Report of the Institute of Control and Computation Engineering Warsaw University of Technology

On Efficient Optimization of the LP Computable Risk Measures for Portfolio Selection

Włodzimierz Ogryczak and Tomasz Śliwiński

February, 2008

Report nr: 08–04

Warsaw University of Technology
Institute of Control & Computation Engineering
ul. Nowowiejska 15/19, 00–665 Warsaw, Poland
phone (48-22)-234-7397, fax (48-22)-825-3719

On Efficient Optimization of the LP Computable Risk Measures for Portfolio Selection

Włodzimierz Ogryczak — *wogrycza@ia.pw.edu.pl*
Tomasz Śliwiński — *tsliwins@ia.pw.edu.pl*

February, 2008

Abstract

The portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem where the expected return is maximized and some (scalar) risk measure is minimized. In the original Markowitz model the risk is measured by the variance while several polyhedral risk measures have been introduced leading to Linear Programming (LP) computable portfolio optimization models in the case of discrete random variables represented by their realizations under specified scenarios. Recently, the second order quantile risk measures have been introduced and become popular in finance and banking. The simplest such measure, now commonly called the Conditional Value at Risk (CVaR) or Tail VaR, represents the mean shortfall at a specified confidence level. The corresponding portfolio optimization models can be solved with general purpose LP solvers. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model with huge number of variables and constraints thus decreasing the computational efficiency of the model. Actually, the number of constraints (matrix rows) is proportional to the number of scenarios. while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments. We show that the computational efficiency can be then dramatically improved with an alternative model taking advantages of the LP duality. In the introduced model the number of structural constraints (matrix rows) is proportional to the number of instruments thus not affecting seriously the simplex method efficiency by the number of scenarios. Moreover, similar reformulation can be applied to more complex quantile risk measures like the Gini's mean difference and the tail Gini's measures as well as to the mean absolute deviation.

Key Words. Risk Measures, Portfolio Optimization, Computability, Linear Programming, Duality, CVaR, MAD.

1 Introduction

Following Markowitz [12], the portfolio selection problem is modeled as a mean-risk bicriteria optimization problem where the expected return is maximized and some (scalar) risk measure is minimized. In the original Markowitz model the risk is measured by the variance while several polyhedral risk measures have been introduced leading to Linear Programming (LP) computable portfolio optimization models in the case of discrete random variables represented by their realizations under specified scenarios. The simplest LP computable risk measures are dispersion measures similar to the variance. Konno and Yamazaki

[6] presented the portfolio selection model with the mean absolute deviation (MAD). Yitzhaki [27] introduced the mean-risk model using Gini's mean (absolute) difference as the risk measure. The Gini's mean difference turn out to be a special aggregation technique of the multiple criteria LP model [16] based on the pointwise comparison of the absolute Lorenz curves. The latter leads the quantile shortfall risk measures which are more commonly used and accepted. Recently, the second order quantile risk measures have been introduced in different ways by many authors [2, 5, 14, 15, 22]. The measure, usually called the Conditional Value at Risk (CVaR) or Tail VaR, represents the mean shortfall at a specified confidence level. The CVaR measures maximization is consistent with the second degree stochastic dominance [18]. Several empirical analyses confirm its applicability to various financial optimization problems [1, 10]. This paper is focused on computational efficiency of the CVaR and related LP computable portfolio optimization models.

For returns represented by their realizations under T scenarios the basic LP model for CVaR portfolio optimization contains T auxiliary variables as well as T corresponding linear inequalities. Actually, the number of structural constraints in the LP model (matrix rows) is proportional to the number of scenarios T, while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments $T + n$. Hence, its dimensionality is proportional to the number of scenarios T. It does not cause any computational difficulties for a few hundreds of scenarios as in computational analysis based on historical data. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios [21]. This may lead to the LP model with huge number of auxiliary variables and constraints thus decreasing the computational efficiency of the model. Actually, in the case of fifty thousand scenarios and one hundred instruments the model may require more than half an hour computation time [8] with the state-of-art LP solver (CPLEX code). We show that the computational efficiency can be then dramatically improved with an alternative model formulation taking advantages of the LP duality. In the introduced model the number of structural constraints is proportional to the number of instruments n while only the number of variables is proportional to the number of scenarios T thus not affecting so seriously the simplex method efficiency. Indeed, the computation time is then below 30 seconds.

Moreover, similar reformulation can be applied to the classical LP portfolio optimization model based on the mean absolute deviation as well as to more complex quantile risk measures. The Tail Gini's measures or the Weighted CVaR measures defined as combinations of CVaR measures for m tolerance levels lead to LP models with the number of structural constraints (matrix rows) proportional to the respectively multiplied number of scenarios mT. In the alternative model taking advantages of the LP duality the number of structural constraints is proportional to the total of the number of instruments and number of tolerance levels $n+m$. This guarantees a high computational efficiency of the dual model even for a very large number of scenarios. The standard LP models for the Gini's mean difference [27] and its downside version [7] require T^2 auxiliary constraints which makes them hard already for medium numbers of scenarios, like a few hundred scenarios given by historical data. The models taking advantages of the LP duality allow one to limit the number of structural constraints making it proportional to the number of scenarios T thus increasing dramatically computational performances for medium numbers of scenario although still remaining hard for very large numbers of scenarios.

2 Portfolio Optimization and Risk Measures

The portfolio optimization problem considered in this paper follows the original Markowitz' formulation and is based on a single period model of investment. At the beginning of a period, an investor allocates the capital among various securities, thus assigning a nonnegative weight (share of the capital) to each security. Let $J = \{1, 2, \ldots, n\}$ denote a set of securities considered for an investment. For each security $j \in J$, its rate of return is represented by a random variable R_j with a given mean $\mu_j = \mathbb{E}\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,2,\ldots,n}$ denote a vector of decision variables x_j expressing the weights defining a portfolio. The weights must satisfy a set of constraints to represent a portfolio. The simplest way of defining a feasible set $\mathcal{P}$ is by a requirement that the weights must sum to one and they are nonnegative (short sales are not allowed), i.e.

$$\mathcal{P} = \{\mathbf{x} : \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n\} \tag{1}$$

Hereafter, we perform detailed analysis for the set $\mathcal{P}$ given with constraints (1). Although the presented results can easily be adapted to a general LP feasible set given as a system of linear equations and inequalities, thus allowing one to include short sales, upper bounds on single shares or portfolio structure restrictions which may be faced by a real-life investor.

Each portfolio $\mathbf{x}$ defines a corresponding random variable $R_{\mathbf{x}} = \sum_{j=1}^{n} R_j x_j$ that represents the portfolio rate of return while the expected value can be computed as $\mu(\mathbf{x}) = \sum_{j=1}^{n} \mu_j x_j$. We consider T scenarios with probabilities p_t (where $t = 1, \ldots, T$). We assume that for each random variable R_j its realization r_{jt} under the scenario t is known. Typically, the realizations are derived from historical data treating T historical periods as equally probable scenarios ($p_t = 1/T$). The realizations of the portfolio return $R_{\mathbf{x}}$ are given as $y_t = \sum_{j=1}^{n} r_{jt} x_j$.

The portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem where the mean $\mu(\mathbf{x})$ is maximized and the risk measure $\varrho(\mathbf{x})$ is minimized. In the original Markowitz model, the standard deviation was used as the risk measure. Several other risk measures have been later considered thus creating the entire family of mean-risk models (see [9] and [10]). These risk measures, similar to the standard deviation, are not affected by any shift of the outcome scale and are equal to 0 in the case of a risk-free portfolio while taking positive values for any risky portfolio. Unfortunately, such risk measures are not consistent with the stochastic dominance order [13] or other axiomatic models of risk-averse preferences [23] and risk measurement [2].

In stochastic dominance, uncertain returns (modeled as random variables) are compared by pointwise comparison of some performance functions constructed from their distribution functions. The first performance function $F_{\mathbf{x}}^{(1)}$ is defined as the right-continuous cumulative distribution function: $F_{\mathbf{x}}^{(1)}(\eta) = F_{\mathbf{x}}(\eta) = \mathbb{P}\{R_{\mathbf{x}} \leq \eta\}$ and it defines the first degree stochastic dominance (FSD). The second function is derived from the first as $F_{\mathbf{x}}^{(2)}(\eta) = \int_{-\infty}^{\eta} F_{\mathbf{x}}(\xi)\, d\xi$ and it defines the second degree stochastic dominance (SSD). We say that portfolio $\mathbf{x}'$ *dominates* $\mathbf{x}''$ *under the SSD* ($R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$), if $F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta)$ for all η, with at least one strict inequality. A feasible portfolio $\mathbf{x}^0 \in \mathcal{P}$ is called *SSD efficient* if there is no $\mathbf{x} \in \mathcal{P}$ such that $R_{\mathbf{x}} \succ_{SSD} R_{\mathbf{x}^0}$.

Stochastic dominance relates the notion of risk to a possible failure of achieving some targets. Note that function $F_{\mathbf{x}}^{(2)}$, used to define the SSD relation, can also be presented as follows [17]: $F_{\mathbf{x}}^{(2)}(\eta) = \mathbb{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\}$ and thereby its values are LP computable for returns represented by their realizations y_t.

When the mean $\mu(\mathbf{x})$ is used instead of the fixed target the value $F^{(2)}_{\mathbf{x}}(\mu(\mathbf{x}))$ defines the risk measure known as the *downside mean semideviation* from the mean

$$\bar{\delta}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = F^{(2)}_{\mathbf{x}}(\mu(\mathbf{x})). \qquad (2)$$

The downside mean semideviation is always equal to the upside one and therefore we refer to it hereafter as to the mean semideviation. The mean semideviation is a half of the mean absolute deviation (MAD) from the mean [17] $\delta(\mathbf{x}) = \mathbb{E}\{|R_{\mathbf{x}} - \mu(\mathbf{x})|\} = 2\bar{\delta}(\mathbf{x})$. Hence the corresponding portfolio optimization model is equivalent to the MAD. Since $\bar{\delta}(\mathbf{x}) = F^{(2)}_{\mathbf{x}}(\mu(\mathbf{x}))$, the mean semideviation (2) is LP computable (when minimized), for a discrete random variable represented by its realizations y_t. Although, due to the use of distribution dependent target value $\mu(\mathbf{x})$, the mean semideviation cannot be directly considered an SSD consistent risk measure. SSD consistency [17] and coherency [10] of the MAD model can be achieved with maximization of for complementary risk measure $\mu_\delta(\mathbf{x}) = \mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) = \mathbb{E}\{\min\{\mu(\mathbf{x}), R_{\mathbf{x}}\}\}$, which also remains LP computable for a discrete random variable represented by its realizations y_t.

An alternative characterization of the SSD relation can be achieved with the so-called *Absolute Lorenz Curves* (ALC) [14, 26] which represent the second quantile functions defined as

$$F^{(-2)}_{\mathbf{x}}(p) = \int_0^p F^{(-1)}_{\mathbf{x}}(\alpha)d\alpha \quad \text{for } 0 < p \leq 1 \quad \text{and} \quad F^{(-2)}_{\mathbf{x}}(0) = 0, \qquad (3)$$

where $F^{(-1)}_{\mathbf{x}}(p) = \inf\,\{\eta : F_{\mathbf{x}}(\eta) \geq p\}$ is the left-continuous inverse of the cumulative distribution function $F_{\mathbf{x}}$. The pointwise comparison of ALCs is equivalent to the SSD relation [18] in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ if and only if $F^{(-2)}_{\mathbf{x}'}(\beta) \geq F^{(-2)}_{\mathbf{x}''}(\beta)$ for all $0 < \beta \leq 1$. Moreover,

$$F^{(-2)}_{\mathbf{x}}(\beta) = \max_{\eta \in R} \left[\beta\eta - F^{(2)}_{\mathbf{x}}(\eta)\right] = \max_{\eta \in R} \left[\beta\eta - \mathbb{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\}\right] \qquad (4)$$

where η is a real variable taking the value of β-quantile $Q_\beta(\mathbf{x})$ at the optimum. For a discrete random variable represented by its realizations y_t problem (4) becomes an LP.

For any real tolerance level $0 < \beta \leq 1$, the normalized value of the ALC defined as

$$M_\beta(\mathbf{x}) = F^{(-2)}_{\mathbf{x}}(\beta)/\beta \qquad (5)$$

is called the *Conditional Value-at-Risk (CVaR)* or Tail VaR or Average VaR. The CVaR measure is an increasing function of the tolerance level β, with $M_1(\mathbf{x}) = \mu(\mathbf{x})$. For $\beta = 0.5$ the CVaR corresponds to the mean absolute deviation from the median [9], the risk measure suggested by Sharpe [25] as the right MAD model. For any $0 < \beta < 1$, the CVaR measure is SSD consistent [18] and coherent [20]. Due to (4), for a discrete random variable represented by its realizations y_t the CVaR measures are LP computable. It is important to notice that although the quantile risk measures (VaR and CVaR) were introduced in banking as extreme risk measures for very small tolerance levels (like $\beta = 0.05$), for the portfolio optimization good results have been provided by rather larger tolerance levels [10].

For β approaching 0, the CVaR measure tends to the Minimax measure

$$M(\mathbf{x}) = \min_{t=1,\ldots,T} y_t \qquad (6)$$

introduced to portfolio optimization by Young [28].

3 Computational LP Models for Basic Risk Measures

Let us consider portfolio optimization problem with security returns given by discrete random variables with realization r_{jt}. Following (4) and (5), the CVaR portfolio optimization model can be formulated as the following LP problem:

$$\begin{array}{ll} \text{maximize} & \eta - \dfrac{1}{\beta}\sum_{t=1}^{T} p_t d_t \\ \text{subject to} & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1,\ldots,n \\ & d_t - \eta + \sum_{j=1}^{n} r_{jt} x_j \geq 0,\ d_t \geq 0 \quad \text{for } t = 1,\ldots,T \end{array} \tag{7}$$

where η is unbounded variable. Except from the core portfolio constraints (1), model (7) contains T nonnegative variables d_t plus single η variable and T corresponding linear inequalities. Hence, its dimensionality is proportional to the number of scenarios T. Exactly, the LP model contains $T + n + 1$ variables and $T + 1$ constraints. It does not cause any computational difficulties for a few hundreds of scenarios as in several computational analysis based on historical data [11]. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model (7) with huge number of variables and constraints thus decreasing the computational efficiency of the model. If the core portfolio constraints contain only linear relations, like (1), then the computational efficiency can easily be achieved by taking advantages of the LP dual to model (7). The LP dual model takes the following form:

$$\begin{array}{ll} \text{minimize} & q \\ \text{subject to} & q - \sum_{t=1}^{T} r_{jt} u_t \geq 0 \quad \text{for } j = 1,\ldots,n \\ & \sum_{t=1}^{T} u_t = 1 \\ & 0 \leq u_t \leq p_t/\beta \quad \text{for } t = 1,\ldots,T \end{array} \tag{8}$$

The dual LP model contains T variables u_t, but the T constraints corresponding to variables d_t from (7) take the form of simple upper bounds (SUB) on u_t thus not affecting the problem complexity. Actually, the number of constraints in (8) is proportional to the total of portfolio size n, thus it is independent from the number of scenarios. Exactly, there are $T + 1$ variables and $n + 1$ constraints. This guarantees a high computational efficiency of the dual model even for vary large number of scenarios. Note that introducing a lower bound on the required expected return in the primal portfolio optimization model (7) result only in a single additional variable in the dual model (8). Similarly, other portfolio structure requirements are modeled with rather small number of constraints thus generating small number of additional variables in the dual model.

We have run computational test on 10 randomly generated test instances developed by Lim et al. [8]. They were originally generated from a multivariate normal distribution for 50 securities with the number of scenarios 50,000 just providing an adequate approximation to the underlying unknown continuous price distribution. Scenarios were generated using the Triangular Factorization Method [24] as recommended in [3]. All computations were performed on a PC with the Pentium 4 2.6GHz processor and 1GB RAM employing the simplex code of the CPLEX 9.1 package. An attempt to solve the primal model (7) resulted in 2600 seconds of computation (much more than reported in [8]). On the other hand, the dual models (8) were solved in 14.3 to 27.7 CPU seconds on average, depending on the tolerance level (see Table 1).

Table 1: Computational times (in seconds) for the dual CVaR model (averages of 10 instances with 50,000 scenarios)

Tolerance level β	0.05	0.1	0.2	0.3	0.4	0.5
CPU time	14.3	18.7	23.6	26.4	27.4	27.7

The Min-max Portfolio optimization model representing a limiting CVaR model for β tending to 0 is even simpler than the general CVaR model. It can be written as the following LP problem:

$$\begin{aligned} \text{maximize} \quad & \eta \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n \\ & -\eta + \sum_{j=1}^{n} r_{jt} x_j \geq 0, \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{9}$$

Except from the portfolio weights x_j, the model contains only one additional variable η. Nevertheless, it still contains T linear inequalities in addition to the core constraints (1). Hence, its dimensionality is $(T+1) \times (n+1)$. The LP dual model takes then the following form:

$$\begin{aligned} \text{minimize} \quad & q \\ \text{subject to} \quad & q - \sum_{t=1}^{T} r_{jt} u_t \geq 0 \quad \text{for } j = 1, \ldots, n \\ & \sum_{t=1}^{T} u_t = 1 \\ & u_t \geq 0 \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{10}$$

with dimensionality $(n+1) \times (T+1)$. This guarantees a high computational efficiency of the dual model even for vary large number of scenarios. Comparing the model to the dual CVaR model (8) one may notice that upper bounds are skipped. Indeed, the upper bounds p_t/β tend to the infinity with β approaching 0. Similar to the CVaR model, introducing a lower bound on the required expected return in the primal portfolio optimization model (9) result only in a single additional variable in the dual model (10).

The Min-max models are computationally very easy. Running computational test on 10 randomly generated test instances of 50 securities with the number of scenarios 50,000 we were able to solve the dual model (10) in 3.5 seconds on average. Actually, even the primal model (9) could be solved in 7.1 seconds on average, despite its huge number of constraints.

The standard MAD model [6], when implemented with the mean semideviation as the risk measure

($\bar{\delta}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\}$), leads to the following LP problem:

$$\begin{aligned} \text{maximize} \quad & -\sum_{t=1}^{T} p_t d_t \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n \\ & d_t - \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \geq 0, \; d_t \geq 0 \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{11}$$

where nonnegative variables d_t represent downside deviations from the mean under several scenarios t. The above LP formulation, similar to the CVaR model (7), uses $T + n$ variables and $T + 1$ constraints to model the mean semideviation. The LP dual model takes then the following form:

$$\begin{aligned} \text{minimize} \quad & q \\ \text{subject to} \quad & q + \sum_{t=1}^{T} (\mu_j - r_{jt}) u_t \geq 0 \quad \text{for } j = 1, \ldots, n \\ & 0 \leq u_t \leq p_t \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{12}$$

with dimensionality $n \times (T + 1)$ which guarantees the high computational efficiency even for vary large number of scenarios.

The SSD consistent and coherent MAD model with complementary risk measure ($\mu_\delta(\mathbf{x}) = \mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) = \mathbb{E}\{\min\{\mu(\mathbf{x}), R_{\mathbf{x}}\}\}$), leads to the following LP problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{n} \mu_j x_j - \sum_{t=1}^{T} p_t d_t \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n \\ & d_t - \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \geq 0, \; d_t \geq 0 \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{13}$$

where nonnegative variables d_t represent downside deviations from the mean under several scenarios t. The above LP formulation, similar to the CVaR model (7), uses $T + n$ variables and $T + 1$ constraints to model the mean semideviation. The LP dual model takes then the following form:

$$\begin{aligned} \text{minimize} \quad & q \\ \text{subject to} \quad & q + \sum_{t=1}^{T} (\mu_j - r_{jt}) u_t \geq \mu_j \quad \text{for } j = 1, \ldots, n \\ & 0 \leq u_t \leq p_t \quad \text{for } t = 1, \ldots, T \end{aligned} \tag{14}$$

with dimensionality $n \times (T + 1)$ which guarantees the high computational efficiency even for vary large number of scenarios. Indeed, the 10 test problems of 50 securities with the number of scenarios 50,000 we were able to solve the dual model (14) in 25.3 seconds on average.

4 Gini's Mean Difference and Related Models

Yitzhaki [27] introduced the GMD portfolio optimization model using Gini's mean (absolute) difference as risk measure. The *Gini's mean difference* (GMD) is given as $\Gamma(\mathbf{x}) = \frac{1}{2}\int\int|\eta - \xi|dF_{\mathbf{x}}(\eta)dF_{\mathbf{x}}(\xi)$ although several alternative formulae exist. For a discrete random variable represented by its realizations y_t, the measure $\Gamma(\mathbf{x}) = \sum_{t'=1}^{T}\sum_{t''\neq t'-1}\max\{y_{t'} - y_{t''}, 0\}p_{t'}p_{t''}$ is LP computable (when minimized) leading to the following portfolio optimization model:

$$\begin{array}{ll} \max & -\sum_{t=1}^{T}\sum_{t'\neq t} p_t p_{t'} d_{tt'} \\ \text{s.t.} & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1,\ldots,n \\ & d_{tt'} \geq \sum_{j=1}^{n} r_{jt}x_j - \sum_{j=1}^{n} r_{jt'}x_j,\ d_{tt'} \geq 0 \quad \text{for } t,t' = 1,\ldots,T;\ t \neq t' \end{array} \tag{15}$$

which contains $T(T-1)$ nonnegative variables $d_{tt'}$ and $T(T-1)$ inequalities to define them. This generate a huge LP problem even for the historical data case where the number of scenarios is 100 or 200. Actually, as shown with the earlier experiments [7], the CPU time of 7 seconds on average for $T = 52$ has increased to above 30 sec. with $T = 104$ and even more than 180 sec. for $T = 156$. However, similar to the CVaR models, variables $d_{tt'}$ are associated with the singleton coefficient columns. Hence, while solving the dual instead of the original primal, the corresponding dual constraints take the form of simple upper bounds (SUB) which are handled implicitly outside the LP matrix. For the simplest form of the feasible set (1) the dual GMD model takes the following form:

$$\begin{array}{ll} \min & v \\ \text{s.t.} & v - \sum_{t=1}^{T}\sum_{t'\neq t}(r_{jt} - r_{jt'})u_{tt'} \geq 0 \quad \text{for } j = 1,\ldots,n \\ & 0 \leq u_{tt'} \leq p_t p_{t'} \quad \text{for } t,t' = 1,\ldots,T; t \neq t' \end{array} \tag{16}$$

where original portfolio variables x_j are dual prices to the inequalities. The dual model contains $T(T-1)$ variables $u_{tt'}$ but the number of constraints (excluding the SUB structure) $n+1$ is proportional to the number of securities. The above dual formulation can be further simplified by introducing variables:

$$\bar{u}_{tt'} = u_{tt'} - u_{t't} \quad \text{for} \quad \text{for } t,t' = 1,\ldots,T; t < t' \tag{17}$$

which allows us to reduce the number of variables to $T(T-1)/2$ by replacing (16) with the following:

$$\begin{array}{ll} \min & v \\ \text{s.t.} & v - \sum_{t=1}^{T}\sum_{t'>t}(r_{jt} - r_{jt'})\bar{u}_{tt'} \geq 0 \quad \text{for } j = 1,\ldots,n \\ & -p_t p_{t'} \leq \bar{u}_{tt'} \leq p_t p_{t'} \quad \text{for } t,t' = 1,\ldots,T; t < t' \end{array} \tag{18}$$

Such a dual approach may dramatically improve the LP model efficiency in the case of larger number of scenarios. Actually, as shown with the earlier experiments [7], the above dual formulations let us to reduce the optimization time below 10 seconds for $T = 104$ and $T = 156$. Nevertheless, the case of really large number of scenarios still may cause computational difficulties, due to huge number of variables ($T(T-1)/2$). This may require some column generation techniques [4] or nondifferentiable optimization algorithms [8].

As shown by Yitzhaki [27] for the SSD consistency of the GMD model one needs to maximize the complementary measure

$$\mu_\Gamma(\mathbf{x}) = \mu(\mathbf{x}) - \Gamma(\mathbf{x}) = \mathbb{E}\{R_\mathbf{x} \wedge R_\mathbf{x}\} \tag{19}$$

where the cumulative distribution function of $R_\mathbf{x} \wedge R_\mathbf{x}$ for any $\eta \in \mathbb{R}$ is given as $F_\mathbf{x}(\eta)(2 - F_\mathbf{x}(\eta))$. Hence, (19) is the expectation of the minimum of two independent identically distributed random variables (i.i.d.r.v.) $R_\mathbf{x}$ [27] thus representing the *mean worse return*. This provides us with another LP model although it is not more compact than that of (15) and its dual (16). Alternatively, the GMD may be expressed with integral of the absolute Lorenz curve as

$$\Gamma(\mathbf{x}) = 2\int_0^1 (\alpha\mu(\mathbf{x}) - F_\mathbf{x}^{(-2)}(\alpha))d\alpha = 2\int_0^1 \alpha(\mu(\mathbf{x}) - M_\alpha(\mathbf{x}))d\alpha$$

and respectively

$$\mu_\Gamma(\mathbf{x}) = \mu(\mathbf{x}) - \Gamma(\mathbf{x}) = 2\int_0^1 F_\mathbf{x}^{(-2)}(\alpha)d\alpha = 2\int_0^1 \alpha M_\alpha(\mathbf{x})d\alpha \tag{20}$$

thus combining all the CVaR measures. In order to enrich the modeling capabilities, one may treat differently some more or less extreme events. In order to model downside risk aversion, instead of the Gini's mean difference, the *tail Gini's* measure [18, 19] can be used:

$$\mu_{\Gamma_\beta}(\mathbf{x}) = \mu(\mathbf{x}) - \frac{2}{\beta^2}\int_0^\beta (\mu(\mathbf{x})\alpha - F_\mathbf{x}^{(-2)}(\alpha))d\alpha = \frac{2}{\beta^2}\int_0^\beta F_\mathbf{x}^{(-2)}(\alpha)d\alpha \tag{21}$$

In the simplest case of equally probable T scenarios with $p_t = 1/T$ (historical data for T periods), the tail Gini's measure for $\beta = K/T$ may be expressed as the weighted combination of CVaRs $M_{\beta_k}(\mathbf{x})$ with tolerance levels $\beta_k = k/T$ for $k = 1, 2, \ldots, K$ and properly defined weights [19]. In a general case, we may resort to an approximation based on some reasonably chosen grid β_k, $k = 1, \ldots, m$ and weights w_k expressing the corresponding trapezoidal approximation of the integral in the formula (21). Exactly, for any $0 < \beta \leq 1$, while using the grid of m tolerance levels $0 < \beta_1 < \ldots < \beta_k < \ldots < \beta_m = \beta$ one may define weights:

$$w_k = \frac{(\beta_{k+1} - \beta_{k-1})\beta_k}{\beta^2}, \quad \text{for } k = 1, \ldots, m-1, \quad \text{and} \quad w_m = \frac{\beta - \beta_{m-1}}{\beta} \tag{22}$$

where $\beta_0 = 0$. This leads us to the Weighted CVaR (WCVaR) measure [11] defined as

$$M_\mathbf{w}^{(m)}(\mathbf{x}) = \sum_{k=1}^m w_k M_{\beta_k}(\mathbf{x}), \quad \sum_{k=1}^m w_k = 1, \quad w_k > 0 \quad \text{for } k = 1, \ldots, m \tag{23}$$

We emphasize that despite being only an approximation to (21), any WCVaR measure itself is a well defined LP computable measure with guaranteed SSD consistency and coherency, as a combination of the CVaR measures. Hence, it needs not to be built on a very dense grid to provide proper modeling of risk averse preferences. While analyzed on the real-life data from the Milan Stock Exchange the weighted CVaR models have usually performed better than the GMD itself, the Minimax or the extremal CVaR models [11].

Here we analyze only computational efficiency of the LP models representing the WCVaR portfolio optimization. For returns represented by their realizations we get the following LP optimization problem:

$$
\begin{aligned}
\max \quad & \sum_{k=1}^{m} w_k \eta_k - \sum_{k=1}^{m} \frac{w_k}{\beta_k} \sum_{t=1}^{T} p_t d_{tk} \\
\text{s.t.} \quad & \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n \\
& d_{tk} - \eta_k + \sum_{j=1}^{n} r_{jt} x_j \geq 0, \; d_{tk} \geq 0 \quad \text{for } t = 1, \ldots, T; \; k = 1, \ldots, m
\end{aligned}
\tag{24}
$$

where η_k (for $k = 1, \ldots, m$) are unbounded variables taking the values of the corresponding β_k-quantiles (in the optimal solution). Except from the core portfolio constraints (1), model (24) contains T nonnegative variables d_{tk} and T corresponding linear inequalities for each k. Hence, its dimensionality is proportional to the number of scenarios T and to the number of tolerance levels m. Exactly, the LP model contains $m \times T + n$ variables and $m \times T + 1$ constraints. It does not cause any computational difficulties for a few hundreds of scenarios and a few tolerance levels, as in our computational analysis based on historical data. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model (24) with huge number of variables and constraints thus decreasing the computational efficiency of the model. If the core portfolio constraints contain only linear relations, like (1), then the computational efficiency can easily be achieved by taking advantages of the LP dual to model (24). The LP dual model takes the following form:

$$
\begin{aligned}
\text{minimize} \quad & q \\
\text{subject to} \quad & q - \sum_{t=1}^{T} r_{jt} \sum_{k=1}^{m} u_{tk} \geq 0 \quad \text{for } j = 1, \ldots, n \\
& \sum_{t=1}^{T} u_{tk} = w_k \quad \text{for } k = 1, \ldots, m \\
& 0 \leq u_{tk} \leq p_t w_k / \beta_k \quad \text{for } t = 1, \ldots, T; \; k = 1, \ldots, m
\end{aligned}
\tag{25}
$$

The dual LP model contains $m \times T$ variables u_{tk}, but the $m \times T$ constraints corresponding to variables d_{tk} from (24) take the form of simple upper bounds (SUB) on u_{tk} thus not affecting the problem complexity. Actually, the number of constraints in (25) is proportional to the total of portfolio size n and the number of tolerance levels m, thus it is independent from the number of scenarios. Exactly, there are $m \times T + 1$ variables and $m + n$ constraints. This guarantees a high computational efficiency of the dual model even for very large number of scenarios. Similar to the CVaR model, introducing a lower bound on the required expected return in the primal portfolio optimization model (24) result only in a single additional variable in the dual model (25).

We have tested computational efficiency of the dual model (25) using the same 10 randomly generated test instances [8] as for testing the CVaR models. Recall that they were originally generated from a multivariate normal distribution for 50 securities with the number of scenarios 50,000. For $m = 3$ with tolerance levels $\beta_1 = 0.1$, $\beta_2 = 0.25$, $\beta_3 = 0.5$ and weights $w_1 = 0.1$, $w_2 = 0.4$ and $w_3 = 0.5$, thus representing the parameters leading to good results on real life data [11], the dual model (25) was solved in 123.2 seconds on average. For $m = 5$ with uniformly distributed tolerance levels $\beta_1 = 0.1$, $\beta_2 = 0.2$, $\beta_3 = 0.3$, $\beta_4 = 0.4$, $\beta_5 = 0.5$ and weights defined according to (22) the dual model was solved in 296.2 seconds on average The corresponding primal models could not be solved in one hour computations.

5 Concluding Remarks

The classical Markowitz model uses the variance as the risk measure, thus resulting in a quadratic optimization problem. There were introduced several alternative risk measures which are computationally attractive as (for discrete random variables) they result in solving linear programming (LP) problems. The LP solvability is very important for applications to real-life financial decisions where the constructed portfolios have to meet numerous side constraints and take into account transaction costs. A gamut of LP computable risk measures has been presented in the portfolio optimization literature although most of them are related to the absolute Lorenz curve and thereby the CVaR measures. We have shown that all the risk measures used in the LP solvable portfolio optimization models can be derived from the SSD shortfall criteria. This allows to guarantee their SSD consistency for any distribution of outcomes.

The corresponding portfolio optimization models can be solved with general purpose LP solvers. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model with huge number of variables and constraints thus decreasing the computational efficiency of the model. For the CVaR model, the number of constraints (matrix rows) is proportional to the number of scenarios. while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments. We have shown that the computational efficiency can be then dramatically improved with an alternative model taking advantages of the LP duality. In the introduced model the number of structural constraints (matrix rows) is proportional to the number of instruments thus not affecting seriously the simplex method efficiency by the number of scenarios and resulting in computation times below 30 seconds. Moreover, similar reformulation can be applied to more complex quantile risk measures like the Gini's mean difference and the tail Gini's measures as well as to the mean absolute deviation.

Acknowledgment

The authors are indebted to Professor Churlzu Lim from the University of North Carolina at Charlotte for providing the test data.

References

[1] Andersson, F., Mausser, H., Rosen, D., Uryasev, S.: Credit Risk Optimization with Conditional Value-at-Risk Criterion. Math. Programming **89** (2001) 273–291.

[2] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent Measures of Risk. Math. Finance **9** (1999) 203–228.

[3] Barr, D.R., Slezak, N.L.: A Comparison of Multivariate Normal Generators. Communications ACM **15** (1972) 1048–1049.

[4] Desrosiers, J., Luebbecke, M.: A primer in column generation. In: Desaulniers, G., Desrosier, J., Solomon, M., (eds.): Column Generation. Springer (2005) 132.

[5] Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extremal Events for Insurance and Finance. Springer, New York, 1997.

[6] Konno, H., Yamazaki, H.: Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market. Management Science **37** (1991) 519–531.

[7] Krzemienowski, A., Ogryczak, W.: On extending the LP computable risk measures to account downside risk. Comput. Optimization and Appls., **32**, (2005), 133–160.

[8] Lim, C., Sherali, H.D., Uryasev, S.: Portfolio Optimization by Minimizing Conditional Value-at-Risk Via Nondifferentiable Optimization. University of North Carolina at Charlotte, Lee College of Engineering, Woking Paper, 2007.

[9] Mansini, R., Ogryczak, W., Speranza, M.G.: On LP Solvable Models for Portfolio Selection. Informatica **14** (2003) 37–62.

[10] Mansini, R., Ogryczak, W., Speranza, M.G.: LP Solvable Models for Portfolio Optimization: A Classification and Computational Comparison. IMA Journal of Management Mathematics **14** (2003) 187–220.

[11] Mansini, R., Ogryczak, W., Speranza, M.G.: Conditional Value at Risk and Related Linear Programming Models for Portfolio Optimization. Annals Oper. Res. **152** (2007) 227–256.

[12] Markowitz, H.M.: Portfolio selection. J. Fin. **7** (1952) 77–91.

[13] Müller, A., Stoyan, D.: Comparison Methods for Stochastic Models and Risks. Wiley, Chichester, 2002.

[14] Ogryczak, W.: Stochastic dominance relation and linear risk measures. In: *Financial Modelling – Proc. 23rd Meeting EURO WG Financial Modelling, Cracow, 1998,* Skulimowski, A.M.J. (ed.), Progress & Business Publ., Cracow, 1999, 191–212.

[15] Ogryczak, W.: Risk measurement: Mean absolute deviation versus Ginis mean difference. In: Decision Theory and Optimization in Theory and Practice – Proc. 9th Workshop GOR WG, Chemnitz, 1999. Wanka, W.G. (ed.), Shaker Verlag, Aachen, 2000, 33–51.

[16] Ogryczak, W.: Multiple criteria linear programming model for portfolio selection. Annals Oper. Res. **97** (2000) 143–162.

[17] Ogryczak, W., Ruszczyński, A.: From stochastic dominance to mean-risk models: semideviations as risk measures. Eur. J. Opnl. Res. **116** (1999) 33–50.

[18] Ogryczak, W., Ruszczyński, A.: Dual stochastic dominance and related mean-risk models. SIAM J. Optimization **13** (2002) 60–78.

[19] Ogryczak, W., Ruszczyński, A.: Dual Stochastic Dominance and Quantile Risk Measures. Intnl. Trans. Opnl. Res. **9** (2002) 661–680.

[20] Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In: Probabilistic Constrained Optimization: Methodology and Applications. Uryasev, S. (ed.), Kluwer AP, Dordrecht, 2000.

[21] Pflug, G.Ch.: Scenario tree generation for multiperiod financial optimization by optimal discretization. Math. Programming **89** (2001) 251-271.

[22] Rockafellar, R.T., Uryasev, S.: Optimization of Conditional Value-at-Risk. J. Risk **2** (2000) 21–41.

[23] Rothschild, M., Stiglitz, J.E.: Increasing risk: I. A definition. J. Econ. Theory **2** (1969) 225–243.

[24] Scheuer, E.M., Stoller, D.S.: On the Generation of Normal Random Vectors. Technometrics **4** (1962) 278–281.

[25] Sharpe, W.F.: Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios. Manage. Sci. **18** (1971) B1–B13.

[26] Shorrocks, A.F.: Ranking Income Distributions. Economica **50** (1983) 3–17.

[27] Yitzhaki, S.: Stochastic dominance, mean variance, and Gini's mean difference. American Econ. Rev. **72** (1982) 178–185.

[28] Young, M.R.: A minimax portfolio selection rule with linear programming solution. Manage. Sci. **44** (1998) 673–683.