

RESEARCH

Open Access

Extragradient thresholding methods for sparse solutions of co-coercive NCPs

Meijuan Shang^{1,2*}, Shenglong Zhou³ and Naihua Xiu¹

*Correspondence:

meijuanshang@163.com

¹Department of Applied Mathematics, Beijing Jiaotong University, Beijing, 100044, P.R. China
²Department of Mathematics, Shijiazhuang University, Shijiazhuang, 050035, P.R. China
Full list of author information is available at the end of the article**Abstract**

In this paper, we aim to find sparse solutions of co-coercive nonlinear complementarity problems (NCPs). Mathematically, the underlying model is NP-hard in general. Thus an ℓ_1 regularized projection minimization model is proposed for relaxation and an extragradient thresholding algorithm (ETA) is then designed for this regularized model. Furthermore, we analyze the convergence of this algorithm and show any cluster point of the sequence generated by ETA is a solution of NCP. Numerical results demonstrate that the ETA can effectively solve the ℓ_1 regularized model and output very sparse solutions of co-coercive NCPs with high quality.

MSC: 90C33; 90C26; 90C90**Keywords:** nonlinear complementarity problems; sparse solution; co-coercive; ℓ_1 relaxation; extragradient; convergence**1 Introduction**

The nonlinear complementarity problem, denoted by the $\text{NCP}(F)$, is to find a vector $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^\top F(x) = 0,$$

where F is a mapping from \mathbb{R}^n into itself. The set of solutions to this problem is denoted by $\text{SOL}(F)$. Throughout this paper, we assume $\text{SOL}(F) \neq \emptyset$.

NCPs have various important applications in economics and engineering, such as Nash equilibrium problems, traffic equilibrium problems, contact mechanics problems, option pricing. Extensive studies of NCPs have been done; see [1–3] and the references therein. Numerical methods for solving NCPs, such as filter method, continuation method, non-smooth Newton's method, smoothing Newton methods, Levenberg-Marquardt method, projection method, descent method, interior-point method have been extensively investigated in the literature. However, it seems that there is a vacant study of sparse solutions for NCPs. In fact, in real applications, it is very necessary to investigate the sparse solution of the NCPs. For example this is so in bimatrix games [1] and portfolio selections [4]. For more details, see [5].

In this paper, we try to compute a sparse solution of the $\text{NCP}(F)$, which has the smallest number of nonzero entries. To be specific, we seek a vector $x \in \mathbb{R}^n$ by solving the ℓ_0 -norm minimization problem

$$\begin{aligned} & \min \|x\|_0 \\ & \text{s.t. } x \geq 0, F(x) \geq 0, x^\top F(x) = 0, \end{aligned} \tag{1}$$

where $\|x\|_0$ stands for the number of nonzero components of x . A solution of (1) is called the sparse solution of $\text{NCP}(F)$.

The above minimization problem (1) is in fact a sparse optimization [6–9] with equilibrium constraints. In the view of the objection function, the problem is ℓ_0 -norm minimization problem, which is combinatorial and generally NP-hard. From the point of view of constraint conditions, it is in fact a mathematical program with equilibrium constraints (MPEC) [10–13]. It is not easy to get solutions due to the equilibrium constraints, even for a continuous objective function.

To overcome the difficulty for the ℓ_0 -norm, many researchers have suggested to relax the ℓ_0 norm and, instead, to consider the ℓ_1 norm; see [8, 9, 14–17]. Motivated by this outstanding work, we consider applying ℓ_1 norm minimization to find the sparse solution of NCPs, and we obtain the following minimization problem to approximate (1):

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} \|x\|_1 \\ & \text{s.t. } x \geq 0, F(x) \geq 0, x^\top F(x) = 0, \end{aligned} \tag{2}$$

where $\|x\|_1 = \sum_{i=1}^n |x_i|$.

To overcome the difficulty for the complementarity constraint, we make use of the C -function $\mathbf{F}_{\min}(x)$ to construct the penalty of violating the complementarity constraints. The C -function \mathbf{F}_{\min} associated with the ‘min’ function can be given by

$$\mathbf{F}_{\min}(x) = x - \Pi_{\mathbb{R}_+^n}(x - F(x)) \triangleq x - [x - F(x)]_+, \tag{3}$$

where F is a mapping from \mathbb{R}^n into itself, and $\Pi_{\mathbb{R}_+^n}$ is the Euclidean metric projector onto the nonnegative orthant.

It is well known [2] that solving $\text{NCP}(F)$ is equivalent to solving the fixed point equation $\mathbf{F}_{\min}(x) = 0$, that is,

$$x \in \text{SOL}(F) \iff x = [x - F(x)]_+, \tag{4}$$

where $[\cdot]_+$ is the Euclidean metric projector onto the nonnegative orthant.

Combining (2) and (4), by introducing a new variable $z \in \mathbb{R}^n$, we obtain the following regularized minimization problem:

$$\begin{aligned} & \min_{x, z \in \mathbb{R}^n} f_\lambda(x, z) := \|x - z\|^2 + \lambda \|x\|_1 \\ & \text{s.t. } z = [x - F(x)]_+, \end{aligned} \tag{5}$$

where $\lambda > 0$ is a regularization parameter and $\|\cdot\|$ is denoted as the Euclidean norm. We call (5) the ℓ_1 regularized projection minimization problem.

This paper is organized as follows. In Section 2, we approximate (1) by the ℓ_1 regularization projection minimization problem (5), and we show theoretically that (5) is a

good approximation. In Section 3, we propose an extragradient thresholding algorithm (ETA) for (5) and also analyze the convergence of this algorithm. Numerical results are demonstrated in Section 4 to show that (5) is promising in providing a sparse solution of co-coercive NCPs.

2 The ℓ_1 regularized approximation

In this section, we study the relation between the solutions of model (5) and those of model (2).

Theorem 2.1 *For any fixed $\lambda > 0$, the solution set of (5) is nonempty and bounded. Let $\{(\widehat{x}_{\lambda_k}, \widehat{z}_{\lambda_k})\}$ be a solution of (5), and $\{\lambda_k\}$ be any positive sequence converging to 0. If $\text{SOL}(F) \neq \emptyset$, then $\{(\widehat{x}_{\lambda_k}, \widehat{z}_{\lambda_k})\}$ has at least one accumulation point, and any accumulation point x^* of $\{\widehat{x}_{\lambda_k}\}$ is a solution of (2).*

Proof For any fixed $\lambda > 0$, it is easy to show the coercivity of $f_\lambda(x, z)$ in (5), namely

$$f_\lambda(x, z) \rightarrow +\infty \quad \text{as } \|(x, z)\| \rightarrow \infty. \tag{6}$$

We also note that for any $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$, $f_\lambda(x, z) \geq 0$. This together with (6) implies the level set

$$L = \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^n \mid f_\lambda(x, z) \leq f_\lambda(x_0, z_0) \text{ and } z = [x - F(x)]_+\}$$

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 (5) is nonempty and bounded since $f_\lambda(x, z)$ is continuous on L .

Now we show the second part of this theorem. Let $\widehat{x} \in \text{SOL}(F)$ and $\widehat{z} = [\widehat{x} - F(\widehat{x})]_+$. From (5), we have $\widehat{x} = \widehat{z}$. Since $(\widehat{x}_{\lambda_k}, \widehat{z}_{\lambda_k})$ is a solution of (5) with $\lambda = \lambda_k$, where $\widehat{z}_{\lambda_k} = [\widehat{x}_{\lambda_k} - F(\widehat{x}_{\lambda_k})]_+$, it follows that

$$\begin{aligned} \max\{\|\widehat{x}_{\lambda_k} - \widehat{z}_{\lambda_k}\|^2, \lambda_k \|\widehat{x}_{\lambda_k}\|_1\} &\leq \|\widehat{x}_{\lambda_k} - \widehat{z}_{\lambda_k}\|^2 + \lambda_k \|\widehat{x}_{\lambda_k}\|_1 \\ &\leq \|\widehat{x} - \widehat{z}\|^2 + \lambda_k \|\widehat{x}\|_1 \\ &= \lambda_k \|\widehat{x}\|_1. \end{aligned} \tag{7}$$

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

$$\|\widehat{x}_{\lambda_k}\|_1 \leq \|\widehat{x}\|_1. \tag{8}$$

Hence the sequence $\{\widehat{x}_{\lambda_k}\}$ is bounded and has at least one cluster point. Note that the sequence $\{\widehat{z}_{\lambda_k}\}$ is also bounded because $\|\widehat{x}_{\lambda_k} - \widehat{z}_{\lambda_k}\|^2 \leq \lambda_k \|\widehat{x}\|_1 \leq \lambda_0 \|\widehat{x}\|_1$ ($\lambda_k \rightarrow 0$).

Let x^* and z^* be any cluster points of $\{\widehat{x}_{\lambda_k}\}$ and $\{\widehat{z}_{\lambda_k}\}$, respectively. Then there exists a subsequence of $\{\lambda_k\}$, say $\{\lambda_{k_j}\}$, such that

$$\lim_{k_j \rightarrow \infty} \widehat{x}_{\lambda_{k_j}} = x^* \quad \text{and} \quad \lim_{k_j \rightarrow \infty} \widehat{z}_{\lambda_{k_j}} = z^*.$$

We can obtain $z^* = [x^* - F(x^*)]_+$ by letting k_j tend to ∞ in $z_{\lambda_{k_j}} = [x_{\lambda_{k_j}} - F(x_{\lambda_{k_j}})]_+$. Letting λ_{k_j} tend to 0 in

$$\|\widehat{x}_{\lambda_{k_j}} - \widehat{z}_{\lambda_{k_j}}\|^2 \leq \lambda_{k_j} \|\widehat{x}\|_1$$

yields $x^* = z^*$. Consequently, $x^* = [x^* - F(x^*)]_+$ follows, which manifests $x^* \in \text{SOL}(F)$. From (8), namely $\|\widehat{x}_{\lambda_{k_j}}\|_1 \leq \|\widehat{x}\|_1$, k_j tending to ∞ , we get $\|x^*\|_1 \leq \|\widehat{x}\|_1$. Then by the arbitrariness of $\widehat{x} \in \text{SOL}(F)$, we know x^* is a solution of problem (2). This completes the proof. \square

3 Algorithm and convergence

In this section, we suggest the extragradient thresholding algorithm (ETA) to solve ℓ_1 regularization projection minimization problem (5) and give the convergence analysis of ETA.

First we state some basic operator concepts as regards monotonicity and some properties of the projection operator. Let $P_K(\cdot)$ denote the projection operator from \mathbb{R}^n onto K , a nonempty closed convex subset of \mathbb{R}^n . From the definition of projection operator, it follows that

$$\langle y - P_K(x), P_K(x) - x \rangle \geq 0, \quad \forall y \in K, x \in \mathbb{R}^n. \tag{9}$$

Consequently, we have

$$\langle P_K(x) - P_K(y), x - y \rangle \geq \|P_K(x) - P_K(y)\|^2, \quad \forall x, y \in \mathbb{R}^n, \tag{10}$$

$$\|P_K(x) - P_K(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n, \tag{11}$$

$$\|P_K(x) - y\|^2 \leq \|x - y\|^2 - \|P_K(x) - x\|^2, \quad \forall y \in K, x \in \mathbb{R}^n. \tag{12}$$

Lemma 3.1 [18] *Define a residue function*

$$e(x, \alpha) = x - P_K[x - \alpha F(x)], \quad \alpha \geq 0.$$

The following statements are valid.

- (a) $\forall \alpha > 0, F(x)^\top e(x, \alpha) \geq \frac{\|e(x, \alpha)\|^2}{\alpha}$;
- (b) for any $\alpha > 0, \frac{\|e(x, \alpha)\|}{\alpha}$ is non-increasing;
- (c) for any $\alpha \geq 0, \|e(x, \alpha)\|$ is non-decreasing.

In this paper, we suppose the mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is co-coercive on a subset K of \mathbb{R}^n . That is, there exists a constant $c > 0$ such that

$$\langle F(x) - F(y), x - y \rangle \geq c \|F(x) - F(y)\|^2, \quad \forall x, y \in K.$$

It is clear that the co-coercive mapping is monotone, namely,

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in K,$$

but not necessarily strongly monotone, i.e., there is a constant $c > 0$ such that

$$\langle F(x) - F(y), x - y \rangle \geq c \|x - y\|^2, \quad \forall x, y \in K.$$

Remark 3.1 Every affine monotone function which is also symmetric must be co-coercive (on \mathbb{R}^n). The Euclidean projector P_K and $I - P_K$ are both ‘co-coercive’ functions [2, 19].

Lemma 3.2 Suppose that $F(\cdot)$ is co-coercive on K with modulus $c > 0$. Then for any given positive real number α , when $c > \alpha/2$, the operator $I - \alpha F$ is nonexpansive, that is, for any $x, y \in K$,

$$\|(I - \alpha F)(x) - (I - \alpha F)(y)\| \leq \|x - y\|.$$

Proof For any $x, y \in K$, when $c > \alpha/2$, using the co-coercivity of F , it follows that

$$\begin{aligned} & \|(I - \alpha F)(x) - (I - \alpha F)(y)\|^2 \\ &= \|(x - y) - \alpha[F(x) - F(y)]\|^2 \\ &= \|x - y\|^2 - 2\alpha\langle x - y, F(x) - F(y) \rangle + \alpha^2 \|F(x) - F(y)\|^2 \\ &\leq \|x - y\|^2 - \alpha(2c - \alpha) \|F(x) - F(y)\|^2 \\ &\leq \|x - y\|^2, \end{aligned}$$

which shows $I - \alpha F$ is nonexpansive. □

For giving $z^k \in \mathbb{R}_+^n$ and $\lambda_k > 0$, we consider an unconstrained minimization subproblem:

$$\min_{x \in \mathbb{R}^n} f_{\lambda_k}(x, z^k) := \|x - z^k\|^2 + \lambda_k \|x\|_1. \tag{13}$$

Evidently, the minimizer x^s of the model (13) must satisfy the corresponding optimality condition

$$x^s = S_{\lambda_k}(z^k), \tag{14}$$

where the shrinkage operator S_λ is defined by

$$(S_\lambda(z))_i = \begin{cases} z_i - \frac{\lambda}{2}, & z_i \geq \frac{\lambda}{2}, \\ 0, & 0 \leq z_i < \frac{\lambda}{2}. \end{cases} \tag{15}$$

Evidently, the shrinkage operator S_λ is component-wise, i.e., $(S_\lambda(z))_i = S_\lambda(z_i)$. Moreover, it is nonexpansive; i.e., $\|S_\lambda(x) - S_\lambda(y)\| \leq \|x - y\|$, for any $x, y \in \mathbb{R}_+^n$, see [20]. It demonstrates that a solution $x \in \mathbb{R}^n$ of the subproblem (13) can be analytically expressed by (14).

By the solution representation, we construct the following extragradient thresholding algorithm (ETA) to solve the ℓ_1 regularized projection minimization problem (5).

Input: c -the co-coercive modulus of F .

Step 0: Choose $0 \neq z^0 \in \mathbb{R}_+^n$, $\lambda_0, \beta > 0$, $\tau, \gamma, \mu \in (0, 1)$, $\beta\gamma < 2c$, $\epsilon > 0$ and integers $n_{\max} > K_0 > 0$. Set $k = 0$.

Step 1: Compute

$$\begin{aligned} x^k &= S_{\lambda_k}(z^k), \\ y^k &= [x^k - \alpha_k F(x^k)]_+, \end{aligned}$$

where $\alpha_k = \beta \gamma^{m_k}$ with m_k being the smallest nonnegative integer satisfying

$$\|F(x^k) - F(y^k)\| \leq \mu \frac{\|x^k - y^k\|}{\alpha_k}. \tag{16}$$

Step 2: If $\|x^k - z^k\| \leq \epsilon$ or the number of iterations is greater than n_{\max} , then return z^k, x^k, y^k and stop. Otherwise, compute

$$z^{k+1} = [x^k - \alpha_k F(y^k)]_+$$

and update λ_{k+1} by

$$\lambda_{k+1} = \begin{cases} \tau \lambda_k, & \text{if } k + 1 \text{ is a multiple of } K_0, \\ \lambda_k, & \text{otherwise,} \end{cases}$$

and $k = k + 1$, then go to Step 1.

Before analyzing the convergence of ETA, we first present a key lemma as regards co-coercive mapping.

Lemma 3.3 *Suppose that mapping F is co-coercive and $\text{SOL}(F) \neq \emptyset$. If x^k generated by ETA is not a solution of $\text{NCP}(F)$, then for any $\hat{x} \in \text{SOL}(F)$, we have*

$$\langle F(y^k), x^k - \hat{x} \rangle \geq \langle F(y^k), x^k - y^k \rangle \geq (1 - \mu) \frac{\|x^k - y^k\|^2}{\beta}. \tag{17}$$

Proof For any $\hat{x} \in \text{SOL}(F)$, we have $F(\hat{x})^\top \hat{x} = 0$. Since $y^k \in \mathbb{R}_+^n$, it follows that $\langle F(\hat{x}), y^k - \hat{x} \rangle \geq 0$. It is clear that the co-coercive mapping is pseudo-monotone, that is,

$$\langle x - y, F(y) \rangle \geq 0 \implies \langle x - y, F(x) \rangle \geq 0, \quad \forall x, y \in K \text{ and } x \neq y.$$

By the definition of pseudo-monotonicity, it follows that $\langle F(y^k), y^k - \hat{x} \rangle \geq 0$. Hence,

$$\begin{aligned} \langle F(y^k), x^k - \hat{x} \rangle &= \langle F(y^k), x^k - y^k + y^k - \hat{x} \rangle \\ &\geq \langle F(y^k), x^k - y^k \rangle \\ &= \langle F(x^k), x^k - y^k \rangle - \langle F(x^k) - F(y^k), x^k - y^k \rangle \\ &\geq \frac{1}{\alpha_k} \|x^k - y^k\|^2 - \frac{\mu}{\alpha_k} \|x^k - y^k\|^2 \\ &\geq \frac{1 - \mu}{\beta} \|x^k - y^k\|^2, \end{aligned}$$

where the last inequality but one follows from Lemma 3.1 and (16). □

We now begin to analyze the convergence of the proposed ETA.

Theorem 3.1 *Suppose that the mapping F is co-coercive with modulus $c > \beta \gamma / 2$ and $\text{SOL}(F) \neq \emptyset$. Let $\{z^k, x^k, y^k\}$ and $\{\lambda_k\}$ be sequences generated by ETA, then*

- (i) *the sequences $\{z^k\}$, $\{x^k\}$, and $\{y^k\}$ are all bounded;*
- (ii) *any cluster point of the sequence $\{x^k\}$ is a solution of $\text{NCP}(F)$.*

Proof (i) Let $\widehat{x} \in \text{SOL}(F)$. By the definition (15) of operator S_λ , we have

$$\|x^k - \widehat{x}\| = \|S_\lambda(z^k) - \widehat{x}\| \leq \|z^k - \widehat{x}\| + \sqrt{n}\lambda_k/2 \leq \|z^k - \widehat{x}\| + \sqrt{n}\lambda_0/2. \tag{18}$$

In view of $\widehat{x} \in \text{SOL}(F)$, we have $\widehat{x} = [\widehat{x} - \alpha_k F(\widehat{x})]_+$. Since $c > \beta\gamma/2 > \alpha_k/2$, by Lemma 3.2, we see that $I - \alpha_k F$ is nonexpansive. Together with the nonexpansive property of the projection operator, it follows that

$$\begin{aligned} \|y^k - \widehat{x}\| &= \|[x^k - \alpha_k F(x^k)]_+ - \widehat{x}\| \\ &= \|[x^k - \alpha_k F(x^k)]_+ - [\widehat{x} - \alpha_k F(\widehat{x})]_+\| \\ &\leq \|(I - \alpha_k F)(x^k - \widehat{x})\| \\ &\leq \|x^k - \widehat{x}\| \\ &\leq \|z^k - \widehat{x}\| + \sqrt{n}\lambda_k/2 \\ &\leq \|z^k - \widehat{x}\| + \sqrt{n}\lambda_0/2. \end{aligned} \tag{19}$$

From (12) and (17), we obtain

$$\begin{aligned} \|z^{k+1} - \widehat{x}\|^2 &= \|[x^k - \alpha_k F(y^k)]_+ - \widehat{x}\|^2 \\ &\leq \|x^k - \alpha_k F(y^k) - \widehat{x}\|^2 - \|z^{k+1} - x^k + \alpha_k F(y^k)\|^2 \\ &= \|x^k - \widehat{x}\|^2 - 2\alpha_k \langle F(y^k), z^{k+1} - \widehat{x} \rangle - \|z^{k+1} - x^k\|^2 \\ &\leq \|x^k - \widehat{x}\|^2 - 2\alpha_k \langle F(y^k), z^{k+1} - y^k \rangle - \|z^{k+1} - x^k\|^2 \\ &= \|x^k - \widehat{x}\|^2 - \|z^{k+1} - y^k\|^2 - \|y^k - x^k\|^2 \\ &\quad + 2\langle x^k - y^k - \alpha_k F(y^k), z^{k+1} - y^k \rangle. \end{aligned} \tag{20}$$

By $y^k = [x^k - \alpha_k F(x^k)]_+$ and (9), it follows that

$$\begin{aligned} &2\langle x^k - y^k - \alpha_k F(y^k), z^{k+1} - y^k \rangle \\ &\leq 2\langle x^k - y^k - \alpha_k F(y^k), z^{k+1} - y^k \rangle + 2\langle y^k - x^k + \alpha_k F(x^k), z^{k+1} - y^k \rangle \\ &= 2\alpha_k \langle F(x^k) - F(y^k), z^{k+1} - y^k \rangle \\ &\leq \alpha_k^2 \|F(x^k) - F(y^k)\|^2 + \|z^{k+1} - y^k\|^2. \end{aligned} \tag{21}$$

Replacing (21) into (20), by (16) and (18), we deduce

$$\begin{aligned} \|z^{k+1} - \widehat{x}\|^2 &\leq \|x^k - \widehat{x}\|^2 - \|y^k - x^k\|^2 + \alpha_k^2 \|F(x^k) - F(y^k)\|^2 \\ &\leq \|x^k - \widehat{x}\|^2 - \|y^k - x^k\|^2 + \mu^2 \|x^k - y^k\|^2 \\ &= \|x^k - \widehat{x}\|^2 - (1 - \mu^2) \|y^k - x^k\|^2 \\ &\leq \|x^k - \widehat{x}\|^2. \end{aligned} \tag{22}$$

Hence, by definition of λ_k , it follows that

$$\begin{aligned} \|z^{k+1} - \widehat{x}\| &\leq \|x^k - \widehat{x}\| \leq \|z^k - \widehat{x}\| + \frac{\sqrt{n}}{2}\lambda_k \leq \|x^{k-1} - \widehat{x}\| + \frac{\sqrt{n}}{2}\lambda_k \\ &\leq \|z^{k-1} - \widehat{x}\| + \frac{\sqrt{n}}{2}(\lambda_k + \lambda_{k-1}) \leq \dots \\ &\leq \|z^0 - \widehat{x}\| + \frac{\sqrt{n}}{2} \sum_{i=0}^k \lambda_i \\ &\leq \|z^0 - \widehat{x}\| + \frac{\sqrt{n} \lambda_0 K_0}{2(1-\tau)} := C, \end{aligned} \tag{23}$$

which shows $\{z^k\}$ is bounded. Together with (18) and (19), we see that $\{x^k\}$ and $\{y^k\}$ are both bounded.

(ii) Now we prove $\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0$. By (22) and (18), it follows that

$$\begin{aligned} (1 - \mu^2) \|y^k - x^k\|^2 &\leq \|x^k - \widehat{x}\|^2 - \|z^{k+1} - \widehat{x}\|^2 \\ &\leq \|x^k - \widehat{x}\|^2 - (\|x^{k+1} - \widehat{x}\| - \sqrt{n}\lambda_{k+1}/2)^2 \\ &= \|x^k - \widehat{x}\|^2 - \|x^{k+1} - \widehat{x}\|^2 + \sqrt{n}\lambda_{k+1} \|x^{k+1} - \widehat{x}\| - n\lambda_{k+1}^2/4 \\ &\leq \|x^k - \widehat{x}\|^2 - \|x^{k+1} - \widehat{x}\|^2 + \sqrt{n}\lambda_{k+1} \|x^{k+1} - \widehat{x}\|, \end{aligned}$$

which leads to the following inequality:

$$\begin{aligned} (1 - \mu^2) \sum_{k=0}^{\infty} \|y^k - x^k\|^2 &\leq \sum_{k=0}^{\infty} (\|x^k - \widehat{x}\|^2 - \|x^{k+1} - \widehat{x}\|^2 + \sqrt{n}\lambda_{k+1} \|x^{k+1} - \widehat{x}\|) \\ &\leq \|x^0 - \widehat{x}\|^2 + \sqrt{n} \sum_{k=0}^{\infty} \lambda_{k+1} \|x^{k+1} - \widehat{x}\| \\ &\leq \|x^0 - \widehat{x}\|^2 + \sqrt{n}C \sum_{k=0}^{\infty} \lambda_{k+1} \\ &= \|x^0 - \widehat{x}\|^2 + \sqrt{n}C \frac{\lambda_0 K_0}{1-\tau} < +\infty, \end{aligned}$$

where the third inequality holds from (23), and thus we have $\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0$.

Since $\{x^k\}$ is bounded, $\{x^k\}$ has at least one cluster point. Let x^* be a cluster point of $\{x^k\}$ and a subsequence $\{x^{k_i}\}$ converge to x^* . Next we will show $x^* \in \text{SOL}(F)$.

If there is a positive low bounded α_{\min} such that $\alpha_{k_i} \geq \alpha_{\min} > 0$, from Lemma 3.1(b) and (c), we get

$$\min\{1, \alpha\} \|e(x, 1)\| \leq \|e(x, \alpha)\| \leq \max\{1, \alpha\} \|e(x, 1)\|, \tag{24}$$

where $e(x, \alpha) = x - [x - \alpha F(x)]_+$. Together with the continuity of $e(x, \alpha)$ for x and $\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0$, we have

$$\begin{aligned} \|e(x^*, 1)\| &= \lim_{k_i \rightarrow \infty} \|e(x^{k_i}, 1)\| \leq \lim_{k_i \rightarrow \infty} \frac{\|e(x^{k_i}, \alpha_{k_i})\|}{\min\{1, \alpha_{k_i}\}} \\ &\leq \lim_{k_i \rightarrow \infty} \frac{\|e(x^{k_i}, \alpha_{k_i})\|}{\min\{1, \alpha_{\min}\}} = \lim_{k_i \rightarrow \infty} \frac{\|x^{k_i} - y^{k_i}\|}{\min\{1, \alpha_{\min}\}} = 0. \end{aligned} \tag{25}$$

If $\lim_{k_i \rightarrow \infty} \alpha_{k_i} = 0$, for enough large k_i , by Lemma 3.1(b) and (16), we get

$$\|e(x^{k_i}, 1)\| \leq \frac{\|e(x^{k_i}, \frac{1}{\beta} \alpha_{k_i})\|}{\frac{1}{\beta} \alpha_{k_i}} < \frac{1}{\mu} \|F(x^{k_i}) - F(\bar{y}^{k_i})\|, \tag{26}$$

where $\bar{y}^{k_i} = [x^{k_i} - \frac{1}{\beta} \alpha_{k_i} F(x^{k_i})]_+$. Taking the limit in the above inequality, we have

$$\|e(x^*, 1)\| = \lim_{k_i \rightarrow \infty} \|e(x^{k_i}, 1)\| \leq \lim_{k_i \rightarrow \infty} \frac{1}{\mu} \|F(x^{k_i}) - F(\bar{y}^{k_i})\| = 0. \tag{27}$$

It means $x^* = [x^* - F(x^*)]_+$. Hence we get $x^* \in \text{SOL}(F)$. The proof is thus complete. \square

4 Numerical experiments

In this section, we present some numerical experiments to demonstrate the effectiveness of our ETA algorithm. All the numerical experiments were performed on a laptop (2.50GHz, 6.00GB of RAM) by utilizing MATLAB R2013a.

We will stimulate three examples to implement the ETA algorithm. They will be ran 100 times for difference dimensions, and thus average results will be recorded. In each experiment, we set $z^0 = e$, $\beta = 2c$, $\gamma = 0.1$, $\mu = 1/c$, $n_{\max} = 2,000$.

4.1 Test for LCPs with Z-matrix [5]

The test is associated with the Z-matrix which has an important property, that is, there is a unique sparse solution of LCPs when M is a kind of Z-matrix [5]. Let us consider $\text{LCP}(q, M)$ where

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

Here I_n is the identity matrix of order n and $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$. 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 $\text{LCP}(q, M)$, since it satisfies

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

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

We choose $z^0 = e$, $c = 1$, $\lambda_0 = 0.2$, $\beta = 2c$, $\tau = 0.75$, $\gamma = 0.1$, $\mu = 1/c$, $\epsilon = 1e - 6$, $n_{\max} = 2,000$, $K_0 = 5$. We will take advantage of the recovery error $\|x - \hat{x}\|$ to evaluate our algorithm. Apart from that, the average cpu time (in seconds), the average number of iteration times and the residual $\|x - z\|$ will also be taken into consideration on judging the performance of the method.

As indicated in Table 1, the ETA algorithm behaves very robust because the average number of times of iteration is identically equal to 205; the recovered error $\|x - \hat{x}\|$ and residual $\|x - z\|$ are basically similar. In addition, the sparsity $\|x\|_0$ of the recovered solution x is in all cases 1s, which means the recover is successful. Most importantly, the

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

n	Iter	$\ x - \hat{x}\ $	$\ x - z\ $	$\ \hat{x}\ _0$	$\ x\ _0$	Time (sec.)
3,000	205	7.7007E-06	7.5424E-07	1	1	2.90
5,000	205	7.6995E-06	7.5424E-07	1	1	7.93
10,000	205	7.6986E-06	7.5424E-07	1	1	36.70
15,000	205	7.6983E-06	7.5424E-07	1	1	78.46
20,000	205	7.6981E-06	7.5424E-07	1	1	148.50
25,000	205	7.6980E-06	7.5424E-07	1	1	232.87

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

n	Iter	$\ x - \hat{x}\ $	$\ \hat{x}\ _0$	$\ x\ _0$	Time (sec.)
100	1,012	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	--	--	--	--	--

ETA algorithm is exceptionally fast, which results in only 36.70 seconds being needed to address the problem with dimension $n = 10,000$.

In order to illustrate the effectiveness of the ETA algorithm we proposed, we introduce another method of tackling the LCPs. In [5], the authors established an l_p ($0 < p < 1$) regularized minimization model:

$$\min_{x \in \mathbb{R}^n} f(x) := \frac{1}{2} \|\Phi_{FB}(x)\|^2 + \lambda \|x\|_p^p \tag{28}$$

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 the recovered solution is in all cases as large as 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 not be appealing when n is relatively large. Contrasted with the SSG method, the ETA algorithm is more outstanding in the cpu time and the size of the solvable problems.

4.2 Test for LCPs with positive semidefinite matrices

In this subsection, we test ETA 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 $M = ZZ^T$. 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,

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

Table 3 Results on randomly created LCPs with positive semidefinite matrices.

n	Iter	$\ x - z\ $	$\ \hat{x}\ _0$	$\ x\ _0$	Time (sec.)
2,000	350	8.0316E-11	20	20	18.36
3,000	210	9.8366E-11	30	30	12.02
4,000	142	8.5188E-11	40	40	29.97
5,000	142	9.5243E-11	50	50	22.29
7,000	144	8.4519E-11	70	70	46.67

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 ETA algorithm and take $z^0 = e$, $c = \max(\text{svd}(M))$, $\lambda_0 = 0.02$, $\beta = 2c$, $\tau = 0.75$, $\gamma = 0.1$, $\mu = 1/c$, $\epsilon = 1e - 10$, $n_{\max} = 2,000$, $K_0 = \max(2, \text{floor}(10,000/n))$. Then ETA will output a solution x . Similarly, the residual $\|x - \hat{x}\|$, average cpu time (in seconds), the average number of iteration times, and the residual $\|x - z\|$ will also be taken into consideration on valuating our ETA algorithm.

As manifested in Table 3, the ETA algorithm performs quite efficiently. Furthermore, the sparsity $\|x\|_0$ of recovered solution x is in all cases equal to the sparsity $\|\hat{x}\|_0$, which means the recover is exact. Likewise, the ETA algorithm is exceptionally fast in this example, which makes that only 46.67 seconds are needed to pursue the sparse solution of LCP when the dimension $n = 7,000$.

4.3 Test for co-coercive nonlinear complementarity problem

We now consider a co-coercive nonlinear complementarity problems (NCP) with

$$F(x) = D(x) + Mx + q, \tag{29}$$

where $D(x)$ and $Mx + q$ are the nonlinear part and the linear part of $F(x)$, respectively. We form $F(x)$ similarly as in [21, 22]. The matrix $M = A^T A + B$, where A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, 5)$, and a skew-symmetric matrix B is generated in the same way. In $D(x)$, the nonlinear part of $F(x)$, the components are

$$D_j(x) = a_j * \arctan(x_j)$$

and a_j is a random variable in $(-1, 0)$. Then the sequent part of generating the sparse vector \hat{x} and vector $q \in \mathbb{R}^n$ such that

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

is similar to the procedure of Section 4.2. Let M and q be the input to our ETA algorithm and take $z^0 = e$, $c = 150 * \log(n)$, $\lambda_0 = 0.2$, $\beta = 2c$, $\tau = 0.75$, $\gamma = 0.1$, $\mu = 1/c$, $\epsilon = 1e - 6$, $n_{\max} = 2,000$, $K_0 = \max(2, \text{floor}(10,000/n))$, and $a = -\text{rand}(n, 1)$. Then ETA will output a solution x . Similarly, the average number of iteration times, the average residual $\|x - z\|$, the average sparsity $\|x\|_0$ of x , and the average cpu time (in seconds) will also be taken into consideration on valuating our ETA algorithm.

It is not difficult to see from Table 4 that the ETA algorithm also performs quite efficiently in such nonlinear complementarity problems. The sparsity $\|x\|_0$ of the recovered solution x are all equal to the sparsity $\|\hat{x}\|_0$, that is, the recover is exact. What is also striking is that the ETA algorithm is exceptionally fast in this example as well, with only 144.99 seconds being needed to tackle the sparse NCP when the dimension $n = 10,000$.

Table 4 Results on co-coercive nonlinear complementarity problems.

n	Iter	$\ x - z\ $	$\ \hat{x}\ _0$	$\ x\ _0$	Time
1,000	450	7.5467E-07	10	10	3.28
3,000	138	9.8034E-07	30	30	7.86
5,000	94	9.4921E-07	50	50	14.85
7,000	96	8.4234E-07	70	70	30.24
10,000	98	7.5510E-07	100	100	143.99

5 Conclusions

In this paper, we concentrate on finding sparse solutions of co-coercive nonlinear complementarity problems (NCPs). An ℓ_1 regularized projection minimization model is proposed for relaxation, and an extragradient thresholding algorithm (ETA) is then designed for this regularized model. Furthermore, we analyze the convergence of this algorithm and show any cluster point of the sequence generated by ETA is a solution of NCP. Preliminary numerical results indicate that the ℓ_1 regularized model as well as the ETA are promising to find sparse solutions of NCPs.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors completed this paper together. All authors read and approved the final manuscript.

Author details

¹Department of Applied Mathematics, Beijing Jiaotong University, Beijing, 100044, P.R. China. ²Department of Mathematics, Shijiazhuang University, Shijiazhuang, 050035, P.R. China. ³School of Mathematics, University of Southampton, Southampton, SO17 1BJ, UK.

Acknowledgements

We would like to thank the two referees for their valuable comments. 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).

Received: 28 September 2014 Accepted: 6 January 2015 Published online: 28 January 2015

References

- Cottle, RW, Pang, JS, Stone, RE: *The Linear Complementarity Problem*. Academic Press, Boston (1992)
- Facchinei, F, Pang, JS: *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research, vols. I and II. Springer, New York (2003)
- Ferris, MC, Mangasarian, OL, Pang, JS: *Complementarity: Applications, Algorithms and Extensions*. Kluwer Academic, Dordrecht (2001)
- Xie, J, He, S, Zhang, S: Randomized portfolio selection with constraints. *Pac. J. Optim.* **4**, 87-112 (2008)
- Shang, M, Zhang, C, Xiu, N: Minimal zero norm solutions of linear complementarity problems. *J. Optim. Theory Appl.* **163**, 795-814 (2014)
- Candès, EJ, Randall, PA: Highly robust error correction by convex programming. *IEEE Trans. Inf. Theory* **54**, 2829-2840 (2006)
- Candès, EJ, Recht, B: Exact matrix completion via convex optimization. *Found. Comput. Math.* **9**, 717-772 (2008)
- Candès, EJ, Romberg, J, Tao, T: Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207-1223 (2006)
- Donoho, DL: Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289-1306 (2006)
- Fukushima, M, Pang, JS: Some feasibility issues in mathematical programs with equilibrium constraints. *SIAM J. Optim.* **8**, 673-681 (1998)
- 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)
- Lin, G, Fukushima, M: New reformulations for stochastic nonlinear complementarity problems. *Optim. Methods Softw.* **21**, 551-564 (2006)
- Luo, ZQ, Pang, JS, Ralph, D: *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge (1996)
- Figueiredo, MAT, Nowak, RD: An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.* **12**, 906-916 (2003)
- Starck, JL, Donoho, DL, Candès, EJ: Astronomical image representation by the curvelet transform. *Astron. Astrophys.* **398**, 785-800 (2003)

16. Daubechies, I, Defrise, M, De Mol, C: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57**, 1413-1457 (2004)
17. Figueiredo, MAT, Nowak, RD, Wright, SJ: Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* **1**, 586-597 (2007)
18. Zhu, T, Yu, ZG: A simple proof for some important properties of the projection mappings. *Math. Inequal. Appl.* **7**, 453-456 (2004)
19. Zhu, DL, Marcotte, P: Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. *SIAM J. Optim.* **6**, 714-726 (1996)
20. Hale, ET, Yin, W, Zhang, Y: Fixed-point continuation for ℓ_1 minimization: methodology and convergence. *SIAM J. Optim.* **19**, 1107-1130 (2008)
21. He, BS, Liao, LZ: Improvements of some projection methods for monotone nonlinear variational inequalities. *J. Optim. Theory Appl.* **112**(1), 111-128 (2002)
22. Yan, X, Han, D, Sun, W: A modified projection method with a new direction for solving variational inequalities. *Appl. Math. Comput.* **211**, 118-129 (2009)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
