

Chapter 10

Sparse and Multiple Risk Measures Approach for Data Driven Mean-CVaR Portfolio Optimization Model

Jianjun Gao and Weiping Wu

Abstract This paper studies the out-of-sample performance of the data driven Mean-CVaR portfolio optimization(DDMC) model, in which the historical data of the stock returns are regarded as the realized returns and used directly in the mean-CVaR portfolio optimization formulation. However, in practical portfolio management, due to a limited number of monthly or weekly based historical data, the out-of-sample performance of the DDMC model is quite unstable. To overcome such a difficulty, we propose to add the penalty on the sparsity of the portfolio weight and combine the variance term in the DDMC formulation. Our experiments demonstrate that the proposed method mitigates the fragility of out-of-sample performance of the DDMC model significantly.

Keywords Conditional value-at-risk • Portfolio optimization • Multiple risk measures • Sparse portfolio • Out-of-sample stability

10.1 Introduction

The mean-variance(MV) portfolio selection model proposed by Markowitz (1952) laid the foundation of the modern investment theory. It suggests to balance the profit and the risk in portfolio decision. Following the spirit of Markowitz's MV model, the framework of mean-risk portfolio analysis has been extended in various directions, e.g., see Li et al. (2006), Kolm et al. (2014), Gao and Li (2013) and the references therein. However, using variance as the risk measure has some drawbacks, i.e., it

J. Gao (✉)

School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai, People's Republic of China
e-mail: gao.jianjun@shufe.edu.cn

W. Wu

Department of Automation, Shanghai Jiao Tong University, Shanghai, People's Republic of China
e-mail: godream@sjtu.edu.cn

© Springer International Publishing AG 2017

T.-M. Choi et al. (eds.), *Optimization and Control for Systems in the Big-Data Era*,
International Series in Operations Research & Management Science 252,
DOI 10.1007/978-3-319-53518-0_10

167

penalizes both profit and loss of the random return symmetrically. Realizing the variance is not a perfect term for risk measure, a large amount of new risk measures have been proposed since the development of the MV portfolio selection model. Among these risk measures, the *Value-at-Risk* (VaR), defined as the quantile of a specified exceeding probability of the loss, becomes popular in the financial industry since the mid-90s. However, the VaR fails to satisfy the axiomatic system of coherent risk measures proposed by Artzner et al. (1999), and it suffers from the non-convexity property in the corresponding portfolio optimization problems. On the other hand, the conditional Value-at-Risk (CVaR), defined as the expected value of the loss exceeding the VaR (Rockafellar and Uryasev 2000, 2002), possesses several good properties, such as convexity, monotonicity and homogeneity, which also proved to be in the class of coherent risk measures (Pflug 2000; Artzner et al. 1999). Rockafellar and Uryasev (2000, 2002) developed an equivalent formulation to compute the CVaR which leads to a convex optimization problem. Due to these nice properties, CVaR has been widely applied in various applications of portfolio selection and risk management, e.g., derivative portfolio (Alexander et al. 2006), credit risk optimization (Andersson et al. 2001), and robust portfolio management (Zhu and Fukushima 2009).

Although the mean-risk portfolio optimization model has been studied extensively in the academic society, translating these models as some useful tools in the real world financial practice is not a trivial task. Even for the classical MV portfolio selection model, it is well known that estimating the expected return and covariance matrix are not an easy task, especially when the size of the portfolio is large (e.g., see Merton 1980; Demiguel et al. 2009a,b). Highly related to the estimation problem of the stock return statistics, the stableness of the out-of-sample performance of the portfolio optimization model is another issue. Demiguel et al. (2009b) checked several portfolio construction methods rooted from the MV portfolio selection formulation. However, these models cannot significantly or consistently outperform the naive portfolio strategy which allocates wealth evenly in all assets. As for the mean-CVaR portfolio optimization model, since CVaR measures just a small portion of the whole distribution, a large number of samples is needed to guarantee the statistical stability. Takeda and Kanamori (2009) and Kondor et al. (2007) showed that the mean-CVaR portfolio optimization model has more serious problems of instability regarding the out-of-sample performance than the MV model. Recently, Lim et al. (2011) reported the similar results that the correspondent portfolio of the mean-CVaR portfolio decision model is extremely unreliable due to the estimation errors. Furthermore, Lim et al. (2011) showed that this problem is even worse when the distribution of the return has a heavy tail. To deal with unstable out-of-sample performance of the mean-CVaR portfolio optimization model, several methods have been proposed. Gotoh and Takeda (2011) introduced the norm-regularity in the mean-risk portfolio decision model to reduce the sparsity of the portfolio decision. Gotoh et al. (2013) further adopted the robust mean-CVaR portfolio optimization technique to overcome such an instability problem.

Motivated by the above research (Lim et al. 2011; Gotoh and Takeda 2011; Gotoh et al. 2013), we propose to use the sparse portfolio and multiple risk measures to

mitigate the fragility of the CVaR based data driven portfolio selection model. More specifically, we add the l_1 -norm penalty of the portfolio decision vector and the variance of the portfolio return in the mean-CVaR portfolio selection model. To enhance the sparsity of the solution, we also adopt the reweighted- l_1 norm method by computing the weights iteratively. Our numerical experiments show that the resulted out-of-sample performance is significantly enhanced comparing with the traditional DDMC portfolio optimization model.

This paper is organized as follows. The alternative formulations of the DDMC portfolio optimization problems are proposed in Sect. 10.2. The out-of-sample performance of these different models is evaluated by using the simulation approach in Sect. 10.3. The paper is concluded in Sect. 10.4.

10.2 The Data Driven Mean-Risk Portfolio Optimization

We consider a portfolio constructed by n candidate risky assets, whose random returns are denoted as $\mathbf{R} \in \mathbb{R}^n$. Let $\mathbf{x} = (x_1, \dots, x_n)' \in \mathbb{R}^n$ be the portfolio decision vector, which represents the weight of the allocation of the wealth in each securities. Let $f(\mathbf{x}, \mathbf{R})$ be the portfolio loss associated with \mathbf{x} and \mathbf{R} , e.g., we can simply set $f(\mathbf{x}, \mathbf{R}) = b - \mathbf{R}'\mathbf{x}$, where b is the benchmark return. To define the CVaR of the loss $f(\mathbf{x}, \mathbf{R})$ for a given confidence level β (i.e., $\beta = 95\%$), we need the cumulative distribution function of $f(\mathbf{x}, \mathbf{R})$,

$$\Psi(y) = \mathbb{P}(f(\mathbf{x}, \mathbf{R}) \leq y),$$

for some number $y \in \mathbb{R}$, the corresponding β -tail distribution for a given confidence level β is

$$\Psi_\beta(y) = \begin{cases} 0, & \text{if } y < \text{VaR}_\beta, \\ \frac{\Psi(y) - \beta}{1 - \beta}, & \text{if } y \geq \text{VaR}_\beta, \end{cases} \quad (10.1)$$

where $\text{VaR}_\beta = \inf\{z \mid \Psi(y) \geq \beta\}$. The CVaR of the loss function $f(\mathbf{x}, \mathbf{R})$ is then given by

$$\text{CVaR}[f(\mathbf{x}, \mathbf{R})] := \int_{f(\mathbf{x}, \mathbf{R}) \geq \text{VaR}_\beta} f(\mathbf{x}, \mathbf{R}) d\Psi_\beta(y), \quad (10.2)$$

where the integration should be understood as a summation when \mathbf{R} is a discrete random vector. Note that the above definition of CVaR is for the general distribution function of the loss function $f(\mathbf{x}, \mathbf{R})$, see, e.g., Rockafellar and Uryasev (2002) for some subtle difference on the definition of the CVaR between the cases of discrete random variable and continuous random variable. Rockafellar and Uryasev (2000) and Rockafellar and Uryasev (2002) showed that the $\text{CVaR}[f(\mathbf{x}, \mathbf{R})]$ can be computed by solving a simple convex optimization problem.

Lemma 2.1 *The CVaR of the loss $f(\mathbf{x}, \mathbf{R})$ of the terminal wealth can be computed as follows:*

$$\text{CVaR}[f(\mathbf{x}, \mathbf{R})] = \min_{\alpha} \left\{ \alpha + \frac{1}{1-\beta} \mathbb{E}[(f(\mathbf{x}, \mathbf{R}) - \alpha)^+] \right\},$$

where α is an auxiliary variable and $(y)^+ := \max\{y, 0\}$.

Let $\mathbf{D} = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ be the data set of the historical returns, where $\mathbf{r}_i \in \mathbb{R}^n$ is the i -th sample of the returns and m is the number of the samples we can observe. Without loss of generality, we assume \mathbf{r}_i and \mathbf{r}_j to be independent for any $i, j \in \{1, \dots, m\}$. The data set \mathbf{D} can also be regarded as m realizations of the random return \mathbf{R} . From Lemma 2.1, if we fix the loss function as $f(\mathbf{R}, \mathbf{x}) = b - \mathbf{R}'\mathbf{x}$, the data driven mean-CVaR portfolio optimization model is given as follows:

$$(\mathcal{P}_1) \quad \min_{\mathbf{x}} \quad \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^m (b - \mathbf{r}'_i \mathbf{x})^+ \quad (10.3)$$

$$\text{Subject to : } \sum_{i=1}^n x_i = 1, \quad (10.4)$$

$$\frac{1}{m} \sum_{i=1}^m \mathbf{r}'_i \mathbf{x} \geq d, \quad (10.5)$$

where d is a pre-given target return level. By introducing some auxiliary variables, problem (\mathcal{P}_1) can be reformulated as a linear programming problem. To overcome the instability of the out-of-sample performance of the DDMC model (\mathcal{P}_1) , we propose to use the following model $(\mathcal{P}_2(\omega))$ with some given weighting vector $\omega \in \mathbb{R}^n$,

$$(\mathcal{P}_2(\omega)) : \quad \min_{\mathbf{x}} \quad \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^m (b - \mathbf{r}'_i \mathbf{x})^+ + \|\mathbf{x}\|_1^{\omega}, \quad (10.6)$$

Subject to : \mathbf{x} satisfies (10.4) and (10.5),

where $\omega = (\omega_1, \dots, \omega_n)'$ with $\omega_i \geq 0$, for $i = 1, \dots, n$ and

$$\|\mathbf{x}\|_1^{\omega} := \sum_{i=1}^n \omega_i |x_i|.$$

When ω is a unit vector with all elements being 1, the weighted l_1 -norm formulation becomes the l_1 -norm formulation, which is denoted by $\|\mathbf{x}\|_1$. Using the l_1 norm as the penalty for the sparsity of the solution is a standard routine in data analysis. The ideal penalty of the sparsity of the solution is l_0 norm, which is defined as

$\|\mathbf{x}\|_0 = \sum_{i=1}^n |\text{Sign}(x_i)|$ with $\text{Sign}(a) = 1$ if $a > 0$, $\text{Sign}(a) = -1$ if $a < 0$ and $\text{Sign}(a) = 0$ if $a = 0$. However, the l_0 norm is highly nonconvex and hard to be optimized directly. It has been proved that the l_1 norm of \mathbf{x} , $\|\mathbf{x}\|_1$, is the convex hull of $\|\mathbf{x}\|_0$ (see Zhao and Li 2012). Thus, it is reasonable to use l_1 norm as the surrogate of l_0 norm to penalize the sparsity. In model $(\mathcal{P}_2(\omega))$, we prefer to use the formulation of weighted- l_1 norm, which further enhances the sparsity by varying the choice of vector ω . Note that problem $(\mathcal{P}_2(\omega))$ can be reformulated as a linear programming problem,

$$(\bar{\mathcal{P}}_2(\omega)) : \min_{\mathbf{x}, \tau, \phi} \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^m \tau_i + \sum_{j=1}^n \phi_j,$$

Subject to :

$$\tau_i \geq 0, \quad i = 1, \dots, m,$$

$$b - \mathbf{r}'_i \mathbf{x} \leq \tau_i, \quad i = 1, \dots, m,$$

$$\omega_j x_j \leq \phi_j, \quad j = 1, \dots, n,$$

$$\omega_j x_j \geq -\phi_j, \quad j = 1, \dots, n,$$

$$\sum_j^n x_j = 1,$$

$$\frac{1}{m} \sum_j^n \mathbf{r}'_j \mathbf{x} \geq d,$$

where τ_i for $i = 1, \dots, m$ and ϕ_j for $j = 1, \dots, n$ are auxiliary decision variables.

In this work, we also consider to integrate the variance term of the portfolio return in model $(\mathcal{P}_2(\omega))$ to further enhance the stability of the out-of-sample performance, i.e.,

$$(\mathcal{P}_3(\omega)) : \min_{\mathbf{x}} \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^m (b - \mathbf{r}'_i \mathbf{x})^+ + \|\mathbf{x}\|_1^\omega + \rho \mathbf{x}' F \mathbf{x} \quad (10.7)$$

Subject to : \mathbf{x} satisfies (10.4) and (10.5),

where $F \in \mathbb{R}^{n \times n}$ is the sample covariance matrix of the asset returns. Note that, similar to problem $(\mathcal{P}_2(\omega))$, problem $(\mathcal{P}_3(\omega))$ can be reformulated as a convex quadratic programming formulation, which can be solved efficiently by a commercial solver like IBM CPLEX (IBM 2015).

10.3 Evaluation and Discussion

10.3.1 Evaluation Methods

To evaluate the out-of-sample performance of the three portfolio optimization models (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$, we mainly adopt the simulation approach with all parameters being estimated from the real historical price data of some stock index. The main reason of using this approach is as follows. The number of the historical data of the monthly return is very limited in real portfolio management. Thus, it is hard to carry on various tests by solely using the true market historical data. On the other hand, by using the simulation approach, different types of test data sets can be generated, which provides us more freedom to evaluate the performances of the three models under different situations. More specifically, we adopt the following procedures.

- (a) **Data Generation:** Generate a data set of returns $\mathbf{D}_{\text{sample}} = \{\mathbf{r}_1, \dots, \mathbf{r}_m\}$ with a sample size being m according to some distributions of the returns.¹ For example, if we assume the random returns follow a mixed distribution of multivariate normal distribution and exponential distribution with given mean vector and covariance matrix, we then generate m samples of the returns according to this distribution.
- (b) **Optimization:** Solve all three problems (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ according to the data set $\mathbf{D}_{\text{sample}}$ to generate the portfolio decisions \mathbf{x}^1 , \mathbf{x}^2 and \mathbf{x}^3 , respectively. If it is necessary, we can vary the target return level d in three models to achieve the portfolio policy $\mathbf{x}^i(d)$, $i = 1, 2, 3$, for different level of d .
- (c) **Evaluation:** Generate 50 data set $\mathbf{D}_{\text{test}}^{(i)}$, $i = 1, \dots, 50$ according to the similar distribution used in step **Data Generation** with the size of the each data set $\mathbf{D}_{\text{test}}^{(i)}$ being m . For each test set $\mathbf{D}_{\text{test}}^{(i)}$, we implement the portfolio policy $\mathbf{x}^i(d)$, $i = 1, 2, 3$ and compute the corresponding empirical expected return and CVaR.

In step **Evaluation**, we actually perform 50 trials of out-of-sample tests and the resulted empirical sample expected return and CVaR are recorded. In each iteration, we use the IBM CPLEX (IBM 2015) as the solver to solve the correspondent linear programming and convex quadratic programming problems of (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$.

¹The detailed discussion of the distribution is given in Sect. 10.3.2.

10.3.2 Data Generation

In this paper, we use the 48 industry portfolios constructed by Fama and French as the basic data set for our test.² We estimate the mean return vector and covariance matrix of monthly return by using the historical monthly returns from Jan 1998 to Dec 2015. Note that there are only 216 samples of the returns, however, we need to estimate 1176 unknown parameters in the covariance matrix,³ which implies that using the sample covariance matrix method may generate a singular matrix. To overcome this difficulty, we adopt the shrinkage estimation method for the covariance matrix proposed by Ledoit and Wolf (2003) by setting the shrinkage coefficient to 0.1. After we have achieved the sample mean vector of the returns, $\hat{\mathbf{R}} := (\hat{R}_1, \dots, \hat{R}_n)'$ and the estimation of the covariance matrix $\hat{\Sigma} := \{\Sigma_{i,j}\}_{i=1,j=1}^{n,n}$, we then use the following method to generate the samples. Adopting a similar setting given by Lim et al. (2011), we construct a hybrid distribution combining the multivariate normal distribution and the exponential distribution. Let $B(\eta)$ be the Bernoulli random variable with parameter η , i.e., $B(\eta) = 1$ with probability η and $B(\eta) = 0$ with probability $1 - \eta$. Let z be the exponential random variable with the probability distribution function being

$$\mathbb{P}(z < a) = \int_0^a \lambda e^{-\lambda s} ds.$$

In this paper, we simply fix $\lambda = 10$. Suppose the random vector $Y \in \mathbb{R}^n$ follows the multivariate normal distribution with mean and covariance matrix being $\hat{\mathbf{R}}$ and $\hat{\Sigma}$, respectively. We assume the random return is captured by the hybrid distribution as follows:

$$\mathbf{R} \sim -B(\eta)(z\mathbf{1} + \mathbf{1}) + (1 - B(\eta))Y,$$

where $\mathbf{c} := (c_1, \dots, c_n)'$ with $c_i := \hat{R}_i - \sqrt{\Sigma_{ii}}$ for $i = 1, \dots, n$ and Σ_{ii} is the i -th diagonal element of Σ . Note that the parameter η controls the tail-loss of the distribution, i.e., the larger the η is, the heavier tail of the distribution will be. Figure 10.1 gives the distribution of one entry of \mathbf{R} for different η .

²The data of 48 industry portfolio can be found in <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data>.

³Since the covariance matrix is symmetrical, we only need to estimate the upper triangle of the matrix. Thus, the total number of unknown parameters is $(48 + 1) \times 48/2 = 1176$.

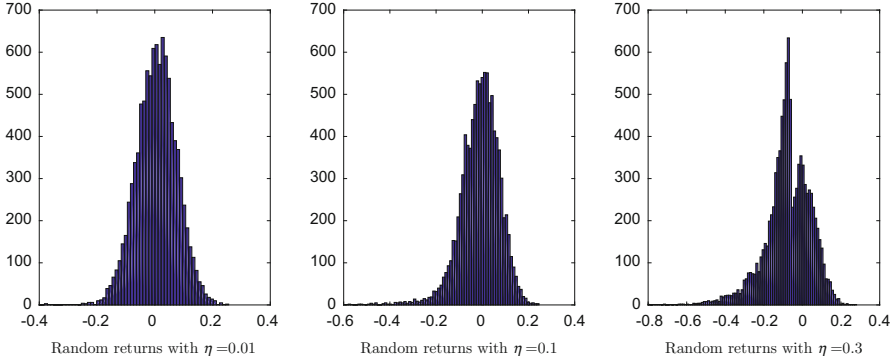


Fig. 10.1 The empirical distribution of hybrid random returns R_1 with different value of η

10.3.3 Re-Weighted Method for Sparse Solution

In portfolio optimization models $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$, we use the weighted- l_1 norm to penalize the sparsity of the solution. However, since the objective function is a weighted summation of the CVaR and the weighted- l_1 norm of the portfolio weight, we need to choose the weighting parameter ω carefully. If $\|\omega\|$ is too large, the optimality of the CVaR will be jeopardized. On the other hand, if $\|\omega\|$ is too small, then the resulted solution will be not sparse enough. To overcome this difficulty, we adopt the iterative reweighted method of the l_1 norm to enhance the sparsity of the solution(see, e.g., Zhao and Li 2012). More specifically, we apply the following iterative procedure to change the weighting parameter ω dynamically and adaptively. Let $\omega^{(k)} \in \mathbb{R}^n$ and $\mathbf{x}^{(k)}$ be the weighting vector and portfolio decision vector in k -th iteration, respectively. We repeat the following steps.

- (1) For any given $\omega^{(k)}$, solve the problem $\mathcal{P}_2(\omega^{(k)})$ (or problem $(\mathcal{P}_3(\omega^{(k)}))$), which gives the solution $\mathbf{x}^{(k)}$. If the stopping criteria is satisfied, e.g., the sparsity of $\mathbf{x}^{(k)}$ does not change any more, we stop the iteration. Otherwise, go to step II.
- (2) Use $\mathbf{x}^{(k)}$ to construct the new weighting parameter $\omega^{(k+1)}$ and let $k = k + 1$. Go to step 1.

There are several ways to construct the new weighting vector $\omega^{(k+1)} = (\omega_1^{(k+1)}, \omega_2^{(k+1)}, \dots, \omega_n^{(k+1)})'$ by using the information of $\mathbf{x}^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})'$. Motivated by Zhao and Li (2012) and based on our numerical experiments, we select the following three methods which perform relatively better than the others. Let $\epsilon > 0$ be a small positive number.

- (a) Method I: Let $\omega_j^{(k+1)} = 1/(|x_j^{(k)}| + \epsilon)$ for $j = 1, \dots, n$.
- (b) Method II: Let $\omega_j^{(k+1)} = 1/(|x_j^{(k)}| + \epsilon)^{(1-p)}$, for $j = 1, \dots, n$ and $p \in (0, 1)$.
- (c) Method III: Let $\omega_j^{(k+1)} = (p + (|x_i^{(k)}| + \epsilon)^{1-p}) / ((|x_i^{(k)}| + \epsilon)^{1-p} [|x_i^{(k)}| + \epsilon + (|x_j^{(k)}| + \epsilon)^p])$ for $j = 1, \dots, n$ with $p \in (0, 1)$.

It is not hard to see that when x_i^k is a small number then the corresponding weighting coefficient ω_i^{k+1} will be large, which will drive x_i^{k+1} to be even smaller in the next round of optimization.

10.3.4 Comparison of the Global Mean-CVaR Portfolio

In this section, we compare the out-of-sample performance of the three models (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ for the special case of finding the *global minimum CVaR* portfolio. More specifically, we consider the problems with ignoring the constraint (10.5) in all three models (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$. Following the evaluation procedure illustrated in Sect. 10.3.1, we generate one data set $\mathbf{D}_{\text{sample}}$ to compute the correspondent portfolio weights and apply such portfolio decision in 50 testing data sets $\mathbf{D}_{\text{test}}^{(j)}$ for $j = 1, \dots, 50$ as the out-of-sample tests. We check three different types of size of $\mathbf{D}_{\text{sample}}$ and $\mathbf{D}_{\text{test}}^{(j)}$ as $m = 200$, $m = 300$ and $m = 400$.

Figures 10.2, 10.3 and 10.4 plot the empirical mean value and CVaR of the global minimum CVaR portfolio return generated from 50 out-of-sample tests. We can observe that the empirical mean and CVaR pair spread in a quite large range for model (\mathcal{P}_1) . However, by using our proposed models $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$, we can see that the range of the resulted empirical mean and CVaR pair are significantly reduced. Table 10.1 records the detail of the above experiments. The column ‘min’, ‘max’ and ‘range’ show the minimum value, the maximum value and the range(i.e., ‘max’-‘min’) of the corresponding data set, respectively. For the case $m = 200$, the minimum and maximum value of resulted mean and CVaR of model (\mathcal{P}_1) is from -0.0037 to 0.541 and 0.2634 to 0.5943 , respectively. That is to say, the relative difference of the out-of-sample CVaR and mean value are 0.33 and 0.0578 for model (\mathcal{P}_1) . In the same row of $m = 200$, we can observe that this range is reduced to 0.1756 and 0.0343 , respectively, for model $(\mathcal{P}_2(\omega))$ and reduce to 0.1394 and

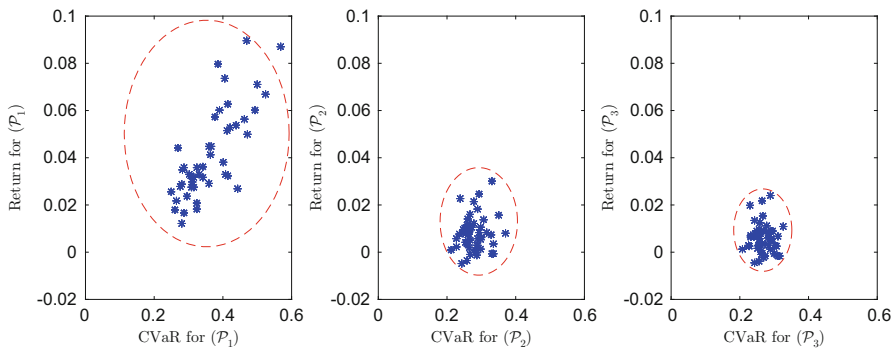


Fig. 10.2 The out-of-sample performance of three models with sample size $m = 200$ and $\epsilon = 0.1$

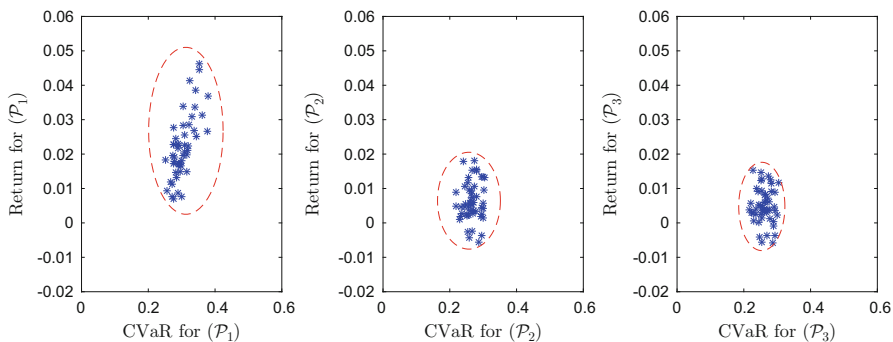


Fig. 10.3 The out-of-sample performance of three models with sample size $m = 300$ and $\epsilon = 0.1$

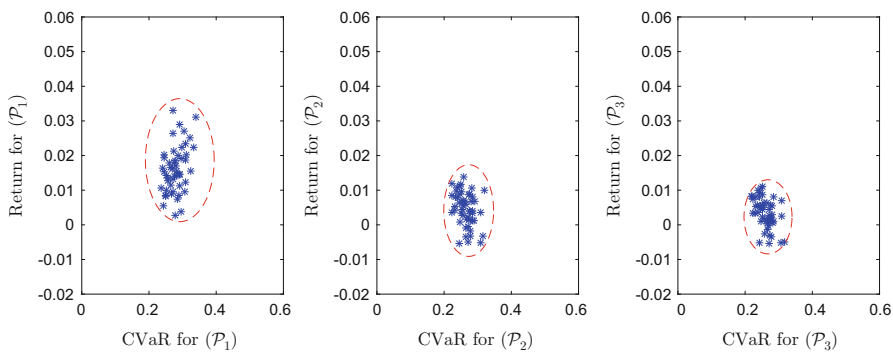


Fig. 10.4 The out-of-sample performance of three models with sample size $m = 400$ and $\epsilon = 0.1$

Table 10.1 The empirical mean value and CVaR of portfolio returns for global minimum CVaR problem generated by different models with $\eta = 0.1$

m	(\mathcal{P}_1)			$\mathcal{P}_2(\omega)$			$\mathcal{P}_3(\omega)$		
	Min	Max	Range	Min	Max	Range	Min	Max	Range
CVaR ($\times 10^{-2}$)									
$m = 200$	26.34	59.43	33.09	20.61	38.17	17.56	19.30	33.24	13.94
$m = 300$	25.25	38.00	12.75	21.80	30.43	8.63	20.63	28.85	8.22
$m = 400$	23.70	34.03	10.33	22.30	32.09	9.79	20.75	30.24	9.48
Exp return ($\times 10^{-2}$)									
$m = 200$	-0.37	5.41	5.78	-1.09	2.34	3.43	-1.13	1.79	2.92
$m = 300$	-0.32	3.93	4.26	-0.57	1.80	2.37	-0.55	1.45	2.00
$m = 400$	-0.35	2.56	2.91	-0.55	1.38	1.93	-0.50	1.05	1.55

Table 10.2 The empirical mean value and CVaR of portfolio returns for global minimum CVaR problem generated by different models with $\eta = 0.2$

m	(\mathcal{P}_1)			$\mathcal{P}_2(\omega)$			$\mathcal{P}_3(\omega)$		
	Mean	Max	Range	Min	Max	Range	Min	Max	Range
CVaR ($\times 10^{-2}$)									
$m = 200$	33.57	118.67	85.10	25.67	46.31	20.65	24.23	40.95	16.72
$m = 300$	30.64	48.43	17.79	27.20	36.74	9.54	23.78	34.89	11.11
$m = 400$	26.99	39.05	12.05	25.62	33.03	7.41	23.89	30.88	6.99
Exp return ($\times 10^{-2}$)									
$m = 200$	-0.79	14.62	15.41	-0.50	3.56	4.06	-0.49	3.07	3.56
$m = 300$	0.24	5.22	4.99	-0.54	3.19	3.73	-0.68	2.43	3.11
$m = 400$	0.19	5.10	4.90	-0.27	2.01	2.27	-0.42	1.75	2.18

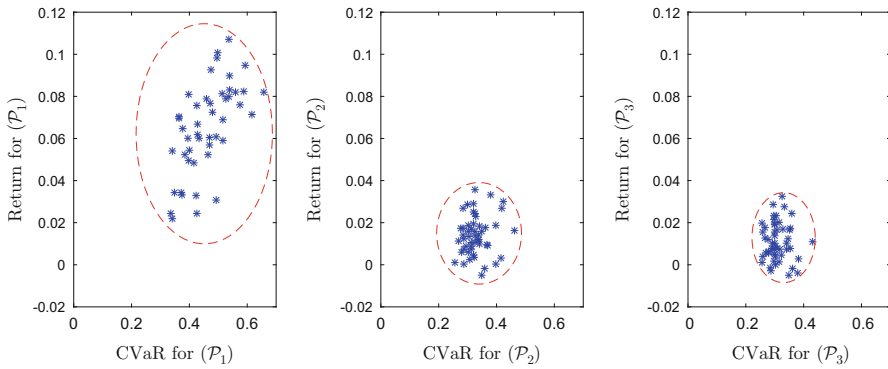


Fig. 10.5 The out-of-sample performance of three models with sample size $m = 200$ and $\epsilon = 0.2$

0.0292, respectively, for model $(\mathcal{P}_3(\omega))$. From Table 10.1, we can see that, as the size of the sample increases, e.g., the case of $m = 300$ and $m = 400$, the variation of the resulted empirical mean and CVaR of model (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ are reduced. However, the performance of the models $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ is better than model (\mathcal{P}_1) .

Table 10.2 and Figs. 10.5, 10.6 and 10.7 show the detailed results of the comparison between the three models when $\eta = 0.2$. As we have illustrated in Sect. 10.3.2, the parameter η controls the shape of the tail distribution of the random returns. Under this case, the stock returns have heavier tails comparing with the previous case with $\eta = 0.1$. However, a similar pattern can be observed that the formulation $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ can better control the variation of the empirical mean return and CVaR.

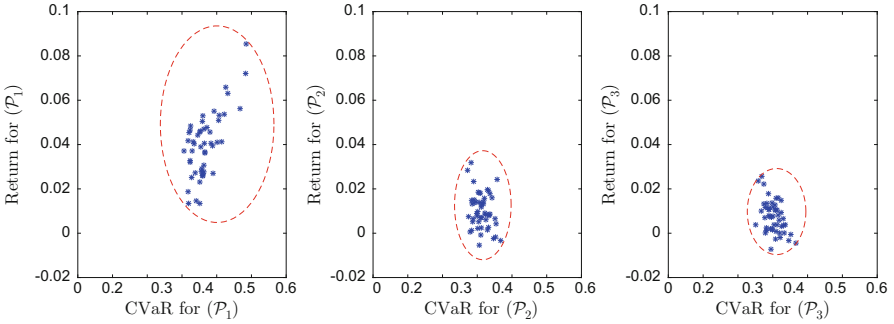


Fig. 10.6 The out-of-sample performance of three models with sample size $m = 300$ and $\epsilon = 0.2$

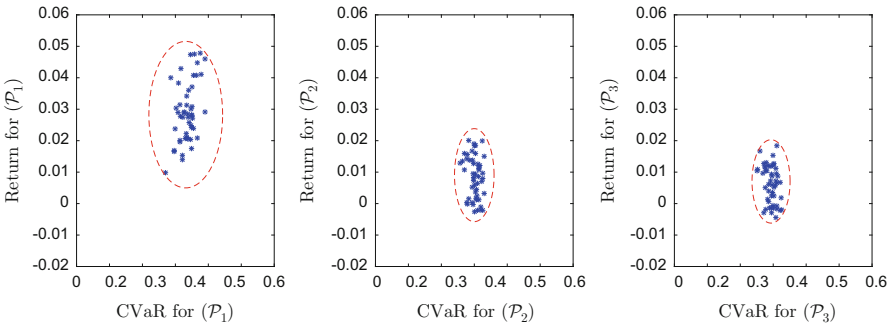


Fig. 10.7 The out-of-sample performance of three models with sample size $m = 400$ and $\epsilon = 0.2$

10.3.5 Comparison of the Empirical Efficient Frontiers

In this section, we compare the mean-CVaR efficient frontiers generated by three models (\mathcal{P}_1) , $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$. The efficient frontiers are generated by varying the target return d from 0.01 to 0.1 in all these models. Figures 10.8, 10.9 and 10.10 plot the out-of-sample empirical mean-CVaR efficient frontier for 50 trials of simulations with $\eta = 0.1$. Table 10.3 shows the detailed statistics of the comparison. In Table 10.3, the columns ‘min dev’, ‘max dev’ and ‘mean dev’ represent the minimum deviation, maximum deviation and average deviation of the out-of-sample CVaR and expected return.⁴ Note that the minimum, maximum and average deviation is computed for all different value of d in 50 trials of simulation.

⁴Given some random samples a_1, \dots, a_m , the maximum, minimum and average deviation is defined as $\max\{|a_i - \bar{a}| \mid i = 1, \dots, m\}$, $\min\{|a_i - \bar{a}| \mid i = 1, \dots, m\}$ and $\frac{1}{m} \sum_{i=1}^m (|a_i - \bar{a}|)$, where $\bar{a} = \frac{1}{m} \sum_{i=1}^m a_i$.

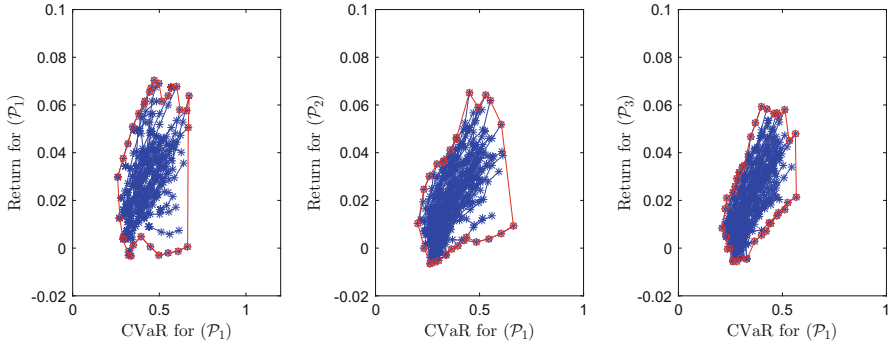


Fig. 10.8 The out-of-sample performance of three models with sample size $m = 200$ and $\epsilon = 0.2$

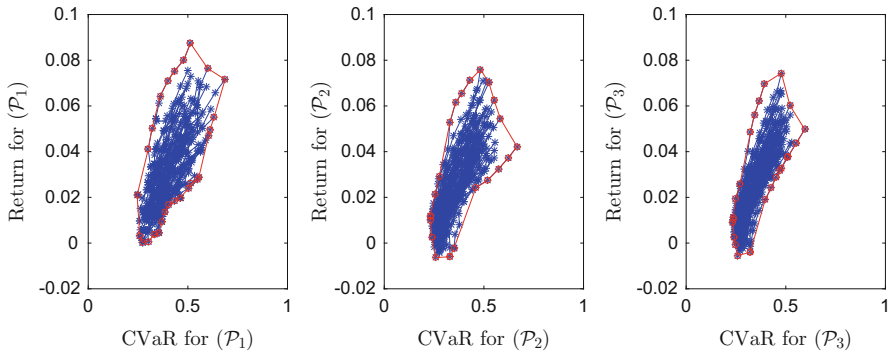


Fig. 10.9 The out-of-sample performance of three models with sample size $m = 300$ and $\epsilon = 0.2$

For all of these tests, we can observe that the proposed formulations $(\mathcal{P}_2(\omega))$ and $(\mathcal{P}_3(\omega))$ perform better than the traditional model (\mathcal{P}_1) . For example, in the row of $m = 200$ in Table 10.3, the maximum deviation of three models are 19.98%, 17.35% and 12.54%, respectively. The average deviation of three models are 5.04%, 2.99%, and 2.68%, respectively. Similar pattern can be observed when we increase the tail part of the distribution of the random return. Figures 10.11, 10.12, and 10.13 and Table 10.4 provide the detail of the improvement under this case.

10.4 Conclusion

In this work, we proposed some methods to reduce the instability issue of the out-of-sample performance for mean-CVaR portfolio optimization model. More specifically, we suggest to add the weighted l_1 norm as a penalty of the sparsity

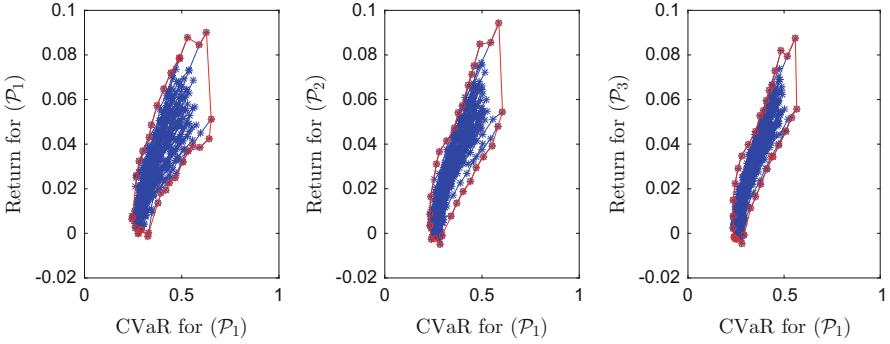


Fig. 10.10 The out-of-sample performance of three models with sample size $m = 400$ and $\epsilon = 0.2$

Table 10.3 Comparison of mean-CVaR efficient frontiers for different models with $\eta = 0.1$

m	\mathcal{P}_1			$\mathcal{P}_2(\omega)$			$\mathcal{P}_3(\omega)$		
	Min dev	Max dev	Mean dev	Min dev	Max dev	Mean dev	Min dev	Max dev	Mean dev
CVaR ($\times 10^{-2}$)									
$m = 200$	0.04	19.98	5.04	0.01	17.35	2.99	0.00	12.54	2.68
$m = 300$	0.01	16.14	3.52	0.01	19.15	2.82	0.00	15.68	2.48
$m = 400$	0.01	13.95	2.30	0.00	13.33	2.12	0.00	12.49	2.04
Exp return ($\times 10^{-2}$)									
$m = 200$	0.00	4.56	1.11	0.00	3.21	0.83	0.00	2.87	0.80
$m = 300$	0.00	3.72	0.85	0.00	2.94	0.67	0.00	3.12	0.64
$m = 400$	0.00	3.36	0.67	0.00	4.04	0.55	0.00	3.68	0.51

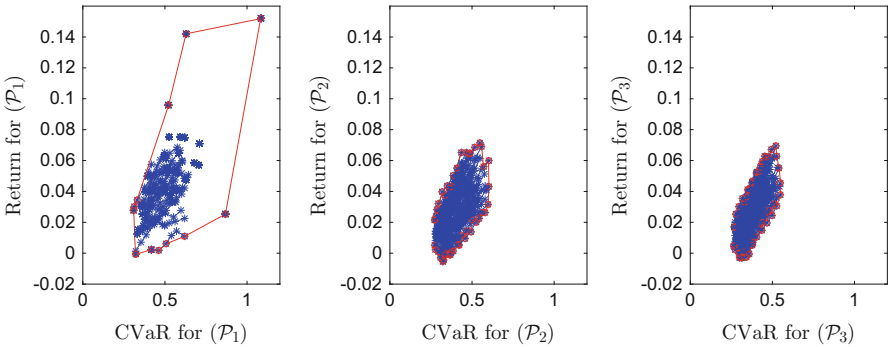


Fig. 10.11 The out-of-sample performance of three models with sample size $m = 200$ and $\epsilon = 0.2$

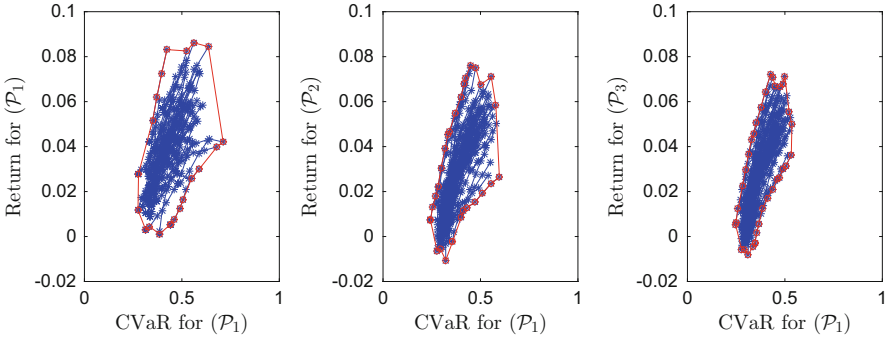


Fig. 10.12 The out-of-sample performance of three models with sample size $m = 300$ and $\epsilon = 0.2$

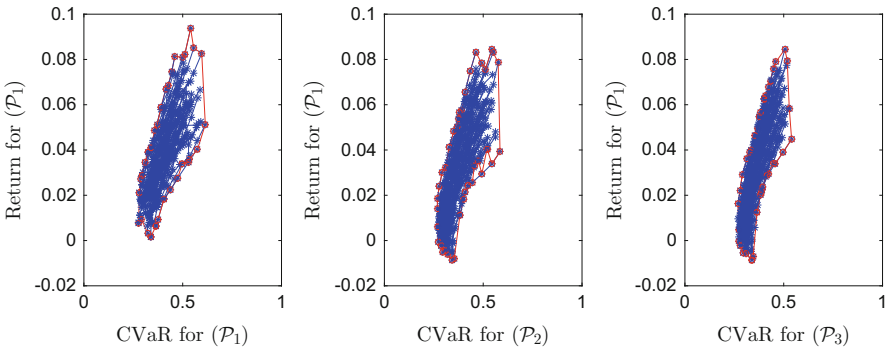


Fig. 10.13 The out-of-sample performance of three models with sample size $m = 400$ and $\epsilon = 0.2$

of the portfolio decision and add the variance term in the objective function to control the total variation in mean-CVaR portfolio formulation. In order to balance the sparsity and optimality of the solution, the reweighted l_1 norm method is adopted to adjust the weighting coefficients. Our simulation based experiments show that the proposed methods reduce the variation of the empirical mean value and the CVaR of the portfolio return in out-of-sample test significantly. However, observing from our experiment, the proposed methods still have some limitations. When the size of the portfolio is large, e.g., when $n = 500$, solely using our methods may not control the variation of the out-of-sample test to a desired level. A possible solution for this case is to increase the number of the samples by using some statistical sampling methods like bootstrap. Another important issue is the computational burden of the proposed methods when n and m are large. For example, for problem $(\mathcal{P}_2(\omega))$, the linear programming formulation (given in Sect. 10.2) has almost $m + 2n$ decision variables and $2(m + n)$ constraints. In the literature, Kunzi-Bay and Janos (2006) have showed that using the dual formulation and decomposition approach may

Table 10.4 Comparison of mean-CVaR efficient frontiers for different models with $\eta = 0.2$

m	\mathcal{P}_1			$\mathcal{P}_2(\omega)$			$\mathcal{P}_3(\omega)$		
	Min dev	Max dev	Mean dev	Min dev	Max dev	Mean dev	Min dev	Max dev	Mean dev
CVaR ($\times 10^{-2}$)									
$m = 200$	0.01	58.24	9.41	0.01	12.73	3.90	0.00	9.46	3.00
$m = 300$	0.01	17.88	4.06	0.00	11.02	2.66	0.01	9.67	2.42
$m = 400$	0.01	10.49	2.68	0.00	10.32	2.54	0.01	8.71	2.31
Exp return ($\times 10^{-2}$)									
$m = 200$	0.00	10.63	1.83	0.00	3.03	0.91	0.00	2.98	0.81
$m = 300$	0.00	3.13	0.93	0.00	2.97	0.71	0.00	2.31	0.68
$m = 400$	0.00	3.25	0.73	0.00	2.96	0.69	0.00	2.30	0.66

enhance the efficiency of the solution procedure. All the models considered in this work belong to the static portfolio optimization formulation, which gives the buy-and-hold type of portfolio policy. Studying the stability issue of the out-of-sample test for multiperiod mean-CVaR portfolio optimization problem is an interesting and challenging topic.

Acknowledgements This research work was partially supported by National Natural Science Foundation of China under grant 71201102 and 61573244.

References

- S. Alexander, T. Coleman, Y. Li, Minimizing cvar and var for portfolio of derivatives. *J. Bank. Financ.* **30**, 583–605 (2006)
- F. Andersson, H. Mausser, D. Rosen, S. Uryasev, Credit risk optimization with conditional value at risk criterion. *Math. Program. Ser. B* **89**, 273–291 (2001)
- P. Artzner, F. Delbaen, J. Eber, D. Heath, Coherent measure of risk. *Math. Financ.* **9**, 203–228 (1999)
- V. Demiguel, L. Garlappi, F.J. Nogales, R. Uppal, A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Manag. Sci.* **55**(5), 798–812 (2009a)
- V. Demiguel, L. Garlappi, R. Uppal, Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? *Rev. Financ. Stud.* **22**(5), 1915–1953 (2009b)
- J. Gao, D. Li, Optimal cardinality constrained portfolio selection. *Oper. Res.* **61**, 745–761 (2013)
- J. Gotoh, A. Takeda, On the role of norm constraints in portfolio selection. *Comput. Manag. Sci.* **8**, 323–353 (2011)
- J. Gotoh, K. Shinozaki, A. Takeda, Robust portfolio techniques for mitigating the fragile of cvar minimization and generalization to coherent risk measures. *Quant. Financ.* **13**(10), 1621–1635 (2013)
- IBM, Reference Manual CPLEX 12.5, IBM, USA, (2015)
- P.N. Kolm, R. Tütüncü, F. Fabozzi, 60 years of portfolio optimization: practical challenges and current trends. *Eur. J. Oper. Res.* **234**(2), 356–371 (2014)

- I. Kondor, S. Pafka, G. Nagy, Noise sensitivity of portfolio selection under various risk measures. *J. Bank. Financ.* **31**, 1545–1573 (2007)
- A. Kunzi-Bay, M. Janos, Computational aspects of minimizing conditional value-at-risk. *Comput. Manag. Sci.* **3**, 3–27 (2006)
- O. Ledoit, M. Wolf, Improved estimation of the covariance matrix of stock return with an application to portfolio selection. *J. Empir. Financ.* **10**, 603–621 (2003)
- D. Li, X. Sun, J. Wang, Optimal lot solution to cardinality constrained mean-variance formulation for portfolio selection. *Math. Financ.* **16**, 83–101 (2006)
- A. Lim, J.G. Shanthikumar, G. Vahn, Conditional value-at-risk in portfolio optimization: coherent but fragile. *Oper. Res. Lett.* **39**, 163–171 (2011)
- H. Markowitz, Portfolio selection. *J. Financ.* **7**, 77–91 (1952)
- R.C. Merton, On estimating the expected return on the market: an exploratory investigation. *J. Financ. Econ.* **8**, 43–54 (1980)
- G. Pflug, Some remarks on the value-at-risk and conditional value-at-risk, in *Probabilistic Constrained optimization: Methodology and Applications* (Springer, Boston, 2000), pp. 272–281
- R. Rockafellar, S. Uryasev, Optimization of conditional value-at-risk. *J. Risk* **2**, 21–41 (2000)
- R. Rockafellar, S. Uryasev, Conditional value-at-risk for general loss distributions. *J. Bank. Financ.* **26**, 1443–1471 (2002)
- A. Takeda, T. Kanamori, A robust approach based on conditional value-at-risk measure to statistical learning problems. *Eur. J. Oper. Res.* **198**(1), 287–296 (2009)
- Y. Zhao, D. Li, Reweighted ℓ_1 -norm minimization for sparse solutions to underdetermined linear systems. *SIAM J. Optim.* **22**(3), 1065–1088 (2012)
- S.S. Zhu, M. Fukushima, Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.* **57**, 1155–1168 (2009)