigma = cAA^{\mathrm{T}}. \tag{14}$$

Recall that for the vector of one period (short term) returns, $\mathbf{Y}$, the loss of the constructed portfolio is given by $-C\mathbf{x}^{\mathrm{T}}\mathbf{Y}$. Therefore, we need to specify and evaluate a risk measure associated with this random loss. One can now show the following important result for a random vector $\mathbf{Y}$ having an elliptical distribution.

Lemma 4.3 *If $\mathbf{Y}$ has an elliptical distribution with covariance matrix Σ and representation $(A,\mu,\mathbf{X})$ and the considered risk measure ρ is positive homogeneous, translation invariant and $\rho(X_1)>0$, then for any two nonzero portfolio vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ satisfying $\mathbf{x}_1^{\mathrm{T}}\boldsymbol{\mu}=\mathbf{x}_2^{\mathrm{T}}\boldsymbol{\mu}$, we have*

$$\rho(-\mathbf{x}_1^{\mathrm{T}}\mathbf{Y}) \le \rho(-\mathbf{x}_2^{\mathrm{T}}\mathbf{Y}) \Leftrightarrow \mathbf{x}_1^{\mathrm{T}}\Sigma\mathbf{x}_1 \le \mathbf{x}_2^{\mathrm{T}}\Sigma\mathbf{x}_2.$$

Proof. Since ρ is translation invariant and $\mathbf{x}_1^{\mathrm{T}}\boldsymbol{\mu}=\mathbf{x}_2^{\mathrm{T}}\boldsymbol{\mu}$ we obtain by Lemma 4.2 that

$$\rho(-\mathbf{x}_1^{\mathrm{T}}\mathbf{Y}) \le \rho(-\mathbf{x}_2^{\mathrm{T}}\mathbf{Y}) \Leftrightarrow \rho(\left\|A^{\mathrm{T}}\mathbf{x}_1\right\|X_1) \le \rho(\left\|A^{\mathrm{T}}\mathbf{x}_2\right\|X_1). \tag{15}$$

Then by the positive homogeneity of ρ and $\rho(X_1)>0$ we have

$$\rho(\left\|A^{\mathrm{T}}\mathbf{x}_1\right\|X_1) \le \rho(\left\|A^{\mathrm{T}}\mathbf{x}_2\right\|X_1) \Leftrightarrow \left\|A^{\mathrm{T}}\mathbf{x}_1\right\| \le \left\|A^{\mathrm{T}}\mathbf{x}_2\right\| \Leftrightarrow \mathbf{x}_1^{\mathrm{T}}AA^{\mathrm{T}}\mathbf{x}_1 \le \mathbf{x}_2AA^{\mathrm{T}}\mathbf{x}_2. \tag{16}$$

Relations (14), (15) and (16) and $c=\dfrac{E\left[\mathbf{R}^2\right]}{n}>0$ yield the desired result. □

As mentioned before, both CVaR and VaR satisfy the assumptions of Lemma 4.3. Therefore, for these important risk measures the portfolio optimization problem (Q) reduces to a mean-variance Markowitz model

$$\min\left\{\mathbf{x}^{\mathrm{T}}\Sigma\mathbf{x} \mid \mathbf{e}^{\mathrm{T}}\mathbf{x}=1,\ \mu^{\mathrm{T}}\mathbf{x}=r,\ \mathbf{x}\ge\mathbf{0}\right\}. \tag{MP-Q}$$

Both problems construct the same optimal portfolio. When $\alpha>\dfrac{1}{2}$ implying $\mathrm{VaR}_{\alpha}(X_1)>0$, it follows from Lemma 4.3 that for the portfolio loss $-C\mathbf{x}^{\mathrm{T}}\mathbf{Y}$ we obtain

$$\mathrm{VaR}_{\alpha}(-C\mathbf{x}^{\mathrm{T}}\mathbf{Y}) = C(\mathrm{VaR}_{\alpha}(-\mathbf{x}^{\mathrm{T}}\mathbf{Y})) = C(\|A^{\mathrm{T}}\mathbf{x}\|\,\mathrm{VaR}_{\alpha}(X_1)-\mathbf{x}^{\mathrm{T}}\mu) \tag{17}$$

and

$$\mathrm{CVaR}_{\alpha}(-C\mathbf{x}^{\mathrm{T}}\mathbf{Y}) = C(\|A^{\mathrm{T}}\mathbf{x}\|\,\mathrm{CVaR}_{\alpha}(X_1)-\mathbf{x}^{\mathrm{T}}\mu). \tag{18}$$

5 Modified Michelot Algorithm.

The algorithm introduced by Michelot (1986) finds in finite steps the projection of a given vector onto a special polytope. The main idea of this algorithm is to use the analytic solutions of a sequence of projections onto canonical simplices and elementary cones. The discussion in Michelot's paper is applicable when the objective function in problem (MP-Q) is perfect quadratic; that is, the covariance matrix Σ is the identity matrix. Unfortunately, the algorithm in Michelot's paper is not clear and difficult to follow. Our next step is to follow the main steps in Michelot's paper and apply necessary modification to solve

$$\min\left\{\mathbf{x}^{\mathrm{T}}\Sigma\mathbf{x} \mid \mathbf{e}^{\mathrm{T}}\mathbf{x}=1,\ \mu^{\mathrm{T}}\mathbf{x}=r,\ \mathbf{x}\geq\mathbf{0}\right\}. \tag{19}$$

To modify Michelot's algorithm according to our problem, we need to introduce several sets. Let

$$\mathcal{V} = \{\mathbf{x}\in\mathbb{R}^n \mid \mathbf{e}^T\mathbf{x}=1,\ \mu^T\mathbf{x}=r\},\quad \mathcal{X}_I=\left\{\mathbf{x}\in\mathbb{R}^n \mid x_i=0, i\in I\right\},\quad \text{and } \mathcal{V}_I=\mathcal{V}\cap\mathcal{X}_I,$$

where $I\subseteq\{1,2,\cdots,n\}$ denotes an index set. Algorithm 1 gives the steps of the Modified Michelot Algorithm. The algorithm starts with obtaining the optimal solution of the following quadratic programming problem

$$P_{\mathcal{V}} := \operatorname{argmin}\{\mathbf{x}^T\Sigma\mathbf{x} : \mathbf{x}\in\mathcal{V}\}. \tag{20}$$

Naturally, some of the components x_i may be negative; otherwise, the solution is optimal. After identifying the most negative component and initializing the index set I, the algorithm iterates between projections of the incumbent solution $\bar{\mathbf{x}}$ onto subspace $\mathcal{X}_I$, and then onto subspace $\mathcal{V}_I$ until none of the components are negative; i.e., the solution is optimal. The first projection is denoted by

$$P_{\mathcal{X}_I}(\bar{\mathbf{x}}) := \operatorname{argmin}\left\{(\bar{\mathbf{x}}-\mathbf{x})^T\Sigma(\bar{\mathbf{x}}-\mathbf{x}) : \mathbf{x}\in\mathcal{X}_I\right\}. \tag{21}$$

Similarly, the second projection is given by

$$P_{\mathcal{V}_I}(\bar{\mathbf{x}}) := \operatorname{argmin}\left\{(\bar{\mathbf{x}}-\mathbf{x})^T\Sigma(\bar{\mathbf{x}}-\mathbf{x}) : \mathbf{x}\in\mathcal{V}_I\right\}. \tag{22}$$

At every iteration one index is added to set I. Since we have a finite number of assets, the modified algorithm terminates within at most n iterations (see also Michelot, 1986).

Notice that all three problems, (20), (21) and (22), are equality constrained quadratic programming (QP) problems. Therefore, we consider a general equality constrained QP problem given by

$$\min\{(\bar{\mathbf{x}}-\mathbf{x})^{\mathrm{T}}\Sigma(\bar{\mathbf{x}}-\mathbf{x}) : T\mathbf{x}=\mathbf{b}\},$$

where T is an $m\times n$ matrix and $\mathbf{b}\in\mathbb{R}^m$. It is easy to show (Bertsekas, 1999) that this general problem has the optimal solution

$$\bar{\mathbf{x}}-\Sigma^{-1}T^{\mathrm{T}}\left(T\Sigma^{-1}T^{\mathrm{T}}\right)^{-1}(T\bar{\mathbf{x}}-b). \tag{23}$$

Algorithm 1: Modified Michelot Algorithm

```
1:  Input Σ, μ, r, I = ∅.
2:  Set x̄ ← P_V.
3:  if x̄ ≥ 0 then
4:      Stop; x̄ is optimal.
5:  else
6:      Select i with most negative x̄_i.
7:      Set I ← i.
8:      while x̄ < 0 do
9:          Set x̄ ← P_{X_I}(x̄).
10:         Set x̄ ← P_{V_I}(x̄).
11:         if x̄ ≥ 0 then
12:             Stop; x̄ is optimal.
13:         else
14:             Select i with most negative x̄_i.
15:             Set I ← I ∪ i.
16: Output: x̄
```

The matrix inversions in (23) constitute the main computational burden of Algorithm 1, since these inversions are required at every iteration. In line 2 of Algorithm 1, we need to find $P_{\mathcal{V}}$. To use relation (23), we set $T = [e \quad \mu]^{\mathrm{T}}$ and $\mathbf{b} = [1 \quad r]^{\mathrm{T}}$. These relations imply that we should compute the inverse of $n \times n$ matrix Σ as well as the inverse of 2×2 symmetric matrix

$$K := \begin{bmatrix} \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mathbf{e} & \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mu \\ \mu^{\mathrm{T}}\Sigma^{-1}\mathbf{e} & \mu^{\mathrm{T}}\Sigma^{-1}\mu \end{bmatrix}.$$

Luckily, the blockwise inverse method (Lancaster, 1885) allows us to complete Algorithm 1 by only these two matrix inversions because at every subsequent iteration, only one index is added to set I. For example, if we denote the i^{th} unit vector by $\mathbf{u}_i$, the first time the algorithm reaches line 10, we set $T = [e \quad \mu \quad \mathbf{u}_i]^{\mathrm{T}}$ and $\mathbf{b} = [1 \quad r \quad 0]^{\mathrm{T}}$ in relation (23). Therefore, we need to compute the inverse of a 3×3 symmetric matrix given by

$$\begin{bmatrix} K & \mathbf{v} \\ & \\ \mathbf{v}^{\mathrm{T}} & c_0 \end{bmatrix} := \begin{bmatrix} \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mathbf{e} & \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mu & \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mathbf{u}_i \\ \mu^{\mathrm{T}}\Sigma^{-1}\mathbf{e} & \mu^{\mathrm{T}}\Sigma^{-1}\mu & \mu^{\mathrm{T}}\Sigma^{-1}\mathbf{u}_i \\ \mathbf{e}^{\mathrm{T}}\Sigma^{-1}\mathbf{u}_i & \mu^{\mathrm{T}}\Sigma^{-1}\mathbf{u}_i & \mathbf{u}_i^{\mathrm{T}}\Sigma^{-1}\mathbf{u}_i \end{bmatrix}$$

Using now the blockwise inverse method yields

$$\begin{bmatrix} K & \mathbf{v} \\ \mathbf{v}^{\mathrm{T}} & c_0 \end{bmatrix}^{-1} = \begin{bmatrix} K^{-1} + c_1 K^{-1}\mathbf{v}\mathbf{v}^{\mathrm{T}}K^{-1} & -c_1 K^{-1}\mathbf{v} \\ -c_1\mathbf{v}^{\mathrm{T}}K^{-1} & c_1 \end{bmatrix},$$

where $c_1 := 1/(c_0 - \mathbf{v}^{\mathrm{T}}K^{-1}\mathbf{v})$. Since we already computed K^{-1}, the new inverse can be obtained without any matrix inversion. Moreover, the vector $\mathbf{v}$ and the scalar c_0 can be computed fast without any matrix multiplications, since the unit vector $\mathbf{u}_i$ is involved in their computations (for example, c_0 is simply the i^{th} diagonal component of Σ^{-1}). One can derive similar results for the projection in line 9 of Algorithm 1, since T grows again by one unit vector at each iteration and $\mathbf{b}$ is simply the zero vector.

6 Computational Results

To analyze the performance of Algorithm 1, **MATLAB** has been as our testing environment. All the computational experiments are conducted on a Intel(R) Core(TM) 2 2.00 GHz personal computer running Windows. First, we have randomly generated a set of test problems for different numbers of assets (n) as follows:

- The components of $n \times n$ matrix $\Sigma^{-1/2}$ are sampled uniformly from interval $(-2.5,5)$.
- The components of vector μ are sampled uniformly from interval $(0.01,0.50)$, and the first two components are sorted in ascending order; i.e., $\mu_1 \le \mu_2$.
- To ensure feasibility, the value r is then sampled uniformly from interval (μ_1, μ_2).
- For each value of n, 10 replications are generated.

Clearly, Problem (MP-Q) can be solved by any quadratic programming solver. In **MATLAB,** the procedure that solves these types of problems is called `quadprog`, which is also used in the financial toolbox. Therefore, to compare the proposed algorithm, we also solved the set of problems with `quadprog`. Table 1 shows the statistics of the computation times out of 10 replications. The second and third columns in Table 1 indicate averages and standard deviations of the computation times obtained by Algorithm 1, respectively. Similarly, columns four and five give the average and the standard deviation of the computation times found by `quadprog`, respectively.

	Algorithm 1		quadprog	
n	Average	Std. Dev.	Average	Std. Dev.
25	0.0030	0.0063	0.0173	0.0048
50	0.0111	0.0077	0.0451	0.0139
100	0.0548	0.0341	0.2783	0.0829
200	0.5672	0.4480	2.0345	0.5530
400	7.7626	6.4093	32.6268	10.6788
500	19.4047	10.3963	87.0015	23.8663
750	126.8378	92.6268	353.1186	71.8344
1000	378.5811	319.0916	1129.4000	251.4301

Table 1: Computation time statistics of quadprog and Algorithm 1 in seconds.

The average computational times in Table 1 demonstrate that Modified Michelot Algorithm is several times faster than is the **MATLAB** procedure `quadprog`. However, it is important to note that the **MATLAB** procedure `quadprog` involves many error checks that may also be the cause of higher computation times. The standard deviation figures in Table 1 do not yield a clear conclusion when we compare Algorithm 1 and `quadprog`. Nevertheless, Algorithm 1 still performs better than does `quadprog` in most of the problems. Overall, these results allow us to claim that Modified Michelot Algorithm is a fast and finite step alternative for solving (MP-Q).

As we presented in Section 5, the Modified Michelot Algorithm takes at most n steps. In Table 2, we report some summary statistics regarding the number of iterations required to solve the problem instances. These figures show that the number of iterations to solve a problem takes, on average, half of the problem dimension (n).

n	Average	Std. Dev.
25	12.1000	3.1429
50	24.0000	8.4853
100	47.4400	17.1995
200	95.4000	39.5058
400	189.5000	65.8707
500	247.8000	63.6375
750	413.1000	159.4926
1000	534.4000	236.9844

Table 2: Number of iterations statistics of quadprog and Algorithm 1.

An illustrative example explains the intuitive idea behind the optimal objective function values (17) and (18). We use the same portfolio optimization example given in Rockafellar and

Uryasev (2000) that involves three instruments. The rates of returns on these instruments have multivariate normal distribution, which simplifies the procedure to calculate the optimal objective function values. The mean return vector (in percentage terms) and the covariance matrix are given as

$$\mu^{\mathrm{T}} = (0.01001110, 0.0043532, 0.0137058)$$

and

$$\Sigma = \begin{bmatrix} 0.00324625 & 0.00022983 & 0.00420395 \\ 0.00022983 & 0.00049937 & 0.00019247 \\ 0.00420395 & 0.00019247 & 0.00764097 \end{bmatrix},$$

respectively. The expected return r is equal to 0.011. Assume that our budget C is 1000 at the beginning of the investment period. We first solve the portfolio problem (19) with Algorithm 1. We then use equations (17) and (18) to obtain the optimal VaR and CVaR values, respectively. Figure (1) shows these values against varying α. As expected, CVaR values are always greater VaR values.

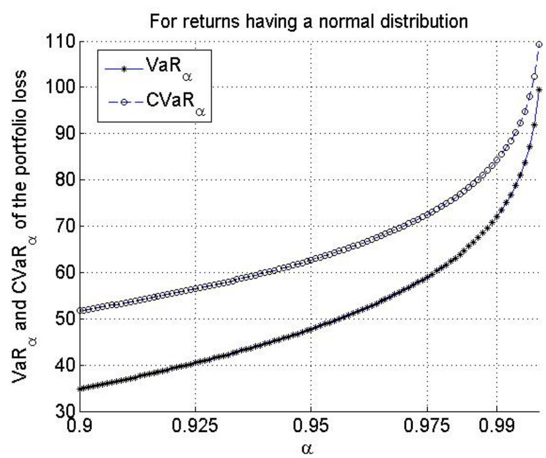
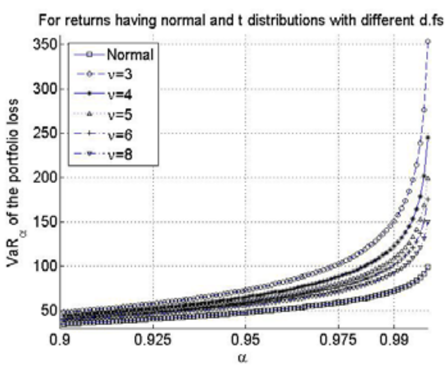


Figure 1: VaR and CVaR values for the elementary example.

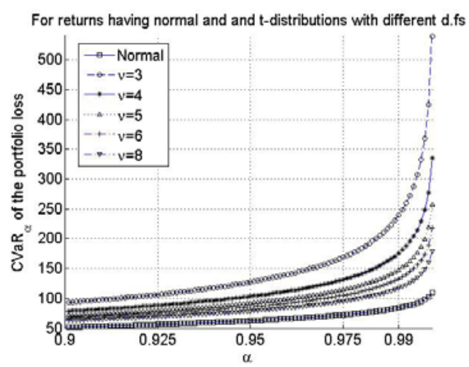
As mentioned in Section 4 the standard multivariate t-distribution with ν degrees of freedom (d.f.) belongs to the class of spherical distributions. Therefore, a random vector $\mathbf{X}$ having a standard multivariate t-distribution with ν degrees can be represented by $\mathbf{X} = S\mathbf{V}$, where $\mathbf{V} : N(\mathbf{0},\mathbf{I})$ and S is a real-valued random variable, independent of $\mathbf{V}$. For standard multivariate t-distribution with ν degrees of freedom, the random variable $\nu^{1/2}S$ has a chi-distribution χ_ν with ν degrees of freedom. According to Definition 4.2, we can obtain a (elliptically distributed) random vector Y with a (nonstandard) multivariate t-distribution by applying an affine mapping $\mathbf{x} \mapsto A\mathbf{x} + \mu$ on the (spherically distributed) random vector X with the standard multivariate t-distribution. Recall (17) and (18), where in our setup the random variable X_1 has a univariate t-distribution with ν degrees of freedom. We have used the **MATLAB** function `tinv` to calculate $\mathrm{VaR}_{\alpha}(X_1)$. Using (7) and the probability density function of X_1 we obtain

$$\mathrm{CVaR}_{\alpha}(X_1) = \frac{1}{1-\alpha} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu}{1-\nu}\left(-\left(1+\frac{\mathrm{VaR}_{\alpha}(X_1)^2}{\nu}\right)^{\frac{1-\nu}{2}}\right).$$

We also consider the same example given in Rockafellar and Uryasev (2000) and provide the optimal VaR and CVaR values of the total portfolio loss for a normal distribution and t-distributions with different degrees of freedom parameters. As mentioned at the end of Section 3, the widely accepted values of degrees of freedom parameter ν according to the literature are between 3 and 8.



a. VaR values



b. CVaR values

Figure 2: Risk values for returns having normal distribution and t-distribution with different degrees of freedom.

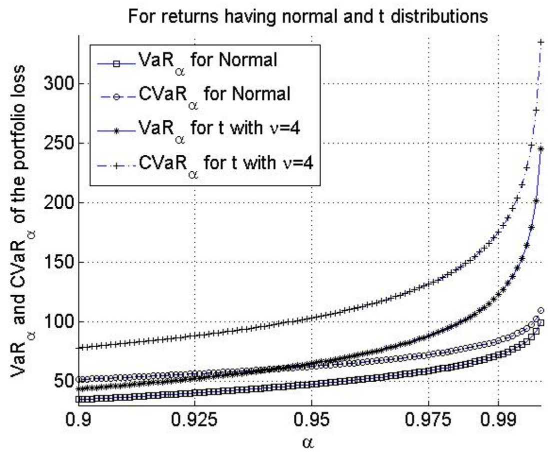


Figure 3: Comparison of VaR and CVaR values for different distributions.

Figure 2 shows VaR and CVaR values for different distributions. Clearly, as the degress of freedom parameter ν increases, the tail of t-distribution becomes less heavy and hence, approaches to the normal distribution. Therefore, we observe that the differences in VaR (Figure 2.a) and CVaR (Figure 2.b) values between t and the normal distributions diminish.

Figure 3 illustrates the difference between VaR and CVaR values for a normal distribution and a particular t-distribution ($\nu = 4$). As it can be seen from the figure, VaR and CVaR values are closer to each other for the normal distribution than the t-distribution. This is an expected observation, since a t-distribution has a heavier tail than a normal distribution.

7 Conclusion

In this paper we first discuss general risk measures and then concentrate on two recent ones, VaR and CVaR. Then we shift our focus to efficiently construct optimal portfolios, where the returns have elliptical distributions and either VaR or CVaR can be used as the risk measure. It is well-known that optimization problems, which are in the form of (Q) with VaR or CVaR as the risk measure, are equivalent to the mean-variance Markowitz model in the form of (MP-Q). In fact we discuss this equivalence holds for a larger class of positive homogeneous and translation invariant risk measures. To solve the resulting special quadratic programming problem, we modify a finite step algorithm from the literature and provide some computational results. To the best of our knowledge, portfolio management literature lacks numerical examples where the returns have distributions other than the normal distribution. Therefore, in addition to the numerical results for normal returns, we also provide results for returns that have multivariate t-distributions.

References

Artzner, P., F. Delbaen, J.M. Eber, and D. Heath. "Coherent Measures of Risk." *Mathematical Finance*, Vol. 9, No. 3 (1999), pp. 203-228.

Bertsekas, D.P. (1999) *Nonlinear Programming (2nd Edition)*. Belmont, CA: Athena Scientific.

Blattberg, R.C. and N.J. Gonedes. "A Comparison of the Stable and Student-t Distributions as Statistical Models for Stock Prices." *Journal of Business*, Vol. 47, No. 2 (1974), pp. 244-280.

Crouhy, M., D. Galai and R. Mark (2001) *Risk Management*. New York, NY: McGraw-Hill.

Embrechts, P., McNeil A. and Straumann, D. (2002) Correlation and Dependence in Risk Management: Properties and Pittfals. In: Dempster, M.A.H. (Ed.), Risk Value Managemant at Risk and Beyond. Cambridge: Cambridge University Press.

Fama, E.F. "The Behavior of Stock Market Prices." *Journal of Business*, Vol.38, No. 1 (1965), pp. 34-105.

Fang, K.T, S. Kotz and K.W. Ng. (1990) *Symmetric Multivariate and Related Distributions*. New York, NY: Chapman and Hall.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications, 2nd Edition*. New York, NY: Wiley.

Glasserman, P., P. Heidelberger and P. Shahabuddin. "Portfolio Value-at-Risk with Heavy Tailed Risk Factors." *Mathematical Finance*, Vol.12, No. 3 (2002), pp. 239-269.

Huisman, R., K.G. Koedijk, and R.A.G. Pownall. "VaR-x: Fat Tails in Financial Risk Management." *Journal of Risk*, Vol. 1, No. 1 (1998), pp. 47-62.

Lancaster P. and M. Tismenetsky. (1985) *The Theory of Matrices of Computer Science and Applied Mathematics*. California, CA: Academic Press.

Mandelbrot, B. "The Variations of Certain Speculative Prices." *Journal of Business*, Vol. 36, No. 4 (1963), pp. 394-419.

Mansini, R., W. Ogryczak and M.G. Speranza. "LP Solvable Models for Portfolio Optimization: A Classification and Computational Comparison." *IMA Journal of Management Mathematics*, Vol. 14, No. 3 (2003), pp. 187-220.

Markowitz, H. "Portfolio Selection." *Journal of Finance*, Vol. 7, No. 1 (1952), pp. 77-91.

Markowitz, H. M. (1959) *Portfolio Selection*. New York, NY: John Wiley & Sons.

Ogryczak, W and A. Ruszczyński. "Dual Stochastic Dominance and Related Mean-Risks Models." *SIAM Journal of Optimization*, Vol. 13, No. 2 (2002), pp. 60-78.

Pflug, G. (2000) Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In: Uryasev, S. (Ed.), Probabilistic Constrained Optimization: Methodology and Applications. Dordrecht: Kluwer Academic Publishers.

Praetz, J. "The Distribution of Share Price Changes." *Journal of Business*, Vol. 45, No. 1 (1972), pp. 49-55.

RiskMetrics Group (1997) *CreditMetrics Technical Document.* New York, NY: J. P. Morgan Inc. www.riskmetrics.com/research/techdocs.

Rockafellar, R.T. (1972) *Convex Analysis*. Princeton, NJ: Princeton University Press.

Rockafellar, R.T and S. Uryasev. "Conditional Value-at-Risk for General Loss Distributions." *Journal of Banking and Finance*, Vol. 26, No. 7 (2002), pp. 1443-1471.

Rockafellar, R.T, S. Uryasev and M. Zabarankin. "Master Funds in Portfolio Analysis with General Deviation Measures." *Journal of Banking and Finance*, Vol. 30, No. 2 (2006), pp. 743-778.

Rockafellar, R.T and S. Uryasev. "Optimization of Conditional Value-at-Risk." *The Journal of Risk*, Vol. 2, No. 3 (2000), pp. 21-41.

Ruszczyński, A. and Shapiro, A. (2006) Probabilistic and Randomized Methods for Design under Uncertainty. In: G. Calafiore and F. Dabbene (Ed.), Optimization of Risk Measures. London: Springer-Verlag.

Sadefo-Kamdem, J. "Value-at-Risk and Expected Shortfall for Linear Portfolios With Elliptically Distributed Risk Factors." *International Journal of Theoretical and Applied Finance*, Vol. 8, No. 5 (2005), pp. 537-551.

Steinbach, M.C. "Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis." *SIAM Review*, Vol. 43, No. 1 (2001), pp. 31-85.