**Robust Mean-Covariance Solutions
for Stochastic Optimization**

by

I. Popescu

2004/04/TM

Robust Mean-Covariance Solutions for Stochastic Optimization

Ioana Popescu [*]

December 2003

Abstract

We provide a method for deriving robust solutions to stochastic optimization problems with general objective, based on mean-covariance information about the distributions underlying the uncertain vector of returns. We prove that for a general class of objective functions, the robust solutions amount to solving a certain deterministic parametric quadratic program. Our approach relies on two key results. First, we prove a general projection property for sets of distributions with given means and covariances, which reduces the problem to a deterministic bi-criteria optimization problem. Second, we adapt a result from Geoffrion [21] to reduce this problem to solving a parametric quadratic program.

In particular we provide closed form solutions for special quantile-based criteria, such as probability and option-like objectives, value at risk and conditional value at risk. We investigate applications of our results in financial portfolio management, generalized regression and multi-product pricing. Finally, we investigate an extension of these results for the case of non-negative returns.

---

[*]INSEAD, Technology Management Area, Blvd. de Constance Fontainebleau 77300, France, ioana.popescu@insead.edu

1 Introduction

In this paper we provide a general approach for deriving robust solutions for stochastic optimization programs, based on mean-covariance information about the distributions underlying the return vector.

Consider a decision maker interested in maximizing his expected utility u of a linear reward function, subject to a given set of convex constraints $\mathcal{P}$ on the vector of decision variables x. The decision maker solves the following general stochastic mathematical program without recourse:

$$
\begin{aligned}
(P) \quad \text{maximize} \quad & E[u(x'\mathbf{R})] \\
\text{subject to} \quad & x \in \mathcal{P},
\end{aligned}
$$

Random variables and random vectors will be denoted in bold, to be distinguished from deterministic quantities. Throughout the paper, we use max and min in the wide sense of sup and inf, to account for the cases when the optimum is not achieved.

When the distribution of the uncertain return vector $\mathbf{R}$ is exactly known, Problem (P) is a general one stage stochastic program. One obvious case when the solution of Problem (P) does not depend on the actual distribution of $\mathbf{R}$, except its mean and covariance, is for quadratic utilities $u(r) = ar^2 + br + c$, in which case it amounts to: $\max_{x \in \mathcal{P}} ax'(\Sigma + \mu\mu')x + bx'\mu + c$. However, this is only the case for quadratic (or linear) utility functions, explaining in part why these classes are particularly favored in economics and financial models.

In many applications of stochastic programming, there is considerable uncertainty regarding the distribution of $\mathbf{R}$. In many instances, only partial information about this distribution is available, or reliable, such as means, variances and covariances. A common approach in the literature is to assume that the underlying distribution is multivariate normal, and solve the corresponding stochastic program. However, this assumption is hard to justify unless the randomness comes from a central limit theorem type of process.

Alternatively, the *robust approach* consists in basing the decision strictly on the available information, and maximizing the worst case expected utility over all possible distributions with the given properties. Even if the underlying distribution is actually known, the robust approach can be used to provide tractable approximations for the solution of Problem (P), which is typically a very difficult problem. The robust problem is the focus of our

1

investigation, and can be formulated as follows:

$$(G) \quad \max_{x \in \mathcal{P}} \min_{\mathbf{R} \sim (\mu, \Sigma)} E[u(x'\mathbf{R})], \tag{1}$$

where throughout the paper we denote $\mathbf{R} \sim (\mu, \Sigma)$ to represent the fact that the random vector $\mathbf{R}$ has mean μ and covariance Σ. One can think about Problem (G) as a strategic game, where the first player chooses a decision x while nature picks the worst possible distribution $\mathbf{R}$ (see for example [4] and references therein).

As an example, consider a portfolio manager interested in choosing a portfolio x among n assets with returns $\mathbf{R} = (\mathbf{R}_1, \ldots, \mathbf{R}_n)$ in order to maximize his expected utility, or payoff, subject to a set of portfolio constraints, such as $\mathcal{P} = \{x \in [0,1]^n \mid x'1_n = 1\}$. In practice, one seldom has full information about the joint distribution of asset returns, which are typically assumed multi-variate normal and interpolated from very few data points (hence are not robust). Very often the only reliable data consists of first and second order moments. For example, RiskMetrics (www.riskmetrics.com) provides very large data set of asset covariances, but no additional distributional information. In this case, if the manager has access only to the mean and covariance matrix of the asset returns, and not to the entire distribution, he may be interested in the best and worst case return of a given portfolio, and in the portfolios that would maximize such a worst case return. There is a considerable amount of literature on robust portfolio optimization (see for example [34, 25, 46, 41, 23]). We contribute to this stream by studying the porfolio example in detail in Section 4.1. While the techniques developed in this paper are quite general, we also explore applications in statistics and pricing.

The robust approach adopted in this paper is very much in the spirit of the max-min expected utility framework axiomatized by Gilboa and Schmeidler [24]. The concept of robustness, or incomplete knowledge about distributions, is captured in various ways in the stochastic optimization literature, the most common ones being uncertainty sets [6] and moments of the underlying distribution. Our model falls in the second category, and is related in that sense to Dupačová [18], Gaivoronski [20], Shapiro and Kleywegt [43], Birge and Maddox [10] etc. While our model has no recourse, there is remarkable literature on max-min stochastic programs with recourse (see [17, 16, 27, 28, 11] and references therein).

An alternative to the robust approach is provided by Bayesian analysis, whereby a *prior* distribution is defined and updated based on additional knowledge. A hybrid approach that

is widely used in statistical decision analysis is Bayesian robustness, see for example Berger [7] and references therein.

Contributions and Structure. There are two main steps to our solution to Problem (G). First, we transform the inner optimization into an equivalent optimization problem over univariate distributions with a given mean and variance. This reduces the original problem to a deterministic bi-criteria optimization problem. Second, we reduce this bi-criteria problem to solving a parametric quadratic problem.

For a given vector x, we refer to the problem

$$(I) \quad U(x) \;=\; \min_{\mathbf{R}\sim(\mu,\Sigma)} E[u(x'\mathbf{R})] \tag{2}$$

as the inner optimization problem and to the function $U(x)$ as the robust objective (or criterion). In Section 2, we show that for a given vector x, we can find sharp upper and lower bounds on the robust objective $U(x)$ by solving a simple one-dimensional optimization problem. This type of sensitivity analysis is of interest in many decision analysis applications, see for example Smith [44].

The first key result of our derivations is a new projection property for classes of multivariate distributions $\mathbf{R}$ with given mean vector μ and covariance matrix Σ. Based on this result, we show in Proposition 1 that for any vector x, Problem (I) reduces to an optimization problem over the class of "projected" univariate distributions $\mathbf{r}_x$ with given mean $\mu_x = x'\mu$ and variance $\sigma_x^2 = x'\Sigma x$, that is:

$$U(x) \;=\; \min_{\mathbf{R}\sim(\mu,\Sigma)} E[u(x'\mathbf{R})] \;=\; \min_{\mathbf{r}_x\sim(\mu_x,\sigma_x^2)} E[u(\mathbf{r}_x)].$$

We show in Proposition 2 that the above problem can be further reduced to optimizing over a restricted class of two-point distributions, which allows to compute the robust objective $U(x)$ as a deterministic optimization problem in a single variable. This procedure reduces the robust stochastic optimization problem (G) to a deterministic bi-criteria problem that only depends on the mean and variance of the random variable $x'\mathbf{R}$.

Our first set of result relies on the classical theory of moments (see for example [29, 26, 1, 31, 39]), and does not require any assumption on the objective function u. Sharper moment type bounds may be obtained in some particular cases, by exploiting functional

3

properties of u, such as Jensen's or Edmundson-Mandanski's inequalities (see for example [15, 19, 13] and references therein).

In Section 3, we use results from Geoffrion [21] to show (Proposition 3) that for very general classes of objective functions the robust Problem (G) is equivalent to solving the following parametric quadratic program (PQP):

$$(P_\lambda) \qquad \max_{x \in \mathcal{P}} \lambda \, x'\mu - (1 - \lambda) \, x'\Sigma x. \tag{3}$$

In particular, this is true for an increasing and concave utility u. Although our model and results are motivated by utility maximization, we do not require any a priori assumptions on the objective function u. In particular, by a simple change of variable $u = -u$ we obtain the min-max version of Problem (G). All the results in this paper hold for both types of problems, but for simplicity we restrict the analysis to the max-min case.

In Section 4 we illustrate two applications: financial portfolio management and generalized regression. These applications motivate the solution of Problem (G) for special quantile-based objectives such as targets (yield, or probability) or options (stop-loss), and non-expectation quantile criteria such as value at risk (VaR) and conditional value at risk (CVaR). We show that these can be cast as deterministic bi-criteria optimization problems and solved efficiently as PQPs. We discover an unexpected equivalence between the robust VaR and CVaR criteria. Also, it turns out that the robust target criterion has the same optimal solution as its multivariate normal counterpart.

In Section 5 we propose two extensions of our results motivated by an application in multi-product pricing. First, we show that our general model and solution naturally extend to the case when the distribution $\mathbf{R}$ (its mean and covariance) is dependent on the decision variable x. Second, we attempt to extend our framework to the case of non-negative random returns. While we prove that bi-criteria solutions are generally not optimal, we provide a method to compute (weak) bi-criteria bounds on the corresponding robust problems, and show how they can be solved using parametric quadratic programming. Our conclusions are summarized in the last section.

2   The Robust Objective

In this section we focus on Problem (I): for a non-zero vector x, we consider the problem $U(x) = \min_{\mathbf{R} \sim (\mu, \Sigma)} E[u(x'\mathbf{R})]$. A lower bound on the robust objective $U(x)$ can be obtained by optimizing over univariate distributions with mean μ_x and variance σ_x^2:

$$U(x) = \min_{\mathbf{R} \sim (\mu, \Sigma)} E[u(x'\mathbf{R})] \geq \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[u(\mathbf{r})] \ . \tag{4}$$

This is because for any feasible random vector $\mathbf{R} \sim (\mu, \Sigma)$ on the left hand side, the corresponding projected random variable $\mathbf{r}_x = x'\mathbf{R}$ has mean $x'\mu = \mu_x$ and variance $x'\Sigma x = \sigma_x^2$, hence $\mathbf{r}_x$ is feasible on the right hand side.

In Proposition 1 below we show that, irrespective of the objective function u, the above result holds with equality. The central idea behind Proposition 1 is that for any non-zero vector x, the optimization over random vectors $\mathbf{R} \sim (\mu, \Sigma)$ in Problem (I) is equivalent to solving the corresponding "projected" problem over the class of univariate distributions with mean μ_x and variance σ_x^2. We provide a constructive proof for the following stronger result: for any non-zero vector x and any distribution $\mathbf{r} \sim (\mu_x, \sigma_x^2)$, there exists a multivariate distribution $\mathbf{R} \sim (\mu, \Sigma)$ such that $\mathbf{r} = x'\mathbf{R}$. Let us denote $\mathcal{M}_{(\mu, \Sigma)}^n$ the set of probability measures on R^n with mean μ and covariance Σ, i.e. the set of possible underlying distributions for $\mathbf{R} \sim (\mu, \Sigma)$ (the superscript $n = 1$ will be omitted by default).

Theorem 1 (General Projection Property for Real-Valued Distributions) *For any full rank $n \times k$ matrix A, the A-projection mapping $\mathbf{r} = A'\mathbf{R}$ from n-variate distributions $\mathbf{R} \sim (\mu, \Sigma)$ to k-variate distributions $\mathbf{r} \sim (\mu_A, \Sigma_A)$ is onto, meaning that every distribution $\mathbf{r}$ in $\mathcal{M}_{(\mu_A, \Sigma_A)}^k$ can be generated by some distribution $\mathbf{R}$ in $\mathcal{M}_{(\mu, \Sigma)}^n$ by $\mathbf{r} = A'\mathbf{R}$.*

Proof:

Consider $\mathbf{z}$ and $\mathbf{Z}$ to be the standardized versions of $\mathbf{r}$ and $\mathbf{R}$, respectively:

$$\begin{aligned} \mathbf{R} &= \mu + \Sigma^{\frac{1}{2}}\mathbf{Z} & , \ \mathbf{Z} \sim (0, I_n) \\ \mathbf{r} &= A'\mu + (A'\Sigma A)^{\frac{1}{2}}\mathbf{z} & , \ \mathbf{z} \sim (0, I_k). \end{aligned}$$

It follows that $\mathbf{z} = B'\mathbf{Z}$, where $B' = (A'\Sigma A)^{-\frac{1}{2}}A'\Sigma^{\frac{1}{2}}$ is a rank k matrix.

We may assume without loss of generality that $B' = (B'_{(k)} B'_{(n-k)})$, where $B_{(k)}$ is a full rank $(k \times k)$ matrix. Notice that $B'B = I_k$. It follows that

$$B'_{(k)} B_{(k)} + B'_{(n-k)} B_{(n-k)} = I_k. \tag{5}$$

Similarly, we write $\mathbf{Z}' = (\mathbf{Z}'_{(k)} \mathbf{Z}'_{(n-k)})$, where $\mathbf{Z}_{(k)}$ and $\mathbf{Z}_{(n-k)}$ are k, respectively $(n-k)$ dimensional vectors. We obtain:

$$\mathbf{z} = B'_{(k)} \mathbf{Z}_{(k)} + B'_{(n-k)} \mathbf{Z}_{(n-k)}. \tag{6}$$

We first construct $\mathbf{Z}_{(n-k)}$ conditionally on $\mathbf{z}$ as follows: for any given value $\hat{z}$ of $\mathbf{z}$, define $\mathbf{Z}_{(n-k)}(\hat{z}) = (\mathbf{Z}_{(n-k)} | \mathbf{z} = \hat{z})$ to follow any specified distribution with mean $B_{(n-k)}\hat{z}$ and second moment matrix I_{n-k}. For example, this could be chosen to be a normal distribution: $\mathbf{Z}_{(n-k)}(\hat{z}) \sim N(B_{(n-k)}\hat{z}, I_{n-k})$. The following relations hold:

$$
\begin{aligned}
E[\mathbf{Z}_{(n-k)}] &= E_{\mathbf{z}}[E[\mathbf{Z}_{(n-k)} | \mathbf{z} = \hat{z}]] &= E[B'_{(n-k)}\mathbf{z}] &= 0 \\
E[\mathbf{Z}_{(n-k)}\mathbf{Z}'_{(n-k)}] &= E_{\mathbf{z}}[E[\mathbf{Z}_{(n-k)}\mathbf{Z}'_{(n-k)} | \mathbf{z} = \hat{z}]] &= E_{\mathbf{z}}[I_{n-k}] &= I_{n-k}
\end{aligned}
$$

We now define $\mathbf{Z}_{(k)}$ conditional on $\mathbf{z} = \hat{z}$, so as to satisfy Equation (6):

$$\mathbf{Z}_{(k)}(\hat{z}) = (B'_{(k)})^{-1}(\hat{z} - B'_{(n-k)}\mathbf{Z}_{(n-k)}(\hat{z})).$$

Because $\mathbf{Z}_{(n-k)}(\hat{z}) \sim (B_{(n-k)}\hat{z}, I_{n-k})$, we obtain that

$$E[\mathbf{Z}_{(k)}] \quad = \quad E_{\mathbf{z}}[(B'_{(k)})^{-1}(I_k - B'_{(n-k)}B_{(n-k)})\,\mathbf{z}] = 0$$

$$
\begin{aligned}
E[\mathbf{Z}_{(k)}\mathbf{Z}'_{(n-k)}] &= E_{\mathbf{z}}[E[(B'_{(k)})^{-1}(\mathbf{z} - B'_{(n-k)}\mathbf{Z}_{(n-k)})\mathbf{Z}'_{(n-k)} | \mathbf{z} = \hat{z}]] \\
&= (B'_{(k)})^{-1} E_{\mathbf{z}}[(\mathbf{z}\mathbf{z}'B'_{(n-k)} - B'_{(n-k)}I_{n-k})] = 0
\end{aligned}
$$

Using equation (5), we also obtain that

$$
\begin{aligned}
E[\mathbf{Z}_{(k)}\mathbf{Z}'_{(k)}] &= (B'_{(k)})^{-1} E_{\mathbf{z}}[E[(\mathbf{z} - B'_{(n-k)}\mathbf{Z}_{(n-k)})(\mathbf{z} - B'_{(n-k)}\mathbf{Z}_{(n-k)})' | \mathbf{z} = \hat{z}]](B_{(k)})^{-1} \\
&= (B'_{(k)})^{-1} E_{\mathbf{z}}[\mathbf{z}\mathbf{z}' - 2\mathbf{z}\mathbf{z}'B'_{(n-k)}B_{(n-k)} + B'_{(n-k)}B_{(n-k)}](B_{(k)})^{-1} \\
&= (B'_{(k)})^{-1}[I_k - B'_{(n-k)}B_{(n-k)}](B_{(k)})^{-1} = I_k.
\end{aligned}
$$

This shows that the random vector $\mathbf{R} = \mu + \Sigma^{\frac{1}{2}}\mathbf{Z}$ where $\mathbf{Z} = (\mathbf{Z}'_{(k)}\mathbf{Z}'_{(n-k)})' \sim (0, I_n)$ is defined as above satisfies the desired properties. $\square$

The one dimensional projection result $(k = 1)$ of Theorem 1 implies the following result:

Proposition 1 *For any real vector x we have the following result:*

$$\min_{\mathbf{R}\sim(\mu,\Sigma)} E[u(x'\mathbf{R})] = \min_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} E[u(\mathbf{r})]. \tag{7}$$

Based on Proposition 1, the general robust stochastic optimization problem (G) amounts to maximizing $U(x) = \min_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} E[u(\mathbf{r})]$, over the domain $x \in \mathcal{P}$. The next result shows how the robust criterion $U(x)$ further simplifies to a deterministic optimization problem in one variable.

Proposition 2 *For any vector x, the robust objective $U(x) = \min_{\mathbf{R}\sim(\mu,\Sigma)} E[u(x'R)]$ equals:*

$$U(x) = \min_{p\in(0,1)} p \cdot u(\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x) + (1-p) \cdot u(\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x). \tag{8}$$

Proof:

Since expectation is a linear operator, optimizing $E[u(\mathbf{r})]$ over $\mathbf{r} \sim (\mu_x, \sigma_x^2)$ is equivalent to optimizing over the extreme points of the set $\mathcal{M}_{(\mu_x,\sigma_x^2)}$ of feasible measures. These are two point distribution with mean μ_x and standard deviation σ_x, which is a consequence of Caratheodory's theorem (see e.g. Rogosinski [39] for a detailed proof). Such extremal two-point distributions assign mass p to the point $\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x$ and mass $1-p$ to the point $\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x$. By optimizing the expected utility over all such distributions we obtain the desired result. $\square$

Example 1. Let $u(x) = x - \ln(1 + x^2)$. By Propositions 1 and 2, we have that $U(x) = \min_{p\in(0,1)} U(p, x)$, where

$$U(p, x) = \mu_x - p \cdot \ln(1 + (\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x)^2) - (1-p) \cdot \ln(1 + (\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x)^2).$$

Remark that $\lim\limits_{p\to 0} U(p,x) = \lim\limits_{p\to 1} U(p,x) = u(\mu_x) = \mu_x - \ln(1+\mu_x^2)$, and $U(x,p)$ is decreasing in p around $p = 0$ and increasing around $p = 1$, so the minimum is achieved somewhere inbetween.

Suppose, for example that for a certain $x = x_0$ we have $\mu_x = 0$ and $\sigma_x = 1$. Then $U(p,x_0) = p\ln p + (1-p)\ln(1-p)$ which is minimized for $p = 1/2$ and $U(x_0) = -\ln 2$. While it is not possible to plot $U(x)$ for high dimensional vectors x, we remark that U only depends on x through the two variables μ_x and σ_x. In Figure 1 we plot the bivariate function $V(\mu_x, \sigma_x) = U(x)$.
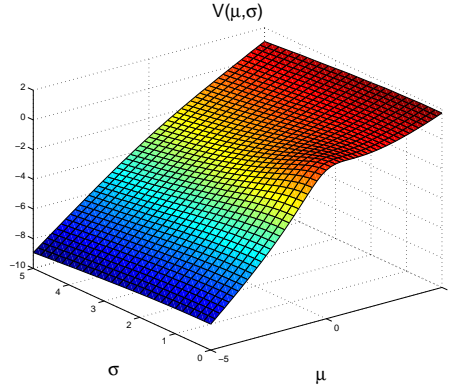

Figure 1: $V(\mu, \sigma)$ for $\mu \in [-5, 5], \sigma \in [0, 5]$.

Example 2.   Consider an exponential utility model $u(x) = 1 - e^{-ax}$, with $a > 0$. By Propositions 1 and 2, we have that $U(x) = \min\limits_{p\in(0,1)} U(p,x)$, where

$$U(p,x) = 1 - p \cdot e^{-a(\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x)} - (1-p) \cdot e^{-a(\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x)}.$$

Furthermore, one can easily calculate:
$$\frac{\partial U}{\partial p}(p,x) = -e^{-a(\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x)} + e^{-a(\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x)} + a[e^{-a(\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x)} + e^{-a(\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x)}]\frac{\sigma_x}{2\sqrt{p(1-p)}}.$$

Note that this has the same sign as $g(r) = 1 - r - (1+r)e^{-2r}$, where we denoted $r = \frac{a\sigma_x}{2\sqrt{p(1-p)}}$. Moreover $g(r)$ is 0-unimodal with $g(0) = 0$. Since $p \in (0,1)$, it follows that $U(x) = \min(\lim\limits_{p\to 0} U(p,x), \lim\limits_{p\to 1} U(p,x))$. We have $\lim\limits_{p\to 0} U(p,x) = u(\mu_x) = 1 - e^{-a\mu_x}$, whereas $\lim\limits_{p\to 1} U(p,x) = u(\mu_x) + \sigma_x^2 \lim\limits_{q\to\infty} u''(\mu_x - q\sigma_x)$, which equals $-\infty$ if $\sigma_x > 0$ and $u(\mu_x)$ otherwise. Hence $U(x) = -\infty$ for $\sigma_x > 0$ and $U(x) = 1 - e^{-a\mu_x}$ otherwise. So for exponential utilities, the robust solution should focus on setting $\sigma_x = 0$.

Further examples are illustrated in the context of several applications in Section 4.

8

3 Bi-criteria Optimization via PQP

Proposition 2 above shows that the inner optimization Problem (I) can be viewed as a bi-criteria optimization problem in μ_x and σ_x, $U(x) \equiv V(\mu_x, \sigma_x)$. In this section we propose a parametric quadratic programming (PQP) approach to solve this type of bi-criteria problems. Based on Geoffrion [21], we provide sufficient conditions on the objective function u under which Problem (G) can be reduced to solving a parametric quadratic program. Our results show that optimizing the robust counterpart of a very general class of objectives is equivalent to a one-dimensional mean-variance optimization problem, for a certain trade-off coefficient.

The following result due to Geoffrion [21] (see also [22]) provides general sufficient conditions for solving bi-criteria problems efficiently as parametric quadratic programs:

Theorem 2 (Geoffrion [21]) *Consider the program*

$$\max_{x \in X} v(p_1(f_1(x)), p_2(f_2(x))), \tag{9}$$

where $X \in \mathbf{R}^n$ is a convex set. Suppose that $f_1, f_2, p_1(f_1), p_2(f_2)$ are concave on X, p_1 and p_2 are increasing on the image of X under f_1, f_2 respectively, and v is non-decreasing and quasi-concave[1] on the convex hull of the image of X under $(p_1(f_1), p_2(f_2))$, and all functions are continuous. Further assume that the optimal solution $x^*(\lambda)$ of the parametric program $\max_{x \in X} \lambda f_1(x) + (1-\lambda) f_2(x)$ is continuous for $\lambda \in [0, 1]$. Then the function $w(\lambda) = v(p_1(f_1(x^*(\lambda))), p_2(f_2(x^*(\lambda))))$ is continuous and unimodal on $[0, 1]$. Furthermore, if λ^* is its maximum, then $x^*(\lambda^*)$ is optimal for (9).

Propositions 1 and 2 reduced Problem (G) to solving a bi-criteria (mean-variance) deterministic optimization problem. Based on Theorem 2, the following Proposition shows that under fairly unrestrictive conditions, Problem (G) can be optimized by solving a parametric quadratic program.

Proposition 3 *Suppose that the robust objective function $V(\mu_x, \sigma_x) = \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[u(\mathbf{r})]$ is continuous, non-decreasing in μ_x, non-increasing in σ_x and quasi-concave. Then Problem*

---

[1]A real function f defined on a convex set S is said to be quasi-concave if and only if for any $\lambda \in [0, 1]$ and $x, y \in S$ we have $f(\lambda x + (1-\lambda)y) \geq \min(f(x), f(y))$. Equivalently, f is quasi-concave whenever for any l, the upper level sets $L_l = \{x \in S \mid f(x) \geq l\}$ are convex.

(G) over a convex set $\mathcal{P}$ is equivalent to solving the following parametric quadratic program:

$$(P_\lambda) \qquad \max_{x \in \mathcal{P}} \lambda\, x'\mu - (1-\lambda)\, x'\Sigma x. \qquad (10)$$

Specifically, if $x^*(\lambda)$ denotes the optimal solution of Problem (P_λ), then the function $w(\lambda) = V(\mu_{x^*(\lambda)}, \sigma_{x^*(\lambda)})$ is continuous and unimodal in $\lambda \in [0,1]$. Moreover, if λ^* is its maximum, then $x^*(\lambda^*)$ is optimal for Problem (G).

Proof:

Let $f_1(x) = \mu'x$ and $f_2(x) = -x'\Sigma x$, $p_1(f_1) = f_1$ and $p_2(f_2) = -(-f_2)^{\frac{1}{2}}$, these are all concave, and the latter two are non-decreasing. In the notation of Theorem 2, we obtain $v(\mu_x, -\sigma_x) = v(p_1(f_1(x)), p_2(f_2(x))) = \min\limits_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[u(\mathbf{r})] = V(\mu_x, \sigma_x)$. Based on Proposition 2, we can rewrite this as:

$$v(\mu_x, -\sigma_x) = V(\mu_x, \sigma_x) = \min_{p \in [0,1]} p \cdot u\left(\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x\right) + (1-p) \cdot u\left(\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x\right). \quad (11)$$

The assumptions on V insure that v is non-decreasing and quasi-concave, so the result follows from Theorem 2.

$\square$

It is reasonable to assume that the robust utility $U(x) = V(\mu_x, \sigma_x)$ is non-decreasing in the mean μ_x and non-increasing in the standard deviation σ_x of the underlying return. This basically says that the decision maker prefers distributions with higher return and lower risk. Quasi-concavity can be interpreted as a diminishing marginal rate of substitution between mean and standard deviation, a property shared by most (robust) utility functions. When the actual utility function u is increasing and concave, as is typically assumed in decision analysis, these conditions are automatically satisfied:

Corollary 1 *Suppose that the function u is non-decreasing and concave. Then the general utility maximization criterion (G) over a convex set $\mathcal{P}$ is equivalent to solving the parametric quadratic program (P_λ) for a suitable value $\lambda \in [0,1]$.*

The proof is straightforward based on the formulation from Proposition 2. More generally, for other types of u, monotonicity of V in μ_x can be ensured by assuming that u is non-decreasing. One can check monotonicity in σ_x as well as quasi-concavity using the

result of Proposition 2. When u is concave then $U(x) = V(\mu_x, \sigma_x)$ is concave in x, so alternatively, other traditional computational methods for convex programming may be suitable to solve Problem (G) more directly, based on the result of Proposition 2.

Several algorithms are available for solving parametric quadratic programs (Wolfe [45], Geoffrion [21, 22], Best [9]). They all yield an optimal solution $x^*(\lambda)$ that is piecewise linear in $\lambda/(1-\lambda)$. Geoffrion [21] solves P_λ with $\lambda = 1$ and decreases λ until the unimodal function $w^*(\lambda) = v(\mu_{x^*(\lambda)}, -\sigma_{x^*(\lambda)})$ achieves its maximum on $[0, 1]$. In our applications, the intermediate solutions are also of interest, as they yield a portion of the mean-variance trade-off curve. Recent algorithms for PQP are due to Best [9] and Arseneau and Best [2].

Example 3. In Problem (G), suppose that the constraint set is given by $\mathcal{P} = \{x | \sum x_i = 1, x \geq 0\}$, and let $u(x) = x - \ln(1 + x^2)$ as in Example 1. Since u is increasing and concave, it follows that

$$V(\mu_x, \sigma_x) = \min_{p \in (0,1)} \mu_x - p \cdot \ln(1 + (\mu_x + \sqrt{\frac{1-p}{p}}\sigma_x)^2) - (1-p) \cdot \ln(1 + (\mu_x - \sqrt{\frac{p}{1-p}}\sigma_x)^2)$$

satisfies the conditions of Proposition 3. This can also be seen by inspecting Figure 1 in the previous section. To solve Problem (G), we first compute $x^*(\lambda)$, the optimal solution of the PQP:

$$
\begin{aligned}
&\max && \lambda\, x'\mu - (1-\lambda)\, x'\Sigma x \\
(P_\lambda) \quad &\text{s.t.} && \sum x_i = 1 \\
& && x \geq 0,
\end{aligned}
$$

and then maximize over $\lambda \in (0, 1)$ the function $w(\lambda) = V(\mu_{x^*(\lambda)}, \sigma_{x^*(\lambda)})$, which is unimodal.

This can be seen from Figure 2, where we display $w(\lambda)$ for the following three sets of data:

$\mu_1 = (1\ 3)'$, $\Sigma_1 = \begin{pmatrix} 1 & -2 \\ -2 & 6 \end{pmatrix}$, $\mu_2 = (0\ 2\ 1\ -2)'$, $\Sigma_2 = \begin{pmatrix} 2 & -2 & 1 & 0 \\ -2 & 9 & -1 & -2 \\ 1 & -1 & 3 & -1 \\ 0 & -2 & -1 & 2 \end{pmatrix}$, respec-

tively $\mu_3 = -\mu_2, \Sigma_3 = \Sigma_2$. The resulting optimal values $w(\lambda^*)$ are the desired optimal values of (G), in our case 0.5491, .1512 and 0.4260, achieved at $\lambda^* = 0.84, 0.67, 1$, respectively. The corresponding solutions of Problem (G) are $x^*(\lambda^*) = (0.25\ 0.75)'$, $(0.2140\ 0.3429\ 0.4431\ 0.00)'$ and $(0, 0, 0, 1)'$. All programs were run in MATLAB 5.3 under Windows NT.
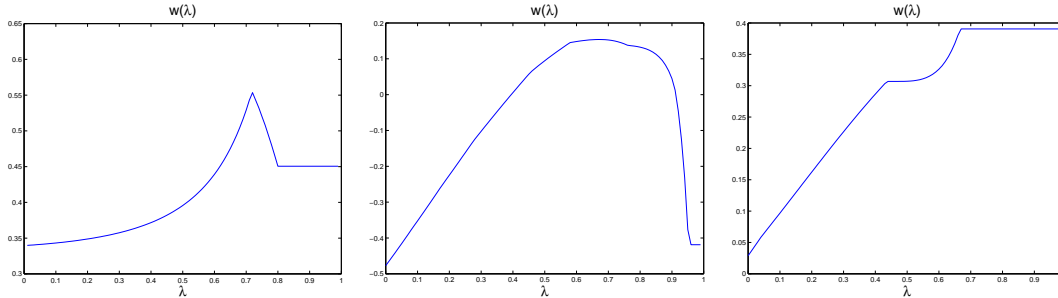
11

Figure 2: Plots of the unimodal function $w(\lambda)$.

4 Applications in Finance and Statistics

In this section we present two applications of our results, one in financial portfolio management, the other in generalized regression. These applications prompt us to analyse a set of special quantile-based criteria, such as probability (target), option-type, VaR (fractile) and CVaR (tail expectation).

The first application considers a risk-sensitive portfolio manager optimizing the expected reward from his real-valued return. In particular we investigate target and commission based reward systems, and combinations thereof.

The second application provides a robust approach of fitting a multiple linear regression among a set of variables, given limited information about their distributions (means and covariances). In a robust setting, we look at general deviation measures (see Rockafellar, Uryasev and Zabarankin [36]) such as VaR and CVaR.

4.1 Financial Portfolio Management and Incentives

Consider the example described in the introduction: a portfolio manager is interested in a robust portfolio among a set of n assets with random returns $\mathbf{R}_1, ..., \mathbf{R}_n$. Suppose that the manager does not know the distribution of the random returns, but estimates them to have means μ and covariance structure Σ. The manager is thus interested in a robust portfolio that maximizes his worst case expected utility of the return $u(\mathbf{r})$, where $\mathbf{r}$ is the achieved return of herhis portfolio at the end of the evaluation period. This problem can be formulated as follows:

$$(G_F) \quad \max_{x \in \mathcal{P}_F} \min_{\mathbf{R} \sim (\mu, \Sigma)} E[u(x'\mathbf{R})].$$

12

The underlying feasible set may have different forms, according to the type of portfolio problem of interest. For example, $\mathcal{P}_F = \{x \in \mathbf{R}^n \mid A'x \leq b\}$ or in particular $\mathcal{P}_F = \{x \in [0,1]^n \mid x'1_n = 1\}$.

According to Proposition 3, whenever the function $V(\mu_x, \sigma_x) = \min_{\mathbf{R} \sim (\mu, \Sigma)} E[u(x'\mathbf{R})]$ is quasi-concave, non-increasing in μ_x and non-decreasing in σ_x, the Problem (G_F) amounts to solving the PQP (10) over the portfolio choice set $\mathcal{P}_F$ for a suitable value λ^*, which is equivalent to solving for a certain parameter $\kappa^* \geq 0$ the following quadratic optimization program:

$$\max_{x \in \mathcal{P}_F} \mu'x - \kappa^* x' \Sigma x.$$

This is a deterministic mean-variance portfolio problem that can be solved by standard methods. This robust optimization approach associates with each utility function a worst case risk-return trade-off factor $k^* = (1 - \lambda^*)/\lambda^*$, obtainable by the methods of Theorem 2 and 3, such that solving the robust optimization problem is equivalent to solving a deterministic problem with quadratic utility $u_{\kappa^*}(x) = \mu'x - \kappa^* x' \Sigma x$.

Consider for comparison purposes the alternative approach, based on the assumption that the underlying distribution of returns is multi-variate normal. In this case the problem does not appear to be tractable by other than numeric methods :

$$(N_F) \quad \max_{x \in \mathcal{P}_F} E[u(\mathbf{r}_x)] = \max_{x \in \mathcal{P}_F} \int u(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - \mu'x)^2}{x' \Sigma x}} dx.$$

We investigate three criteria of interest with respect to the portfolio managers' incentive scheme: targets, commissions and target plus commission based systems.

4.1.1 The Target Criterion

Portfolio managers, and generally sales managers, are often incentivised by a fixed salary plus a bonus, contingent on the realization of a certain preset performance target, denoted α. In this case, the manager's objective is to maximize the probability of achieving the given target α.

When the distribution of returns is known, this problem, often referred to as the satisficing problem or the stochastic problem of maximum probability, is simply $\max_{x \in \mathcal{P}} P(x'\mathbf{R} \geq \alpha)$.

13

In the case of normal returns and linear constraints, this problem has received significant attention, especially in the 1960's operations research literature, see Markovitz [33], Kataoka [30], Charnes and Cooper [12], Geoffrion [21, 22], Dragomirescu [14]. One can easily prove (see [12]) in this case that the target criterion is equivalent to solving the following standardized **Z**-statistic criterion:

$$T^N(\alpha) \; = \; \max_{x \in \mathcal{P}} \; \frac{x'\mu - \alpha}{\sqrt{x'\Sigma x}} = \max_{x \in \mathcal{P}} \frac{\mu_x - \alpha}{\sigma_x}.$$

This is a linear fractional programming problem, for which several solution procedures are known (see for example Bazaara, Sherali and Shetty [5]).

In the robust setting, the manager maximizes his worst case probability of achieving the target α, over all distributions with given mean μ and covariance structure Σ. The corresponding maximization problem can be formulated as the following target-objective robust stochastic program:

$$T(\alpha) \; = \; \max_{x \in \mathcal{P}} \min_{\mathbf{R} \sim (\mu, \Sigma)} P(x'\mathbf{R} \geq \alpha) = \max_{x \in \mathcal{P}} \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} P(\mathbf{r} \geq \alpha), \tag{12}$$

where the last equality follows by Proposition 1. The proof of the following result is provided in the Appendix.

Proposition 4 *The robust target criterion (12) is equivalent to* $\displaystyle \max_{x \in \mathcal{P}} \frac{(\mu_x - \alpha)_+^2}{(\mu_x - \alpha)^2 + \sigma_x^2}$, *which has the same solution as* $\displaystyle \max_{x \in \mathcal{P}} \frac{(\mu_x - \alpha)_+}{\sigma_x}$. *These objectives satisfy the conditions of Proposition 3.*

It is interesting to notice that, if $\mu_x \geq \alpha$ is feasible, then the robust criterion produces the same optimal solution as the normal distribution assumption. This can be interpreted as follows: consider two portfolio managers incentivised by the same target α, who have access to the same set of assets with known means μ and covariance Σ. The first manager believes that stocks follow a multivariate normal distribution with these parameters, whereas the other manager ignores the distribution and adopts a worst case approach. This result says that the two managers, despite their different beliefs, will choose the same portfolio (provided that they have access to at least one portfolio whose return exceeds the target level α).

4.1.2 Option-type Criteria

It is often argued that fixed target bonus systems are unappealing because managers have no direct incentive to over-perform the target. As a remedy, managers are offered a commission for over-achieving a preset target level α, which amounts to an option-type payoff $u(x) = (x - \alpha)_+$. The corresponding robust criterion of interest involves maximizing the (worst case) expected over-achievement of a given target α, and is mathematically formulated as follows:

$$O^+(\alpha) = \max_{x \in \mathcal{P}} \min_{\mathbf{R} \sim (\mu, \Sigma)} E[(x'\mathbf{R} - \alpha)_+] = \max_{x \in \mathcal{P}} \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[(\mathbf{r} - \alpha)_+], \qquad (13)$$

where the last equality follows by Proposition 1.

Similarly, Ross [40] suggests an incentive system whereby managers can be penalized for not reaching a lower benchmark target level α, which amounts to a different option-type payoff $u(x) = (\alpha - x)_+$. The corresponding robust problem amounts to minimizing the worst case penalty from not meeting the target:

$$O^-(\alpha) = \min_{x \in \mathcal{P}} \max_{\mathbf{R} \sim (\mu, \Sigma)} E[(\alpha - x'\mathbf{R})_+] = \min_{x \in \mathcal{P}} \max_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[(\alpha - \mathbf{r})_+]. \qquad (14)$$

These kinds of objectives are ubiquitous in operations research and economics, appearing for example in option-prices in finance (Lo [32]), the classical newsboy problem in inventory control (Scarf [42]), PERT in reliability theory. The proof of the following result can be found in the Appendix.

Proposition 5 *The robust option criteria (13) and (14) are equivalent to solving:*

$$\max_{x \in \mathcal{P}} (\mu_x - \alpha)_+, \quad and\ respectively \quad \max_{x \in \mathcal{P}} (\alpha - \mu_x) + \sqrt{\sigma_x^2 + (\alpha - \mu_x)^2}.$$

In particular, based on Proposition 3, the latter can be solved as a PQP.

An alternative reward system is a mix of a bonus b and commission rate a for output above the target α, which amounts to $u(x) = a(x - \alpha)_+ + b\mathbf{1}_{(x \geq \alpha)}$. The corresponding

robust criterion of interest is mathematically formulated as follows:

$$
\begin{aligned}
TO(\alpha) &= \max_{x \in \mathcal{P}} \min_{\mathbf{R} \sim (\mu, \Sigma)} aE[(x'\mathbf{R} - \alpha)_+] + bP(x'\mathbf{R} \geq \alpha) \\
&= \max_{x \in \mathcal{P}} \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} aE[(\mathbf{r} - \alpha)_+] + bP(\mathbf{r} \geq \alpha),
\end{aligned}
\tag{15}
$$

where the last equality follows by Proposition 1. The proof of the following result is provided in the Appendix.

Proposition 6 *The robust target-option criterion (15) is equivalent to optimizing the corresponding linear combination of the robust target and option criteria, i.e.:*

$$
\max_{x \in \mathcal{P}} a(\mu_x - \alpha)_+ + b \frac{(\mu_x - \alpha)_+^2}{\sigma_x^2 + (\mu_x - \alpha)^2}.
$$

This problem can be solved as a parametric quadratic program.

4.2 Robust Regression with General Deviation Measures

In multiple linear regression, a random variable $\mathbf{X}_0$ is approximated in terms of other variables $\mathbf{X}_1, \dots, \mathbf{X}_n$ by a linear function $a + b_1 \mathbf{X}_1 + \dots + b_n \mathbf{X}_n$. Typically, the regression coefficients are determined by minimizing the mean squared error, i.e. $E[(X_0 - (a + b_1 \mathbf{X}_1 + \dots + b_n \mathbf{X}_n))^2]$. Rockafellar, Uryasev and Zabarankin [36] propose a concept of generalized regression, whereby other, more general deviation measures than the standard deviation σ are optimized. They argue why such general measures may be more appropriate in certain application settings, by better reflecting losses, fat tails and distinguishing between upside and downside risk. The generalized regression problem for a general deviation measure $\mathcal{D}$ can be stated as follows:

$$
\begin{aligned}
\min_{a, b_i} \quad & \mathcal{D}(X_0 - (a + b_1 \mathbf{X}_1 + \dots + b_n \mathbf{X}_n)) \\
\text{s.t.} \quad & E[a + b_1 \mathbf{X}_1 + \dots + b_n \mathbf{X}_n] = E[X_0].
\end{aligned}
\tag{16}
$$

In particular, the classical regression problem amounts to taking $\mathcal{D} = \sigma$ above.

Denote $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_n)$ and $b = (-1, b_1, \dots, b_n)$, and suppose that only the mean and covariance of the random vector $\mathbf{X}$ are known. For a general deviation measure $\mathcal{D}$, the problem of determining a robust set of regression coefficients (a, b) can be formulated in our framework as follows:

$$\min_{\substack{a,b \\ a+b'\mu=0}} \max_{\mathbf{X}\sim(\mu,\Sigma)} \mathcal{D}(b'\mathbf{X}-a) = \min_{b} \max_{\mathbf{X}\sim(\mu,\Sigma)} \mathcal{D}(b'(\mathbf{X}-\mu)) = \min_{b} \max_{\mathbf{x}\sim(0,\sigma_b^2)} \mathcal{D}(\mathbf{x}), \quad (17)$$

where the last relation follows by Proposition 1.

The case when $\mathcal{D}$ represents the standard deviation is trivial. For the purpose of this paper, we will focus on two quantile-based deviation measures of interest: value at risk (VaR) and conditional value at risk (CVaR). Notice that the corresponding objectives are not expectations (although they involve expectations), thereby extending in some sense our general framework.

4.2.1 The Value at Risk (VaR) Criterion

Value at risk is a percentile-based deviation measure defined as the negative of the fractile criterion for the corresponding loss distribution:

$$VaR(\alpha;\mathbf{x}) = \min\{z|P(\mathbf{x}\leq -z)\leq\alpha\} = -\max\{z|P(\mathbf{x}\geq z)\geq 1-\alpha\}.$$

For normal returns, Charnes and Cooper [12] prove that this is equivalent to optimizing:

$$VaR^N(\alpha) = -\mu_x - \Phi^{-1}(\alpha)\sigma_x,$$

where Φ is the standard normal cdf.

The general robust problem (G) with VaR criterion can be formulated as follows:

$$\begin{aligned} VaR(\alpha) = -\max_{x\in\mathcal{P}} \quad & z \\ \text{s.t.} \quad & \min_{\mathbf{R}\sim(\mu,\Sigma)} P(x'\mathbf{R}\geq z)\geq 1-\alpha. \end{aligned} \quad (18)$$

Proposition 7 *The robust VaR criterion is equivalent to solving* $\min\limits_{x\in\mathcal{P}} \sqrt{\dfrac{1-\alpha}{\alpha}}\,\sigma_x - \mu_x$. *This objective satisfies the conditions of Proposition 3.*

This result has been proved by El Ghaoui, Oustry and Oks [23] in the context of robust portfolio value at risk. Our approach yields the same solution, provided in the Appendix.

The robust regression problem (17) with VaR objective can be formulated as follows:

$$\min_{a,b} \quad z \qquad\qquad = \quad \min_{b} \quad z$$
$$\text{s.t.} \quad \max_{\mathbf{X}\sim(\mu,\Sigma)} P(b'\mathbf{X} - a \leq -z) \leq \alpha \qquad \text{s.t.} \quad \max_{\mathbf{x}\sim(0,\sigma_b^2)} P(\mathbf{x} \leq -z) \leq \alpha \qquad (19)$$
$$a + b'\mu = 0$$

This is a special case of the robust VaR criterion (18). By Proposition 7, the optimal robust regression coefficients are $(a = -\mu'b, b)$, where b minimizes $\sigma_b^2 = b'\Sigma b$.

4.2.2 The Conditional Value at Risk (CVaR) Criterion

The VaR criterion, while commonly used in the financial industry, has been criticised in the academic literature for a set of considerable drawbacks: it is non-convex, it is also not a coherent deviation measure in the sense of Rockafellar, Uryasev and Zabarankin [37], nor a coherent measure of risk as argued in Artzner, Delbaen, Eber and Heath [3]. Rockafellar and Uryasev [38] propose an alternative tail measure, called the α-conditional value at risk and defined as the negative expectation of the lower α-tail of a risk $\mathbf{x}$:

$$CVaR(\alpha; \mathbf{x}) = -E[\mathbf{x}|\mathbf{x} \leq z], \; P(\mathbf{x} \leq z) = \alpha].$$

We define the concept of robust CVaR as follows:

$$CVar(\alpha) = \min_{x\in\mathcal{P}} \quad \max_{\mathbf{R}\sim(\mu,\Sigma)} \quad E[-x'\mathbf{R}|x'\mathbf{R} \leq z]$$
$$\text{s.t.} \quad P(x'\mathbf{R} \leq z) = \alpha. \qquad (20)$$

The proof of the following result is provided in the Appendix:

Proposition 8 *In the robust case, the CVaR criterion is equivalent to the VaR criterion, and amounts to solving* $\min_{x\in\mathcal{P}} \sqrt{\dfrac{1-\alpha}{\alpha}} \, \sigma_x - \mu_x$. *This satisfies the conditions of Proposition 3.*

Recall that deviation measures are defined with respect to centered random variables $\mathbf{x} - E[\mathbf{x}]$. Our results show that the robust VaR and CVaR deviation measures are equivalent to the standard deviation measure, showing in some sense that standard deviation is a generic robust deviation measure.

5   A Multi-Product Pricing Application and Extensions

In this section we consider the optimal pricing decision of a firm (or sales manager) setting prices for a line of products so as to maximize the expected utility of profits (respectively, of his expected compensation). In this case the random returns are the demands for each product, which are non-negative and depend on the set prices. This setting indicates two possible extensions of our original problem, which we address in this section: (1) we incorporate dependence of the random vector $\mathbf{R}$ on the decision variable x, and (2) we analyse the case of non-negative random vectors $\mathbf{R}$. We extend Proposition 3 to address the first issue, and we provide weak bounds for non-negative random variables to address the second.

Consider a firm setting a vector of prices p for a set of n products with random demands $\mathbf{D}(p) = (\mathbf{D}_1(p), ..., \mathbf{D}_n(p))$ in order to maximize the expected utility of profits, given by $u(\mathbf{r})$, where $\mathbf{r}$ is the total profit. Suppose that the demands for the n products are correlated due to substitution and/or complementarity effects, and depend on the prices of all products. Demand is estimated to have mean $\mu(p)$ and covariance structure $\Sigma(p)$. The variable cost per product is given by the vector c. The problem of the firm looking for a "robust" pricing scheme for the n products can be formulated as follows:

$$(G_P) \quad \max_{p \in \mathcal{P}_P} \min_{\mathbf{D} \sim (\mu(p), \Sigma(p))} E[u((p-c)'\mathbf{D}(p))].$$

The marketing constraints on the feasible set of prices can be assumed to be of the following form: $\mathcal{P}_P = \{p \in R_+^n \mid A'p \leq b, \, p_i^{min} \leq p_i \leq p_i^{max}\}$.

5.1   Returns Dependent on the Decision Variable

The above problem (G_P) is *not* a special case of the robust problem (G) studied in this paper, since the random return (in this case the demand), together with its mean and variance, actually depend on the decision variable (in this case the price p). While the results of this section are presented in the context of the specific pricing application, one can naturally extend them to the general setting.

First, it can be easily seen, that the results of Propositions 1 and 2 do not change. Denote the mean and variance of the profit by $\mu_p = (p-c)'\mu(p)$ and $\sigma_p^2 = (p-c)'\Sigma(p)(p-c)$. The robust solution can be obtained by solving $\max_{p \in \mathcal{P}_P} \min_{\mathbf{r} \sim (\mu_p, \sigma_p^2)} E[u(\mathbf{r})]$.

We now provide conditions for solving the above problem as a parametric quadratic program. Based on Theorem 2, it is easy to see that the conditions described below are sufficient for extending Proposition 3, and thereby solving Problem (G_P) as a parametric quadratic program:

(C1) the expected profit $\mu_p = (p - c)'\mu(p)$ is concave in price;

(C2) the standard deviation of the profit $\sigma_p = \sqrt{(p - c)'\Sigma(p)(p - c)}$ is convex in price;

(C3) the robust objective function $V(\mu_p, \sigma_p) = \min\limits_{\mathbf{r}\sim(\mu_p,\sigma_p^2)} E[u(\mathbf{r})]$ is quasi-concave, non-decreasing in the first argument and non-increasing in the second.

Alternatively, conditions (C2) and (C3) may be structured to involve the variance instead of standard deviation:

(C2') the variance of the profit $\sigma_p^2 = (p - c)'\Sigma(p)(p - c)$ is convex in price;

(C3') the robust objective function $V(\mu_p, \sigma_p^2) = \min\limits_{\mathbf{r}\sim(\mu_p,\sigma_p^2)} E[u(\mathbf{r})]$ is quasi-concave, non-decreasing in the first argument and non-increasing in the second.

Let us argue why these assumptions make practical sense. The concavity of the (expected) profit function with respect to price is a universally accepted assumption in the pricing literature. The convexity of the standard deviation (or variance) intuitively means that revenue is more variable under randomized or differential pricing. In particular, this is true when demand variability is not affected by price i.e. $\Sigma(p) = \Sigma$, or when the demands are uncorrelated among products $\Sigma(p) = \sigma^2(p)I_n$ and demand variability for each product $\sigma_i^2(p_i)$ is increasing and convex in price. Finally, the quasi-concavity condition requires that the marginal rate of substitution between the mean and variance (or standard deviation) of the profit decreases with price.

5.2 The Problem with Non-Negative Returns

In the above pricing application, we are ignoring a salient feature of demand distributions, namely the fact that demand may only take non-negative values. A natural question to investigate is how our results for real-valued returns extend when the underlying return vector is restricted to the non-negative orthant. The following problem is the focus of this

section:

$$(G^+) \quad \max_{x \in \mathcal{P}} \min_{\mathbf{R} \sim (\mu, \Sigma)^+} E[u(x'\mathbf{R})], \tag{21}$$

where $\mathbf{R} \sim (\mu, \Sigma)^+$ means that the inner optimization is over all non-negative random vectors $\mathbf{R}$ with mean μ and covariance Σ. We denote $\mathcal{M}^n_{(\mu, \Sigma)^+}$ the set of measures on R^n_+ with mean μ and covariance Σ (for $n = 1$, the superscript will be omitted by default). Let us first notice that the problem of deciding whether the set $\mathcal{M}^n_{(\mu, \Sigma)^+}$ is non-empty is already a difficult problem (NP-hard, see [35]). To avoid this issue, we explicitly assume that (μ, Σ) are given so that $\mathcal{M}^n_{(\mu, \Sigma)^+}$ is non-empty.

The following result extends Propositions 1 and 2, by providing bounds on the inner part of problem (G^+):

Proposition 9 *Suppose $\mathbf{R} \sim (\mu, \Sigma)^+$, and for any non-zero vector $x \geq 0$ denote $\mu_x = x'\mu$ and $\sigma_x^2 = x'\Sigma x$. We have the following result:*

$$\begin{aligned} \min_{\mathbf{R} \sim (\mu, \Sigma)^+} E[u(x'\mathbf{R})] &\geq \min_{\mathbf{r} \sim (\mu_x, \sigma_x^2)^+} E[u(\mathbf{r})] \\ \max_{\mathbf{R} \sim (\mu, \Sigma)^+} E[u(x'\mathbf{R})] &\leq \max_{\mathbf{r} \sim (\mu_x, \sigma_x^2)^+} E[u(\mathbf{r})] \end{aligned}. \tag{22}$$

Furthermore, the right hand side bounds are equivalent to optimizing over $h \geq \frac{\sigma_x^2}{\mu_x}$ the objective:

$$\frac{\sigma_x^2}{h^2 + \sigma_x^2} u(\mu_x + h) + \frac{h^2}{h^2 + \sigma_x^2} u(\mu_x - \frac{\sigma_x^2}{h}). \tag{23}$$

Proof :

Consider the first inequality: for the minimizing $\mathbf{R}^*$ on the left hand side, the corresponding projection $\mathbf{r}^* = x'\mathbf{R}^*$ is feasible and yields the same objective value on the right hand side. Similarly, the reverse inequality is trivial for the second inequality.

The right hand side bounds are equivalent to optimizing over the class of extreme measures of $\mathcal{M}_{(\mu_x, \sigma_x^2)^+}$. Following a similar approach to Rogosinskii [39] (see [35] for a full proof), one can show that these are non-negative 2-point distributions with mean $\mu_x = x'\mu$ and

variance $\sigma_x^2 = x'\Sigma x$. In particular, they can be parametrized by $h \geq \frac{\sigma_x^2}{\mu_x}$ as follows:

$$\delta_{\mu_x,\sigma_x}(h) = \begin{cases} \mu_x + h & \text{with probability} \quad \frac{\sigma_x^2}{h^2 + \sigma_x^2} \\ \mu_x - \frac{\sigma_x^2}{h} & \text{with probability} \quad \frac{h^2}{h^2 + \sigma_x^2} \end{cases}.$$

$\square$

A natural question to ask is whether the bounds in Proposition 9 are exact, i.e. whether set of measures $\mathcal{M}^n_{(\mu,\Sigma)^+}$ projects through a non-negative linear mapping $\mathbf{R} \xrightarrow{x \geq 0} \mathbf{r} = x'\mathbf{R}$ **onto** the set $\mathcal{M}_{(\mu_x,\sigma_x^2)^+}$ of univariate measures with non-negative support, mean μ_x and variance σ_x^2. It turns out that this is not true. Popescu [35] completely characterized the set of projected measures corresponding to non-negative random vectors with given mean and covariance structure. This characterization unfortunately indicates that the robust objective under non-negative returns is not generally expressible as a bi-criteria problem, since the projected set of measures depends not only on μ_x and σ_x, but also on the values of $\frac{x'\Sigma_i}{\mu_i}$. Nevertheless, based on Proposition 9, one can compute simple (non-sharp) bi-criteria bounds for problems with non-negative returns.

6   Conclusions

This paper demonstrated a general approach for solving robust max-min and min-max stochastic optimization problems (without recourse), based only on mean-covariance information on the random vector of returns. We provided a two-step approach that reduced this problem to solving a parametric quadratic program. First, we prove that the inner optimization problem is equivalent to a problem over univariate distributions, due to a new projection result for multivariate distributions. This result holds for arbitrary objective functions u, and simplifies the initial problem to solving a bi-criteria optimization problem. Second, we provided some intuitive conditions on the robust objective function that reduce the problem to solving a parametric quadratic program.

We provided extensions to handle dependence of the return distribution on the decision variable, as well as non-negative distributions. We also showed how our results can be used to solve problems with non-differentiable quantile-based objectives, such as probability, option, VaR and CVaR criteria. While we illustrated our results with applications in finance, statistics and pricing, we expect our approach to be relevant in a wide variety of contexts.

Appendix

Proof of Proposition 4:   By Propositions 1 and 2 (or alternatively by applying the optimal one-sided Chebyshev inequality ([29], [8])), we obtain the following bound:

$$\min_{\mathbf{R}\sim(\mu,\Sigma)} P(x'\mathbf{R} \geq \alpha) = 1 - \max_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} P(\mathbf{r} \leq \alpha) = \begin{cases} \dfrac{(\mu_x - \alpha)^2}{(\mu_x - \alpha)^2 + \sigma_x^2} & \text{, if } \alpha < \mu_x \\ 0 & \text{, otherwise} \end{cases}. \quad (24)$$

If there is no feasible $x \in \mathcal{P}$ with $\mu_x \geq \alpha$, then the objective is zero. Otherwise, the target criterion amounts to optimizing $\frac{\mu_x - \alpha}{\sigma_x}$ for $\mu_x \geq \alpha, x \in \mathcal{P}$. This gives the desired result. Both functions $\dfrac{(\mu_x - \alpha)^2}{(\mu_x - \alpha)^2 + \sigma_x^2}$ and $\dfrac{\mu_x - \alpha}{\sigma_x}$ are non-decreasing in μ_x, non-increasing in σ_x and quasi-concave on $[\alpha, \infty) \times [0, \infty)$, so the results of Proposition 3 apply. $\quad\square$

Proof of Proposition 5:   For the maximization problem, it is easy to see that $\min\limits_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} E[(\mathbf{r} - \alpha)_+] = (\mu_x - \alpha)_+$.

The minimization case is a consequence of Proposition 2. Using the change of variable $h = \sigma_x\sqrt{\frac{p}{1-p}} \geq 0$ in the formulation (8), the robust minimization bound amounts to minimizing over $x \in \mathcal{P}$ the objective :

$$\max_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} E[(\mathbf{r} - \alpha)_+] = \max_{h\geq 0} \frac{h^2}{h^2 + \sigma_x^2}\left(\mu_x + \frac{\sigma_x^2}{h} - \alpha\right)_+ + \frac{\sigma_x^2}{h^2 + \sigma_x^2}(\mu_x - h - \alpha)_+. \quad (25)$$

First, if h is such that $\alpha \leq \mu_x - h$ or $\alpha \geq \mu_x + \frac{\sigma_x^2}{h}$, the bound we obtain is $(\mu_x - \alpha)_+$. In the remaining case, only the first term in (25) is non-zero. By optimizing $\frac{h^2}{h^2 + \sigma_x^2}(\mu_x + \frac{\sigma_x^2}{h} - \alpha)$ we obtain $h = \mu_x - \alpha + \sqrt{\sigma_x^2 + (\mu_x - \alpha)^2} > (\mu_x - \alpha)_+$. The desired bound is obtained:

$$\max_{\mathbf{r}\sim(\mu_x,\sigma_x^2)} E[(\mathbf{r} - \alpha)_+] = \frac{1}{2}\left[(\mu_x - \alpha) + \sqrt{\sigma_x^2 + (\mu_x - \alpha)^2}\right]. \quad (26)$$

Letting $z_1 = -\mu_x$ and $z_2 = -\sigma_x$, minimizing the criterion (26) amounts to maximizing $v(z_1, z_2) = z_1 + \alpha - \sqrt{(z_1 + \alpha)^2 + z_2^2}$. This is non-decreasing in both arguments when $z_2 \leq 0$.

We also show that v is quasi-concave, i.e. the upper level sets $L_l = \{(z_1, z_2)|v(z_1, z_2) \geq l\}$ are convex. We want to show that $\lambda(z_1, z_2) + (1-\lambda)(w_1, w_2) \in L_l$ for any $(z_1, z_2), (w_1, w_2) \in L_l$ and $\lambda \in [0, 1]$, which amounts, after some simplification, to checking

$$l(l - 2(\lambda z_1 + (1 - \lambda)w_1 + \alpha)) \geq (\lambda z_2 + (1 - \lambda)w_2)^2.$$

The conditions $(z_1, z_2), (w_1, w_2) \in L_l$ imply that $l(l - 2(z_1 + \alpha)) \geq z_2^2$ and respectively $l(l - 2(w_1 + \alpha)) \geq w_2^2$. From these two conditions and convexity of $f(x) = x^2$ we obtain:

$$l(l - 2(\lambda z_1 + (1 - \lambda)w_1 + \alpha)) \geq \lambda z_2^2 + (1 - \lambda)w_2^2 \geq (\lambda z_2 + (1 - \lambda)w_2)^2.$$

Therefore, Proposition 3 applies. $\qquad \square$

Proof of Proposition 6: The robust target-option maximization problem amounts to optimizing over $x \in \mathcal{P}$ the criterion $\min\limits_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} E[a\mathbf{1}_{(\mathbf{r} \geq \alpha)} + b(\mathbf{r} - \alpha)_+]$, which equals

$$\min_{q \geq 0} \frac{a}{1 + q^2}[\mathbf{1}_{(\mu_x + q\sigma_x \geq \alpha)} + q^2\mathbf{1}_{(\mu_x - \frac{1}{q}\sigma_x \geq \alpha)}] + \frac{b}{1 + q^2}[(\mu_x + q\sigma_x - \alpha)_+ + q^2(\mu_x - \frac{1}{q}\sigma_x - \alpha)_+].$$

This follows by Proposition 2 with the change of variable $q = \sqrt{(1 - p)/p}$. There are three possible cases, depending on the value of α relative to $\mu_x - \frac{1}{q}\sigma_x$ and $\mu_x + q\sigma_x$. The only relevant case (that yields a non-zero value) is when α falls between the two and $\mu_x \leq \alpha$, producing the desired result. Monotonicity and quasi-concavity in the non-trivial case $(\mu_x \leq \alpha)$ can be shown by inspection. $\qquad \square$

Proof of Proposition 7: Using the proof of Proposition 4, the robust VaR minimization problem (18) amounts to solving:

$$-\max_{x \in \mathcal{P}} \quad z$$
$$\text{s.t.} \quad \frac{(\mu_x - \alpha)^2}{(\mu_x - z)^2 + \sigma_x^2} \geq 1 - \alpha$$
$$z \leq \mu_x.$$

The optimum is achieved when equality holds in the first constraint. By solving for z, we obtain the equivalent robust VaR criteria $-z^* = -\mu_x + \sqrt{\frac{1-\alpha}{\alpha}}\sigma_x$. Minimizing this over $x \in \mathcal{P}$ is the same as maximizing $\mu_x - \sqrt{\frac{1-\alpha}{\alpha}}\sigma_x$, which is non-decreasing in μ_x, non-increasing in σ_x and quasi-concave. $\qquad \square$

Proof of Proposition 8: Based on Rockafellar and Uryasev [38], the CVaR for a random

24

variable $\mathbf{x}$ can be alternatively computed as follows:

$$CVaR(\alpha; \mathbf{x}) = -E[\mathbf{x}|\mathbf{x} \leq z\,,\ P(\mathbf{x} \leq z) = \alpha] = \min_{\beta} \beta + \frac{1}{\alpha} E[(-\mathbf{x} - \beta)_+].$$

This shows that the robust CVaR can be calculated as:

$$
\begin{aligned}
CVaR(\alpha) \quad &= \max_{\mathbf{r} \sim (\mu_x, \sigma_x^2)} \min_{\beta} \{\beta + \frac{1}{\alpha} E[(-\mathbf{r} - \beta)_+]\} \\
&= \max_{q \geq 0} \min_{\beta} \{\beta + \frac{1}{\alpha(1 + q^2)}[(-\mu_x - q\sigma_x - \beta)_+ + q^2(-\mu_x + \frac{1}{q}\sigma_x - \beta)_+]\},
\end{aligned}
$$

where the last equality follows by Proposition 2 with the change of variable $q = \sqrt{(1-p)/p}$. There are three possible cases, depending on the value of β relative to $-\mu_x + \frac{1}{q}\sigma_x$ and $-\mu_x - q\sigma_x$. The only relevant case is when β falls between the two. In this case, the desired result is obtained. $\qquad\square$

Acknowledgements.

References

[1] N. I. Akhiezer. *The Classical Moment Problem.* Hafner, New York, NY, 1965.

[2] L. Arseneau and M.J. Best. Resolution of degenerate critical parameter values in parametric quadratic programming. CORR 99-47, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, 1999.

[3] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[4] J. Dupačová (as Žáčková). On minimax solutions of stochastic linear programming problems. *Časopis pro Pěstování Matematiky*, 91:423–429, 1966.

[5] M.S. Bazaara, H.D. Sherali, and C.M. Shetty. *Nonlinear Programmin - Theory and Algorithms.* John Wiley, 2 edition, 1993.

[6] A. Ben-Tal and A. Nemirovski. Robust optimization - methodology and applications. *Math. Program., Ser. B*, 92:453–480, 2002.

[7] J. Berger. The robust bayesian viewpoint. In J. Kadane (eds.), editor, *Robustness of Bayesian Analysis*, pages 63–124. Elsevier Science, 1984.

[8] D. Bertsimas and I. Popescu. Optimal inequalities in probability: A convex programming approach. Working paper, INSEAD, http://faculty.insead.edu/popescu/ioana, 1998.

[9] M.J. Best. An algorithm for the solution of the parametric quadratic programming problem. In B. Riedmller H. Fischer and S. Schffler, editors, *Applied Mathematics and Parallel Computing - Festschrift for Klaus Ritter*, pages 57–76. Heidelburg: Physica-Verlag, 1996.

[10] J.R. Birge and M.J. Maddox. Bounds on expected project tardiness. *Opns. Res.*, 43(5):838–850, 1995.

[11] J.R. Birge and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Math. Oper. Res.*, 12(1):149–162, 1987.

[12] A. Charnes and W. Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11(1):18–39, January-February 1963.

[13] S.P. Dokov and D.P. Morton. Higher-order upper bounds on the expectation of a convex function. UT Austin working paper, January 2002.

[14] M. Dragomirescu. An algorithm for the minimum risk problem of stochastic programming. *Operations Research*, (20):154–164, 1972.

[15] J.H. Dulá and R.V. Murthy. A tchebysheff-type bound on the expectation of sublinear polyhedral functions. *Opns. Res.*, 40(5):914–922, 1992.

[16] J. Dupaçová, A. Prekopa, J.R. Birge, and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Math. OR.*, 12:149–162, 1987.

[17] J. Dupačová. The minimax aproach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[18] J. Dupačová. Stochastic programming: Minimax aproach. In C.A. Floudas and P.M. Pardalos editors, editors, *Encyclopedia of Optimization*, volume 5, pages 327–330. Kluwer Acedemic Publishers, Kluwer, 2001.

[19] N.C.P. Edirisinghe. New second order bounds on the expectation of saddle functions with applications to stochastic linear programming. *Ann. Opns. Res.*, 64:143–178, 1996.

[20] A.A. Gaivoronski. A numerical method for solving stochastic programming problems with moment constraints on a distribution function. *Ann. Opns. Res.*, 31:347–369, 1991.

[21] A. Geoffrion. Solving bicriterion mathematical programs. *Operations Research*, 15(1):39–54, January-February 1967.

[22] A. Geoffrion. Stochastic programming with aspiration or fractile criteria. *Management Science*, 13(9):672–679, May 1967.

[23] L. El Ghaoui, F. Oustry, and M. Oks. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, July-August 2003.

[24] I. Gilboa and D. Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–173, 1989.

[25] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. to appear in Math O.R., CORC Report 2002–03, www.corc.ieor.columbia.edu, 2002.

[26] K. Isii. On tchebycheff type inequalities. *Proc. Inst. Stat. Math.*, 7:123–143, 1959.

[27] P. Kall. Stochastic programming with recourse: Upper bounds and moment problems, a review. In J. Guddat et al. (eds.), editor, *Advances in Mathematical Optimization*. Akademie-Verlag, Berlin, 1988.

[28] P. Kall. An upper bound for slp using first and total second moments. *Ann. Opns. Res.*, 30:267–276, 1991.

[29] S. Karlin and W.J. Studden. *Tchebycheff Systems: with Applications in Analysis and Statistics.* Pure and Applied Mathematics, A Series of Texts and Monographs. Interscience Publishers, John Wiley and Sons, 1966.

[30] S. Kataoka. Stochastic programming, maximum probability model. *Hitosubasi Journal of Arts Sci.*, 8:51–59, 1967.

[31] J. H. B. Kempermann. Geometry of the moment problem. In ed. H.J. Landau, editor, *Moments in Mathematics*, number 37 in Proc. Sympos. Appl. Math., Series III, pages 16–53, Providence, RI, 1987. American Mathematical Society.

[32] A. Lo. Semiparametric upper bounds for option prices and expected payoffs. *Journal of Financial Economics*, 19:373–388, 1987.

[33] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, March 1952.

[34] H. M. Markowitz. *Portfolio Selection.* Wiley, 1959.

[35] I. Popescu. Optimal mean-covariance bounds for sums of positive risks: Applications in finance and insurance. INSEAD working paper, March 2003.

[36] R.T. Rockafellar, Uryasev S., and M. Zabarankin. Deviation measures in generalized linear regression. Research Report 2002-9. ISE Dept., University of Florida, http://www.ise.ufl.edu/uryasev/pubs.html, December 2002.

[37] R.T. Rockafellar, Uryasev S., and M. Zabarankin. Deviation measures in risk analysis and optimization. Research Report 2002-7. ISE Dept., University of Florida, http://www.ise.ufl.edu/uryasev/pubs.html, December 2002.

[38] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3):21–41, 2000.

[39] W.W. Rogosinsky. Moments of non-negative mass. *Proc. Roy. Soc. London. Ser. A.*, (245):1–27, 1958.

[40] S. Ross. Compensation, incentives, and the duality of risk aversion and riskiness. *The Journal of Finance*, 59(1), 2004.

[41] B. Rustem, R. Becker, and W. Marty. Robust min-max portfolio strategies for rival forecast and risk scenarios. *Fournal of Economic Dynamics and Control*, 24:1591–1621, 2000.

[42] H. Scarf. A min-max solution of an inventory problem. In *Studies in the mathematical theory of inventory and production*, pages 201–209, Stanford University Press, Stanford, CA, 1958. K.J. Arrow, S. Karlin and H. Scarf, eds.

[43] A. Shapiro and A.J. Kleywegt. Minimax analysis of stochastic problems. *Optimization Methods and Software*, 17:523–542, 2002.

[44] J. Smith. Generalized chebyshev inequalities: Theory and applications in decision analysis. *Operations Research*, 43(5):807–825, 1995.

[45] P. Wolfe. The simplex method for quadratic programming. *Econometrica*, 27:382–398, 1959.

[46] M. Young. A minimax portfolio selection rule with linear programming solution. *Management Science*, 44(5):637–683, 1998.