AMO- Advanced Modeling and Optimization, Volume 20, Number 2, 2018*

A spectral gradient-Newton two phase method for solving nonlinear equations

Yuezhe Zhang†, Zhensheng Yu

(College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Systems of semismooth equations arises in various applications, great efforts have been made to find a solution of semismooth e quations. In this paper, we proposed a spectral gradient-Newton two phase method for solving semismooth equations and obtain the global convergence of the algorithm. In the first stage, we use the spectral gradient method to obtain the global convergence of the algorithm, and then use the final point obtained in the first stage as a new initial point to turn to Newton method for fast convergence speed. Numerical results show that the new initial point can accelerate the convergence speed effectively.

Keywords: semismooth equations; spectral gradient method; semismooth Newton method; two-phase method.

1 Introduction

In this paper, we consider the nonlinear semismooth equations problem: finding a vector $x_* \in R^n$ such that

$$H(x_*) = 0, \tag{1.1}$$

where $H : R^n \to R^n$ is a semismooth mapping. The notion of semismoothness was introduced for the functions by Mifflin in [Mifflin, 1977] and extend to vector functions by Qi and Sun in [Qi and Sun, 1993].

*AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

†Correspondence: yuezhezhang@126.com

Systems of semismooth equations arise in various applications, some variational inequalities, complementarity problems and mathematical programming can be converted into semismooth equations (see [Facchinei and Pang, 2003] and references therein) . Besides more and more semismooth equations come from the problems such as the semiinfinite programming problem see [Qi et al, 2003], optimal control problem see [Yousept, 2008], operator equations problems see [Ulbrich, 2002] and optimal power flow problems see [Tong et al, 2 017]. Many numerical methods for problem (1.1) have been proposed, These metoods include inexact Newton Methods [Facchinei et al, 1996], Quasi-Newton Methods [Sun and Han, 1997], interior-point affine-scaling trust-region methods [Kanzow and Klug, 2007] and so on.

The solution of nonlinear equations (1.1) can be transformed into solving the following unconstrained problem:

$$min f(x) = \frac{1}{2}\|H(x)\|^2, \tag{1.2}$$

where $f : R^n \to R^n$ is continuously differentiable and its gradient denoted by $\nabla f(x)$. The methods of optimization problem involve the first-order methods and second-order methods. The main advantage of the first-order method is its small storage, which is particularly suitable for large-scale problem. However, the first-order method cannot meet the requirement of high precision, and its convergence speed is linear at most. For the second-order method, the convergence speed is fast. Under certain conditions, it can achieve superlinear convergence and even quadratic convergence, but the drawback is the need for a sufficiently good initial point. Motivated by this, in this paper, we combine the spectral gradient method and semismooth Newton method to solve the problem (1.2) , for any given initial point, the spectral gradient method is used to calculate. When the iteration point is close to the optimal solution, the Newton method is used to speed up the convergence speed.

The organization of this paper is as follows. In section 2, we present some preliminaries that are useful to our main results. In section 3, we propose our algorithm and analyze the global convergence of the proposed algorithm. The application and numerical tests are given in section 4.

2 Preliminaries

In this section, we present some definitions and theorems that are useful to our main results, they can be found in [Qi and Sun, 1993].

Suppose $H : R^n \to R^n$ is a locally Lipschitzian. According to Rademacher theorem, H is differentiable almost everywhere. Denote the set of points at which H is differentiable by D_H . We write $H'(x_k)$ for the usual $n \times m$ Jacobian matrix of partial derivatives whenever x is a point at which the necessary partial derivatives exists. Let $\partial H(x)$ be the

generalized Jacobian defined by Clarke in [Clarke, 1983]. Then

$$\partial H(x) = C_0\{\lim_{\substack{x_j \to x \\ x_j \in D_H}} H'(x_j)\}.$$

Definition 2.1 (see [Qi and Sun, 1993]) Suppose $H : R^n \to R^m$ is a locally Lipschitzian function, we say that H is semismooth at x is $\lim_{\substack{V \in \partial H(x+th') \\ h' \to h, t \downarrow 0}} \{Vh'\}$ exists for any $h \in R^n$.

Theorem 2.1 *(see [Qi and Sun, 1993]) Suppose $H : R^n \to R^m$ is a locally Lipschitzian function, the following statements are equivalent:*

(a) H is semismooth at x;

(b) for any $V \in \partial H(x + h), h \to 0$,

$$Vh - H'(x; h) = o(\|h\|).$$

Theorem 2.2 *(see [Qi and Sun, 1993]) Suppose $H : R^n \to R^m$ is a locally Lipschitzian function, if each component of H is semismooth at x, then H is semismooth at x.*

Definition 2.2 (see [Qi and Sun, 1993]) Suppose $H : R^n \to R^m$ is a locally Lipschitzian function, If for any $V \in \partial H(x + h), h \to 0$,

$$Vh - H'(x; h) = O(\|h\|)^{1+p},$$

where $0 < p \leq 1$, then we call H is p-order semismooth at x.

Definition 2.3 (see [Qi and Sun, 1993]) Suppose $H : R^n \to R^m$ is a locally Lipschitzian function, we say H is strongly BD-regular at x if all $V \in \partial_B H(x)$ are nonsingular.

3 Algorithm and global convergence

Spectral gradient method was first introduced by Barzilai and Borwein in [Barzilai and Borwein, 1988]. Due to its simplicity and numerical efficiency, the spectral as well as the spectral projected gradient method has been applied successfully to find local minimizers of large scale problems see [Birgin et al, 2003, Birgin et al, 2008, Dai and Fletcher, 2005, Diniz-Ehrhardt, 2004, La Cruz et al, 2006, Raydan, 2003, Molina and Raydan, 1996, Zhang and Zhou, 2006, Yu, 2008].

For the problem (1.2), at kth iteration, denote $s_{k-1} = x_k - x_{k-1}, y_{k-1} = \nabla f(x_k) - \nabla f(x_{k-1})$. The Quasi-Newton secant equation is

$$B_{k+1} s_k = y_k.$$

Let I denote the identical matrix in $R^{n \times n}$ and assume that we want a matrix B_k with a very simple structure that satisfies the secant equation. More precisely, we wish $B_k = \lambda_k I$ with $\lambda_k \in R$.

Then the secant equation becomes:

$$\lambda_k s_{k-1} = y_{k-1}.$$

In general, this equation may not be consistent. However, accepting the least-squares solution that minimizers

$$\|\lambda_k S_{k-1} - y_{k-1}\|^2,$$

we obtain

$$\lambda_k = \frac{\langle s_{k-1}, y_{k-1}\rangle}{\langle s_{k-1}, s_{k-1}\rangle},$$

here $\prec, \succ$ denotes the inner product for given two vectors. This formular defines the most popular spectral gradient method for unconstrained optimization with the search direction:

$$d_k = -\frac{1}{\lambda_k} g_k,$$

where g_k denote the gradient of f at x_k .

Newton's method

$$x_{k+1} = x_k - [H'(x_k)]^{-1} H(x_k) \tag{3.1}$$

is a classic method for solving the nonlinear equation, when H is not a smooth function, but a locally Lipschitzian function, the formula cannot be used. In this case, insted of (3.1), one may use

$$x_{k+1} = x_k - V_k^{-1} H(x_k), \tag{3.2}$$

where $V_k \in \partial H_B(x_k)$, to solve (1.1).

In order to obtain the global convergence of the algorithm, we use the spectral gradient method see [Raydan, 1997] in the first stage.

Algorithm 3.1 (GBB)

Given x_0 ,α_0 ,integer $M \geq 0$, $\gamma \in (0, 1)$, $\delta > 0, 0 < \sigma_1 < \sigma_2 < 1, 0 < \varepsilon < 1$. Set $k = 0$.

Step 1: If $\|g_k\| = 0$ stop;

Step 2: If $\alpha_k \leq \varepsilon$ or $\alpha_k \geq \frac{1}{\varepsilon}$, then set $\alpha_k = \delta$, the parameter δ was chosen in the following way:

$$\delta = \begin{cases} 1, & \|g_k\|_2 > 1, \\ \|g_k\|_2^{-1}, & 10^{-5} \leq \|g_k\|_2 \leq 1, \\ 10^5, & \|g_k\|_2 \leq 10^{-5}. \end{cases}$$

Step 3: Set $\lambda = \frac{1}{\alpha_k}$;

Step 4:(nonmonotone line search)

If $f(x_k - \lambda g_k) \leq max_{0\leq j\leq min(k,M)}(f_{k-j}) - \gamma\lambda g_k^t g_k$,

then set $\lambda_k = \lambda$, $x_{k+1} = x_k - \lambda_k g_k$, and go to Step 6;

Step 5: Choose $\sigma \in [\sigma_1, \sigma_2]$, set $\lambda = \sigma\lambda$, and go to Step 4;

Step 6: Set $\alpha_{k+1} = -\frac{g_k^t y_k}{\lambda_k g_k^t g_k}$, $k = k + 1$,and go to Step 1.

Using the final point obtained in the first stage as a new initial point to turn to the semismooth Newton method see [Qi and Sun, 1993].

Algorithm 3.2

Set $x_0 = x_N$, where x_N is the output point of the last stage. Given $\rho > 0$, $p > 2$, $\theta \in (0, \frac{1}{2})$, $0 \leq \varepsilon_2 < \varepsilon_1$. Set $k = 0$.

Step 1: If $\|g_k\| \leq \varepsilon_2$, stop;

Step 2: Select an element V_k in $\partial_B H(x_k)$.

Find the solution d_k of the system

$$V_k d = -H(x_k). \tag{3.3}$$

If system (3.3) is not solvable or if the condition

$$g_k^T d_k \leq -\rho\|d_k\|^p \tag{3.4}$$

is not satisfied, set $d_k = -g_k$;

Step 3: Find the smallest $i_k = 0, 1, 2, ...$ such that

$$f(x_k + 2^{-i}d_k) \leq f(x_k) + \theta 2^{-i} g_k^T d_k.$$

Set $x_{k+1} = x_k + 2^{-i_k}d_k$, $k := k + 1$, and go to Step 1;

To establish the global convergence of algorithm, We make the following assumptions.

Assumption

($A1$) The solution set of (1.1) is nonempty, denoted by S;

($A2$) The mapping $H : R^n \to R^n$ is a locally Lipschitzian function on R^n.

The convergence properties of the GGB algorithm are stated in the following theorem. The proof of the theorem can be found in [Raydan, 1997].

Theorem 3.1 *Assume that $\Omega_0 = \{x : f(x) \leq f(x_0)\}$ is a bounded set. Let $f : R^n \to R$ be continuously differentiable in some neighborhood N of Ω_0 . Let $\{x_k\}$ be the sequence generated by the GGB algorithm. Then either $g(x_j) = 0$ for some finite j, or the following properties hold:*

(a) $lim_{k\to 0}\|g_k\| = 0$;

(b) no limit point of x_k is a local maximum of f;

(c) if the number of stationary point of f in Ω_0 is finite, then the sequence x_k converges.

The following theorem states the local convergence of the algorithm 3.2 see [Qi, 1993].

Theorem 3.2 *Suppose that $H(x_*) = 0$, and that H is semismooth strongly BD-regular at x_*. Then the iteration method (3.2) is well defined and convergent to x_* superlinearly in a neighborhood of x_*; Besides, if $H(x_k) \neq 0$ for all k, then $\{\|H(x_k)\|\}$ decrease superlinearly.*

4 Application and numerical results

We consider the nonlinear complementarity problem, NCP(F) for short, which is to find a vector in R^n satisfying the conditions

$$x \geq 0, F(x) \geq 0, x^T F(x) = 0, \tag{4.1}$$

where $F : R^n \to R^n$ is a continuously differentiable function. We can reformulate NCP(F) as a system of equations based on the following function:

$$\phi(a, b) := \sqrt{a^2 + b^2} - (a + b)$$

The most interesting property of this function is

$$\phi(a, b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0 \tag{4.2}$$

Exploiting (4.2), it is readily seen that the nonlinear complementarity problem is equivalent to the following system of nonsmooth equations:

$$\Phi(x) := \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix} = 0 \tag{4.3}$$

It is then also obvious that the nonnegative function

$$\Psi(x) := \frac{1}{2}\|\Phi(x)\|^2 \tag{4.4}$$

is zero at a point x if and only if x is a solution of NCP(F), so that solving NCP(F) (4.1) is equivalent to find the unconstrained global solution of the problem

$$min\Psi(x) := \frac{1}{2}\|\Phi(x)\|^2 \tag{4.5}$$

We use two phase algorithm to solve the problem (4.5).

To test the efficiency of our proposed algorithm, we conducted the numerical experiments on some test problem. The algorithm were coded in MATLAB 7.14. We compare the results obtained by our method with that obtained by the semismooth Newton method see[Qi ang Sun, 1993] and GGB algorithm see [MR, 1997]. The parameters are set as follows: $\gamma = 10^{-4}, \varepsilon = 10^{-10}, \sigma_1 = 0.1, \sigma_2 = 0.5, \alpha_0 = 1, M = 10, \rho = 0.2, p = 2.2, \theta = 0.4, \varepsilon_1 = 10^{-3}, \varepsilon_2 = 10^{-7}$.

The Fail denotes the iterations are more than 1000. The test problems are taken from [Jiang and Qi, 1997, Yu and Qin, 2011, Zhang and Zhang, 2006], where the function $F(x)$ is as follows:

LCP1: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

q=(-1,-1), initial point: (a)=(0,0), (b)=(100,120), (c)=(500,700).

NCP2:

$$F(x) = \begin{pmatrix} x_1 - 2 \\ x_2^3 + x_2 - x_3 - 3 \\ x_2 + 2x_3^3 + x_3 - 3 \end{pmatrix},$$

initial point: (a)=(0,0,0), (b)=(10,10,10), (c)=(100,100,100).

LCP3: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix},$$

q=(1,0,-1), initial point: (a)=(1,2,1), (b)=(1000,1000,1000), (c)=(0.001,0.001,0.001).

LCP4: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix},$$

q=(0,0,1), initial point: (a)=(2,2,2), (b)=(5,5,5), (c)=(4,5,6).

NCP5:

$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix},$$

initial point: (a)=(0,0,0,0), (b)=(0,0,1,1), (c)=(100,100,100,100).

NCP6:

$$F(x) = \begin{pmatrix} x_1^3 - 8 \\ x_2 - x_3 + x_2^3 + 3 \\ x_2 + x_3 + 2x_3^3 - 3 \\ x_4 + 2x_4^3 \end{pmatrix},$$

initial point: (a)=(0,0,0,0), (b)=(10,10,10,10), (c)=(100,100,100,100).

LCP7: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 0 & 0 & 10 & 20 \\ 0 & 0 & 30 & 15 \\ 10 & 20 & 0 & 0 \\ 30 & 15 & 0 & 0 \end{pmatrix},$$

q=(-1,-1,-1,-1), initial point: (a)=(0,0,0,0), (b)=(1,2,1,2), (c)=(100,100,100,100).

Table 1

problem	initial point	two-phase	first stage	second stage	new initial point	[15]	[17]
LCP1	(a)	26	23	3	(0.4999,0.4999)	6	48
dim=2	(b)	50	47	3	(-0.0002,1.001)	7	82
	(c)	50	47	3	(-0.0006,1.0015)	7	93
NCP2	(a)	82	79	3	(1.9992,1.3428,0.7643)	7	174
dim=3	(b)	136	133	3	(2.0008,1.3428,0.7643)	9	258
	(c)	212	209	3	(2.0007,1.3429,0.7643)	8	313
LCP3	(a)	59	56	3	(0.0009,0.0670,0.2667)	7	164
dim=3	(b)	77	74	3	(0.0007,0.0669,0.2667)	7	173
	(c)	50	47	3	(-0.0007,0.0665,0.2666)	7	137
LCP4	(a)	34	30	4	(0.0008,0.9634,0)	11	51
dim=3	(b)	28	25	3	(2.2617,0.0008,0)	9	49
	(c)	49	46	3	(0,1.0018,0.0013)	10	85
NCP5	(a)	240	237	3	(1.224,0.0009,-0.0009,0.5013)	67	499
dim=4	(b)	237	234	3	(1.2241,0.0008,-0.0008,0.5012)	752	527
	(c)	261	258	3	(1.224,0.0009,-0.0009,0.5014)	9	495
NCP6	(a)	248	245	3	(2,0.0006,0.9999,0)	8	580
dim=4	(b)	291	289	2	(2,0,1,0.0023)	9	966
	(c)	406	404	2	(2,0,1,0.0026)	9	856
LCP7	(a)	434	430	4	(0.0976,0.0012,0.0976,0.0012)	Fail	Fail
dim=4	(b)	131	128	3	(0.0111,0.0444,0.0111,0.0444)	97	259
	(c)	360	356	4	(0.0975,0.0012,0.0975,0.0012)	7	Fail

LCP8: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 1 \end{pmatrix},$$

q=(0,0,1), initial point: $x_0 = (1, 1, 1)$.

NCP9:

$$F(x) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - \frac{4.5x_3+2.7x_4}{x_2+1} \\ 5 - x_1 - \frac{0.5x_3+0.3x_4}{x_3+1} \\ 3 - x_1 \end{pmatrix},$$

initial point: $x_0 = (2, 1, 1, 1)$.

LCP10: $F(x) = Mx + q$, where

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

q=(-8,-6,-4,3), initial point: $x_0 = (1, 1, 1, 1)$.

Table 2

problem	dim	two-phase	first stage	second stage	new initial point	[15]	[17]
LCP8	3	43	32	11	(0.0005,0.52,0)	12	55
NCP9	4	82	78	4	(0.8049,0.0006,0.0004,0.0003)	8	114
LCP10	4	165	162	3	(1.3328,0.7780,0.4448,0.2226)	7	286

LCP11:

$$F(x) = \begin{pmatrix} 2x_1 - x_3 + x_5 + 3x_6 - 1 \\ x_2 + 2x_5 + x_6 - x_7 - 3 \\ -x_1 + 2x_5 + x_6 - x_7 - 3 \\ x_3 + x_4 + x_5 - x_6 - 1 \\ -x_1 - 2x_2 - x_3 - x_4 + 5 \\ -3x_1 - x_2 - 2x_3 + x_4 + 4 \\ x_2 + 4x_3 - 1.5 \end{pmatrix},$$

initial point: $x_0 = (-1, -2, -1, -2, -1, -2, -2)$.

LCP12: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

q=-e, n=16, initial point: $x_0 = (0, \cdots, 0)$.

LCP13: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ 0 & 1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix},$$

q=-e, initial point: $x_0 = (0, \cdots, 0)$.

LCP14: $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 4 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 4 \end{pmatrix},$$

q=-e, initial point: $x_0 = (0, \cdots, 0)$.

Table 3

problem	dim	two phase	first stage	second stage	[15]	[17]
LCP11	7	341	127	214	482	352
LCP12	16	50	47	3	6	134
LCP13	100	39	36	3	6	57
LCP13	300	38	36	2	6	57
LCP13	500	26	23	3	6	55
LCP14	100	32	30	2	6	68
LCP14	300	38	36	2	6	73
LCP14	500	41	39	2	6	65

From Tables 1, 2 and 3, we can see that the new initial point obtained in the first stage can accelerate the convergence speed of the second stage . Two phase algorithm overcome the drawback of the second order method for a sufficiently good initial point. We can also see that our method is better than the GBB algorithm. The numerical results show that our proposed algorithm works well.

Reference

[1] Barzilai, J., Borwein, J.M., (1988) Two-point step size gradient methods[J]. Ima J. Numer Anal, vol. 8, nr. 1, pp. 141-148.

[2] Birgin, E.G., Martłnez, J.M., Raydan, M., (2008) Spectral projected gradient methods[M]. Springer US.

[3] Clarke, F.H., (1983) Optimization and Nonsmooth Analysis[J]. vol. 5, pp. 847-853.

[4] Dai, Y.H., Fletcher, R., (2005) Projected barzilai-borwein methods for large-scale box-constrained quadratic programming[J]. Numerische Mathematik, vol. 100, nr. 1, pp. 21-47.

[5] Diniz-Ehrhardt, M.A., Gomes-Ruggiero, M.A., Martłnez, J.M., (2004) Augmented lagrangian algorithms based on the spectral projected gradient method for solving nonlinear programming problems[J]. Journal of Optimization Theory & Applications, vol. 123, nr 3, pp. 497-517.

[6] Facchinei, F., Fischer, A., Kanzow, C., (1996) Inexact Newton Methods for Semismooth Equations with Applications to Variational Inequality Problems[M] & Nonlinear Optimization and Applications. Springer US, pp. 125-139.

[7] Facchinei, F., Pang, J.S., (2003) Finite-dimensional variational inequalities and complementarity problems. New York: Springer, Berline.

[8] Jiang, H., Qi, L., (1997) A New Nonsmooth Equations Approach To Nonlinear Complementarity Problems[J]. Siam Journal on Control & Optimization, vol. 35, nr. 1, pp. 178-193.

[9] Kanzow, C., Klug, A., (2007) An interior-point affine-scaling trust-region method for semismooth equations with box constraints[J]. Computational Optimization & Applications, vol. 37, nr. 3, pp. 329-353.

[10] La Cruz, W., Martłnez, J.M., Raydan, M., (2006) Spectral residual method without gradient information for solving large-scale nonlinear systems of equations[J]. Mathematics of Computation, vol. 75, nr. 255, pp. 1429-1448.

[11] Mart, E.G., Nez, J., (2003) Nonmonotone spectral projected gradient methods on convex sets[J]. Ima Journal of Numerical Analysis, vol. 23, nr. 23, pp. 539-559.

[12] Mifflin, R., (1977), Semismooth and semiconvex functions in constrained optimization[J]. Siam Journal on Control & Optimization, vol. 15, nr. 6, pp. 959-972.

[13] Molina, B., Raydan, M., (1996) Preconditioned Barzilai-Borwein method for the numerical solution of partial differential equations[J]. Numerical Algorithms, vol. 13, nr. 1, pp. 45-60.

[14] Qi, L.Q., (1993) Convergence analysis of some algorithms for solving nonsmooth equations[J]. Mathematics of Operations Research, vol. 18, nr. 1, pp. 227-244.

[15] Qi, L.Q., Sun, J., (1993) A nonsmooth version of Newton's method.[J]. Mathematical Programming, vol. 58, nr. 1-3, pp. 353-367.

[16] Qi, L., Wu, S.Y., Zhou, G., (2003) Semismooth newton methods for solving semi-infinite programming problems[J]. Journal of Global Optimization, vol. 27, nr. 2-3, pp. 215-232.

[17] Raydan, M., (1997) The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem[M]. Society for Industrial and Applied Mathematics, vol. 7, pp. 26-33.

[18] Raydan, M., (2003) Nonmonotone spectral methods for large-scale nonlinear systems[J]. Optimization Methods & Software, vol. 18, nr. 5, pp. 583-599.

[19] Sun, D., Han, J., (1997) Newton and Quasi-Newton Methods for a Class of Nonsmooth Equations and Related Problems[J]. Siam Journal on Optimization, vol. 7, nr. 2, pp. 463-480.

[20] Tong, X.J., (2017) A semismooth newton method for solving optimal power flow[J]. Journal of Industrial & Management Optimization, vol. 3, nr. 3, pp. 553-567.

[21] Ulbrich, M., (2002) Semismooth newton methods for operator equations in function spaces[M]. Society for Industrial and Applied Mathematics.

[22] Yu, Z., (2008) Solving bound constrained optimization via a new nonmonotone spectral projected gradient method[M]. Appl Number Math, vol. 58, pp. 1340-1348.

[23] Yu, Z., Qin, Y., (2011) A cosh-based smoothing Newton method for P 0 nonlinear complementarity problem[J]. Nonlinear Analysis Real World Applications, vol. 12, nr. 2, pp. 875-884.

[24] Yousept, I., (2008) Mesh-independence of semismooth Newton methods for Lavrentiev-regularized state constrained nonlinear optimal control problems[J]. Numerische Mathematik, vol. 108, nr. 4, pp. 571-603.

[25] Zhang, J.L., Zhang, X., (2006) A smoothing Levenberg-Marquardt method for NCP.[J]. Applied Mathematics & Computation, vol. 178, nr. 2, pp. 212-228.

[26] Zhang, L., Zhou, W., (2006) Spectral gradient projection method for solving nonlinear monotone equations[M]. Comput Appl Math, vol. 196, pp. 478-484.