RAIRO-Oper. Res. 55 (2021) 247–260
https://doi.org/10.1051/ro/2021011

RAIRO Operations Research
www.rairo-ro.org

ON THE MINIMUM-NORM SOLUTION OF CONVEX QUADRATIC PROGRAMMING

Saeed Ketabchi[1], Hossein Moosaei[2,*] and Milan Hladík[3]

Abstract. We discuss some basic concepts and present a numerical procedure for finding the minimum-norm solution of convex quadratic programs (QPs) subject to linear equality and inequality constraints. Our approach is based on a theorem of alternatives and on a convenient characterization of the solution set of convex QPs. We show that this problem can be reduced to a simple constrained minimization problem with a once-differentiable convex objective function. We use finite termination of an appropriate Newton's method to solve this problem. Numerical results show that the proposed method is efficient.

Mathematics Subject Classification. 90C05, 90C30, 15A39.

Received May 31, 2020. Accepted January 24, 2021.

1. Introduction

Linearly constrained quadratic programs (QPs) arise from numerous areas of application, such as manufacturing, economics, and social, and public planning. A QP can have a unique optimal solution, multiple solutions, or no solution. When such a problem has multiple optimal solutions, one can be interested in computation all of them or computation of the most favourable one (in some point of view). For the former approach, interval methods [10, 12, 23] are suitable for determining a tight approximation of the set of all optimal solutions. Our focus is on the second approach. The choice of a particular solution can be important, and a natural choice is a solution with the minimum norm [21]. In fact, computation of the minimum-norm solution, which is the solution with the smallest norm among the infinite possible solutions, is a research line in various fields, including linear systems, absolute value equations or variational inequalities [4, 20, 25].

Motivated by [13, 15, 19], the aim of this paper is finding the minimum norm solution of QPs that have multiple solutions. In other words, we want to solve the following problem:

$$\min \ \frac{1}{2}\|x\|^2 \ \text{ s.t. } \ x \in X^*,$$

where X^* is the optimal solution set of QP.

Keywords. Solution set of convex problems, minimum-norm solution of convex quadratic programs, Newton's method, theorems of alternative.

[1] Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran.

[2] Department of Mathematics, Faculty of Science, University of Bojnord, Bojnord, Iran.

[3] Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic.

[*]Corresponding author: hmoosaei@gmail.com, moosaei@ub.ac.ir

Here, we suggest a novel method for finding the minimum norm solution of the QPs based on alternative theorems. According to this issue, instead of solving a complicated constrained quadratic programming like the above problem, we deal with a simple constrained minimization problem which objective function is convex and only once diffrentiable (not twice). We suggest a type of Newton's method to solve it. We also compare our method to the Tikhonov regularization approach [22], which is a standard approach for finding the minimum norm solution of QPs, and with another standard method that will be described in this paper. The numerical results show that the proposed method performs better than the other two methods, and also our methodology applies on a practical example arising in finance.

As for our notations, a few points need to be made. We denote the n-dimensional real space by $\mathbb{R}^n$, and A^T, $\|\cdot\|$ mean the transpose of matrix A and Euclidean norm, respectively. The subplus function a_+ replaces negative components of the vector a by zeros; in fact it represents the positive part of vector a.

The paper is organized as follows. A characterization of the solution set and the minimum-norm solution of a convex QPs subject to linear equality and inequality constraints are described in Section 2. In Section 3, we describe our numerical algorithm. In Section 4, numerical experiments are reported to illustrate the efficiency of the proposed method and finally Section 5 concludes the paper.

2. Minimum-norm solution of convex QPs

In this section, we first characterize the solution set of the following convex QPs subject to linear equality and inequality constraints. Then, we introduce a method for finding its minimum-norm solution.

$$\begin{aligned} \min_{x_1,x_2}\ & f(x) = \frac{1}{2}x_1^T Q_1 x_1 + \frac{1}{2}x_2^T Q_2 x_2 + d_1^T x_1 + d_2^T x_2, \\ \text{s.t.}\ & A_{11}x_1 + A_{12}x_2 \geq b_1,\ A_{21}x_1 + A_{22}x_2 = b_2, \\ & x_1 \in \mathbb{R}^{n_1},\ x_2 \in \mathbb{R}^{n_2},\ x_1 \geq 0_{n_1}. \end{aligned} \tag{2.1}$$

Herein, Q_1 and Q_2 are $n_1 \times n_1$ and $n_2 \times n_2$ positive semidefinite matrices, respectively, and $d = [d_1^T, d_2^T]^T \in \mathbb{R}^n$, $b = [b_1^T, b_2^T]^T \in \mathbb{R}^m$ are given vectors such that $n = n_1 + n_2$, $m = m_1 + m_2$ ($d_1 \in \mathbb{R}^{n_1}$, $d_2 \in \mathbb{R}^{n_2}$, $b_1 \in \mathbb{R}^{m_1}$ and $b_2 \in \mathbb{R}^{m_2}$). Furthermore, let A_{11}, A_{12}, A_{21}, and A_{22} be $m_1 \times n_1$, $m_1 \times n_2$, $m_2 \times n_1$ and $m_2 \times n_2$ matrices, respectively. Here, we considered the problem in the general case. That's why the variables in problem (2.1) are split to free variables and nonnegative variables.

Problem (2.1) might have multiple optimal solutions, and its the minimum norm solution problem can be written as follows

$$\min\ \frac{1}{2}\|x\|^2 \ \text{ s.t. }\ x \in X^*, \tag{2.2}$$

where X^* is the solution set of problem (2.1). The first approach for solving problem (2.2) is using the problem $\min \frac{1}{2}\|x\|^2$ such that x is a KKT point for problem (2.1). In this case, we need to solve a nonlinear optimization problem in which there are $m+n$ variables, so this method is expensive and we avoid doing it.

We note that the standard approach for finding the minimum norm solution of a convex program is the Tikhonov regularization, see [13, 22]. More precisely, for linearly constrained quadratic programs, the Tikhonov regularization generates a sequence of iterates $\{x_k\}$ with x_k being the unique solution of the following regularized program called Tikhonov regularized problem [22]:

$$\begin{aligned} \min_{x_1,x_2}\ & f(x) + \mu_k\|x\|^2 \\ \text{s.t.}\ & A_{11}x_1 + A_{12}x_2 \geq b_1,\ A_{21}x_1 + A_{22}x_2 = b_2, \\ & x_1 \in \mathbb{R}^{n_1},\ x_2 \in \mathbb{R}^{n_2},\ x_1 \geq 0_{n_1}, \end{aligned} \tag{2.3}$$

where $\mu_k > 0$ is a positive parameter and the sequence $\{\mu_k\}$ tends to zero.

From the computational point of view, although the solution of problem (2.3) with respect to a sufficiently small and positive μ_k is considered as an approximation of the minimum norm solution of (2.2), the Tikhonov regularization is rather costly. This is because solving a sequence of quadratic programs is a tough problem, and it is also sometimes not clear how to choose an appropriate sequence of regularization parameters $\{\mu_k\} \to 0^+$. To be more precise, choosing the regularization parameter in Tikhonov regularization is still an open research line [8, 16].

The main contribution of this paper is to suggest an alternative approach for finding the minimum norm solution of linearly constrained quadratic programs based on Gale's theorem of alternatives. First, we briefly characterize the solution set of convex QPs. Let $S \subseteq \mathbb{R}^n$ be an open convex subset, $f\colon S \to \mathbb{R}$ a convex differentiable function, and $X \subseteq S$ any convex subset. The following theorem gives a precise characterization of solution sets of convex programs with twice continuously differentiable convex objective functions (see Thm. 1 in [17]). It was adapted for once differentiable functions in Theorem 2.1 of [14].

Theorem 2.1. *Let $S \subseteq \mathbb{R}^n$ be an open convex subset and $X \subseteq S$ any convex subset and also f be a convex differentiable function. Consider the following convex problem:*

$$\min_{x\in X}\ f(x), \tag{2.4}$$

and suppose that its solution set X^* is nonempty. Then,

$$X^* = \left\{x \in X : \nabla f(x^*)^T x^* = \nabla f(x^*)^T x,\ \nabla f(x^*) = \nabla f(x)\right\}, \tag{2.5}$$

where $x^* \in X^*$ is arbitrarily chosen.

This theorem provides a precise characterization of solution sets of convex programs. For an extensive discussion of this theorem, see [17]. The formulation adapted for our problem is as follows.

Lemma 2.2. *Let X^* be the solution set for the problem (2.1) and assume $x^* = [x_1^{*T}, x_2^{*T}]^T \in X^*$. Then $x = [x_1^T, x_2^T]^T \in X^*$ if and only if x satisfies the following system*

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 &\geq b_1,\\ A_{21}x_1 + A_{22}x_2 &= b_2,\\ Q_1x_1 &= Q_1x_1^*,\\ Q_2x_2 &= Q_2x_2^*,\\ d_1^Tx_1 + d_2^Tx_2 &= d_1^Tx_1^* + d_2^Tx_2^*,\\ x_1 \geq 0,\ x_1 &\in \mathbb{R}^{n_1},\ x_2 \in \mathbb{R}^{n_2}. \end{aligned} \tag{2.6}$$

Proof. By Theorem 2.1, we have $x \in X^*$ if and only if x is feasible and $\nabla f(x^*)^T x^* = \nabla f(x^*)^T x,\ \ \nabla f(x^*) = \nabla f(x)$. This yields $d^T x = d^T x^*$ and $Qx = Qx^*$, where

$$Q = \begin{bmatrix} Q_1 & 0_{n_1\times n_2}\\ 0_{n_2\times n_1} & Q_2 \end{bmatrix}.$$

Hence, $Q_1x_1 = Q_1x_1{}^*$, $Q_2x_2 = Q_2x_2{}^*$ and $d_1{}^Tx_1 + d_2{}^Tx_2 = d_1{}^Tx_1{}^* + d_2{}^Tx_2{}^*$. This means that $x \in X^*$ if and only if x satisfies in (2.6). □

To solve problem (2.2), we consider an alternative system to (2.6), and a constrained minimization problem for its residual vector. The next theorem may be found in Theorem 2.7 from [7], and for more details see also Theorem 1 from [5].

Theorem 2.3 (Gale's Theorem of Alternatives)**.** *Let matrix $A \in \mathbb{R}^{m\times n}$ and vector $b \in \mathbb{R}^m$ be given. Let $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. Exactly one of the following systems* (I), (II) *is feasible. In other words, the systems* (I), (II) *are alternative, which means that exactly one of them is consistent.*

$$Ax \geq b, \quad \text{(I)} \qquad \text{and} \qquad A^T u = 0_n, \quad b^T u = 1, \quad u \geq 0_m. \quad \text{(II)}$$

Now, consider the following system

$$\begin{aligned} A_{11}^T u_1 + A_{21}^T u_2 + Q_1 u_3 + d_1 u_5 &\leq 0_{n_1}, \\ A_{12}^T u_1 + A_{22}^T u_2 + Q_2 u_4 + d_2 u_5 &= 0_{n_2}, \\ b_1^T u_1 + b_2^T u_2 + {x_1^*}^T Q_1 u_3 + {x_2^*}^T Q_2 u_4 + d_1^T x_1^* u_5 + d_2^T x_2^* u_5 &= \rho, \\ u_1 &\geq 0_{m_1}, \end{aligned} \tag{2.7}$$

where $u_1 \in \mathbb{R}^{m_1}$, $u_2 \in \mathbb{R}^{m_2}$, $u_3 \in \mathbb{R}^{n_1}$, $u_4 \in \mathbb{R}^{n_2}$, and $u_5 \in \mathbb{R}$. This implies that the linear systems (2.6) and (2.7) are alternative for any positive value of ρ (where $\rho > 0$ as in the Farkas lemma or $\rho = 1$ as in the Gale's theorem), which means that exactly one of them is consistent.

Now, we introduce the constrained minimization problem as follows:

$$\begin{aligned} \min_{u_1 \geq 0_{m_1},\, u \in \mathbb{R}^{m_2+n+1}} \ & \frac{1}{2}\left\| \left(A_{11}^T u_1 + A_{21}^T u_2 + Q_1 u_3 + d_1 u_5\right)_+ \right\|^2 \\ & + \frac{1}{2}\left\| A_{12}^T u_1 + A_{22}^T u_2 + Q_2 u_4 + d_2 u_5 \right\|^2 \\ & + \frac{1}{2}\Big| \rho - b_1^T u_1 - b_2^T u_2 - {x_1^*}^T Q_1 u_3 - {x_2^*}^T Q_2 u_4 \\ & \quad - d_1^T x_1^* u_5 - d_2^T x_2^* u_5 \Big|^2, \end{aligned} \tag{2.8}$$

where $u = [u_2^T, u_3^T, u_4^T, u_5]^T$ and the objective function is the residual vector of the system (2.7) (notice that the residual vector to the inequality $Ax \leq b$ is $(Ax-b)_+$, and it is obvious that $x \in \{x : Ax \leq b\}$ if and only if $(Ax-b)_+ = 0_m$. For more details about residual vector see [9]).

Remark 2.4. In order to be more precise about the problem (2.8), this problem is not an unconstrained minimization problem due to involving a positive variable $u_1 \geq 0_{m_1}$ as a constraint. However, the unconstrained optimization methods can easily be modified for this type of problems, then we keep this term for the problem (2.8).

Remark 2.5. The problem (2.8) is always solvable because of the quadratic objective function defined on nonempty feasible set $u_1 \geq 0_{m_1}$. Also if the optimal value of problem (2.8) is zero, then the system (2.7) is solvable, and if it is non-zero, then it means that the system (2.7) is unsolvable.

Suppose that X^* is the same as in Theorem 2.1 and let $\widetilde{x}^* = [{\widetilde{x}_1^*}^T, {\widetilde{x}_2^*}^T]^T$ denote the minimum-norm solution to the problem (2.1). Then we have the following theorem, which shows existence of the minimum norm solution $\widetilde{x}^*$ and it also proposes a method for finding it. This is a modification of Theorem 3 from [6] and Theorem 3 from [9], adapted to the QP case.

Theorem 2.6. *Let $u^* = [{u_1^*}^T, {u_2^*}^T, {u_3^*}^T, {u_4^*}^T, u_5^*]^T$ be the solution of problem (2.8). Then there exists a vector $[{w_1^*}^T, {w_2^*}^T, w_3^*]^T$, $w_1^* \geq 0_{n_1}$, $w_2^* \in \mathbb{R}^{n_2}$, and $w_3^* > 0$ for which*

$$\widetilde{x}_1^* = w_1^*/w_3^*, \quad \widetilde{x}_2^* = w_2^*/w_3^*. \tag{2.9}$$

Proof. Since the linear systems (2.6) and (2.7) are alternative and the system (2.6) is consistent, so the system (2.7) is inconsistent, and we deduce that the optimal value of the problem (2.8) must be nonzero. Also, note that problem (2.8) is convex with a nonnegative variable as a constraint. Thus, KKT conditions are satisfied at the minimum point. Therefore, $u^* = [u_1^{*T}, u_2^{*T}, u_3^{*T}, u_4^{*T}, u_5^*]^T$ minimizes this problem if and only if

$$\begin{aligned} A_{11}w_1^* + A_{12}w_2^* - b_1 w_3^* &\geq 0_{m_1}, \\ u_1^{*T}\left(A_{11}w_1^* + A_{12}w_2^* - b_1 w_3^*\right) &= 0, \\ A_{21}w_1^* + A_{22}w_2^* - b_2 w_3^* &= 0_{m_2}, \\ Q_1 w_1^* - Q_1 x_1^* w_3^* &= 0_{n_1}, \\ Q_2 w_2^* - Q_2 x_2^* w_3^* &= 0_{n_2}, \\ {d_1}^T w_1^* + d_2^T w_2^* - \left(d_1^T x_1^* + d_2^T x_2^*\right) w_3^* &= 0, \end{aligned} \tag{2.10}$$

where,

$$\begin{aligned} w_1^* &= \left(A_{11}^T u_1^* + A_{21}^T u_2^* + Q_1 u_3^* + d_1 u_5^*\right)_+, \\ w_2^* &= A_{12}^T u_1^* + {A_{22}}^T u_2^* + Q_2 u_4^* + d_2 u_5^*, \\ w_3^* &= \rho - b_1^T u_1^* - b_2^T u_2^* - x_1^{*T} Q_1 u_3^* - x_2^{*T} Q_2 u_4^* - d_1^T x_1^* u_5^* - d_2^T x_2^* u_5^*. \end{aligned} \tag{2.11}$$

The vector $w^* = [w_1^{*T}, w_2^{*T}, w_3^*]^T$ determined from (2.11) satisfies the condition $w^* = [w_1^{*T}, w_2^{*T}, w_3^*]^T \neq 0$ and $w_3^* > 0$ by Theorem 2 in [9]. From (2.10) we obtain

$$\begin{aligned} A_{11}(w_1^*/w_3^*) + A_{12}(w_2^*/w_3^*) &\geq b_1, \\ A_{21}(w_1^*/w_3^*) + A_{22}(w_2^*/w_3^*) &= b_2, \\ Q_1(w_1^*/w_3^*) &= Q_1 x_1^*, \\ Q_2(w_2^*/w_3^*) &= Q_2 x_2^*, \\ d_1^T(w_1^*/w_3^*) + d_2^T(w_2^*/w_3^*) &= d_1^T x_1^* + d_2^T x_2^*, \end{aligned} \tag{2.12}$$

therefore, w_1^*/w_3^*, and w_2^*/w_3^* satisfy in (2.6) and, we have $[w_1^{*T}/w_3^*, w_2^{*T}/w_3^*]^T \in X^*$ (see [6]).

Consider the following constrained quadratic problem

$$\begin{aligned} \min_{w_1, w_2, w_3} \quad & \frac{1}{2}\left(\|w_1\|^2 + \|w_2\|^2 + |w_3|^2 - \rho w_3\right) \\ \text{s.t.} \quad & A_{11}w_1 + A_{12}w_2 - b_1 w_3 \geq 0_{m_1}, \\ & A_{21}w_1 + A_{22}w_2 - b_2 w_3 = 0_{m_2}, \\ & Q_1 w_1 - Q_1 {x_1}^* w_3 = 0_{n_1}, \\ & Q_2 w_2 - Q_2 {x_2}^* w_3 = 0_{n_2}, \\ & {d_1}^T w_1 + d_2^T w_2 - \left(d_1^T x_1^* + d_2^T x_2^*\right) w_3 = 0. \\ & w_1 \geq 0_{n_1},\ w_2 \in \mathbb{R}^{n_2},\ w_3 \geq 0. \end{aligned} \tag{2.13}$$

Applying part 3 of Theorem 3 from [9] and considering (2.12), we conclude that $w^* = [w_1^{*T}, w_2^{*T}, w_3^*]^T$ is the solution to the problem (2.13) and $\widetilde{x}_1^* = w_1^{*T}/w_3^*$, $\widetilde{x}_2^* = w_2^{*T}/w_3^*$ is the minimum-norm solution of problem (2.1), completing the proof. □

3. Numerical algorithm

In this section, we introduce a method for solving problem (2.8); the method briefly is described in Algorithm 2. It is based on a modified Newton's method, described in detail in Algorithm 1. The modified

Newton's method employs a generalized Hessian matrix since since the objective function of (2.8) is only once-differentiable. The modification of the Newton method also means a certain kind of regularization, which is justified by the fact that the generalized Hessian matrix can be singular.

For a fixed positive ρ consider the system (2.7). This system can be written as follows:

$$\left\{u \in \mathbb{R}^{m+n+1} : S_1^T u \leq 0,\ S_2^T u = 0,\ S_3^T u = \rho\right\}, \tag{3.1}$$

where

$$\begin{aligned} S_1^T &= \begin{bmatrix} A_{11}^T & A_{21}^T & Q_1 & 0_{n_1\times n_2} & d_1 \\ -1_{1\times m_1} & 0_{1\times m_2} & 0_{1\times n_1} & 0_{1\times n_2} & 0 \end{bmatrix}, \\ S_2^T &= \begin{bmatrix} A_{12}^T & A_{22}^T & 0_{m_1\times n_1} & Q_2 & d_2 \end{bmatrix}, \\ S_3^T &= \begin{bmatrix} b_1^T & b_2^T & {x_1^*}^T Q_1 & {x_2^*}^T Q_2 & d_1^T x_1^* + d_2^T x_2^* \end{bmatrix}. \end{aligned}$$

Considering (3.1), we can therefore rewrite (2.8) as:

$$\min_{u\in\mathbb{R}^{m+n+1}} g(u) = \frac{1}{2}\left[\|\left(S_1^T u\right)_+\|^2 + \|S_2^T u\|^2 + |\rho - S_3^T u|^2\right],$$

where $U^* = \{u \in \mathbb{R}^{m+n+1} : g(u) = \min_{t\in\mathbb{R}^{m+n+1}} g(t)\} \neq \emptyset$, $g(u)$ is convex, piecewise quadratic, and differentiable, but it does not have a conventional Hessian matrix. Indeed, the gradient

$$\nabla g(u) = S_1\left(S_1^T u\right)_+ + S_2\left(S_2^T u\right) - S_3\left(\rho - S_3^T u\right)$$

of $g(u)$ is not differentiable since the term $S_1(S_1^T u)_+$ is not differentiable. However, for the term $\frac{1}{2}\|\left(S_1^T u\right)_+\|^2$, we can define the generalized Hessian matrix [11], an $(m+n+1)\times(m+n+1)$ symmetric positive semidefinite matrix of the form $S_1 D^*(z) S_1^T$. Here, $D^*(z)$ denotes the $(n_1+1)\times(n_1+1)$ diagonal matrix whose ith diagonal entry z_i is equal to one, if $(S_1^T u)_i > 0$, and to zero if $(S_1^T u)_i \leq 0$ $(i = 1, 2, \ldots, n_1)$. Therefore, the generalized Hessian matrix for $g(u)$ can be defined as follows:

$$\partial^2 g(u) = S_1 D^*(z) S_1^T + S_2 S_2^T + S_3 S_3^T.$$

Since the generalized Hessian matrix can be singular, the following modified Newton's direction is used [2].

$$-(\partial^2 g(u) + \delta I_{m+n+1})^{-1}\nabla g(u),$$

where δ is a small positive number (in our numerical experiments we used $\delta = 10^{-4}$), and I_{m+n+1} is the identity matrix of order $m+n+1$. In this case, the modified Newton's method has the form

$$u_{n+1} = u_n - \left(\partial^2 g(u_n) + \delta I_{m+n+1}\right)^{-1}\nabla g(u_n).$$

In addition, our stopping criterion for this method is as follows: $\|u_{n+1} - u_n\| \leq \mathrm{tol}$ (in our numerical experiments we used $\mathrm{tol} = 10^{-9}$).

By employing the stepsize Armijo rule [1], we can derive the global finite-step convergence of the modified Newton's method starting from any point. By combining the generalized Newton's method with a line-search based on the Armijo rule, we arrive at Algorithm 1. We can guarantee a global finite termination of this method by using the following theorem which comes from [18], Theorem 2 on page 925.

Theorem 3.1. *The sequence* $\{u_n\}$ *of Algorithm 1 terminates at the global minimum solution of the problem (2.8).*

The proposed method is summarized in Algorithm 2.

Algorithm 1. Modified Newton's method with the Armijo rule.

Input: Choose any u_0 as a starting point, let tol > 0 be an error tolerance and δ be a small positive number.

$n := 0$,
$U := \infty$,
while $\|U\| \geq$ tol or $\|\nabla g(u_n)\|_\infty \geq$ tol **do**
$d_n := -(\partial^2 g(u_n) + \delta I_{m+n+1})^{-1} \nabla g(u_n)$,
$\lambda_n = \max\{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$ such that: $g(u_n) - g(u_n + \lambda_n d_n) \geq -\gamma \lambda_n \nabla g(u_n)^T d_n$
for some $\gamma \in (0, \frac{1}{2})$,
$U := u_{n+1} - u_n$,
$u_{n+1} := u_n + \lambda_n d_n$,
$n := n + 1$.
end while

Algorithm 2. Minimum norm solution.

Input: $A_{11}, A_{12}, A_{21}, A_{22}, Q_1, Q_2, d_1, d_2$, b_1, and b_2.
Solve the problem (2.8) by using modified Newton's method (Algorithm 1) and obtain u^*.
Compute w_1^*, w_2^*, w_3^* by using the equations (2.11).
Set

$$\widetilde{x}_1^* = w_1^*/w_3^*, \quad \widetilde{x}_2^* = w_2^*/w_3^*.$$

return Minimum norm solution $\widetilde{x}^* = [\widetilde{x}_1^{*T}, \widetilde{x}_2^{*T}]^T$ of the problem (2.1).

Remark 3.2. The modified Newton's method is related to the the quasi-Newton methods. If evaluation and use of the Hessian matrix is impractical or costly, the quasi-Newton methods is to use an approximation to the inverse Hessian. Consider the following steepest descent iteration

$$x_{n+1} = x_n - a_n T_n \nabla f(x_n),$$

where T_n as an approximation to the inverse of the Hessian. In this paper, we define $T_n = S_1 D^*(z) S_1^T$, where $D^*(z)$ denotes the $(n_1+1) \times (n_1+1)$ diagonal matrix whose ith diagonal entry z_i is equal to one if $(S_1^T u)_i > 0$, and to zero if $(S_1^T u)_i \leq 0$ $(i = 1, 2, \ldots, n_1)$.

Remark 3.3. We note that the objective function

$$g(u) = \frac{1}{2}\left[\| \left(S_1^T u\right)_+ \|^2 + \|S_2^T u\|^2 + |\rho - S_3^T u|^2\right]$$

is a piecewise quadratic function and its generalized Hessian matrix $\partial^2 g(u) = S_1 D^*(z) S_1^T + S_2 S_2^T + S_3 S_3^T$ is positive definite. Also, our method is a quasi-Newton methods. Furthermore, for a convex quadratic program a single Newton step captures its solution. Perhaps these points can be an intuitive justification of the finite termination converge to the global solution of problem

$$\min_{u \in \mathbb{R}^{m+n+1}} g(u).$$

4. Numerical testing

In this section, we first give a small numerical example to illustrate the feasibility and effectiveness of the theory of the previous sections. We then present numerical results on various randomly generated convex QPs subject to linear equality and inequality constraints. To further analyse our suggested method, we compare the suggested method with the Tikhonov regularization method and a standard method which will describe more in Section 4.3. Finally, we illustrate the application of the proposed method on a practical problem arising from finance. We have tested the algorithm using MATLAB 7.11.0 on a *Core i7* 2.59 GHz with main memory 16 GB. Besides our algorithm, we implemented the MS method, too.

4.1. A simple example

In this subsection, a simple example is discussed to illustrate the correctness of our proposed method and the process of the suggested method. This problem can be found in [24] with $x^* = [2, 1, -6]^T$ as its solution.

$$\begin{aligned} \min_{x\in\mathbb{R}^3} \quad & f(x) = 11x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 + 6x_1x_3 - 4x_1 \\ \text{s.t.} \quad & 2x_1 + 2x_2 + x_3 = 0, \\ & -x_1 + x_2 \le -1, \\ & 3x_1 + x_3 \le 4, \\ & -6 \le x_1, x_2, x_3 \le 6. \end{aligned} \tag{4.1}$$

If we substitute $y_1 = x_1 + 6$, $y_2 = x_2 + 6$, $y_3 = x_3 + 6$, then the above problem takes the following form:

$$\begin{aligned} \min_{y\in\mathbb{R}^3} \quad & f(y) = 11{y_1}^2 + {y_2}^2 + {y_3}^2 - 2y_1y_2 + 6y_1y_3 - 152y_1 - 48y_3 + 636 \\ \text{s.t.} \quad & 2y_1 + 2y_2 + y_3 = 30, \\ & -y_1 + y_2 \le -1, \\ & 3y_1 + y_3 \le 28, \\ & y_1, y_2, y_3 \le 12, \\ & y_1, y_2, y_3 \ge 0. \end{aligned} \tag{4.2}$$

If we put

$$Q = \begin{bmatrix} 22 & -2 & 6 \\ -2 & 2 & 0 \\ 6 & 0 & 2 \end{bmatrix}, \; d = \begin{bmatrix} -152 \\ 0 \\ -48 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & -1 & 0 \\ -3 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 2 & 2 & 1 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 1 \\ -28 \\ -12 \\ -12 \\ -12 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 30 \end{bmatrix},$$

then we have:

$$\begin{aligned} \min_{y\in\mathbb{R}^n} \quad & f(y) = \frac{1}{2}y^TQy + d^Ty + c \\ \text{s.t.} \quad & A_1y \ge b_1, \\ & A_2y = b_2, \\ & y \ge 0, \end{aligned} \tag{4.3}$$

where $y = (y_1, y_2, y_3)^T$, $c = 636$, and its solution is $y^* = [8, 7, 0]^T$. Since Q is positive definite, the above problem has a unique optimal solution, and we must show that $\widetilde{y}^* = y^*$, where $\widetilde{y}^*$ denotes the minimum-norm solution to problem (4.3).

```
%  Generate a convex QPs subject to linear equality and inequality
   constraints with infinitely many solutions (The norm of the
   minimum norm solutions are zero.)
R=(rand(m,n)-rand(m,n));
Q1=R'*R;
xx=(10*(randn(n,1)));
Q2=diag((xx),0,n,n);
Q2(n,n)=0;
xxx=null(R);
x1=xxx(:,1);
x2=zeros(n,1);  x2(n,1)=100;
A11=(rand(m,n)-rand(m,n));
A12=zeros(m,n);
A21=diag(ones(n,1));
A22=zeros(n,n);
b1=-10*ones(m,1);
b2=zeros(n,1);
```

FIGURE 1. Code generation for the first data set.

Considering (2.8), it yields the following constrained minimization problem:

$$\begin{aligned}\min_{u_1\geq 0_5,\, u\in\mathbb{R}^5} \ & \left\| \left(A_1^T u_1 + A_2^T u_2 + Q^T u_3 + d u_4\right)_+ \right\|^2 \\ & + \left|\rho - b_1^T u_1 - b_2^T u_2 - {y^*}^T Q u_3 - d^T y^* u_4\right|^2,\end{aligned} \tag{4.4}$$

where $u = [u_2^T, u_3^T, u_4^T]^T$. Solving problem (4.4), we obtain

$$\begin{aligned} u^* &= \begin{bmatrix} 1.075733763299230 \\ -0.000000038596423 \\ -0.000000038596423 \\ -0.000000048245529 \\ -0.000000115789270 \\ 0.283499418477488 \\ 0.560392759269804 \\ 0.848532068598588 \\ 1.269502139420115 \\ 0.130356113891751 \end{bmatrix}, \quad w_1^* = \begin{bmatrix} 0.077192846352968 \\ 0.067543740558844 \\ 0 \end{bmatrix}, \\ w_3^* &= 0.009649105794074, \end{aligned} \tag{4.5}$$

and from (2.9) we have

$$\widetilde{y}^* = w_1^*/w_3^* = \begin{bmatrix} 8.000000000039028 \\ 7.000000000033827 \\ 0 \end{bmatrix},$$

which means that $\widetilde{y}^* \approx y^*$. To solve this problem, we use the numerical algorithm based on a new fast Newton method analyzed in Section 3. The presented example illustrates the effectiveness and correctness of our proposed method $\left(\|\widetilde{y}^* - y^*\| < 10^{-10}\right)$.

4.2. Random artificial problems

In this subsection, to show that the proposed method is efficient and reasonable, some random problems are generated. The code in Figure 1 generates problems for which the norm of the minimum norm solution is 0.

Table 1 reports the following information for each test problem:

TABLE 1. The norm of the minimum norm solutions for x_1, x_2 are zero (Sect. 4.2).

m, n	$norm(df)$	$\|\widetilde{x_1}^*\|$	$\|\widetilde{x_2}^*\|$	cpu$_s$
10, 50	2.5358e-07	2.4790e-07	2.7077e-20	0.18
20, 100	4.3384e-12	2.9334e-11	3.5458e-66	0.24
30, 200	1.0493e-06	3.7596e-07	1.0559e-09	0.57
40, 500	7.0839e-06	2.6414e-06	4.8391e-46	2.86
50, 600	3.6662e-06	7.2360e-07	1.3595e-37	3.51
70, 700	5.1892e-08	1.4043e-08	3.9409e-09	21.63
80, 800	7.0751e-06	1.1848e-06	2.0022e-57	28.92
90, 900	5.8725e-07	7.2566e-08	4.6829e-09	35.30
100, 1000	1.1376e-06	1.8327e-06	4.1234e-19	41.66
150, 1500	3.1558e-05	4.0659e-06	1.6654e-16	62.32
200, 2000	3.8155e-05	2.3090e-06	2.4007-09	333.60

```
%Generate random  convex QPs subject to linear equality and inequality constraints with infinitely many optimal solutions (the norm of the minimum norm solution is zero.)
x=sparse(10*(randn(n,1)));
xx=sparse(10*(randn(n,1)));
x1=spdiags((ones(n,1)-sign(x)),0,n,n)*10*ones(n,1);
x2=spdiags((ones(n,1)-sign(xx)),0,n,n)*10*ones(n,1);
Q1=spdiags((x),0,n,n);  Q2=spdiags((xx),0,n,n);
A11=Q1;  A12=Q2;  A21=Q1;  A22=Q2;
b1=b2=d1=d2=zeros(n,1);
```

FIGURE 2. Code generation for the second data set.

– m, n: the size of data $A_{11}, A_{12}, A_{21}, A_{22}, Q_1, Q_2, d_1, d_2$
– $norm(df)$: the norm of the gradient of problem (2.8)
– $\|\widetilde{x}_1^*\|$, $\|\widetilde{x}_2^*\|$: the minimum norm solution.
– cpu$_s$: CPU time in seconds.

Table 1 shows that the norm of the minimum norm solutions of all test problems are close to 0 and $norm(df)$ are near to zero as well. It is shown that our proposed method obtains good results in a reasonable time.

4.3. More experiments and comparing with two approaches

In this subsection, more experiments are conducted to compare the performance of the proposed method with Tikhonov regularization (TR) and a standard method which describe as follows. The Tikhonov regularization method has been popular for finding the minimum norm solution of an optimization problem. In this paper, we characterized the solution set of problem (2.1) as a system (2.6), then for finding the minimum norm solution of the problem (2.1), the following convex quadratic programming problem must be solved; this method is denoted by MS.

$$\min \ \frac{1}{2}\|x\|^2 \quad \text{s.t.} \quad x \text{ satisfies (2.6)}. \tag{4.6}$$

In order to compare the presented method with the TR and MS, and also in order to determine the computational behavior of our method, a program was written to randomly generate sets of problems to be solved. The code in Figure 2 generates random problems for which the norm of the minimum norm solution is 0.

Table 2 reports the following information for each test problem:

TABLE 2. Comparison of the numerical results of the AN, TR, and MS (Sect. 4.3).

n	Algorithms	KKT	$\|\widetilde{x}^*\|$	cpu$_s$
1000	AN	1.5534e-022	7.1524e-019	4.12
	TR	4.6531e-032	9.3567e-023	2.94
	MS	3.7778e-09	3.7778e-09	1.14
2000	AN	1.1567e-025	3.3091e-023	4.92
	TR	4.4326e-028	7.6481e-023	4.51
	MS	3.7778e-09	3.7778e-09	2.73
4000	AN	3.0225e-024	3.2351e-024	6.11
	TR	2.3562e-030	2.0452e-026	8.56
	MS	3.7778e-09	3.7778e-09	10.28
5000	AN	8.4027e-024	6.6401e-022	7.31
	TR	4.4326e-028	7.6481e-023	10.12
	MS	3.7778e-09	3.7778e-09	16.60
10000	AN	4.0306e-025	3.6383e-021	8.89
	TR	5.0396e-031	7.6481e-029	12.34
	MS	3.7778e-09	3.7778e-09	107.82
12000	AN	4.1126e-023	4.5093e-023	9.91
	TR	9.4112e-032	4.1780e-020	14.57
	MS	3.7778e-09	3.7778e-09	178.07
15000	AN	4.4326e-022	7.6481e-019	14.55
	TR	4.9032e-030	8.6488e-020	15.84
	MS	3.7778e-09	3.7778e-09	346.03
18000	AN	4.4006e-021	7.6481e-020	16.09
	TR	5.5727e-031	3.4088e-022	18.01
	MS	3.7778e-09	3.7778e-09	642.88
25000	AN	2.0320e-014	7.7086e-012	29.43
	TR	Out of memory	–	–
	MS	Out of memory	–	–
30000	AN	1.1006e-013	3.2071e-011	92.71
	TR	Out of memory	–	–
	MS	Out of memory	–	–

- AN stands for our approach based on theorem of alternatives method and Newton's method;
- TR stands for the Tikhonov regularization method;
- MS stands for solving the problem (4.6) by the interior point method for convex problem, which is a standard method for solving this type of problems;
- n: the size of data $A_{11}, A_{12}, A_{21}, A_{22}, Q_1, Q_2, d_1, d_2$.
- KKT: the norm of homogeneous system related to first-order optimally conditions;
- $\|\widetilde{x}^*\|$: the norm of the minimum norm solution of x;
- cpu$_s$: CPU time in seconds.

Table 2 shows that the norm of the minimum norm solution of all test problems is close to zero. Also the first-order optimally measure must be zero at a minimum; here for all problems it is near to zero. This indicates that we have successfully obtained the minimum norm solution for all test problems.

In order to more analyse these results, we note that AN, TR and MS work well for small- and medium-scale problems, and when the size of problems increases, we see that AN is faster than two other methods and also TR is faster than MS. Another advantage of the proposed method is finding minimum norm of larger-scale problems in an appropriate time while the other two methods are not able to find the minimum norm because the memory cannot support the requirements for the algorithms in larger-scale problems (we stands "Out of memory" in Table 2).

4.4. Markowitz' Mean-Variance portfolio optimization model

In this section, we apply our theory on a practical problem arising in finance. The standard portfolio optimization problem model known as the Markowitz' Mean-Variance portfolio optimization model can be formulated as follows:

$$\begin{aligned} \min_{x} \quad & x^T \Sigma x \\ \text{s.t.} \quad & e^T x = 1, \\ & \mu^T x \geq r, \\ & x \geq 0. \end{aligned} \tag{4.7}$$

Herein, e is a vector of ones and a vector of expected returns μ and the covariance matrix of the returns of the asset Σ are known.

When the covariance matrix is positive semidefinite and rank deficient, the problem (4.7) has multiple optimal solutions, and in this case a natural choice is finding the minimum norm solution. This situation, *i.e.*, the covariance matrix is rank deficient, can be acquired in many times because the covariance matrix is estimated from the past trading price data and when the number of sampled periods is smaller than the number of assets, the covariance matrix is rank deficient. The data for the problem (4.7) can be found at https://vanderbei.princeton.edu/ampl/nlmodels/markowitz/ and we have used the data between the years 1987 and 1990. Thus we have:

$$\Sigma = \begin{bmatrix} 0.0001 & 0.0012 & 0.0008 & 0.0008 & 0.0005 & 0.0005 & -0.0015 & -0.0012 \\ 0.0012 & 0.0139 & 0.0136 & 0.0141 & 0.0135 & 0.0057 & -0.0041 & -0.0132 \\ 0.0008 & 0.0136 & 0.0227 & 0.0236 & 0.0251 & 0.0051 & 0.0167 & -0.0052 \\ 0.0008 & 0.0141 & 0.0236 & 0.0250 & 0.0272 & 0.0052 & 0.0194 & -0.0072 \\ 0.0005 & 0.0135 & 0.0251 & 0.0272 & 0.0311 & 0.0046 & 0.0273 & -0.0082 \\ 0.0005 & 0.0057 & 0.0051 & 0.0052 & 0.0046 & 0.0024 & -0.0032 & -0.0050 \\ -0.0015 & -0.0041 & 0.0167 & 0.0194 & 0.0273 & -0.0032 & 0.0555 & 0.0120 \\ -0.0012 & -0.0132 & -0.0052 & -0.0072 & -0.0082 & -0.0050 & 0.0120 & 0.0285 \end{bmatrix},$$

$$\mu = \begin{bmatrix} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \end{bmatrix}^T, \quad r = 1.05.$$

The covariance matrix Σ can be computed by the following known formula:

$$\Sigma := \frac{1}{N-1} R \left(I - \frac{1}{N} 11^T \right) R^T,$$

where N is the number of periods (here $T = 4$), 11^T is a square matrix of ones, and R is the 8×4 matrix containing the assets' returns for each of the 4 years [3].

The portfolio optimization problem (4.7) has multiple optimal solutions because the rank of matrix Σ is at most 4 and so it is rank deficient. We solved the problem (4.7) by using "quadprog.m" in MATLAB. The norm of optimal solution is $\|x\| = 0.8888$ and the optimal value of the objective function is 2.7692×10^{-4}. The minimum norm solution through our proposed method is $\|x\| = 0.7978$ in 0.07 s with the same value of the objective function. We computed a similar result for the norm of the optimal solution and the optimal value by the MS method, but in 0.53 s. So we can conclude that although both methods work well, our method is faster.

5. Conclusion

In this paper, based on theorems of alternative, we introduced a new way to find the minimum norm solution of convex quadratic programs. In fact, in contrast to the classical Tikhonov regularization method and convex interior point methods, which have to solve a quadratic programming problem, we introduced a reduced simple constrained minimization problem to find the minimum norm solution. Since the objective function of the reduced problem is once-differentiable and convex, to obtain the solution of it, we proposed an extension of Newton's method.

We presented an example to illustrate the effectiveness and correctness of our proposed method, and we also examined different types of problems including a practical problem arising in finance. The numerical results show that, compared to the Tikhonov regularization method and to a standard convex interior point method, our method behaves more efficiently than the two other methods for randomly generated problems, in particular when the dimension grows.

As pointed out by one reviewer, quadratic programs can be reduced to the linear complementarity problem (LCP). Therefore, approaching the minimum norm solution *via* LCP might be an interesting research problem.

Acknowledgements. H. Moosaei and M. Hladík were supported by the Czech Science Foundation Grant P403-18-04735S.

References

[1] L. Armijo, Minimization of functions having lipschitz continuous first partial derivatives. *Pac. J. Math.* **16** (1966) 1–3.

[2] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, Nonlinear Programming: Theory and Algorithms. John Wiley & Sons, New York (2013).

[3] A. Beck and S. Sabach, A first order method for finding minimal norm-like solutions of convex optimization problems. *Math. Program.* **147** (2014) 25–46.

[4] A. Callejo, F. Gholami, A. Enzenhöfer and J. Kövecses, Unique minimum norm solution to redundant reaction forces in multibody systems. *Mech. Mach. Theory* **116** (2017) 310–325.

[5] Y.G. Evtushenko and A. Golikov, Theorems of alternative and optimization, in Optimization and Applications, edited by N. Olenev *et al.* In Vol. 12422 of *Lecture Note in Computer Science*. Springer, Cham (2020) 86–96.

[6] Y.G. Evtushenko and A. Golikov, New perspective on the theorems of alternative. In: High Performance Algorithms and Software for Nonlinear Optimization, edited by G. Di Pillo and A. Murli. Vol.82 of: *Applied Optimization.* Kluwer Academic Publishers Springer (2003) 227–241.

[7] D. Gale, The Theory of Linear Economic Models. University of Chicago Press (1989).

[8] M.S. Gockenbach and E. Gorgin, On the convergence of a heuristic parameter choice rule for Tikhonov regularization. *SIAM J. Sci. Comput.* **40** (2018) A2694–A2719.

[9] A.I. Golikov and Y.G. Evtushenko, Theorems of the alternative and their applications in numerical methods. *Comput. Math. Math. Phys.* **43** (2003) 338–358.

[10] E.R. Hansen and G.W. Walster, Global Optimization Using Interval Analysis, 2nd edition. Marcel Dekker, New York (2004).

[11] J.-B. Hiriart-Urruty, J.-J. Strodiot and V.H. Nguyen, Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data. *Appl. Math. Optim.* **11** (1984) 43–56.

[12] L. Jaulin, M. Kieffer, O. Didrit and É. Walter, Applied Interval Analysis. Springer, London (2001).

[13] C. Kanzow, H. Qi, and L. Qi, On the minimum norm solution of linear programs. *J. Optim. Theory App.* **116** (2003) 333–345.

[14] S. Ketabchi and E. Ansari-Piri, On the solution set of convex problems and its numerical application. *J. Comput. Appl. Math.* **206** (2007) 288–292.

[15] S. Ketabchi and H. Moosaei, Minimum norm solution to the absolute value equation in the convex case. *J. Optim. Theory App.* **154** (2012) 1080–1087.

[16] S. Kindermann and K. Raik, Convergence of heuristic parameter choice rules for convex Tikhonov regularization. *SIAM J. Numer. Anal.* **58** (2020) 1773–1800.

[17] O.L. Mangasarian, A simple characterization of solution sets of convex programs. *Oper. Res. Lett.* **7** (1988) 21–26.

[18] O.L. Mangasarian, A Finite newton method for classification. *Optim. Methods Softw.* **17** (2002) 913–929.

[19] P.M. Pardalos, S. Ketabchi and H. Moosaei, Minimum norm solution to the positive semidefinite linear complementarity problem. *Optimization* **63** (2014) 359–369.

[20] O. Prokopyev, On equivalent reformulations for absolute value equations. *Comput. Optim. App.* **44** (2009) 363–372.

[21] J.B. Rosen, Minimum norm solution to the linear complementarity problem. In: Functional Analysis, Optimization and Mathematical Economics, edited by L.J. Leifman *et al.* Oxford University Press, Oxford (1990) 208–216.

[22] A.N. Tikhonov and V.I. Arsenin, Solutions of Ill-Posed Problems. Winston, Washington, DC (1977).

[23] X.-H. Vu, D. Sam-Haroud and M.-C. Silaghi, Numerical constraint satisfaction problems with non-isolated solutions. In: Global Optimization and Constraint Satisfaction, edited by C. Bliek, C. Jermann and A. Neumaier. Springer, Berlin-Heidelberg (2003) 194–210.

[24] J. Wang and Y. Xia, A dual neural network solving quadratic programming problems. In: Vol. 1 of *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339)*. IEEE (1999) 588–593.

[25] Y. Yao, R. Chen and H.-K. Xu, Schemes for finding minimum-norm solutions of variational inequalities. *Nonlinear Anal.: Theory Methods App.* **72** (2010) 3447–3456.