

# A Newton-Based Perturbation Method for a Robust Inverse Optimization Problem

Zhiqiang JIA<sup>1</sup>, Jian GU<sup>2,\*</sup>, Xiantao XIAO<sup>1</sup>

1. *School of Mathematical Sciences, Dalian University of Technology, Liaoning 116024, P. R. China;*
2. *School of Sciences, Dalian Ocean University, Liaoning 116023, P. R. China*

**Abstract** In this paper, we aim to solve an inverse robust optimization problem, in which the parameters in both the objective function and the robust constraint set need to be adjusted as little as possible so that a known feasible solution becomes the optimal one. We formulate this inverse problem as a minimization problem with a linear equality constraint, a second-order cone complementarity constraint and a linear complementarity constraint. A perturbation approach is constructed to solve the inverse problem. An inexact Newton method with Armijo line search is applied to solve the perturbed problem. Finally, the numerical results are reported to show the effectiveness of the approach.

**Keywords** inverse optimization; robust linear programming; perturbation approach; inexact Newton method

**MR(2010) Subject Classification** 90C30; 90C33

## 1. Introduction

A typical optimization problem is usually called a forward problem, in which all parameters are given and we need to find an optimal solution from all feasible solutions. However, there are many instances in practice, in which we only have some estimates for parameter values, but we may know certain optimal solutions from experience, observations or experiments. An inverse optimization problem is to find values of parameters which make the known solutions optimal and which differ from the given estimates as little as possible.

Burton and Toint [1] first investigated an inverse shortest path problem. Since then, a number of inverse combinatorial optimization problems have been studied, see the survey paper [2] and the references therein. Recently, several inverse continuous optimization problems have been studied, among which are [3] for inverse linear programming, [4,5] for inverse quadratic programming, [6–8] for inverse semidefinite quadratic programming and [9] for inverse second-order programming.

To the best of our knowledge, there are very few studies to discuss the inverse robust optimization in the literature, except that in [10] the authors used a robust inverse optimization

---

Received August 30, 2015; Accepted September 23, 2016

Supported by the National Natural Science Foundation of China (Grant No. 11571059) and the Fundamental Research Funds for the Central Universities (Grant No. DUT16LK30).

\* Corresponding author

E-mail address: gujian@dlou.edu.cn (Jian GU)

framework to handle a portfolio problem. Inspired by [10], in this paper we study a kind of inverse robust conic optimization problem, which is more general and numerically challenged than that of [10]. We formulate this inverse problem as a linear conic complementarity problem and apply a perturbation approach to solve it. By solving a sequence of perturbed subproblems, we obtain an approximate solution of the inverse robust conic optimization problem.

Firstly, we show that the study of inverse robust optimization problem is significant in practice by considering a portfolio allocation problem with value-at-risk (VaR) constraint in the financial industry. Consider a market with  $n$  risky assets and one riskless asset where investors seek to maximize their expected return asset subject to a threshold level of risk. If the risk is taken by the standard deviation of returns, the investors would solve the well-known Markowitz portfolio allocation problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & u^T x + (1 - e^T x)r_f, \\ \text{s.t.} \quad & \sqrt{x^T \Sigma x} \leq L, \\ & e^T x \leq 1, x \geq 0, \end{aligned} \tag{1}$$

where  $r \in \mathbb{R}^n$  is the random vector of the risky asset returns,  $u = \mathbb{E}[r]$  is the vector of mean asset returns,  $\Sigma \in \mathbb{R}^{n \times n}$  is the covariance matrix of asset returns,  $r_f \in \mathbb{R}_+$  is the return on the riskless asset,  $x \in \mathbb{R}^n$  is the fraction of wealth invested in each risky asset,  $L$  is an investor-specific threshold level of risk and  $e \in \mathbb{R}^n$  refers to the vector of all ones.

Alternative measures of risk have been suggested in risk management, such as value-at-risk (VaR) and conditional value-at-risk (CVaR). Given a random variable  $Z$ , its value-at-risk is defined by

$$\text{VaR}_\alpha(Z) := \inf\{t \in \mathbb{R} | \mathbb{P}(t + Z \geq 0) \leq 1 - \alpha\} \text{ for any } \alpha \in (0, 1).$$

The corresponding optimization problem has the form

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & u^T x + (1 - e^T x)r_f, \\ \text{s.t.} \quad & \text{VaR}_\alpha((r - r_f e)^T x) \leq L, \\ & e^T x \leq 1, x \geq 0. \end{aligned} \tag{2}$$

For normally distributed random variables, VaR is proportional to the standard deviation. For discrete distributions,  $\text{VaR}_\alpha(Z)$  is a nonconvex, discontinuous function. A popular alternative risk that maintains the convexity is conditional value-at-risk (CVaR), which is a *coherent* risk measure. Therefore, the corresponding portfolio allocation problem is in the form of

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & u^T x + (1 - e^T x)r_f, \\ \text{s.t.} \quad & \text{CVaR}_\alpha((r - r_f e)^T x) \leq L, \\ & e^T x \leq 1, x \geq 0. \end{aligned} \tag{3}$$

Consider the following robust linear conic programming problem:

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & u^T x + (1 - e^T x)r_f \\ \text{s.t.} \quad & (r - r_f e)^T x \geq -L, \quad \forall r \in \mathcal{U}, \\ & e^T x \leq 1, x \geq 0 \end{aligned} \tag{4}$$

for some set  $\mathcal{U}$ . The following results show that the portfolio allocation problems (1), (2) and (3) can be casted into the framework of (4).

**Lemma 1.1** ([10, Proposition 3]) *Consider the following uncertainty sets:*

$$\begin{aligned} \mathcal{U}_1 &= \{r : (r - r_f e)^T \Sigma^{-1} (r - r_f e) \leq 1\}, \\ \mathcal{U}_2 &= \{r : (r - u + r_f e)^T \Sigma^{-1} (r - u + r_f e) \leq z_\alpha^2\}, \\ \mathcal{U}_3 &= \{r : (r - u + r_f e)^T \Sigma^{-1} (r - u + r_f e) \leq 2\pi\alpha^2 e^{-z_\alpha^2/2}\}, \end{aligned}$$

where  $z_\alpha$  is the  $\alpha$ -quantile.

(i) Problem (4) with  $\mathcal{U} = \mathcal{U}_1$  is equivalent to the Markowitz problem (1).

If  $r$  is distributed as a multivariate Gaussian,  $r \sim N(u, \Sigma)$ :

(ii) Problem (4) with  $\mathcal{U} = \mathcal{U}_2$  is equivalent to the VaR problem (2).

(iii) Problem (4) with  $\mathcal{U} = \mathcal{U}_3$  is equivalent to the CVaR problem (3).

One of the primary difficulties in solving portfolio allocation problems is finding a stable estimation to the expectation  $u$  of returns and the investor-specific threshold level of risk  $L$ . In practice, we can assume that  $\hat{x}$  is an optimal solution observed from the stock market,  $u_0$  is a given historical return and  $L_0$  is an estimate level from the investors. Let  $\Phi(\hat{x})$  denote the set of all  $(u, L)$  which make  $\hat{x}$  optimal. Therefore, we could find a good estimate by solving the following problem

$$\begin{aligned} \min \quad & \|(u, L) - (u_0, L_0)\|^2, \\ \text{s.t.} \quad & (u, L) \in \Phi(\hat{x}), \\ & u \in \mathbb{R}^n, L \in \mathbb{R}_+, \end{aligned}$$

which is a typical robust inverse optimization problem.

In this paper, we consider a type of robust linear conic programming (RLCP) problem of the form

$$\begin{aligned} \text{(RLCP)} \quad \min \quad & c^T x, \\ \text{s.t.} \quad & Ax \geq d, \\ & r^T x \geq b, \forall r \in U, \end{aligned}$$

where  $U := \{r | \exists v \in \mathbb{R}^n \text{ such that } Fr + Gv - g \in Q_{m+1}\}$ ,  $A \in \mathbb{R}^{l \times n}$ ,  $F \in \mathbb{R}^{(m+1) \times n}$ ,  $G \in \mathbb{R}^{(m+1) \times n}$  and  $Q_{m+1}$  is a second-order cone defined by

$$Q_{m+1} := \{s = (s_0; \bar{s}) \in \mathbb{R} \times \mathbb{R}^m | s_0 \geq \|\bar{s}\|\},$$

with  $\|\cdot\|$  being the Euclidean norm. It is easy to verify that the portfolio allocation problems (1), (2) and (3) can be rewritten in the form of Problem (RLCP).

Given a point  $x^0$ , which is supposed to be an optimal solution to Problem (RLCP) and a pair  $(c^0, b^0) \in \mathbb{R}^n \times \mathbb{R}$  which is an estimate to  $(c, b)$ . The inverse robust linear programming problem considered in this paper is to find  $(c, b) \in \mathbb{R}^n \times \mathbb{R}$  to solve the problem

$$\begin{aligned} \min \quad & \frac{1}{2} \|(c, b) - (c^0, b^0)\|^2, \\ \text{s.t.} \quad & x^0 \in \text{SOL}(\text{RLCP}), \end{aligned} \tag{5}$$

where SOL(RLCP) is the set of optimal solutions to Problem (RLCP).

The following notations and results on second-order cone will be used throughout the paper. Let  $Q_{m+1}$  be a second-order cone defined to be

$$Q_{m+1} := \{s = (s_0; \bar{s}) \in \mathbb{R} \times \mathbb{R}^m \mid s_0 \geq \|\bar{s}\|\}.$$

The topological interior part and the boundary of  $Q_{m+1}$  denoted by  $\text{int}Q_{m+1}$  and  $\text{bd}Q_{m+1}$ , respectively, are given by

$$\begin{aligned} \text{int}Q_{m+1} &:= \{s = (s_0; \bar{s}) \in \mathbb{R} \times \mathbb{R}^m \mid s_0 > \|\bar{s}\|\}, \\ \text{bd}Q_{m+1} &:= \{s = (s_0; \bar{s}) \in \mathbb{R} \times \mathbb{R}^m \mid s_0 = \|\bar{s}\|\}. \end{aligned}$$

For any  $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^m$  and  $y = (y_1; y_2) \in \mathbb{R} \times \mathbb{R}^m$ , we define their Jordan product as

$$x \circ y = (x^T y; y_1 x_2 + x_1 y_2).$$

We write  $x^2$  to mean  $x \circ x$ . For any  $x = (x_1; x_2) \in \mathbb{R} \times \mathbb{R}^m$ , we define its determinant as  $\det(x) = x_1^2 - \|x_2\|^2$  and the linear mapping  $L_x$  from  $\mathbb{R}^{m+1}$  to  $\mathbb{R}^{m+1}$  as

$$L_x y = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I_m \end{bmatrix} y,$$

where  $I_m \in \mathbb{R}^{m \times m}$  is an identity matrix. It can be easily verified that  $x \circ y = L_x y$ , and  $L_x$  is positive definite if and only if  $x \in \text{int}Q_{m+1}$ . Also, we have

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \frac{\det(x)}{x_1} I_m + \frac{x_2 x_2^T}{x_1} \end{bmatrix}.$$

In order to avoid any confusion, in the sequel we let  $\hat{x}$  denote that it is a vector in a second-order cone  $Q_{m+1}$ . We denote  $I_n$  and  $O_n$  as the identity matrix and zero matrix in  $\mathbb{R}^{n \times n}$ . For two matrices  $A$  and  $B$ , we write  $A \succeq B$  ( $A \succ B$ ) to mean that the matrix  $A - B$  is positive semidefinite (positive definite). For a differentiable mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  and a vector  $x \in \mathbb{R}^n$ , let  $\mathcal{J}F(x)$  be the Jacobian of  $F$  at  $x$  and  $\nabla F(x) := \mathcal{J}F(x)$ .

The rest of this paper is organized as follows. In Section 2, we first formulate the robust linear programming problem as a linear programming, then reformulate the inverse robust linear programming problem as a minimization problem with a linear equality constraint, a second-order cone complementarity constraint and a linear complementarity constraint. In Section 3, we use a perturbation approach to solve the inverse problem. An inexact Newton method is applied to solve the perturbed problem. The numerical results are shown in Section 4.

## 2. Problem reformulation

Recall that  $U = \{r \mid \exists v \in \mathbb{R}^n \text{ such that } Fr + Gv - g \in Q_{m+1}\}$ . According to the duality theory, we can obtain

$$\min_{r \in U} r^T x = \max_{\hat{p}} \{\hat{p}^T g \mid F^T \hat{p} = x, G^T \hat{p} = 0_n, \hat{p} \in Q_{m+1}\}.$$

Denote

$$H = [g \quad FA^T \quad G \quad -G]^T \in \mathbb{R}^{(2n+l+1) \times (m+1)},$$

$$q = [b \quad d^T \quad 0_n^T \quad 0_n^T]^T \in \mathbb{R}^{2n+l+1},$$

and  $\hat{u} = Fc - H^T\lambda$  and  $v = H\hat{p} - q$ . By using the classical theory of duality and optimality conditions, Problem (5) can be rewritten as

$$\begin{aligned} \min \quad & \frac{1}{2} \| (c, b) - (c^0, b^0) \|^2, \\ \text{s.t.} \quad & F^T \hat{p} = x^0, \\ & e_1^T q = b, \\ & Fc - H^T \lambda = \hat{u}, \\ & H\hat{p} - q = v, \\ & \hat{u} \circ \hat{p} = 0, \\ & 0 \leq \lambda \perp v \geq 0, \\ & \hat{u} \in Q_{m+1}, \hat{p} \in Q_{m+1}. \end{aligned} \tag{6}$$

First, we consider a subproblem of (6) as follows

$$\begin{aligned} \min \quad & \frac{1}{2} \| (c, b) - (c^0, b^0) \|^2, \\ \text{s.t.} \quad & e_1^T q = b, \\ & Fc - H^T \lambda = \hat{u}, \\ & H\hat{p} - q = v, \end{aligned} \tag{7}$$

which is a convex program parameterized by  $(\hat{p}, \hat{u}, \lambda, v)$ . Since the Slater constraint qualification is satisfied, by the classical duality theory, there is no duality gap between problem (7) and its dual. Moreover, for the remainder of the paper, we assume that  $F$  is of full column rank, so we can get its optimal value  $f(\hat{p}, \hat{u}, \lambda, v)$ , i.e.,

$$f(\hat{p}, \hat{u}, \lambda, v) = \frac{1}{2} \| (F^T F)^{-1} F^T (\hat{u} - H^T \lambda) - c^0 \|^2 + \frac{1}{2} \| e_1^T (H\hat{p} - v) - b^0 \|^2.$$

Therefore, problem (6) can be equivalently expressed as

$$\begin{aligned} \min \quad & f(\hat{p}, \hat{u}, \lambda, v), \\ \text{s.t.} \quad & F^T \hat{p} = x^0, \\ & \hat{u} \circ \hat{p} = 0, \\ & 0 \leq \lambda \perp v \geq 0, \\ & \hat{u} \in Q_{m+1}, \hat{p} \in Q_{m+1}. \end{aligned} \tag{8}$$

The above problem is a minimization problem with a vector complementarity constraint and a second-order cone complementarity constraint, which is a special type of mathematical program with complementarity constraints (MPCC). For vector MPCCs, many algorithms have been proposed [11]. For second-order cone MPCCs, there are only a few references. The smoothing Newton method for second-order cone MPCCs was studied in [12,13] and a perturbation approach was proposed in [9].

From the duality theory of convex programming, if there exists an optimal solution  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)$  to Problem (8), then we have that

$$(c^*, b^*) = ((F^T F)^{-1} F^T (\hat{u}^* - H^T \lambda^*), e_1^T (H\hat{p}^* - v^*))$$

is an optimal solution to the original problem (5).

### 3. A perturbation approach

In this section, we discuss how to solve Problem (8). Due to the complementarity constraints in this problem, the basic constraint qualification does not hold, then the KKT conditions may fail at any local minimizer [11]. To overcome this difficulty, we choose a smoothing function  $\varphi_\varepsilon(\hat{p}, \hat{u}) = 0$  to approximate the second-order cone complementarity relation  $\hat{u} \circ \hat{p} = 0$ ,  $\hat{u} \in Q_{m+1}$ ,  $\hat{p} \in Q_{m+1}$ , where  $\varphi_\varepsilon(\hat{p}, \hat{u})$  is defined by

$$\varphi_\varepsilon(\hat{p}, \hat{u}) = \hat{p} + \hat{u} - \sqrt{(\hat{u} - \hat{p})^2 + 4\varepsilon^2 e}$$

with  $e = (1, 0, \dots, 0) \in \mathbb{R}^{m+1}$  and  $\varepsilon > 0$ . It is obvious that

$$\lim_{\varepsilon \searrow 0} \varphi_\varepsilon(\hat{p}, \hat{u}) = \varphi_0(\hat{p}, \hat{u})$$

and  $\varphi_0(\hat{p}, \hat{u}) = 0$  if and only if  $\hat{u} \circ \hat{p} = 0$ ,  $\hat{u} \in Q_{m+1}$ ,  $\hat{p} \in Q_{m+1}$ . Similarly, we adopt a smoothing function  $\Psi_\varepsilon(\lambda, v) = 0$  to approximate the linear complementarity relation  $0 \leq \lambda$ ,  $0 \leq v$ ,  $\langle \lambda, v \rangle = 0$ , where  $\Psi_\varepsilon(\lambda, v)$  is defined by

$$\Psi_\varepsilon(\lambda, v) = \begin{bmatrix} \psi_\varepsilon(\lambda_1, v_1) \\ \vdots \\ \psi_\varepsilon(\lambda_{2n+l+1}, v_{2n+l+1}) \end{bmatrix},$$

where  $\psi_\varepsilon(\lambda_i, v_i) = \lambda_i + v_i - \sqrt{\lambda_i^2 + v_i^2 + 2\varepsilon^2}$ ,  $i = 1, \dots, 2n + l + 1$ . Obviously,  $\Psi_0(\lambda, v) = 0$  is equivalent to  $0 \leq \lambda$ ,  $0 \leq v$ ,  $\langle \lambda, v \rangle = 0$ .

Then, we construct a perturbation problem of (8) with parameter  $\varepsilon > 0$  as follows

$$\begin{aligned} \min \quad & f(\hat{p}, \hat{u}, \lambda, v), \\ \text{s.t.} \quad & F^T \hat{p} = x^0, \\ \text{(P}_\varepsilon) \quad & \varphi_\varepsilon(\hat{p}, \hat{u}) = 0, \\ & \Psi_\varepsilon(\lambda, v) = 0. \end{aligned} \tag{9}$$

The function  $\varphi_\varepsilon(\hat{p}, \hat{u})$  and  $\Psi_\varepsilon(\lambda, v)$  are continuously differentiable with respect to  $\hat{p}$ ,  $\hat{u}$ ,  $\lambda$  and  $v$  when  $\varepsilon > 0$ .

Define the following sets

$$\begin{aligned} \Omega_1^0 &= \{(\hat{p}, \hat{u}) \in Q_{m+1} \times Q_{m+1} \mid \hat{u} \circ \hat{p} = 0, F^T \hat{p} = x^0\}, \\ \Omega_2^0 &= \{(\lambda, v) \in \mathbb{R}_+^{2n+l+1} \times \mathbb{R}_+^{2n+l+1} \mid \langle \lambda, v \rangle = 0\}, \\ \Omega_1(\varepsilon) &= \{(\hat{p}, \hat{u}) \in Q_{m+1} \times Q_{m+1} \mid \varphi_\varepsilon(\hat{p}, \hat{u}) = 0, F^T \hat{p} = x^0\}, \\ \Omega_2(\varepsilon) &= \{(\lambda, v) \in \mathbb{R}_+^{2n+l+1} \times \mathbb{R}_+^{2n+l+1} \mid \Psi_\varepsilon(\lambda, v) = 0\}, \\ \Omega^0 &= \Omega_1^0 \times \Omega_2^0, \\ \Omega(\varepsilon) &= \Omega_1(\varepsilon) \times \Omega_2(\varepsilon). \end{aligned}$$

Clearly, the sets  $\Omega(\varepsilon)$  and  $\Omega^0$  are the feasible sets of Problem (P<sub>ε</sub>) and Problem (8), respectively.

The following lemmas show that the convergence of the set-value mapping  $\Omega(\varepsilon)$  at  $\varepsilon = 0$  with respect to  $\mathbb{R}_+$  implies that the feasible set of  $(P_\varepsilon)$  converges to the feasible set of the original problem (8).

**Lemma 3.1** ([9, Proposition 4.1]) *Let  $\Omega_1^0$  and  $\Omega_1(\varepsilon)$  be defined as above. Then we have*

$$\lim_{\varepsilon \searrow 0} \Omega_1(\varepsilon) = \Omega_1^0.$$

**Lemma 3.2** ([3, Lemma 2.3]) *Let  $\Omega_2^0$  and  $\Omega_2(\varepsilon)$  be defined as above. Then we have*

$$\lim_{\varepsilon \searrow 0} \Omega_2(\varepsilon) = \Omega_2^0.$$

**Lemma 3.3** *Let  $\Omega(\varepsilon)$  and  $\Omega^0$  be defined as above. Then we have*

$$\lim_{\varepsilon \searrow 0} \Omega(\varepsilon) = \Omega^0.$$

**Proof** Noting that both  $\Omega(\varepsilon)$  and  $\Omega^0$  are the Cartesian product of finite sets, so, according to Lemmas 3.1 and 3.2, the conclusion obviously holds.

Let  $\hat{w} = \sqrt{(\hat{u} - \hat{p})^2 + 4\varepsilon^2 e}$ . For any  $\hat{p} \in Q_{m+1}$  and  $\hat{u} \in Q_{m+1}$ , we get  $\hat{w}^2 = (\hat{u} - \hat{p})^2 + 4\varepsilon^2 e \in \text{int}Q_{m+1}$ , hence  $\hat{w} \in \text{int}Q_{m+1}$ . According to [14, Lemma 3.5], we have that

$$L_{\hat{w}} - L_{\hat{u}-\hat{p}} \succ 0, \quad L_{\hat{w}} + L_{\hat{u}-\hat{p}} \succ 0, \quad L_{\hat{w}} \succ 0.$$

Then, we obtain the Jacobian of  $\varphi_\varepsilon$  as follows,

$$\mathcal{J}_{(\hat{p}, \hat{u})} \varphi_\varepsilon(\hat{p}, \hat{u}) = [I_{m+1} + L_{\hat{w}}^{-1} L_{\hat{u}-\hat{p}} \quad I_{m+1} - L_{\hat{w}}^{-1} L_{\hat{u}-\hat{p}}].$$

Define

$$\Phi_\varepsilon(\hat{p}, \hat{u}, \lambda, v) := \begin{bmatrix} F^T \hat{p} - x^0 \\ \varphi_\varepsilon(\hat{p}, \hat{u}) \\ \Psi_\varepsilon(\lambda, v) \end{bmatrix}.$$

**Lemma 3.4**  $\mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}, \hat{u}, \lambda, v)$  is of full row rank.

**Proof** Note that  $\varphi_\varepsilon(\hat{p}, \hat{u})$  and  $\Psi_\varepsilon(\lambda, v)$  are continuously differentiable with respect to  $\hat{p}$ ,  $\hat{u}$ ,  $\lambda$ , and  $v$  when  $\varepsilon > 0$ . Thus, we obtain

$$\mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}, \hat{u}, \lambda, v) = \begin{bmatrix} F^T & 0 & 0 & 0 \\ \mathcal{J}_{\hat{p}} \varphi_\varepsilon(\hat{p}, \hat{u}) & \mathcal{J}_{\hat{u}} \varphi_\varepsilon(\hat{p}, \hat{u}) & 0 & 0 \\ 0 & 0 & \mathcal{J}_\lambda \Psi_\varepsilon(\lambda, v) & \mathcal{J}_v \Psi_\varepsilon(\lambda, v) \end{bmatrix},$$

where

$$\mathcal{J}_\lambda \Psi_\varepsilon(\lambda, v) = \begin{bmatrix} 1 - \frac{\lambda_1}{\sqrt{\lambda_1^2 + v_1^2 + 2\varepsilon^2}} & & & \\ & \ddots & & \\ & & 1 - \frac{\lambda_{2n+l+1}}{\sqrt{\lambda_{2n+l+1}^2 + v_{2n+l+1}^2 + 2\varepsilon^2}} & \end{bmatrix},$$

and

$$\mathcal{J}_v \Psi_\varepsilon(\lambda, v) = \begin{bmatrix} 1 - \frac{v_1}{\sqrt{\lambda_1^2 + v_1^2 + 2\varepsilon^2}} & & & \\ & \ddots & & \\ & & 1 - \frac{v_{2n+l+1}}{\sqrt{\lambda_{2n+l+1}^2 + v_{2n+l+1}^2 + 2\varepsilon^2}} & \end{bmatrix}.$$

As for any  $\varepsilon > 0$ ,  $\mathcal{J}_{\hat{p}}\varphi_\varepsilon(\hat{p}, \hat{u})$ ,  $\mathcal{J}_{\hat{u}}\varphi_\varepsilon(\hat{p}, \hat{u})$ ,  $\mathcal{J}_\lambda\Psi_\varepsilon(\lambda, v)$  and  $\mathcal{J}_v\Psi_\varepsilon(\lambda, v)$  are invertible, and  $F$  is of full column rank, we have that  $\mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)}\Phi_\varepsilon(\hat{p}, \hat{u}, \lambda, v)$  is of full row rank. The proof is completed.  $\square$

Let  $L(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta)$  be the Lagrange function for  $(P_\varepsilon)$ :

$$L(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta) = f(\hat{p}, \hat{u}, \lambda, v) + \langle \eta, F^T \hat{p} - x^0 \rangle + \langle \theta, \varphi_\varepsilon(\hat{p}, \hat{u}) \rangle + \langle \delta, \Psi_\varepsilon(\lambda, v) \rangle.$$

By a similar calculation as in [9, Lemma 5.2], we can obtain the Hessian matrix of  $L(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta)$ . We omit the details here.

**Lemma 3.5** Denote  $s = L_{\hat{u}}^{-1}\theta$ , then

$$\nabla_{(\hat{p}, \hat{u}, \lambda, v)}^2 L(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta) = \begin{bmatrix} L_{\hat{p}\hat{p}} & L_{\hat{p}\hat{u}} & L_{\hat{p}\lambda} & L_{\hat{p}v} \\ L_{\hat{p}\hat{u}}^T & L_{\hat{u}\hat{u}} & L_{\hat{u}\lambda} & L_{\hat{u}v} \\ L_{\hat{p}\lambda}^T & L_{\hat{u}\lambda}^T & L_{\lambda\lambda} & L_{\lambda v} \\ L_{\hat{p}v}^T & L_{\hat{u}v}^T & L_{\lambda v}^T & L_{vv} \end{bmatrix},$$

where

$$\begin{aligned} L_{\hat{p}\hat{p}} &= H^T e_1 e_1^T H - L_s + L_{\hat{u}-\hat{p}} L_{\hat{u}}^{-1} L_s L_{\hat{u}}^{-1} L_{\hat{u}-\hat{p}}, \\ L_{\hat{p}\hat{u}} &= L_s - L_{\hat{u}-\hat{p}} L_{\hat{u}}^{-1} L_s L_{\hat{u}}^{-1} L_{\hat{u}-\hat{p}}, \\ L_{\hat{p}\lambda} &= 0_{(m+1) \times (2n+l+1)}, \\ L_{\hat{p}v} &= -H^T e_1 e_1^T, \\ L_{\hat{u}\hat{u}} &= F(F^T F)^{-2} F^T - L_s + L_{\hat{u}-\hat{p}} L_{\hat{u}}^{-1} L_s L_{\hat{u}}^{-1} L_{\hat{u}-\hat{p}}, \\ L_{\hat{u}\lambda} &= -F(F^T F)^{-2} F^T H^T, \\ L_{\hat{u}v} &= 0_{(m+1) \times (2n+l+1)} \end{aligned}$$

and

$$\begin{aligned} L_{\lambda\lambda} &= HF(F^T F)^{-2} F^T H^T + \begin{bmatrix} -\frac{v_1^2+2\varepsilon^2}{(\lambda_1^2+v_1^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_1 & & & \\ & \ddots & & \\ & & & -\frac{v_{2n+l+1}^2+2\varepsilon^2}{(\lambda_{2n+l+1}^2+v_{2n+l+1}^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_{2n+l+1} \end{bmatrix}, \\ L_{\lambda v} &= \begin{bmatrix} \frac{\lambda_1 v_1}{(\lambda_1^2+v_1^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_1 & & & \\ & \ddots & & \\ & & \frac{\lambda_{2n+l+1} v_{2n+l+1}}{(\lambda_{2n+l+1}^2+v_{2n+l+1}^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_{2n+l+1} & \\ & & & \end{bmatrix}, \\ L_{vv} &= e_1 e_1^T + \begin{bmatrix} -\frac{\lambda_1^2+2\varepsilon^2}{(\lambda_1^2+v_1^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_1 & & & \\ & \ddots & & \\ & & & -\frac{\lambda_{2n+l+1}^2+2\varepsilon^2}{(\lambda_{2n+l+1}^2+v_{2n+l+1}^2+2\varepsilon^2)^{\frac{3}{2}}} \delta_{2n+l+1} \end{bmatrix}. \end{aligned}$$

As the linear independence constraint qualification (LICQ) holds at any feasible solution of  $(P_\varepsilon)$ , we have that the Lagrange multiplier associated with a KKT point is unique. The following lemma gives the second order sufficient condition of Problem  $(P_\varepsilon)$  at a KKT point.

**Lemma 3.6** Assume that  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)$  is a KKT point of  $(P_\varepsilon)$  and  $(\bar{\eta}, \bar{\theta}, \bar{\delta})$  is the corresponding

Lagrange multiplier. Suppose the following condition holds,

$$\langle d, \nabla_{(\hat{p}, \hat{u}, \lambda, v)}^2 L(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})d \rangle > 0$$

$$\text{for } d \neq 0 \text{ satisfying } \mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)d = 0.$$

Then, the second order growth condition holds at  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)$ , namely there exist positive numbers  $\gamma > 0$  and  $\rho > 0$  such that

$$f(\hat{p}, \hat{u}, \lambda, v) - f(\hat{p}^*, \hat{u}^*, \lambda^*, v^*) \geq \gamma \| (\hat{p}, \hat{u}, \lambda, v) - (\hat{p}^*, \hat{u}^*, \lambda^*, v^*) \|^2,$$

$$\forall (\hat{p}, \hat{u}, \lambda, v) \in \Omega(\varepsilon) \cap B_\rho(\hat{p}^*, \hat{u}^*, \lambda^*, v^*).$$

Let us introduce some notations

$$\kappa(\varepsilon) = \inf\{f(\hat{p}, \hat{u}, \lambda, v) \mid (\hat{p}, \hat{u}, \lambda, v) \in \Omega(\varepsilon)\},$$

$$s(\varepsilon) = \operatorname{argmin}\{f(\hat{p}, \hat{u}, \lambda, v) \mid (\hat{p}, \hat{u}, \lambda, v) \in \Omega(\varepsilon)\},$$

$$\hat{f}_\varepsilon(\hat{p}, \hat{u}, \lambda, v) = f(\hat{p}, \hat{u}, \lambda, v) + \delta_{\Omega(\varepsilon)}(\hat{p}, \hat{u}, \lambda, v),$$

where  $\delta_{\Omega(\varepsilon)}$  is the indicator function of  $\Omega(\varepsilon)$ .

The convergence behavior of the optimal set of  $(P_\varepsilon)$  is obtained in the following theorem.

**Theorem 3.7** *The function  $\kappa(\varepsilon)$  is continuous at  $\varepsilon = 0$  with respect to  $\mathbb{R}_+$  and the set-valued mapping  $s(\varepsilon)$  is outer semicontinuous at  $\varepsilon = 0$  with respect to  $\mathbb{R}_+$ .*

**Proof** As  $f(\hat{p}, \hat{u}, \lambda, v)$  is convex and bounded, we have that  $\kappa(\varepsilon)$  is finite and  $s(\varepsilon)$  is nonempty for any  $\varepsilon \geq 0$ . From Lemma 3.3,  $\Omega(\varepsilon) \rightarrow \Omega^0$  as  $\varepsilon \searrow 0$ , thus  $\hat{f}_\varepsilon$  epi-converges to  $\hat{f}_0$ . The level-boundedness of  $\hat{f}_\varepsilon$  is easily verified for  $\varepsilon \geq 0$ . Therefore, we have from [15, Theorem 7.41] that the function  $\kappa(\varepsilon)$  is continuous at 0 with respect to  $\mathbb{R}_+$  and the set-valued mapping  $s(\varepsilon)$  is outer semi-continuous at 0 with respect to  $\mathbb{R}_+$ . The proof is completed.  $\square$

From Theorem 3.7, we have that the optimal solution set of  $(P_\varepsilon)$  is outer semicontinuous at  $\varepsilon = 0$ . Therefore, in this section we focus on how to solve  $(P_\varepsilon)$  with a sufficiently small  $\varepsilon > 0$ .

Define

$$F_\varepsilon(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta) := \begin{bmatrix} \nabla_{(\hat{p}, \hat{u}, \lambda, v)} L(\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta) \\ \Phi_\varepsilon(\hat{p}, \hat{u}, \lambda, v) \end{bmatrix}.$$

If  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)$  is a local solution for  $(P_\varepsilon)$ , then as the LICQ holds, there is a unique Lagrange multiplier  $(\bar{\eta}, \bar{\theta}, \bar{\delta})$  such that the KKT conditions are satisfied at  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})$ , hence  $F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta}) = 0$ . In fact, finding KKT points of  $(P_\varepsilon)$  is equivalent to solving  $F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta}) = 0$ . The following result shows that the Jacobian of  $F_\varepsilon$  at a KKT point of  $(P_\varepsilon)$  is nonsingular.

**Proposition 3.8** *Assume that  $(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})$  is a KKT point of  $(P_\varepsilon)$  satisfying the second order sufficient conditions shown in Lemma 3.6. Then,  $\mathcal{J}F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})$  is nonsingular.*

**Proof** It is easy to calculate that  $\mathcal{J}F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})$  equals to

$$\begin{bmatrix} \nabla_{(\hat{p}, \hat{u}, \lambda, v)}^2 L(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta}) & \mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)^T \\ \mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*) & 0 \end{bmatrix}.$$

From Lemma 3.4, we know that  $\mathcal{J}_{(\hat{p}, \hat{u}, \lambda, v)} \Phi_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*)$  is of full row rank. Thus, the second order sufficient conditions can derive the non-singularity of  $\mathcal{J}F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta})$ .

In the following, we give a well-known Line Search Inexact Newton Algorithm to solve  $F_\varepsilon(\hat{p}^*, \hat{u}^*, \lambda^*, v^*, \bar{\eta}, \bar{\theta}, \bar{\delta}) = 0$ , in which the Armijo line search is carried out. Let  $z = (\hat{p}, \hat{u}, \lambda, v, \eta, \theta, \delta)$  and

$$g_\varepsilon(z) = \frac{1}{2} \|F_\varepsilon(z)\|^2.$$

Then

$$\nabla g_\varepsilon(z) = \mathcal{J}F_\varepsilon(z)^T F_\varepsilon(z).$$

### Algorithm 3.1

Step 0. Choose  $z^0 = (\hat{p}^0, \hat{u}^0, \lambda^0, v^0, \eta^0, \theta^0, \delta^0) \in \Omega(\varepsilon) \times \mathbb{R}^n \times \mathbb{R}^{m+1} \times \mathbb{R}^{(2n+l+1)}$ ,  $\rho_0 > 0$ ,  $\kappa > 2$ ,  $\sigma \in (0, 1/2)$ ,  $\tau_0 \geq 0$ , and set  $k = 0$ .

Step 1. If  $\|\nabla g_\varepsilon(z^k)\| = 0$ , stop. Otherwise, go to Step 2.

Step 2. Find a solution  $d^k \in \mathbb{R}^{7n+3m+3l+6}$  of the system

$$\mathcal{J}F_\varepsilon(z^k)d = -F_\varepsilon(z^k) + r^k,$$

where the residual vector  $r^k \in \mathbb{R}^{7n+3m+3l+6}$  satisfies the condition

$$\|r^k\| \leq \tau_k \|F_\varepsilon(z^k)\|.$$

If the above system is not compatible or if the condition

$$\langle \nabla g_\varepsilon(z^k), d^k \rangle \leq -\rho_0 \|d^k\|^\kappa$$

is not satisfied, set  $d^k = -\nabla g_\varepsilon(z^k)$ .

Step 3. Find the smallest  $i_k \in \{0, 1, 2, \dots\}$  such that

$$g_\varepsilon(z^k + 2^{-i_k} d^k) \leq g_\varepsilon(z^k) + \sigma 2^{-i_k} \langle \nabla g_\varepsilon(z^k), d^k \rangle.$$

Step 4. Set  $z^{k+1} := z^k + 2^{-i_k} d^k$ , choose  $\tau_{k+1} \geq 0$ ,  $k = k+1$  and go to Step 1.

The global convergence of Algorithm 3.1 is described in the following theorem. The technical proof is omitted here, since the results are classical in references.

**Theorem 3.9** *Assume that Algorithm 3.1 does not terminate within a finite number of iterations. Let  $\{z_k\}$  be generated by Algorithm 3.1. Assume that  $\tau_k \leq \bar{\tau}$  with  $\bar{\tau} \in (0, 1)$ . Then*

(a) *Each accumulation point  $\bar{z}$  of  $\{z_k\}$  satisfies  $\nabla g_\varepsilon(\bar{z}) = 0$ .*

(b) *If  $z^k \rightarrow \bar{z}$ , where  $\bar{z}$  satisfies  $F_\varepsilon(\bar{z}) = 0$  and  $\mathcal{J}F_\varepsilon(\bar{z})$  is nonsingular, then the rate of convergence is  $Q$ -superlinear if  $\tau_k \rightarrow 0$ . Furthermore, if  $\tau_k = O(\|F_\varepsilon(z_k)\|)$ , then the rate of convergence is  $Q$ -quadratic.*

## 4. Preliminary numerical experiments

In this section, we report our numerical experiments. We implemented the algorithm in MATLAB R2012a running on a PC Intel Pentium T4400 of 2.20 GHz CPU and 2 GB of

RAM. The stopping criterion chosen for Algorithm 3.1 is  $\|\nabla g_\varepsilon(z^k)\| < 10^{-3}$  in Example 4.1. For Example 4.2, the stopping criterion is that the relative residual is less than  $10^{-5}$ . We use SYMMLQ MATLAB package to solve the system in Step 2 of Algorithm 3.1. The initial  $(\hat{p}^0, \hat{u}^0, \lambda^0, v^0, \eta^0, \theta^0, \delta^0)$  are chosen to be the zero vectors.

**Example 4.1** Considering the portfolio problem in section 1 with  $n = 10$ , where worst-case losses  $\alpha\% = 2.22\%$ , risk-free rate  $r_f = 0$ , and risk threshold  $L = 0.1287$ . The optimal portfolio  $x_0$ , mean return  $u$  and covariance matrix  $\Sigma$  are based on historical data [10]. Data  $(u_0, L_0)$  are generated by perturbing parameters  $u$  and  $L$  with MATLAB code:

$$u_0 = u + 0.01 * \text{rand}(n, 1), \quad L_0 = L + 0.01 * \text{rand}(1, 1).$$

We set the perturbed parameter  $\varepsilon = 1.0e - 6$ . The data  $x_0$ ,  $u$  and  $\Sigma$  are respectively presented in Tables 1–3. The numerical result is presented in Figure 1.

0.0715	0.0454	0.1092	0.1077	0.1347	0.1356	0.1581	0.1516	0.0588	0.0274
--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Table 1 Optimal portfolio  $x_0$

0.2347	0.1321	0.2292	0.2396	0.2622	0.2941	0.3767	0.3016	0.2358	0.139
--------	--------	--------	--------	--------	--------	--------	--------	--------	-------

Table 2 Mean return  $u$

261.59	81.98	86.24	80.21	22.34	45.74	83.89	170.22	6.83	35.26
81.98	157.88	136.29	125.81	84.26	135.23	132.26	254.68	76.81	24.5
86.24	136.29	179.58	138.87	147.63	169.65	155.77	255.9	149.41	76.54
80.21	125.81	138.87	158.92	102.72	146.82	134.2	211.12	129.38	63.66
22.34	84.26	147.63	102.72	197.46	159.9	158.1	154.78	162.13	98.47
45.74	135.23	169.65	146.82	159.9	239.64	166.05	239.02	171.73	65.17
83.89	132.26	155.77	134.2	158.1	166.05	206.45	226.23	141.88	91.13
170.22	254.68	255.9	211.12	154.78	239.02	226.23	665.95	134.99	2.14
6.83	76.81	149.41	129.38	162.13	171.73	141.88	134.99	280.96	126.22
35.26	24.5	76.54	63.66	98.47	65.17	91.13	2.14	126.22	131.4

Table 3 Covariance matrix  $\Sigma$  ( $10^{-4}$ )

0.2292	0.1325	0.2277	0.237	0.2538	0.2916	0.3686	0.2992	0.2265	0.1337
--------	--------	--------	-------	--------	--------	--------	--------	--------	--------

Table 4 Optimal solution  $u^*$

**Example 4.2** The matrix  $F$  is a randomly generated  $(m + 1) \times n$  full column rank matrix with entries in  $-[1, 1]$  by MATLAB code:

$$F = 2.0 * \text{rand}(m + 1, n) - \text{ones}(m + 1, n).$$

Then we randomly generate an  $l \times n$  matrix  $A$ ,  $n \times 1$  vector  $c$ ,  $l \times 1$  vector  $d$ ,  $(m + 1) \times 1$  vector  $g$ , and real number  $b$ . We let matrix  $G$  be zero in  $\mathbb{R}^{m+1} \times \mathbb{R}^n$ . The parameter set  $(c^0, b^0)$  required in (5) is generated by perturbing  $(c, b)$  and  $x^0 \in \text{SOL}(\text{RLCP})$ . We solve the corresponding perturbed problem  $(P_\varepsilon)$  with  $\varepsilon = 1.0e - 5, 1.0e - 6, 1.0e - 7,$  and  $1.0e - 8,$  respectively. The numerical results are shown in Figure 2.

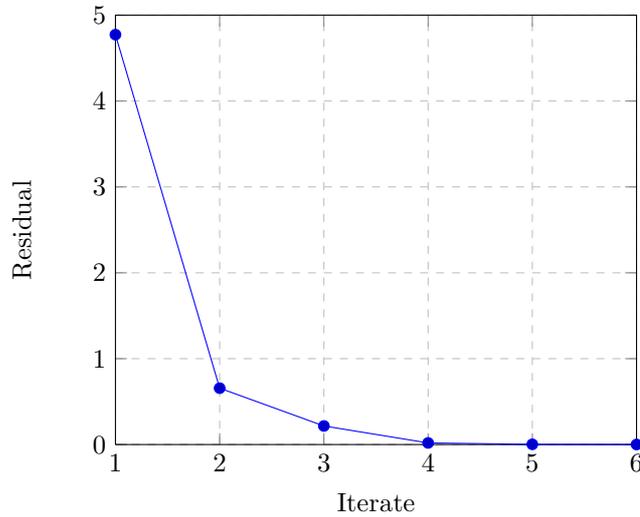


Figure 1 Numerical result of Example 4.1

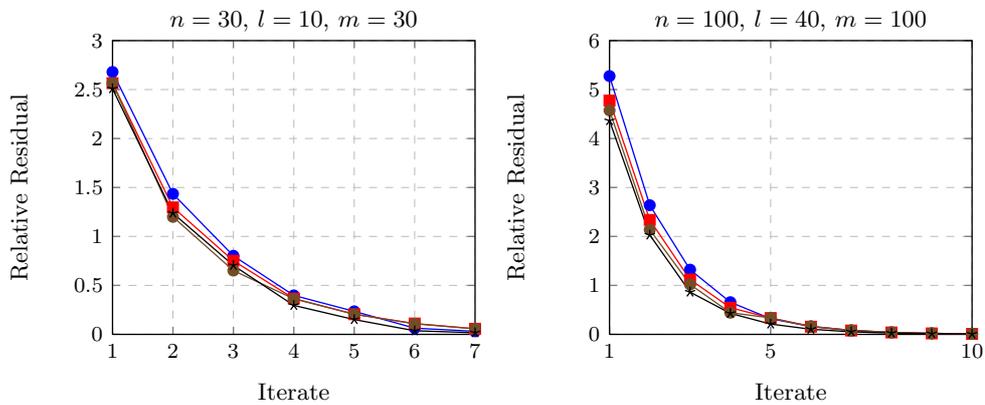


Figure 2 Numerical result of Example 4.2

The numerical results are reported in Figures 1 and 2. “Iterate” and “Residual” stand for the number of iterations and the residuals at the iterated point, respectively. Figure 1 demonstrates the implementation and the residual of each iteration, which show that the convergence is stable and rapid. The optimal solution of the inverse problem of problem (4) is  $(u^*, L^*)$ , where  $u^*$  can

be found in Table 4 and  $L^* = 0.1267$ . Based on the data in Table 2 and Table 4, we can get  $\|(u^*, L^*) - (u, L)\|_\infty < 0.009$ , which can illustrate that the algorithm proposed in this paper is of practical significance. As can be seen in Figure 2, the convergence of random problems is also stable and fast. Furthermore, the accuracy for the random problems with the same dimension does not change so much when the parameter  $\varepsilon$  changes from  $1.0e - 5$  to  $1.0e - 8$ . The numerical experiments show that the algorithm proposed in this paper is implementable and effective, and the inverse Robust linear programming problem is computable.

**Acknowledgments** We thank the referees for their time and comments.

## References

- [1] D. BURTON, P. L. TOINT. *On an instance of the inverse shortest paths problem*. Math. Programming, Ser. A, 1992, **53**(1): 45–61.
- [2] C. HEUBERGER. *Inverse combinatorial optimization: a survey on problems, methods, and results*. J. Comb. Optim., 2004, **8**(3): 329–361.
- [3] Yong JIANG, Xiantao XIAO, Liwei ZHANG, et al. *A perturbation approach for a type of inverse linear programming problems*. Int. J. Comput. Math., 2011, **88**(3): 508–516.
- [4] Xiantao XIAO, Liwei ZHANG. *Solving a class of inverse QP problems by a smoothing Newton method*. J. Comput. Math., 2009, **27**(6): 787–801.
- [5] Jianzhong ZHANG, Liwei ZHANG, Xiantao XIAO. *A perturbation approach for an inverse quadratic programming problem*. Math. Methods Oper. Res., 2010, **72**(3): 379–404.
- [6] Yue LU, Yingen GE, Liwei ZHANG. *An alternating direction method for solving a class of inverse semi-definite quadratic programming problems*. J. Ind. Manag. Optim., 2016, **12**(1): 317–336.
- [7] Xiantao XIAO, Liwei ZHANG, Jianzhong ZHANG. *On convergence of augmented Lagrangian method for inverse semi-definite quadratic programming problems*. J. Ind. Manag. Optim., 2009, **5**(2): 319–339.
- [8] Xiantao XIAO, Liwei ZHANG, Jianzhong ZHANG. *A smoothing Newton method for a type of inverse semi-definite quadratic programming problem*. J. Comput. Appl. Math., 2009, **223**(1): 485–498.
- [9] Yi ZHANG, Liwei ZHANG, Jia WU, et al. *A perturbation approach for an inverse quadratic programming problem over second-order cones*. Math. Comp., 2015, **84**(291): 209–236.
- [10] D. BERTSIMAS, V. GUPTA, I. C. PASCHALIDIS. *Inverse optimization: a new perspective on the Black-Litterman model*. Oper. Res., 2012, **60**(6): 1389–1403.
- [11] Zhiquan LUO, Jongshi PANG, D. RALPH. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- [12] Jia WU, Liwei ZHANG, Yi ZHANG. *A smoothing Newton method for mathematical programs governed by second-order cone constrained generalized equations*. J. Global Optim., 2013, **55**(2): 359–385.
- [13] Yi ZHANG, Jia WU, Liwei ZHANG. *First order necessary optimality conditions for mathematical programs with second-order cone complementarity constraints*. J. Global Optim., 2015, **63**(2): 253–279.
- [14] M. FUKUSHIMA, Zhiquan LUO, P. TSENG. *Smoothing functions for second-order-cone complementarity problems*. SIAM J. Optim., 2001, **12**(2): 436–460.
- [15] R. T. ROCKAFELLAR, R. J.-B. WETS. *Variational Analysis*. Springer-Verlag, Berlin, 1998.