Scientific
Research
Publishing

Spectral Gradient Algorithm Based on the Generalized Fiser-Burmeister Function for Sparse Solutions of LCPS

Chang Gao, Zhensheng Yu, Feiran Wang

College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: gaochang45622@163.com

Abstract

This paper considers the computation of sparse solutions of the linear complementarity problems LCP(q, M). Mathematically, the underlying model is NP-hard in general. Thus an $l_p(0 < p < 1)$ regularized minimization model is proposed for relaxation. We establish the equivalent unconstrained minimization reformation of the NCP-function. Based on the generalized Fiser-Burmeister function, a sequential smoothing spectral gradient method is proposed to solve the equivalent problem. Numerical results are given to show the efficiency of the proposed method.

1. Introduction

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

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

The set of solutions to this problem is denoted by $SOL(M, q)$. Throughout the paper, we always suppose $SOL(M, q) \neq \phi$. The LCP has many wide applications in physics, engineering, mechanics, and economics design [1] [2]. Numerical methods for solving LCPs, such as the Newton method, the interior point method and the non-smooth equation methods, have been extensively investigated in the literature. However, it seems that there are few methods to solve the sparse solutions for LCPs. In fact, it is very necessary to research the sparse solu-

tion of the LCPs such as portfolio selection [3] [4] and bimatrix games [5] in real applications.

In this paper, we consider the sparse solutions of the LCP. We call $x \in SOL(q,M)$ a sparse solution of LCP (q, M) if x is a solution of the following optimization problem

$$\min \quad \|x\|_0$$
$$s.t \quad x \geq 0, \ Mx + q \geq 0, \ x^{\mathrm{T}}(Mx + q) = 0 \tag{1}$$

To be more precise, we seek a vector $x \in R^n$ by solving the l_0 norm minimization problem, where $\|x\|_0$ stands for the number of nonzero components of *x*. A solution of (1) is called the sparsest solution of the LCP.

Recently, Meijuan Shang, Chao Zhang and Naihua Xiu design a sequential smoothing gradient method to solve the sparse solution of LCP [6]. We inspire by the model and use the spectral method based on the generalized Fischer-Burmeister function to solve our new model (3). The spectral method is proposed by Barzilai and Borwein [7] and further analyzed by Raydan [8] [9]. The advantage of this method is that it requires little computational work and greatly speeds up the convergence of gradient methods. Therefore, this technique has received successful applications in unconstrained and constrained optimizations [10]-[13].

In fact, the above minimization problem (1) is a sparse optimization with equilibrium constraints. From the problem of constraint conditions, as well as the non-smooth objective function, it is difficult to get solutions due to the equilibrium constraints to overcome the difficultly, and we use the NCP-functions to construct the penalty of violating the equilibrium constraints.

A function $\varphi: R^2 \to R^1$ is called a NCP-function, if for any pair $(a,b)^{\mathrm{T}} \in R^2 \ \varphi(a,b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0$.

A popular NCP-functions is the Fischer-Burmeister (FB), which is defined as

$$\varphi_{FB}(a,b) = (a+b) - \sqrt{a^2 + b^2}$$

The Fischer-Burmeister function has many interesting properties. However, it has limitations in dealing with monotone complementarity problems since it is too flat in the positive orthant, the region of main interest for a complementarity problem. In terms of the above disadvantage of the Fischer-Burmeister function, we consider the following generalized Fischer-Burmeister function [10].

$$\varphi_P(a,b) = \|(a,b)\|_P - (a+b) \tag{2}$$

where *p* is any fixed real number from $P \in (1, +\infty)$ and $\|(a,b)\|_P$ denotes the *p*-norm, *i.e.*

$$\|(a,b)\|_P = \sqrt[p]{|a|^P + |b|^P}$$

In other words, in the function φ_P, we replace the 2-norm of (a,b) in the FB function φ_{FB} by a more general *p*-norm of (a,b). The function φ_P is still an NCP-function.

Define $\Phi_P(x): R^n \to R^n$ by

$$\Phi_P(x) = \begin{pmatrix} \varphi_P(x_1, F_1(x)) \\ \vdots \\ \varphi_P(x_n, F_n(x)) \end{pmatrix}$$

where $F(x) = Mx + q$ with $q \in R^n$ and $M \in R^{n \times n}$. Obviously, $x \in SOL(M,q)$ if and only if $\Phi_P(x) = 0$. By further employing the l_p regularization term for seeking sparsity, we obtain the following unconstrained minimization problem to approximate

$$\min_{x \in R^n} f(x) = \frac{1}{2} \|\Phi_P(x)\|^2 + \lambda \|x\|_p^p \tag{3}$$

where $\lambda \in (0, \infty)$ is a given regularization parameter, and $\|x\|_p^p = \sum_{i=1}^n |x_i|^p$ for any $0 < p < 1$. We call (3) as l_p regularized minimization problem.

Let us denote the first term of (3) by the function $\Psi_\lambda(x): R^n \to R_+$. That is

$$\Psi_P(x) = \frac{1}{2}\|\Phi_P(x)\|^2 \tag{4}$$

For any given $P > 1$, the function $\Psi_P(x)$ is shown to possess all favorable properties of the FB function; we can see [8]. It plays an important part in our study throughout the paper. We observe in [14] [15] that P has a great influence on the numerical performance of certain descent-type methods; a larger P yields a better convergence rate, whereas a small P often gives a better global convergence.

The paper is organized as follows: In Section 2, we present absolute lower bounds for nonzero entries in local solution of (3). In section 3, we approximate the minimal zero norm solutions of the LCP. In section 4, we give a sequential smoothing spectral gradient method to solve the model. In Section 5, numerical results are given to demonstrate the effectiveness of the sequential smoothing spectral gradient method.

2. The l_p Regularized Approximation

In this section, we consider the minimizers of (3). We study the relation between the original model (1) and the l_p regularized model (3), which indicates the regularized model is a good approximation. We use a threshold lower bound L [6] for nonzero entries in local minimizers and the choice of the l_p minimization problem (3).

2.1. Relation between (3) and (1)

The following result is given in [6], which is essentially based on some results given by Chen Xiaojun [14].

Lemma 2.1. [6] for any fixed $\lambda > 0$ the solution set of (3) is nonempty and bounded. Let x_λ be a solution of (3), and $\{\lambda_k\}$ be any positive sequence converging to 0. If $SOL(M, q) \neq \phi$, then $\{x_{\lambda_k}\}$ has at least one accumulation point, and any accumulation point $\bar{x}$ of $\{x_{\lambda_k}\}$ is a solution of (3). That is, for any $\tilde{x} \in SOL(q, M)$ satisfied

$$\bar{x} \in SOL(q, M) \quad \text{and} \quad \|\bar{x}\|_p^p \leq \|\tilde{x}\|_p^p.$$

2.2. Lower Bounds for Nonzero Entries in Solutions

In this section, we extend the above result to the l_p norm regularization model (2) for approximating minimal l_0 norm solutions of the LCP. We provide a threshold lower bound $L > 0$ for any local minimizer, and show that any nonzero entries of local minimizers must exceed L. Since $f(x) \geq \lambda \|x\|_p^p$, the objective function $f(x)$ is bound below and $f(x) \to +\infty$ if $\|x\| \to \infty$. Moreover, the set χ_p^* of local minimizers of (3) is nonempty and bounded.

Lemma 2.2. [6] let x^* be any local minimizer of (3) satisfying $f(x^*) \leq f(x^0)$ for an arbitrarily given point x^0. Set

$$L = \left(\frac{\lambda p}{2\sqrt{2}(1 + \|M\|)\sqrt{f(x^0)}}\right)^{\frac{1}{1-p}} \tag{5}$$

Then we have: for any $i \in \{1, 2, \cdots, n\}, x_i \in [-L, L] \Rightarrow x_i^* = 0$.

Moreover, the number of nonzero entries in x^* is bounded.

$$\|x^*\|_0 \leq \frac{f(x^0)}{\lambda L^p}$$

Let us denote the first term of (3) by the function $\Psi_P(x) : R^n \to R_+$. That is

$$\Psi_P(x) = \frac{1}{2}\|\Phi_P(x)\|^2 \tag{6}$$

First we present some properties of φ_P and $\Phi_P(x)$.

Lemma 2.3. [12] let $\varphi_P : R^2 \rightarrow R^1$ be given by (3). Then, the following properties hold:

1) φ_P is a positive homogeneous and sub-additive NCP-function.

2) φ_P is strongly semismooth.

3) If $\left\{\left(a^k, b^k\right)\right\} \subseteq R \times R$ with $a^k \rightarrow -\infty$, or $b^k \rightarrow -\infty$, or $a^k \rightarrow \infty$, $b^k \rightarrow \infty$, then $\left|\varphi_P\left(a^k, b^k\right)\right| \rightarrow \infty$ when $k \rightarrow \infty$.

4) Given a point $(a,b) \in R \times R$, very element in the generalized gradient $\partial\varphi_P(a,b)$ has the representation $(\xi - 1, \zeta - 1)$, where

$$\xi = \frac{\mathrm{sgn}(a) \cdot |a|^{P-1}}{\|(a,b)\|_P^{P-1}} \quad \text{and} \quad \zeta = \frac{\mathrm{sgn}(b) \cdot |b|^{P-1}}{\|(a,b)\|_P^{P-1}} \quad \text{for} \ (a,b) \neq (0,0)$$

sgn(.) represents the sign function; ξ and ζ are real numbers that satisfy $|\xi|^{\frac{P}{P-1}} + |\zeta|^{\frac{P}{P-1}} \leq 1$.

Theorem 2.1. The function $\Psi_P(x)$ is continuously differentiable everywhere and the gradient of $\Psi_P(x)$ can be obtained by

$$\nabla\Psi_P(x) = \left[D_a(x) + M^{\mathrm{T}} D_b(x)\right]\Phi_P(x) \tag{7}$$

where $D_a(x) = \mathrm{diag}\{a_i(x)\}$ and $D_b(x) = \mathrm{diag}\{b_i(x)\}$ are diagonal matrices whose diagonal element is given by

$$a_i(x) = \begin{cases} \dfrac{\mathrm{sgn}(x_i) \cdot |x_i|^{P-1}}{\left\|\left(x_i, (Mx+q)_i\right)\right\|_P^{P-1}} - 1 & \text{if } \left(x_i, (Mx+q)_i\right) \neq 0 \\ \xi_i - 1 & \text{otherwise} \end{cases}$$

$$b_i(x) = \begin{cases} \dfrac{\mathrm{sgn}(x_i) \cdot \left|(Mx+q)_i\right|^{P-1}}{\left\|\left(x_i, (Mx+q)_i\right)\right\|_P^{P-1}} - 1 & \text{if } \left(x_i, (Mx+q)_i\right) \neq 0 \\ \zeta_i - 1 & \text{otherwise} \end{cases}$$

where (ξ, ζ) is any vector satisfying $|\xi|^{\frac{P}{P-1}} + |\zeta|^{\frac{P}{P-1}} \leq 1$.

3. Smoothing Method for l_p Regularization

Most optimization algorithms are efficient only for convex and smooth problems. However, some algorithms for Non-smooth and non-convex optimization problems have been developed recently. Note that the term $\|x\|_p^p$ ($0 < p < 1$) in (3) is neither convex nor Lipschitz continuous in R^n. Solving the non-convex, non-Lipschitz continuous minimization problem is not easy. We use some approximation methods to surmount the non-Lipschitz continuity problem in solving (3).

Smoothing Counterpart for (3)

For $\mu \in [0, +\infty)$, let

$$s_\mu(t) = \begin{cases} |t|, & |t| > \mu, \\ \mu \ln\left(\exp\left(\dfrac{t}{\mu}\right) + \exp\left(-\dfrac{t}{\mu}\right)\right), & |t| \leq \mu. \end{cases} \tag{8}$$

It is clear to see that, for any $t \in R$,

$$0 \leq s_\mu^p(t) - |t|^p \leq (\mu \ln 2)^p$$

and $s_\mu^p(t)$ is continuously differentiable with

$$\left(s_\mu^p(t)\right)' = \begin{cases} p|t|^{p-1} sign(t), & |t| > \mu \\ p\left[\mu\ln\left(\exp\left(\dfrac{t}{\mu}\right)+\exp\left(-\dfrac{t}{\mu}\right)\right)\right]^{p-1}\dfrac{\exp\left(\dfrac{t}{\mu}\right)-\exp\left(-\dfrac{t}{\mu}\right)}{\exp\left(\dfrac{t}{\mu}\right)+\exp\left(-\dfrac{t}{\mu}\right)} & |t| \le \mu \end{cases}$$

We can construct a smoothing approximation of (2) as

$$\min_{x\in R^n} f_\mu(x) = \frac{1}{2}\left\|\Phi_P(x)\right\|^2 + \lambda\sum_{i=1}^n s_\mu^p(x_i) \tag{9}$$

by noting that $f_\mu(x)$ is continuously differentiable, and $\lim_{\mu\to 0} f_\mu(x) = f(x)$ since

$$0 \le f_\mu(x) - f(x) \le \lambda n(\mu\ln 2)^p \tag{10}$$

for any $x \in R^n$.

Let $\chi_{p,\mu}^*$ denote the set of local minimizers of (10). We have the similar results as in **Lemmma 2.1** and **2.2**, corresponding to the smoothing counterpart (10).

Theorem 3.1. Let $\left\{x_{\mu_k}\right\}$ be a sequence of vectors being global minimizers of (10) with $\mu_k \to 0$ as $k \to \infty$.

Then, any accumulation point of $\left\{x_{\mu_k}\right\}$ is a global minimizer of (3).

Proof. Let x^* be a global minimizer of (3) and $\overline{x}$ be an accumulation point of $\left\{x_{\mu k}\right\}$.
We can deduce from (11) that

$$f\left(x_{\mu k}\right) \le f_{\mu k}\left(x_{\mu k}\right) \le f_{\mu k}\left(x^*\right) \le f\left(x^*\right) + \lambda n\left(\mu_k \ln 2\right)^p$$

On the other hand, we have $f\left(x^*\right) \le f\left(x_{\mu k}\right)$, and consequently

$$f\left(x^*\right) \le f\left(x_{\mu k}\right) \le f\left(x^*\right) + \lambda n\left(\mu_k \ln 2\right)^p$$

Which indicates $\overline{x}$ is a global minimizer of (3).

Lemma 3.2. [6] for any $\mu > 0$, let $x_\mu^* \in \chi_{p,\mu}^*$ be any local minimizer of (14) satisfying $f\left(x_\mu^*\right) \le f\left(x^0\right)$ for an arbitrarily given initial points x^0. Let L be defined in **Lemma 2.2**. Then, we have for any

$$i \in \{1,2,\cdots,n\}, \left(x_\mu^*\right)_i \in [-L,L] \Rightarrow \left|\left(x_\mu^*\right)\right| \le \mu$$

4. SS-SG Algorithm

We suggest a sequential Smoothing Spectral Gradient (SS-SG) Method to solve (3). With the SS-SG method, we need the Spectral Gradient method as the main step for decreasing the objective value. The smoothing method is very easy to implement and efficient to deal with optimization; see [15].

We first introduce the spectral projected gradient method in [8] as follow.

Algorithm 1. Smoothing Spectral Gradient Method

Step 0: Choose an initial point $x_0 \in R^n$, and parameters $\sigma \in [0,0.5]$, $\beta,\rho,\mu_0 \in [0,1]$, $\eta = 10^{30}$. Let $\overline{\alpha}_1 = 1$, $C_1 = f_{\mu_k}(x_1)$, $k = 0$.

Step1: Let $g_k = \nabla f_{\mu k}(x_k)$, $d_k = -\overline{\alpha}_k g_k(\overline{\alpha}_k)$, $\alpha_k = \dfrac{s_{k-1}^{\mathrm{T}}s_{k-1}}{s_{k-1}^{\mathrm{T}}y_{k-1}}$, where $s_{k-1} = x_k - x_{k-1}$, $y_{k-1} = g_k - g_{k-1}$. If $\|g_k\| = 0$, then stop.

Step 2: Compute the step size v_k by the Armijo line search, where $v_k = \max\left\{\rho^0,\rho^1,\cdots\right\}$ satisfies

$$f\left(x_k + v_k d_k\right) \le C_k + \sigma v_k \overline{\alpha}_k g_k^{\mathrm{T}} g_k\left(\mu v_k - 1\right)$$

Set $x_{k+1} = x_k + v_k d_k$.

Step 3: If $\left\| \nabla f_{\mu k} \left(x_{k+1} \right) \right\| \geq n \mu_k$, then set $\mu_{k+1} = \mu_k$; otherwise, choose $\mu_{k+1} = \beta \mu_k$.

Algorithm 2. Sequential Smoothing Spectral Gradient Method

Step 1: Find x by using the algorithm 1 to solve

$$\min_{x \in R^n} f(x) = \frac{1}{2} \left\| \Phi_P(x) \right\|^2 + \lambda \|x\|_p^p$$

Step 2: Compute

$$L_\lambda = \left(\frac{\lambda p}{2\sqrt{2}\left(1+\|M\|\right)\sqrt{f\left(x^0\right)}} \right)^{\frac{1}{1-p}}$$

Use the lower bound L_λ to set the entries of x with small values to zeros and obtain the computed solution x_λ with

$$\left(x_\lambda \right)_i = \begin{cases} x_i & |x_i| \geq L_\lambda \\ 0 & \text{otherwise} \end{cases}$$

Step 3: Decrease the parameter λ and set $x_0 := x_\lambda$.

5. Numerical Experiments

In this section, we test some numerical experiments to demonstrate the effectiveness of our SG algorithm. In order to illustrating the effectiveness of the SS-SG algorithm we proposed, we introduce another algorithm of talking the LCPs. In [6], the authors designed a sequential smoothing (SSG) method to solve the l_p regularized model and get a sparse solution of LCP(q, M). Numerical experiments show that our algorithm is more effective than (SSG) algorithm.

The program code was written in and run in MATLAB R2013 an environment. The parameters are chooses as $\sigma = 0.5$, $\beta = 0.25$ and $\mu_0 = 0.01$. The maximum number of iterations in step 1 is set to be 2000. We end the SS-SG algorithm in Step 1, if $\left\| \nabla f_{\mu k} \left(x_k \right) \right\| < 10^{-5}$ and $\mu_k < 10^{-4}$, or it reaches the maximum number of iterations.

5.1. Test for LCPs with Positive Semidefinite Matrices

Example 1. We consider the LCP(q, M) with

$$M = \begin{pmatrix} 0.4 & -0.3 & 0.1 \\ -0.3 & 0.3 & -0.3 \\ 0.1 & -0.3 & 0.7 \end{pmatrix}, \qquad q = \begin{pmatrix} -0.4 \\ 0.3 \\ -0.1 \end{pmatrix}$$

The solution set is $SOL(q,M) = \left\{ (1,0,0)^{\mathrm{T}} + a(2,3,1)^{\mathrm{T}} : a \geq 0 \right\}$.

When $a = 0$, the vector $x^* = (1,0,0)^{\mathrm{T}}$ is the sparse solution of LCP(q, M). We choose $P = 10, p = 0.1$ in (3) for this small example, and use our SS-SG algorithm with the regularization parameter $\lambda = 0.01$. We use the initial point $x^0 = (3,3,1)^{\mathrm{T}}$, we get a minimal l_p norm solution $\bar{x} = (1.000,0,0)$ and the distance $\left\| x^* - \bar{x} \right\| = 2.452 \times 10^{-4}$.

Example 2. We consider the LCP(q, M) with

$$M = \begin{pmatrix} 5 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ 0 \\ -2 \end{pmatrix}$$

The solution set is

$$SOL(q,M) = \left\{ (x_1, x_2, x_3)^{\mathrm{T}} : x_1 = a + (1-a)\frac{2}{3}, x_2 = a, x_3 = (1-a)\frac{2}{3}, 0 \leq a \leq 1 \right\}$$

When $a = 0$, the vector $x_1^* = \left(\dfrac{2}{3}, 0, \dfrac{2}{3}\right)^{\mathrm{T}}$ is the sparse solution of LCP(q, M). When $a = 1$, the vector $x_2^* = (1, 1, 0)$ is the sparse solution of LCP(q, M). We choose the same parameters as **Example 1**. We use the initial point $x^0 = (2, 1, 2)^{\mathrm{T}}$, we get a minimal l_p norm solution $\bar{x} = (0.667, 0, 0.667)$ and the distance $\|x_1^* - \bar{x}\| = 1.341 \times 10^{-4}$. We use the initial point $x^0 = (2, 2, 1)^{\mathrm{T}}$, we get a minimal l_p norm solution $\bar{x} = (1.000, 1.000, 0)$ and the distance $\|x_2^* - \bar{x}\| = 1.079 \times 10^{-4}$.

These examples show that, given the proper initial point, our algorithm can effectively find an approximate sparse solution.

5.2. Test for LCPs with Z-Matrix [6]

Let us consider LCP(q, M) where

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

Here I_n is the identity matrix of order n and $e = (1, 1, \cdots, 1)^{\mathrm{T}} \in R^n$. Such a matrix M is widely used in statistics. It is clear that Mis 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

$$x \geq 0, \qquad Mx + q = Me_1 + q = 0, \qquad x^{\mathrm{T}}(Mx + q) = 0$$

Among all the solutions, the vector $x^* = e_1 = (1, 0, \cdots, 0)^{\mathrm{T}}$ is the unique sparsest solution. We test the SS-SG algorithm for different dimensions with n = 100, 300, 500, 1000, 1300, respectively. In this set of experiments, we set $P = 10, p = 0.01, \lambda = 0.01$. The results are displayed in **Table 1**.

In **Table 1**, "$\|x - \hat{x}\|$" denotes the Euclidean distance between x and the true sparsest $\hat{x}$, and "time" denotes the computational time in seconds. Form **Table 1**, we can see that the SS-SG algorithm is effective to find the sparse solution of LCPs.

In order to test the effectiveness of the SS-SG algorithm, we compare with the SSG algorithm of talking the LCPs. In [10], the authors use the Fiser-Burmeister function established a l_p ($0 < p < 1$) regularized minimization model and designed a SSG method to solve the LCPs. The results are displayed in **Table 2**, where "_" denotes the method is invalid. Although the sparsity $\|x\|_0$ is same and the recovered errors $\|x - \hat{x}\|$ are pretty small, the average cpu time less than the SSG algorithm.

Table 1. SS-SG's computation results on LCPs with Z-matrices.

n	*Iter*	$\|x - \hat{x}\|$	$\|x\|_0$	$\|\hat{x}\|_0$	Time(s)
100	789	2.71E−3	1	1	2.25
200	452	5.22E−3	1	1	2.27
500	14	3.91E−4	1	1	4.02
800	11	4.21E−4	1	1	4.73
1000	3	1.64E−5	1	1	4.73
1300	29	2.16E−5	1	1	25.41

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

n	*Iter*	$\|x - \hat{x}\|$	$\|x\|_0$	$\|\hat{x}\|_0$	Time(s)
100	2184	3.41E−3	1	1	9.66
200	549	4.20E−3	1	1	17.44
500	18	5.11E−3	1	1	28.77
800	9	2.23E−3	1	1	63.01
1000	3	1.24E−4	1	1	4.13
1300	−	−	−	−	−

6. Conclusion

In this paper, we have studied a l_p $(0 < p < 1)$ model based on the generalized FB function defined as in (2) to find the sparsest solution of LCPs. Then, an l_p normregularized and unconstrained minimization model is proposed for relaxation, and we use a sequential smoothing spectral gradient method to solve the model. Numerical results demonstrate that the method can efficiently solve this regularized model and gets a sparsest solution of LCP with high quality.

Acknowledgements

References

[1] Facchinei, F. and Pang, J.S. (2003) Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research, Vol. I & II, Springer, New York.

[2] Ferris, M.C., Mangasarian, O.L. and Pang, J.S. (2001) Complementarity: Applications, Algorithms and Extensions. Kluwer Academic Publishers, Dordrecht.

[3] Gao, J.D. (2013) Optimal Cardinality Constrained Portfolio Selection. *Operations Research*, **61**, 745-761. http://dx.doi.org/10.1287/opre.2013.1170

[4] Xie, J., He, S. and Zhang, S. (2008) Randomized Portfolio Selection with Constraints. *Pacific Journal of Optimization*, **4**, 87-112.

[5] Cottle, R.W., Pang, J.S. and Stone, R.E. (1992) The Linear Complementarity Problem. Academic Press, Boston.

[6] Shang, M.J., Zhang, C. and Xiu, N.H. (2014) Minimal Zero Norm Solutions of Linear Complementarity Problems. *Journal of Optimization Theory and Applications*, **163**, 795-814. http://dx.doi.org/10.1007/s10957-014-0549-z

[7] Barzilai, J. and Borwein, J.M. (1988) Two Point Step Size Gradient Method. *IMA Journal of Numerical Analysis*, **8**, 174-184. http://dx.doi.org/10.1093/imanum/8.1.141

[8] Raydan, M. (1997) The Barzilai-Borwein Gradient Method for the Large Scale Unconstrained Optimization Problem. *SIAM Journal on Optimization*, **7**, 26-33. http://dx.doi.org/10.1137/S1052623494266365

[9] Raydan, M. (1993) On the Barzilai and Borwein Choice of Step Length for the Gradient Method. *IMA Journal of Numerical Analysis*, **13**, 321-326. http://dx.doi.org/10.1093/imanum/13.3.321

[10] Chen, J.-S. and Pan, S.-H. (2010) A Neural Network Based on the Generalized Fischer-Burmeister Function Fornonlinear Complementarity Problem. *Information Sciences, 180*, 697-711. http://dx.doi.org/10.1016/j.ins.2009.11.014

[11] Chen, J.-S. and Pan, S.-H. (2008) A Family of NCP Functions and a Descent Method for the Nonlinear Complementarity Problem. *Computation Optimization and Application*, **40**, 389-404. http://dx.doi.org/10.1007/s10589-007-9086-0

[12] Chen, J.-S. and Pan, S.-H. (2008) A Regularization Semismooth Newton Method Based on the Generalized Fischer-Burmeister Function for P_0-NCPs. *Journal of Computation and Applied Mathematics*, **220**, 464-479. http://dx.doi.org/10.1016/j.cam.2007.08.020

[13] Chen, J.-S., Gao, H.-T. and Pan, S.-H. (2009) A Derivative-Free R-Linearly Convergent Derivative-Free Algorithm for the NPCs Based on the Generalized Fischer-Burmeister Merit Function. *Journal of Computation and Applied Mathematics*, **232**, 455-471.

[14] Chen, X.J., Xu, F. and Ye, Y. (2010) Lower Bound Theory of Nonzero Entries in Solutions of l_2-l_p Minimiization. *SIAM*: *SIAM Journal on Scientific Computing*, **32**, 2832-2852. http://dx.doi.org/10.1137/090761471

[15] Chen, X. and Xiang, S. (2011) Implicit Solution Function of P_0 and Z Matrix Linear Complementarity Constrains. *Mathematical Programming*, **128**, 1-18. http://dx.doi.org/10.1007/s10107-009-0291-8