Optim Lett (2015) 9:1231–1245
DOI 10.1007/s11590-014-0834-7

ORIGINAL PAPER

A half thresholding projection algorithm for sparse solutions of LCPs

Meijuan Shang · Chao Zhang · Dingtao Peng · Shenglong Zhou

Received: 25 October 2013 / Accepted: 13 November 2014 / Published online: 3 December 2014

Abstract In this paper, we aim to find sparse solutions of the linear complementarity problems (LCPs), which has many applications such as bimatrix games and portfolio selection. Mathematically, the underlying model is NP-hard in general. Thus, an $\ell_{1/2}$ regularized projection minimization model is proposed for relaxation. A half thresholding projection (HTP) algorithm is then designed for this regularization model, and the convergence of HTP algorithm is studied. Numerical results demonstrate that the HTP algorithm can effectively solve this regularization model and output very sparse solutions of LCPs with high quality.

Keywords Linear complementarity problems · Sparse solutions · $\ell_{1/2}$ regularized minimization · Half thresholding projection algorithm · Convergence

M. Shang · C. Zhang
Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, People's Republic of China
e-mail: zc.njtu@163.com

M. Shang (✉)
Department of Mathematics, Shijiazhuang University, Shijiazhuang 050035, People's Republic of China
e-mail: meijuanshang@163.com

D. Peng
College of Science, Guizhou University, Guiyang 550025, People's Republic of China
e-mail: dingtaopeng@126.com

S. Zhou
School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK
e-mail: longnan_zsl@163.com

1 Introduction

Given a matrix $M \in \mathbb{R}^{n\times n}$ and a vector $q \in \mathbb{R}^n$, the linear complementarity problem, denoted by LCP(q, M), is to find $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad Mx + q \geq 0, \quad x^T(Mx + q) = 0.$$

The set of solutions to this problem is denoted by SOL(q, M). Throughout this paper, we always suppose SOL$(q, M) \neq \emptyset$.

The LCP has attracted much attention due to its wide applications in operations research, economic equilibrium, and engineering design [4,6,7]. However it seems that there is a vacant study of sparse solutions for LCP, which has many applications such as bimatrix game [4] and portfolio selection [10,18,20]. For more details, see [15].

In this paper, we try to find a sparse solution to the LCP, which has the smallest number of nonzero entries. This can be described as the following minimization problem:

$$\begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \ \ & x \geq 0, \quad Mx + q \geq 0, \quad x^T(Mx + q) = 0, \end{aligned} \tag{1.1}$$

where $\|x\|_0$ stands for the number of nonzero components of x. A solution of (1.1) is called a sparse solution of the LCP.

The above minimization problem (1.1) is in fact a sparse optimization [1–3,5,13,16] with equilibrium constraints. In the view of the objection function, the problem is an ℓ_0 norm minimization problem, which is combinational and generally NP-hard. In the view of constraint conditions, it is in fact a mathematical program with equilibrium constraints (MPEC) [8,9,11,12]. It is not easy to get solutions due to the equilibrium constraints, even for a continuous objective function.

To overcome the difficulty for the complementarity constraints, we make use of the $F_{\min}(x)$ NCP-function to construct the penalty of violating the complementarity constraints. Recall that a function $\psi : \mathbb{R}^2 \to \mathbb{R}$ is called a NCP-function, if for any pair $(a, b) \in \mathbb{R}^2$,

$$\psi(a, b) = 0 \Leftrightarrow a \geq 0, \quad b \geq 0, \quad ab = 0.$$

Then the C-function $\mathbf{F}_\psi(x) : \mathbb{R}^n \to \mathbb{R}^n$ associated with ψ is defined as

$$\mathbf{F}_\psi(x) = \begin{pmatrix} \psi(x_1, F_1(x)) \\ \vdots \\ \psi(x_n, F_n(x)) \end{pmatrix}.$$

According to the above definition, the C-function $\mathbf{F}_{\min}$ associated with the 'min' function can be given by

$$\mathbf{F}_{\min}(x) = \begin{pmatrix} \min(x_1, F_1(x)) \\ \vdots \\ \min(x_n, F_n(x)) \end{pmatrix} = x - \Pi_{\mathbb{R}^n_+}(x - F(x)) \triangleq x - [x - F(x)]_+, \tag{1.2}$$

where $F(x) = Mx + q$ with $q \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$, and $\Pi_{\mathbb{R}^n_+}$ is the Euclidean metric projector onto the nonnegative orthant. It is well known [6] that solving LCP(q, M) is equivalent to solving the fixed point equation $\mathbf{F}_{\min}(x) = 0$, that is

$$x \in \text{SOL}(F) \Leftrightarrow x = [x - F(x)]_+. \tag{1.3}$$

On the other hand, to overcome the difficulty for ℓ_0 norm, many researchers have suggested to relax the ℓ_0 norm and instead, to consider the ℓ_1 norm. Recently, Xu et al. [19] developed a thresholding representation theory for the $\ell_{1/2}$ regularization problem, and proposed a fast iterative half thresholding algorithm for the $\ell_{1/2}$ regularization problem as well as the convergence analysis of this algorithm being given. The experiments show that the $\ell_{1/2}$ regularization bears a significantly stronger sparsity-promoting property than the ℓ_1 regularization. Hence, in this paper we consider applying the $\ell_{1/2}$ norm to find the sparse solution of LCPs, that is, we use the following model

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \|x\|_{1/2}^{1/2} \\ \text{s.t.} \quad & x = [x - F(x)]_+ \end{aligned} \tag{1.4}$$

to approximate problem (1.1), where $\|x\|_{1/2}^{1/2} = \sum_{i=1}^n |x_i|^{1/2}$ is the $\ell_{1/2}$ norm of $x \in \mathbb{R}^n$ and $F(x) = Mx + q$. By introducing a new vector of variables $z \in \mathbb{R}^n$, problem (1.4) can be changed as follows

$$\begin{aligned} \min_{x,z\in\mathbb{R}^n} \quad & f_\lambda(x, z) := \|x - z\|_2^2 + \lambda\|x\|_{1/2}^{1/2} \\ \text{s.t.} \quad & z = [x - F(x)]_+, \end{aligned} \tag{1.5}$$

where $\lambda > 0$ is a regularization parameter. We call (1.5) the $\ell_{1/2}$ regularization projection minimization problem.

This paper is organized as follows. In Sect. 2, we show that problem (1.4) can be solved by solving problem (1.5) in the sense that the solution sequence of problem (1.5) can arbitrarily approximate the solution of problem (1.4). In Sect. 3, we give a closed-form solution associated with the half thresholding function for the key subproblem of (1.5). In Sect. 4, we propose a half thresholding projection (HTP) algorithm for problem (1.5). Numerical results are demonstrated in Sect. 5 to verify the effectiveness of the proposed algorithm for finding the sparse solutions of LCPs.

2 The $\ell_{1/2}$ regularized approximation

In this section, we study the relation between the solutions of model (1.5) and the solutions of model (1.4).

Theorem 2.1 *For any fixed* $\lambda > 0$*, the solution set of* (1.5) *is nonempty and bounded. Let* $\{(x_{\lambda_k}, z_{\lambda_k})\}$ *be a solution of* (1.5)*, and* $\{\lambda_k\}$ *be any positive sequence converging to 0. If* $\mathrm{SOL}(q, M) \neq \emptyset$*, then* $\{(x_{\lambda_k}, z_{\lambda_k})\}$ *has at least one accumulation point, and any accumulation point* x^* *of* $\{x_{\lambda_k}\}$ *is a solution of* (1.4).

Proof For any fixed $\lambda > 0$, it is easy to show the coercivity of $f_\lambda(x, z)$ in (1.5), which relies on the property that

$$f_\lambda(x, z) \to +\infty \quad \text{as} \quad \|(x, z)\| \to \infty. \tag{2.1}$$

We also note that for any $x \in \mathbb{R}^n$ and any $z \in \mathbb{R}^n$, $f_\lambda(x, z) \geq 0$. This together with (2.1) implies that, the level set

$$\mathcal{L} = \{(x, z) \mid f_\lambda(x, z) \leq f_\lambda(x_0, z_0) \text{ and } z = [x - F(x)]_+ \}$$

is nonempty and compact, where $x_0 \in \mathbb{R}^n$ and $z_0 = [x_0 - F(x_0)]_+$ are given points. The solution set of (1.5) is nonempty and bounded since $f_\lambda(x, z)$ is continuous on $\mathcal{L}$.

Now we show the second part of this theorem. Let $\hat{x} \in \mathrm{SOL}(q, M)$ and $\hat{z} = [\hat{x} - F(\hat{x})]_+$. From (1.3), we have $\hat{x} = \hat{z}$. Note that $(x_{\lambda_k}, z_{\lambda_k})$ is a solution of (1.5) with $\lambda = \lambda_k$, where $z_{\lambda_k} = [x_{\lambda_k} - F(x_{\lambda_k})]_+$. Then we have

$$\begin{aligned} \max\left\{\|x_{\lambda_k} - z_{\lambda_k}\|^2, \lambda_k \|x_{\lambda_k}\|_{1/2}^{1/2}\right\} &\leq \|x_{\lambda_k} - z_{\lambda_k}\|^2 + \lambda_k \|x_{\lambda_k}\|_{1/2}^{1/2} \\ &\leq \|\hat{x} - \hat{z}\|^2 + \lambda_k \|\hat{x}\|_{1/2}^{1/2} \\ &= \lambda_k \|\hat{x}\|_{1/2}^{1/2}. \end{aligned} \tag{2.2}$$

From the above inequality, we derive that for any $\lambda_k > 0$,

$$\|x_{\lambda_k}\|_{1/2}^{1/2} \leq \|\hat{x}\|_{1/2}^{1/2}. \tag{2.3}$$

Hence the sequence $\{x_{\lambda_k}\}$ is bounded, and has at least one accumulation point. Note that the sequence $\{z_{\lambda_k}\}$ is also bounded since $\|x_{\lambda_k} - z_{\lambda_k}\|^2 \leq \lambda_k \|\hat{x}\|_{1/2}^{1/2}$ and $\lambda_k \to 0$. Let x^* and z^* be any accumulation points of $\{x_{\lambda_k}\}$ and $\{z_{\lambda_k}\}$, respectively. Then there exists a subsequence of $\{\lambda_k\}$, say $\{\lambda_{k_j}\}$, such that

$$\lim_{k_j \to \infty} x_{\lambda_{k_j}} = x^* \text{ and } \lim_{k_j \to \infty} z_{\lambda_{k_j}} = z^*.$$

We can obtain $z^* = [x^* - F(x^*)]_+$ by letting k_j tending to ∞ in $z_{\lambda_{k_j}} = [x_{\lambda_{k_j}} - F(x_{\lambda_{k_j}})]_+$. Letting k_j tend to ∞ in

$$\|x_{\lambda_{k_j}} - z_{\lambda_{k_j}}\|^2 \leq \lambda_{k_j} \|\hat{x}\|_{1/2}^{1/2},$$

we get $x^* = z^*$. Consequently, we get $x^* = [x^* - F(x^*)]_+$, that is, $x^* \in \mathrm{SOL}(q, M)$. From $\|x_{\lambda_{k_j}}\|_{1/2}^{1/2} \leq \|\hat{x}\|_{1/2}^{1/2}$ with k_j tending to ∞, we get $\|x^*\|_{1/2}^{1/2} \leq \|\hat{x}\|_{1/2}^{1/2}$. Then by the arbitrariness of $\hat{x} \in \mathrm{SOL}(q, M)$, we know x^* is a solution of problem (1.4). This completes the proof. □

3 Solution representation of the subproblem

Given $z^k \in \mathbb{R}_+^n$, the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} f_\lambda(x, z^k) := \|x - z^k\|_2^2 + \lambda \|x\|_{1/2}^{1/2} \tag{3.1}$$

is a key subproblem in our proposed algorithm. In what follows, we will show that this subproblem has a closed-form solution associated with a so-called half thresholding function, which is obtained from Theorem 2 in [19] and Lemma 3.1 in [14].

Lemma 3.1 *Let $t \geq 0$ and $\lambda > 0$ be any two given real numbers. Suppose that $x^* \in \mathbb{R}$ is a solution of the following problem*

$$\min_{x \in \mathbb{R}} f(x) := (x - t)^2 + \lambda x^{1/2}.$$

Then x^* can be analytically expressed by

$$x^* = h_\lambda(t) := \begin{cases} h_{\lambda,1/2}(t), & \text{if } t > \frac{\sqrt[3]{54}}{4}\lambda^{2/3}, \\ \{h_{\lambda,1/2}(t),\ 0\}, & \text{if } t = \frac{\sqrt[3]{54}}{4}\lambda^{2/3}, \\ 0, & \text{if } 0 \leq t < \frac{\sqrt[3]{54}}{4}\lambda^{2/3}, \end{cases}$$

where

$$h_{\lambda,1/2}(t) = \frac{2}{3}t\left(1 + \cos\left(\frac{2\pi}{3} - \frac{2}{3}\phi_\lambda(t)\right)\right)$$

with

$$\phi_\lambda(t) = \arccos\left(\frac{\lambda}{8}\left(\frac{|t|}{3}\right)^{-\frac{3}{2}}\right).$$

The involved function $h_\lambda(\cdot)$ in the above lemma is always called a half thresholding function, see [14,19].

An important observation is that

$$\begin{aligned}&\min_{x\in\mathbb{R}^n}\|x-z^k\|_2^2+\lambda\|x\|_{1/2}^{1/2}\\ \Longleftrightarrow\ &\min_{x_1,x_2,\ldots,x_n\in\mathbb{R}}\sum_{i=1}^{n}[(x_i-z_i^k)^2+\lambda|x_i|^{1/2}]\\ \Longleftrightarrow\ &\min_{x_i\in\mathbb{R}}(x_i-z_i^k)^2+\lambda|x_i|^{1/2},\ \text{for each}\ i=1,\ldots,n.\end{aligned}$$

This allows us to directly obtain the closed-form solution of subproblem (3.1), as stated in the following theorem.

Theorem 3.2 *Given $z^k\in\mathbb{R}^n_+$ and $\lambda>0$, the solution x^s of subproblem (3.1) can be analytically given by*

$$x^s=H_\lambda(z^k),$$

where

$$H_\lambda(z^k):=\Big(h_\lambda(z_1^k),h_\lambda(z_2^k),\ldots,h_\lambda(z_n^k)\Big)$$

with $h_\lambda(\cdot)$ the thresholding function given in Lemma 3.1.

To simplify the iterative process and for the aim of finding sparse solutions, we make a slight adjustment of h_λ as follows

$$h_\lambda(t)=\begin{cases}h_{\lambda,1/2}(t), & \text{if}\ \ t>\frac{\sqrt[3]{54}}{4}\lambda^{2/3},\\ 0, & \text{if}\ \ 0\le t\le\frac{\sqrt[3]{54}}{4}\lambda^{2/3}.\end{cases}\tag{3.2}$$

4 Algorithm and convergence

Now we give our half thresholding projection (HTP) algorithm as below for solving problem (1.5). From (1.1), we can easily notice that $x=0$ is just the minimizer of (1.1), when $q\geq 0$. Hence, we design the HTP algorithm to solve the problem (1.5) when $q\ngeq 0$, that is, there exists $i\in\{1,2,\ldots,n\}$ such that $q_i<0$.

In HTP algorithm, $\underline{\lambda}$ is a small positive number near zero, and $n_{\max}$ is a large integer that forces the HTP algorithm to stop if it does not satisfy the convergent criteria $\|z^{k+1}-x^{k+1}\|\leq\epsilon$. To fasten the computational speed, we do not really obtain $\{x_{\lambda_k},z_{\lambda_k}\}$ by exactly solving (1.5) as mentioned in Theorem 2.1, but only approximately solving it according to (4.2). Numerical experiments demonstrate that this modification will not affect the solution quality but save computational cost.

We require the step size α_{k+1} satisfies the inequality and $z^{k+1}\neq 0$ in (4.1). In fact, $z^{k+1}\neq 0$ guarantees that the right-hand side is positive for the next iterate, and consequently the step size α_{k+2} can be found to meet the inequality in (4.1) in the next iterate. See the proof of Lemma 4.2 below. Later, the inequality in (4.1) is essential to

HTP Algorithm
Input: $M \in \mathbb{R}^{n\times n}$, $q \in \mathbb{R}^n$ and $q \ngeq 0$, $F(x) := Mx + q$.
Step 0. Initialize $0 \neq z^0 \in \mathbb{R}^n_+$, $z^0 \neq x^0 \in \mathbb{R}^n_+$, $\lambda_0 > \underline{\lambda} > 0, \beta > 0, \epsilon > 0$, $\gamma, \tau \in (0, 1)$, and integers $n_{\max} > K > 0$. Set $k = 0$.
Step 1. Update $x^{k+1} = H_{\lambda_k}(z^k)$.
Update $z^{k+1} = [x^{k+1} - \alpha_{k+1}F(x^{k+1})]_+$, where $\alpha_{k+1} = \beta\gamma^{m_{k+1}}$ and m_{k+1} is the smallest nonnegative integer satisfying

$$\|x^{k+1}-[x^{k+1}-\beta\gamma^{m_{k+1}}F(x^{k+1})]_+\|^2+\beta\gamma^{m_{k+1}}(\|x^{k+1}-x^k\|^2+\|x^k-z^k\|^2)<\|x^{k+1}-z^k\|^2 \text{ and } [x^{k+1}-\beta\gamma^{m_{k+1}}F(x^{k+1})]_+\neq 0. \tag{4.1}$$

Step 2. Update

$$\lambda_{k+1} = \begin{cases} \max\{\underline{\lambda}, \tau\lambda_k\}, & \text{if } k \text{ is the integral multiple of } K, \\ \lambda_k, & \text{otherwise.} \end{cases} \tag{4.2}$$

Stop rule: If $\|z^{k+1} - x^{k+1}\| \leq \epsilon$ or the number of iterations is greater than $n_{\max}$, stop.
Otherwise, set $k := k + 1$ and return to Step **1**.

show the main convergent result in Theorem 4.3 that $f_{\lambda_k}(x^k, z^k)$ is strictly decreasing and

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0 \quad \text{and} \quad \lim_{k\to\infty} \|z^{k+1} - z^k\| = 0.$$

Moreover, any accumulation point of $\{x^k\}$ is a solution of LCP(q, M), under the assumption of $\inf_k \alpha_k = \alpha > 0$. We also mention that in numerical experiments, if we just choose fixed α_k, e.g., $\alpha_k \equiv 1$, the algorithm fails for the case of positive semidefinite matrix in Sect. 5.

Now we will show that the HTP algorithm is well defined, that is, (4.1) is implementable. Before doing this, we need the following lemma.

Lemma 4.1 [17] *Let P_Ω be a metric projection operator onto a nonempty closed convex set $\Omega \in \mathbb{R}^n$. Given $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$, define*

$$H(\alpha) = P_\Omega[x - \alpha d], \quad \alpha \geq 0,$$

then $\|H(\alpha) - x\|$ is nondecreasing with respect to α.

Lemma 4.2 *The stepsize α_{k+1} in (4.1) must exist.*

Proof We now prove α_{k+1} in (4.1) must exist by induction.

Denote

$$g(\alpha) := \|x^{k+1} - [x^{k+1} - \alpha F(x^{k+1})]_+\|^2 + \alpha(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2). \tag{4.3}$$

From Lemma 4.1, we get $\|x^{k+1} - [x^{k+1} - \alpha F(x^{k+1})]_+\|$ is nondecreasing with respect to α. When $x^k \neq z^k$, it is obvious that $\alpha(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2)$ is strictly increasing with respect to α. It follows that $g(\alpha)$ is strictly increasing with respect to α. Thus $g(\beta\gamma^l)$ is strictly decreasing with respect to the nonnegative integer l.

For $k = 0$, using the fact that $0 \neq z^0 \in \mathbb{R}^n_+$, we show below that

(a) $\exists\, \tilde{\alpha}_1 = \beta\gamma^l > 0$ such that $g(\tilde{\alpha}_1) < \|x^1 - z^0\|^2$;
(b) $\exists\, \alpha_1 = \beta\gamma^{m_1} \in [0, \tilde{\alpha}_1]$ such that $[x^1 - \alpha_1 F(x^1)]_+ \neq 0$.

One can check that $h_\lambda(t) < t$ for $\lambda > 0$ and any $t > 0$ by image of $h_\lambda(t)$. It follows with $0 \neq z^0 \in \mathbb{R}^n_+$ that

$$\|x^1 - z^0\|^2 = \|H_{\lambda_0}(z^0) - z^0\|^2 > 0.$$

Note that $x^1 = H_{\lambda_0}(z^0) \in \mathbb{R}^n_+$, then we have

$$g(0) = \|x^1 - [x^1]_+\| = 0 < \|x^1 - z^0\|^2.$$

It is then clear that (a) holds since $g(\beta\gamma^l)$ is strictly decreasing with respect to the nonnegative integer l.

Next we prove (b). We claim that there must exist $\bar{\alpha} \in [0, \tilde{\alpha}_1]$ such that $[x^1 - \alpha F(x^1)]_+ \neq 0$ for all $\alpha \in [0, \bar{\alpha}]$. Otherwise there exists an infinite sequence $\{\bar{\alpha}^j\} \subset [0, \bar{\alpha}]$ with $\bar{\alpha}^j \downarrow 0$ as $j \to \infty$, such that $[x^1 - \bar{\alpha}^j F(x^1)]_+ = 0$. This indicates that $x^1 = 0$ and $-\bar{\alpha}^j F(x^1) = -\bar{\alpha}^j q \leq 0$. Hence, $q \geq 0$ which contradicts $q \ngeq 0$ in HTP algorithm. Thus we can find $\alpha_1 = \beta\gamma^{m_1} \in [0, \bar{\alpha}]$ since $g(\alpha)$ is strictly increasing for α. Also note that $g(\bar{\alpha}) \leq g(\tilde{\alpha}_1) < \|x^1 - z^0\|^2$. Therefore the stepsize α_1 in (4.1) exists for $k = 0$.

Suppose that the stepsize α_r exists in (4.1) for $k = r - 1$. It follows that

$$0 \neq z^r = [x^r - \beta\gamma^{m_r} F(x^r)]_+ \in \mathbb{R}^n_+.$$

Similar to the proof for the case $k = 0$, we can show α_{r+1} in (4.1) must exist for $k = r$. By induction, the stepsize α_k in (4.1) must exist for all k. □

We now begin to analyze the convergence of the proposed HTP Algorithm.

Theorem 4.3 *Let $\{(x^k,\ z^k)\}$ be a sequence generated by the HTP Algorithm, then*

(i) *$\{f_{\lambda_k}(x^k, z^k)\}$ is strictly decreasing and converges to a constant C^*;*
(ii) *Both $\{x^k\}$ and $\{z^k\}$ are bounded;*
(iii) *Suppose $\inf_k \alpha_k = \alpha > 0$, then $\{x^k\}$ and $\{z^k\}$ are both asymptotically regular, i.e.,*

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0 \text{ and } \lim_{k\to\infty} \|z^{k+1} - z^k\| = 0.$$

Moreover, any accumulation point of $\{x^k\}$ is a solution of LCP(q, M).

Proof (i) By $x^{k+1} = H_{\lambda_k}(z^k)$ and Theorem 3.2, we have

$$x^{k+1} = \arg\min_{x\in\mathbb{R}^n} f_{\lambda_k}(x, z^k).$$

It follows that

$$f_{\lambda_k}(x^{k+1}, z^k) \le f_{\lambda_k}(x^k, z^k). \tag{4.4}$$

By (4.2), λ_k is nonincreasing, from which and (4.1), we get

$$\begin{aligned} &f_{\lambda_{k+1}}(x^{k+1}, z^{k+1}) - f_{\lambda_k}(x^{k+1}, z^k) \\ &\quad = \|x^{k+1} - z^{k+1}\|^2 - \|x^{k+1} - z^k\|^2 + (\lambda_{k+1} - \lambda_k)\|x^{k+1}\|_{1/2}^{1/2} \\ &\quad \le \|x^{k+1} - [x^{k+1} - \alpha_{k+1}F(x^{k+1})]_+\|^2 - \|x^{k+1} - z^k\|^2 \\ &\quad < -\alpha_{k+1}(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2). \end{aligned}$$

Hence,

$$f_{\lambda_{k+1}}(x^{k+1}, z^{k+1}) < f_{\lambda_k}(x^{k+1}, z^k) - \alpha_{k+1}(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2). \tag{4.5}$$

Combining (4.4) with (4.5), we get

$$\begin{aligned} f_{\lambda_{k+1}}(x^{k+1}, z^{k+1}) &< f_{\lambda_k}(x^k, z^k) - \alpha_{k+1}(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2) \\ &\le f_{\lambda_k}(x^k, z^k), \end{aligned} \tag{4.6}$$

which shows that $\{f_{\lambda_k}(x^k, z^k)\}$ is strictly decreasing. Since $\{f_{\lambda_k}(x^k, z^k)\}$ is bounded from below, $\{f_{\lambda_k}(x^k, z^k)\}$ converges to a constant C^*.

(ii) Noting that $\lambda_k \ge \bar{\lambda} > 0$, we get

$$(x^k, z^k, \lambda_k) \in \{(x, z, \lambda) \in \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R}_+ : f_{\underline{\lambda}}(x, z) \le f_\lambda(x, z) \le f_{\lambda_0}(x^0, z^0)\},$$

which indicates that $\{x^k\}$ and $\{z^k\}$ are bounded.

(iii) From (4.6) and $\inf_k \alpha_k = \alpha > 0$, we have

$$\begin{aligned} f_{\lambda_k}(x^k, z^k) - f_{\lambda_{k+1}}(x^{k+1}, z^{k+1}) &> \alpha_{k+1}(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2) \\ &\ge \alpha(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2). \end{aligned}$$

This then implies for any positive integer N,

$$\begin{aligned} \sum_{k=0}^{N}(\|x^{k+1} - x^k\|^2 + \|x^k - z^k\|^2) &< \frac{1}{\alpha}\sum_{k=0}^{N}[f_{\lambda_k}(x^k, z^k) - f_{\lambda_{k+1}}(x^{k+1}, z^{k+1})] \\ &= \frac{1}{\alpha}[f_{\lambda_0}(x^0, z^0) - f_{\lambda_{N+1}}(x^{N+1}, z^{N+1})] \\ &\le \frac{1}{\alpha} f_{\lambda_0}(x^0, z^0). \end{aligned}$$

Thus we get $\sum_{k=0}^{\infty}(\|x^{k+1}-x^k\|^2+\|x^k-z^k\|^2)<+\infty$, which yields

$$\|x^{k+1}-x^k\|\to 0 \quad \text{and} \quad \|x^k-z^k\|\to 0 \quad \text{as} \quad k\to +\infty. \tag{4.7}$$

The above two limitations and the following inequality

$$\|z^{k+1}-z^k\|\le \|z^{k+1}-x^{k+1}\|+\|x^{k+1}-x^k\|+\|x^k-z^k\|$$

yield

$$\|z^{k+1}-z^k\|\to 0 \quad \text{as} \quad k\to +\infty.$$

Since $\{x^k\}$ is bounded, $\{x^k\}$ has at least one accumulation. Let x^* be an accumulation point of $\{x^k\}$ and a subsequence $\{x^{k_j}\}$ converge to x^*. Since $\{\alpha_{k_j}\}\subset[\alpha,\beta]$, without loss of generality, we suppose $\alpha_{k_j}\to\bar{\alpha}\in[\alpha,\beta]$ as $k_j\to\infty$. It follows that

$$z^{k_j}=[x^{k_j}-\alpha_{k_j}F(x^{k_j})]_+\to z^*:=[x^*-\bar{\alpha}F(x^*)]_+ \quad \text{as} \quad k_j\to\infty. \tag{4.8}$$

Combining (4.8) with (4.7), we get $x^*=z^*$, which gives $x^*=[x^*-\bar{\alpha}F(x^*)]_+$ and means $x^*\in \mathrm{SOL}(q,M)$. The proof is thus complete.

□

Remark 4.1 According to the result (iii) of the above theorem, a solution of LCP(q,M) can be obtained and $\|x_k-z_k\|\to 0$, which means we find a feasible solution of (1.1) under the assumption $\inf_k \alpha_k=\alpha>0$. Also from (i), our HTP algorithm can guarantee $f_{\lambda_k}(x^k,z^k)$ is strictly decreasing and converges to a constant, which means $\lambda_k\|x^k\|_{1/2}^{1/2}$ is strictly decreasing. As λ_k is not more than a small positive constant $\underline{\lambda}$, the sparsity of any accumulation point of $\{x^k\}$ is promising.

Remark 4.2 The convergent result in (iii) is obtained under the assumption $\inf_k \alpha_k=\alpha>0$ which is relatively strict. Furthermore, under which circumstances this assumption holds is not known yet. However, the numerical results in the next section are very well in spite of the above weakness in convergent theory of HTP algorithm. We will consider to further strengthen the convergent theory in our future work.

5 Numerical experiments

In this section, we will present some numerical experiments to demonstrate the effectiveness of our HTP algorithm. All the numerical experiments were performed on a laptop (2.60GHz, 7.82GB of RAM) by utilizing MATLAB R2013a.

We will stimulate two examples to implement the HTP algorithm. The first one associated with the Z-matrix is from literature [15], in which we initialize $\gamma=0.99$, $\beta=1$, $\tau=1/7$, $K=5$ and $\lambda_0=10$. Since there is a unique sparse solution $\hat{x}$ of LCPs when M is a kind of Z-matrix, we will take advantage of recovery error $\|x-\hat{x}\|$

to evaluate our algorithm. Apart from that, cpu time (in seconds), iteration times and the residual $\|x - z\|$ will also be taken into consideration of judging the performance of the method.

However, in the second example, the unique solution of LCPs seems not to be evident. We hence will synthesize a sparse optimal solution $\hat{x}$ of LCPs such that

$$\hat{x} \geq 0, \;\; F(\hat{x}) \geq 0, \;\; \hat{x}^T F(\hat{x}) = 0, \quad \text{and} \quad \|\hat{x}\|_0 = 0.01 * n.$$

Through the stimulated data sets, we again consider the items above to check the performance of the HTP algorithm. In the latter example $\gamma = 0.1$, $\beta = 0.75$, $\tau = 1/7$, $K = 5$, $\lambda_0 = 5$ and $\underline{\lambda} = 1.0e - 5$.

For the first fixed example, it will only be run once. However, for the second example which is randomly generated, 100 times will be run for each dimension $n = 1{,}000, 3{,}000, 5{,}000, 7{,}000, 10{,}000$, and thus average results will be recorded. In each experiment, we set z^0 to the e vector and x^0 to the zero vector as the initial points and the maximum number of iterations n_{max} as 200 and the stop criteria as $\epsilon = 10^{-6}$.

5.1 Test for LCPs with Z-matrix [15]

Let us consider LCP(q, M) where

$$M = I_n - \frac{1}{n} ee^T = \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1 - \frac{1}{n} \end{pmatrix} \quad \text{and} \quad q = \begin{pmatrix} \frac{1}{n} - 1 \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}.$$

Here I_n is the identity matrix of order n and $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$. Such a matrix M is widely used in statistics. It is clear that M is a positive semidefinite Z-matrix. For any scalar $a \geq 0$, we know that the vector $x = ae + e_1$ is a solution to LCP(q, M), since it satisfies that

$$x \geq 0, \quad Mx + q = Me_1 + q = 0, \quad x^T(Mx + q) = 0.$$

Among all the solutions, the vector $\hat{x} = e_1 = (1, 0, \ldots, 0)^T$ is the unique sparse solution.

As indicated in Table 1, the HTP algorithm behaves very robust because the times of iteration are identically equal to 35, the recovered error $\|x - \hat{x}\|$ and residual $\|x - z\|$ are basically similar, with 4.15E−06 and 3.94E−06 respectively. In addition, the sparsity $\|x\|_0$ of recovered solution x are all 1s, which means the recover is successful. Most importantly, the HTP algorithm is exceptionally fast, which results in only 3.16 s are needed to address the problem with dimension $n = 10{,}000$.

In order to illustrating the effectiveness of the HTP algorithm we proposed, we introduce another method of tackling the LCPs. In [15], the authors established an l_p $(0 < p < 1)$ regularized minimization model, for more details one can discern [15],

Table 1 HTP's computational results on LCPs with Z-matrices

n	Iter	$\|x-\hat{x}\|$	$\|x-z\|$	$\|\hat{x}\|_0$	$\|x\|_0$	Time (s)
1,000	35	4.15E−06	3.94E−06	1	1	0.08
3,000	35	4.13E−06	3.93E−06	1	1	0.41
5,000	35	4.13E−06	3.93E−06	1	1	0.85
7,000	35	4.13E−06	3.93E−06	1	1	1.63
10,000	35	4.13E−06	3.93E−06	1	1	3.16

Table 2 SSG's computational results on LCPs with Z-matrices

n	Iter	$\|x-\hat{x}\|$	$\|\hat{x}\|_0$	$\|x\|_0$	Time (s)
100	1012	1.86E−05	1	1	2.46
200	972	1.70E−05	1	1	5.11
500	118	2.67E−06	1	1	3.88
1,000	118	1.58E−06	1	1	23.04
2,000	117	1.05E−06	1	1	139.15
3,000	117	8.69E−07	1	1	401.33
5,000	–	–	–	–	–

$$\min_{x\in R^n} f(x) := \frac{1}{2}\|\Phi_{FB}(x)\|^2 + \lambda\|x\|_p^p. \tag{5.1}$$

and designed a sequential smoothing gradient (SSG) method to solve the l_p regularized model and get a sparse solution of LCP(q, M). The results are displayed in Table 2.

It can be discerned in Table 2, where '–' denotes the method is invalid. Although the sparsity $\|x\|_0$ of recovered solution are all as large as $\|x-\hat{x}\| = 1$ and the recovered errors $\|x-\hat{x}\|$ are pretty small, the average cpu time dramatically ascends with the matrix dimension n, which manifests that SSG method for LCPs is appropriate for the small dimensional data set and thus SSG will be not appealing when n is relatively large.

5.2 Test for LCPs with positive semidefinite matrices

In this subsection, we test HTP for randomly created LCPs with positive semidefinite matrices. First, we state the way of constructing LCPs and their solutions. Let a matrix $Z \in \mathbb{R}^{n\times r}$ ($r < n$) be generated with the standard normal distribution and let $M = ZZ^T$. Let the sparse vector $\hat{x}$ be produced by choosing randomly the $s = 0.01 * n$ nonzero components whose values are also randomly generated from a standard normal distribution. After the matrix M and the sparse vector $\hat{x}$ have been generated, a vector $q \in \mathbb{R}^n$ can be constructed such that $\hat{x}$ is a solution of the LCP(q, M). Then $\hat{x}$ can be regarded as a sparse solution of the LCP(q, M). Namely,

Table 3 Results on randomly created LCPs with positive semidefinite matrices

n	Iter	$\|x-\hat{x}\|$	$\|x-z\|$	$\|\hat{x}\|_0$	$\|x\|_0$	Time (s)
1,000	35.4	7.97E−06	5.61E−06	10	10	0.07
3,000	65.4	4.61E−05	8.97E−06	30	30	0.75
5,000	42.8	2.56E−05	8.47E−06	50	50	1.16
7,000	39.6	1.58E−05	7.96E−06	70	70	1.81
10,000	39.2	9.56E−06	6.88E−06	100	100	3.59

$$\hat{x} \geq 0,\ M\hat{x}+q \geq 0,\ \hat{x}^T(M\hat{x}+q)=0, \quad \text{and} \quad \|\hat{x}\|_0 = 0.01 * n.$$

To be more specific, if $\hat{x}_i > 0$ then choose $q_i = -(M\hat{x})_i$, if $\hat{x}_i = 0$ then choose $q_i = |(M\hat{x})_i| - (M\hat{x})_i$. Let M and q be the input to our HTP algorithm, then HTP will output a solution x. Similarly, $\|x-\hat{x}\|$, average cpu time (in seconds), average iteration times and the residual $\|x-z\|$ will also be taken into consideration of valuating our HTP algorithm.

As manifested in Table 3, the HTP algorithm performs quite efficiently. For one thing, the average times of iteration, the recovered error $\|x-\hat{x}\|$ and residual $\|x-z\|$ are all decreasing with the rising of $n \geq 3{,}000$, which indicates that the HTP algorithm probably behaves relatively effectively in high dimensional cases. For another, the sparsity $\|x\|_0$ of recovered solution x are all equal to the sparsity $\|\hat{x}\|_0$, which means the recover is exact. Likewise, the HTP algorithm is exceptionally fast in this example, which makes that only 3.59 seconds are needed to pursue the sparse solution of LCP when the dimension $n = 10{,}000$.

Remark 5.1 There are several comments can be concluded after the numerical experiments.

- HTP algorithm performs exceptional well, particularly in high dimensional data set.
- Since only relatively small times of iteration are needed to pursue the sparse optimal solutions, our HTP algorithm behaves extremely robust. What is more, the average cpu time are approximately close to each other in the two examples, nearly 3.5 s when $n = 10{,}000$ for instance, which implies the method we proposed is very effective and fast.
- We can also conclude that the HTP algorithm does not much depend on the choice of λ_0, as we have tested that the results from $\lambda_0 = 1, 5, 10, 20$ were basically similar (because of this, we omitted the information derived from different λ_0 for simplicity).

6 Conclusions

In this paper, we concentrate on finding sparse solutions of the linear complementarity problems (LCPs). Mathematically, this problem is a sparse optimization with complementary constraints, which is NP-hard in general. Thus, an $l_{1/2}$ regularized projection minimization model is proposed for relaxation. A half thresholding projection (HTP) algorithm is then designed for this regularization model. Through developing a closed-form solution associated with the half thresholding function for a key subproblem of $l_{1/2}$ regularized projection minimization problem, our algorithm can efficiently handle large scale problems. By applying HTP algorithm, the objective value monotonically decreases and converges to a constant, and the accumulation point of $\{x^k\}$ generated by HTP algorithm is a solution of LCP(q, M). Numerical results demonstrate that the HTP algorithm can solve this regularization model efficiently.

Acknowledgments This research was supported by the National Natural Science Foundation of China (71271021,11001011,11431002), the Fundamental Research Funds for the Central Universities of China (2013JBZ005) and the Scientific Research Fund of Hebei Provincial Education Department (No. QN20132030)s. The authors are grateful to the referees for their valuable comments which improve the presentation of this paper.

References

1. Candès, E.J., Randall, P.A.: Highly robust error correction by convex programming. IEEE Trans. Inf. Theory **54**, 2829–2840 (2006)
2. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. of Comput. Math. **9**, 717–772 (2008)
3. Candès, E.J., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. **59**, 1207–1223 (2006)
4. Cottle, R.W., Pang, J.S., Stone, R.E.: The Linear Complementarity Problem. Academic Press, Boston (1992)
5. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)
6. Facchinei, F., Pang, J.S.: Finite-dimensional variational inequalities and complementarity problems. Springer Series in Operations Research, vols. I and II. Springer, New York (2003)
7. Ferris, M.C., Mangasarian, O.L., Pang, J.S.: Complementarity: Applications, Algorithms and Extensions. Kluwer Academic Publishers, Dordrecht (2001)
8. Fukushima, M., Pang, J.S.: Some feasibility issues in mathematical programs with equilibrium constraints. SIAM J. Optim. **8**, 673–681 (1998)
9. Fukushima, M., Tseng, P.: An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints. SIAM J. Optim. **12**, 724–739 (2002)
10. Gao, J., Li, D.: Optimal cardinality constrained portfolio selection. Oper. Res. **61**, 745–761 (2013)
11. Lin, G., Fukushima, M.: New reformulations for stochastic nonlinear complementarity problems. Optim. Methods Softw. **21**, 551–564 (2006)
12. Luo, Z. Q., Pang, J. S., Ralph, D.: Mathematical programs with equilibrium constraints. Cambridge University Press, New York (1996)
13. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**, 607–609 (1996)
14. Peng, D., Xiu, N., Yu, J.: $S_{1/2}$ Regularization methods and fixed point algorithms for affine rank minimization problems. http://www.optimization-online.org/DB_FILE/2013/01/3727
15. Shang, M., Zhang, C., Xiu, N.: Minimal zero norm solutions of linear complementarity problems. J. Optim. Theory Appl. **163**, 795–814 (2014)
16. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. **46**, 431–439 (1996)
17. Toint, P.H.L.: Global convergence of a class of trust region methods for nonconvex minimization in Hilbert space. IMA J Numer. Anal. **8**, 231–252 (1988)
18. Xie, J., He, S., Zhang, S.: Randomized portfolio selection with constraints. Pac. J. Optim. **4**, 87–112 (2008)
19. Xu, Z., Chang, X., Xu, F., Zhang, H.: $L_{1/2}$ regularization: a thresholding representation theory and a fast solver. IEEE Trans. Neural Netw. Learn. Syst. **23**, 1013–1027 (2012)

20. Zheng, X., Sun, X., Li, D.: Improving the performance of MIQP solvers for quadratic programs with cardinality and minimum threshold constraints: a semidefinite program approach. http://www.optimization-online.org/DB_FILE/2010/11/2797