Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 386030, 6 pages
http://dx.doi.org/10.1155/2014/386030

*Research Article*

# A Globally Convergent Matrix-Free Method for Constrained Equations and Its Linear Convergence Rate

**Min Sun[1] and Jing Liu[2]**

[1] School of Mathematics and Statistics, Zaozhuang University, Shandong 277160, China
[2] School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou 310018, China

Correspondence should be addressed to Min Sun; ziyouxiaodou@163.com

Received 24 January 2014; Accepted 8 May 2014; Published 19 May 2014

Academic Editor: Vladimir Danilov



A matrix-free method for constrained equations is proposed, which is a combination of the well-known PRP (Polak-Ribière-Polyak) conjugate gradient method and the famous hyperplane projection method. The new method is not only derivative-free, but also completely matrix-free, and consequently, it can be applied to solve large-scale constrained equations. We obtain global convergence of the new method without any differentiability requirement on the constrained equations. Compared with the existing gradient methods for solving such problem, the new method possesses linear convergence rate under standard conditions, and a relax factor $\gamma$ is attached in the update step to accelerate convergence. Preliminary numerical results show that it is promising in practice.

## 1. Introduction

Let $F : R^n \rightarrow R^n$ be a continuous nonlinear mapping and $C$ a nonempty closed convex set of $R^n$. In this paper, we consider the problem of finding $x \in C$ such that

$$F(x) = 0. \tag{1}$$

Nonlinear constrained equations (1), denoted by CES $(F, C)$, arise in various applications, for instance, ballistic trajectory computation and vibration systems [1], the power flow equations [2], chemical equilibrium systems [3], and so forth.

In recent years, many numerical methods have been proposed to find a solution of nonsmooth CES $(F, C)$, which include the trust region methods [4, 5], the Levenberg-Marquardt method [6], and the projection methods [7–9]. Compared with the trust region method and the Levenberg-Marquardt method, the projection method is more efficient for solving large-scale CES $(F, C)$. Noting this, Wang et al. [7] proposed a projection method for solving CES $(F, C)$, which possesses global convergence property without the differentiability. A drawback of this method is that it needs to solve a linear equation inexactly at each iteration, and its variants [8, 10] also have this drawback.

It is well-known that the spectral gradient method and the conjugate gradient method are two efficient methods for solving large-scale unconstrained optimization problems due to their simplicity and low storage. Recently, La Cruz and Raydan [11] successfully applied the famous spectral gradient method to solve unconstrained equations by using some merit function. Then, Zhang and Zhou [12] presented a spectral gradient projection method (SGP) for solving unconstrained monotone equations, which does not utilize any merit function. Later, the SGP was extended by Yu et al. [9] to solve monotone constrained equations. However, the study of conjugate gradient methods for large-scale (un)constrained equations is relatively rare. Cheng [13] proposed a PRP type method (PRPT) for systems of monotone equations, which is a combination of the well-known PRP method and the hyperplane projection method, and the numerical results in [13] show that the PRPT method performs better than the SGP method in [12].

Different from the methods in [7, 8, 10], the methods in [9, 11–13] do not need to solve a linearized equation at each iteration; however, the latter do not investigate the convergent rate, and even we do not know whether they possess the linear convergence rate. In this paper, motivated by the projection methods in [7, 8, 10] and the gradient methods in [9, 12, 13], we propose a matrix-free method for solving nonlinear constrained equations, which can be viewed as a combination of the well-known PRP conjugate gradient method and the famous hyperplane projection method, and it possesses linear convergence rate under standard conditions. The remainder of this paper is organized as follows. Section 2 describes the new method and presents its global convergence analysis. The linear convergence rate of the new method is established in Section 3. Numerical results are reported in Section 4. Finally, some final remarks are included in Section 5.

## 2. Algorithm and Convergence Analysis

Let $C^*$ denote the solution set of CES $(F, C)$. Throughout this paper, we assume that $C^*$ is nonempty and $F(\cdot)$ is monotone; that is,

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in R^n, \tag{2}$$

which implies that the solution set $C^*$ is closed. Then let $P_C(x)$ denote the orthogonal projection of a point $x \in R^n$ onto the convex set $C$, which has the following nonexpansive property:

$$\| P_C(x) - P_C(y) \| \leq \| x - y \|, \quad \forall x, y \in R^n. \tag{3}$$

Now, we describe the matrix-free method for nonlinear constrained equations.

*Algorithm 1.* Consider the following.

*Step 0.* Given an arbitrary initial point $x_0 \in C$, the parameters $0 < \rho < 1$, $0 < \sigma < r \leq 1$, $0 < \gamma < 2$, and $0 < \beta_{\min} < \beta_{\max}$. Given the initial steplength $\beta_0 = 1$ and set $k := 0$.

*Step 1.* If $F(x_k) = 0$, then stop; otherwise go to Step 2.

*Step 2.* Compute $d_k$ by

$$d_k = \begin{cases} -F(x_k), & \text{if } k = 0, \\ -F(x_k) + \beta_k d_{k-1} - \theta_k (F(x_k) - F(x_{k-1})), & \text{if } k \geq 1, \end{cases} \tag{4}$$

where

$$\beta_k^{\mathrm{PRP}} = \frac{\langle F(x_k), F(x_k) - F(x_{k-1}) \rangle}{\| F(x_{k-1}) \|^2}, \qquad \theta_k = \frac{F(x_k)^\top d_{k-1}}{\| F(x_{k-1}) \|^2}, \qquad \forall k \geq 1. \tag{5}$$

If $\| d_k \| > \| F(x_k) \| / r$, set $d_k = -F(x_k)$.

*Step 3.* Find the trial point $y_k = x_k + \alpha_k d_k$, where $\alpha_k = \beta_k \rho^{m_k}$ with $m_k$ being the smallest nonnegative integer $m$ such that

$$-\langle F(y_k), d_k \rangle \geq \sigma \| d_k \|^2. \tag{6}$$

*Step 4.* Compute

$$x_{k+1} = P_C [x_k - \gamma \xi_k F(y_k)], \tag{7}$$

where

$$\xi_k = \frac{\langle F(y_k), x_k - y_k \rangle}{\| F(y_k) \|^2}. \tag{8}$$

Choose an initial steplength $\beta_{k+1}$ such that $\beta_{k+1} \in [\beta_{\min}, \beta_{\max}]$. Set $k := k + 1$ and go to Step 1.

*Remark 2.* Obviously $d_k$, defined by (4), is motivated by [14], and it is not difficult to deduce that $d_k$ satisfies

$$F(x_k)^\top d_k = -\| F(x_k) \|^2. \tag{9}$$

Therefore, by Cauchy-Schwartz inequality, we have $\| d_k \| \geq \| F(x_k) \|$. This together with Step 2 of Algorithm 1 implies

$$\| F(x_k) \| \leq \| d_k \| \leq \frac{\| F(x_k) \|}{r}. \tag{10}$$

*Remark 3.* In (7), we attach a relax factor $\gamma \in (0, 2)$ (better when close to 2) to $F(y_k)$ based on numerical experiences.

*Remark 4.* Line search (6) is different from that of [12, 13], which is well-defined by the following Lemma.

**Lemma 5.** *For all $k \geq 0$, there exists a nonnegative number $m_k$ satisfying* (6).

*Proof.* In fact, if $d_k = 0$, then from (10), we have $\| F(x_k) \| = 0$, which means that Algorithm 1 terminates with $x_k$ being a solution of CES $(F, C)$. Now, we consider $d_k \neq 0$ for all $k$. For the sake of contradiction, we suppose that there exists $k_0 \geq 0$ such that (6) is not satisfied for any nonnegative integer $m$; that is,

$$-\left\langle F\left(x_{k_0} + \beta_{k_0} \rho^m d_{k_0}\right), d_{k_0} \right\rangle < \sigma \left\| d_{k_0} \right\|^2, \quad \forall m \geq 1. \tag{11}$$

Letting $m \to \infty$ and using the continuity of $F(\cdot)$ yield

$$-\left\langle F\left(x_{k_0}\right), d_{k_0} \right\rangle \leq \sigma \left\| d_{k_0} \right\|^2. \tag{12}$$

On the other hand, by (10), we obtain

$$-\left\langle F\left(x_{k_0}\right), d_{k_0} \right\rangle = \left\| F\left(x_{k_0}\right) \right\|^2 \geq r \left\| d_{k_0} \right\|^2, \tag{13}$$

which together with (12) means that $\sigma \geq r$; however, this contradicts the fact that $r > \sigma > 0$. Therefore the assertion holds. This completes the proof. □

**Lemma 6.** *Suppose that $F(\cdot)$ is monotone and let $\{x_k\}$ and $\{y_k\}$ be the sequences generated by Algorithm 1; then $\{x_k\}$ and $\{y_k\}$ are both bounded; furthermore, it holds that*

$$\lim_{k \to \infty} \alpha_k \| d_k \|^2 = 0. \tag{14}$$

*Proof.* From (6), we have

$$\left\langle F\left(y_k\right), x_k - y_k\right\rangle \geq \sigma\alpha_k\left\|d_k\right\|^2 > 0. \tag{15}$$

For any $x^* \in C^*$, from (3), the nonexpansiveness of the projection operator, it holds that

$$\begin{aligned}
&\|x_{k+1} - x^*\|^2 \\
&= \|P_C[x_k - \gamma\xi_k F(y_k)] - x^*\|^2 \\
&\leq \|x_k - \gamma\xi_k F(y_k) - x^*\|^2 \\
&= \|x_k - x^*\|^2 - 2\gamma\xi_k\left\langle F\left(y_k\right), x_k - x^*\right\rangle + \gamma^2\xi_k^2\|F(y_k)\|^2.
\end{aligned} \tag{16}$$

By the monotonicity of mapping $F(\cdot)$, we have

$$\begin{aligned}
&\left\langle F\left(y_k\right), x_k - x^*\right\rangle \\
&\quad = \left\langle F\left(y_k\right), x_k - y_k\right\rangle + \left\langle F\left(y_k\right), y_k - x^*\right\rangle \\
&\quad \geq \left\langle F\left(y_k\right), x_k - y_k\right\rangle + \left\langle F\left(x^*\right), y_k - x^*\right\rangle \\
&\quad = \left\langle F\left(y_k\right), x_k - y_k\right\rangle.
\end{aligned} \tag{17}$$

Substituting (15) and (17) into (16), we have

$$\begin{aligned}
&\|x_{k+1} - x^*\|^2 \\
&\leq \|x_k - x^*\|^2 - 2\gamma\xi_k\left\langle F\left(y_k\right), x_k - y_k\right\rangle + \gamma^2\xi_k^2\|F\left(y_k\right)\|^2 \\
&= \|x_k - x^*\|^2 - \gamma\left(2-\gamma\right)\frac{\left\langle F\left(y_k\right), x_k - y_k\right\rangle^2}{\|F\left(y_k\right)\|^2} \\
&\leq \|x_k - x^*\|^2 - \gamma\left(2-\gamma\right)\frac{\sigma^2\alpha_k^2\|d_k\|^4}{\|F\left(y_k\right)\|^2},
\end{aligned} \tag{18}$$

which together with $\gamma \in (0, 2)$ indicates that, for all $k$,

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|, \tag{19}$$

which shows that the sequence $\{x_k\}$ is bounded. By (10), it holds that $\{d_k\}$ is bounded and so is $\{y_k\}$. Then, by the continuity of $F(\cdot)$, there exists a constant $M > 0$ such that $\|F(y_k)\| \leq M$ for all $k$. Therefore it follows from (18) that

$$\begin{aligned}
&\gamma\left(2-\gamma\right)\frac{\sigma^2}{M^2}\sum_{k=0}^{\infty}\alpha_k^2\|d_k\|^4 \\
&\quad \leq \sum_{k=0}^{\infty}\left(\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2\right) < \infty,
\end{aligned} \tag{20}$$

which implies that the assertion (14) holds. The proof is completed. □

Now, we prove the global convergence of Algorithm 1.

**Theorem 7.** *Suppose that the conditions in Lemma 6 hold. Then the sequence $\{x_k\}$ generated by Algorithm 1 globally converges to a solution of CES $(F, C)$.*

*Proof.* We consider the following two possible cases.

*Case 1.* Consider $\liminf_{k\to\infty}\|d_k\| = 0$. Thus, by (10), we have $\liminf_{k\to\infty}\|F(x_k)\| = 0$. This together with the continuity of $F(\cdot)$ implies that the sequence $\{x_k\}$ has some accumulation point $\bar{x}$ such that $F(\bar{x}) = 0$. From (19), it holds that $\{\|x_k - \bar{x}\|\}$ converges, and since $\bar{x}$ is an accumulation point of $\{x_k\}$, it must hold that $\{x_k\}$ converges to $\bar{x}$.

*Case 2.* Consider $\liminf_{k\to\infty}\|d_k\| > 0$. Then by (14), it follows that $\lim_{k\to\infty}\alpha_k = 0$. Therefore, from the line search (6), for sufficiently large $k$, we have

$$-\left\langle F\left(x_k + \beta_k\rho^{m_k-1}d_k\right), d_k\right\rangle < \sigma\|d_k\|^2. \tag{21}$$

Since $\{x_k\}$, $\{d_k\}$ are both bounded, we can choose a sequence and letting $k \to \infty$ in (21), we can obtain

$$-\left\langle F\left(\bar{x}\right), \bar{d}\right\rangle \leq \sigma\left\|\bar{d}\right\|^2, \tag{22}$$

where $\bar{x}$ and $\bar{d}$ are limit points of corresponding subsequences. On the other hand, by (10), we obtain

$$-\left\langle F\left(x_k\right), d_k\right\rangle = \|F(x_k)\|^2 \geq r\|d_k\|^2. \tag{23}$$

Letting $k \to \infty$ in the above inequality, we obtain

$$-\left\langle F\left(\bar{x}\right), \bar{d}\right\rangle \geq r\left\|\bar{d}\right\|^2. \tag{24}$$

Thus, by (22) and (24), we get $r \leq \sigma$, and this contradicts the fact that $r > \sigma > 0$. Therefore, $\liminf_{k\to\infty}\|d_k\| > 0$ does not hold. This completes the proof. □

## 3. Convergence Rate

By Theorem 7, we know that the sequence $\{x_k\}$ generated by Algorithm 1 converges to a solution of CES $(F, C)$. In what follows, we always assume that $x_k \to x^*$ as $k \to \infty$, where $x^* \in C^*$. To establish the local convergence rate of the sequence generated by Algorithm 1, we need the following assumption.

*Assumption 8.* For $x^* \in C^*$, there exist three positive constants $\delta$, $c$, and $L$ such that

$$c\,\mathrm{dist}\left(x, C^*\right) \leq \|F\left(x\right)\|, \quad \forall x \in N\left(x^*, \delta\right), \tag{25}$$

$$\|F\left(x\right) - F\left(y\right)\| \leq L\|x - y\|, \quad \forall x, y \in N\left(x^*, \delta\right), \tag{26}$$

where $\mathrm{dist}(x, C^*)$ denotes the distance from $x$ to the solution set $C^*$, and

$$N\left(x^*, \delta\right) = \left\{x \in R^n \mid \|x - x^*\| \leq \delta\right\}. \tag{27}$$

Now, we analyze the convergence rate of the sequence $\{x_k\}$ generated by Algorithm 1 under conditions (25) and (26).

**Lemma 9.** *If the conditions in Assumption 8 hold, then the sequence* $\{\alpha_k\}$ *generated by line search* (6) *has a positive bound from below.*

*Proof.* We only need to prove that for sufficiently large $k$, $\alpha_k$ has a positive bound from below. If $\alpha_k \leq \beta_k$, then by the construction of $\alpha_k$, we have

$$-\left\langle F\left(x_k+\beta_k\alpha_k\rho^{-1}d_k\right),d_k\right\rangle<\sigma\|d_k\|^2. \tag{28}$$

In addition, by (10), we have

$$-\langle F(x_k),d_k\rangle=\|F(x_k)\|^2\geq r\|d_k\|^2. \tag{29}$$

Then, by the above two inequalities, we can obtain

$$\left\langle F\left(x_k+\beta_k\alpha_k\rho^{-1}d_k\right)-F(x_k),d_k\right\rangle\geq(r-\sigma)\|d_k\|^2. \tag{30}$$

On the other hand, from (26), we have

$$\left\langle F\left(x_k+\beta_k\alpha_k\rho^{-1}d_k\right)-F(x_k),d_k\right\rangle\leq\frac{L\beta_k\alpha_k}{\rho}\|d_k\|^2. \tag{31}$$

By (30) and (31), for $k$ sufficiently large we obtain

$$\alpha_k\geq\frac{\rho(r-\sigma)}{L\beta_k}\geq\frac{\rho(r-\sigma)}{L\beta_{\max}}. \tag{32}$$

Therefore, there is a positive constant $\alpha$, such that

$$\alpha_k\geq\alpha, \tag{33}$$

for all $k$. The proof is completed. □

**Theorem 10.** *In addition to the assumptions in Theorem 7, if conditions* (25) *and* (26) *hold, then the sequence* $\{\text{dist}(x_k,C^*)\}$ *Q-linearly converges to 0; hence the whole sequence* $\{x_k\}$ *converges to* $x^*$ *R-linearly.*

*Proof.* Let $z_k\in C^*$ be the closest solution to $x_k$. That is, $\|x_k-z_k\|=\text{dist}(x_k,C^*)$. By (18), we have

$$\|x_{k+1}-z_k\|^2\leq\|x_k-z_k\|^2-\gamma(2-\gamma)\frac{\langle F(y_k),x_k-y_k\rangle^2}{\|F(y_k)\|^2}. \tag{34}$$

For sufficiently large $k$, it follows from (10) and (26) that

$$\begin{aligned}\|F(y_k)\|&=\|F(y_k)-F(z_k)\|\\&\leq L\|y_k-z_k\|\\&\leq L(\|x_k-y_k\|+\|x_k-z_k\|)\\&\leq L(\beta_{\max}\|d_k\|+\|x_k-z_k\|)\\&\leq L\left(\frac{\beta_{\max}\|F(x_k)\|}{r}+\|x_k-z_k\|\right)\\&=L\left(\frac{\beta_{\max}\|F(x_k)-F(z_k)\|}{r}+\|x_k-z_k\|\right)\\&\leq L\left(\frac{\beta_{\max}L}{r}+1\right)\|x_k-z_k\|\\&=L\left(\frac{\beta_{\max}L}{r}+1\right)\text{dist}(x_k,C^*).\end{aligned} \tag{35}$$

Thus, from (6), (10), (25), and (33), for sufficiently large $k$, it holds that

$$\begin{aligned}\langle F(y_k),x_k-y_k\rangle&\geq\sigma\alpha_k\|d_k\|^2\geq\sigma\alpha\|d_k\|^2\\&\geq\sigma\alpha\|F(x_k)\|^2\geq\sigma\alpha c^2\text{dist}^2(x_k,C^*).\end{aligned} \tag{36}$$

Substituting the above two inequalities into (34), we have

$$\begin{aligned}\text{dist}^2(x_{k+1},C^*)&\leq\|x_{k+1}-z_k\|^2\\&\leq\left(1-\frac{\sigma\alpha r^2c^2\gamma(2-\gamma)}{L^2(\beta_{\max}L+r)^2}\right)\text{dist}^2(x_k,C^*),\end{aligned} \tag{37}$$

which implies that the sequence $\{\text{dist}(x_k,C^*)\}$Q-linearly converges to 0. Therefore, the whole sequence $\{x_k\}$ converges to $x^*$ $R$-linearly. The proof is completed. □

## 4. Numerical Results

In this section, we test Algorithm 1 and compared it with the projection method in [7] and the spectral gradient projection method in [9]. We give the following two simple problems to test the efficiency of the three methods.

*Problem 11.* The mapping $F(\cdot)$ is taken as $F(x)=(f_1(x), f_2(x),\ldots,f_n(x))^\top$, where

$$f_i(x)=e^{x_i}-1,\quad\text{for } i=1,2,\ldots,n \tag{38}$$

and $C=R^n_+$. Obviously, this problem has a unique solution $x^*=(0,0,\ldots,0)^\top$.

*Problem 12.* The mapping $F(\cdot)$ is taken as $F(x)=(f_1(x), f_2(x),\ldots,f_n(x))^\top$, where

$$f_i(x)=x_i-\sin|x_i-1|,\quad\text{for } i=1,2,\ldots,n \tag{39}$$

and $C=\{x\in R^n_+\mid\sum_{i=1}^n x_i\leq n, x_i\geq-1, i=1,2,\ldots,n\}$. Obviously, Problem 12 is nonsmooth at $x=(1,1,\ldots,1)^\top$.

The codes are written in Mablab7.0 and run on a personal computer with 2.0 GHZ CPU processor. The parameters used in Algorithm 1 are set as $\rho=0.6$, $r=10^{-4}$, $\sigma=5\times10^{-5}$, and $\gamma=1.65$. The initial steplength in Step 2 of Algorithm 1 is set to be the spectral coefficient

$$\beta_{k+1}=\frac{s_k^\top s_k}{s_k^\top z_k}, \tag{40}$$

where $s_k=x_{k+1}-x_k$ and $z_k=F(x_{k+1})-F(x_k)+0.01(x_{k+1}-x_k)$. By the monotonicity and the Lipschitz continuity of $F(\cdot)$, it is not difficult to show that

$$0.01s_k^\top s_k\leq s_k^\top z_k\leq(L+0.01)s_k^\top s_k, \tag{41}$$

where $L$ is the Lipschitz constant. If $\beta_k\notin[\beta_{\min},\beta_{\max}]$, we replace the spectral coefficient by

$$\beta_k=\begin{cases}1, & \text{if } \|F(x_k)\|\geq1,\\ \|F(x_k)\|^{-1}, & \text{if } 10^{-5}\leq\|F(x_k)\|\leq1,\\ 10^5, & \text{if } \|F(x_k)\|\leq10^{-5},\end{cases} \tag{42}$$

Table 1: Numerical results with different dimensions of Problem 11.

| Dimension | Method | Iter. | Fn | CPU |
|---|---|---|---|---|
| 50 | Algorithm 1 | 1 | 5 | 0.0000 |
| | WPM | 614 | 3703 | 0.5313 |
| | YSGP | 6 | 15 | 0.0000 |
| 500 | Algorithm 1 | 1 | 5 | 0.0156 |
| | WPM | 657 | 3959 | 1.1406 |
| | YSGP | 6 | 15 | 0.0156 |
| 5000 | Algorithm 1 | 1 | 5 | 0.4219 |
| | WPM | 692 | 4157 | 8.4063 |
| | YSGP | 7 | 17 | 0.4375 |
| 50000 | Algorithm 1 | 1 | 5 | 52.4688 |
| | WPM | 724 | 4336 | 148.9219 |
| | YSGP | 7 | 17 | 53.6719 |

Table 2: Numerical results with different initial points of Problem 12 with $n = 64$.

| Initial point | Method | Iter. | Fn | CPU |
|---|---|---|---|---|
| $(1, 1, \ldots, 1)$ | Algorithm 1 | 10 | 115 | 0.3438 |
| | WPM | 330 | 1981 | 13.1094 |
| | YSGP | 16 | 49 | 0.6094 |
| $(2, 2, \ldots, 2)$ | Algorithm 1 | 9 | 91 | 0.2031 |
| | WPM | 331 | 1985 | 11.7969 |
| | YSGP | 16 | 47 | 0.6875 |
| $(3, 3, \ldots, 3)$ | Algorithm 1 | 7 | 88 | 0.2031 |
| | WPM | 331 | 1987 | 13.5781 |
| | YSGP | 16 | 47 | 0.6250 |
| $(4, 4, \ldots, 4)$ | Algorithm 1 | 11 | 117 | 0.2500 |
| | WPM | 331 | 1987 | 11.5469 |
| | YSGP | 16 | 48 | 0.6094 |
| $(5, 5, \ldots, 5)$ | Algorithm 1 | 9 | 79 | 0.3281 |
| | WPM | 331 | 1987 | 11.2031 |
| | YSGP | 16 | 48 | 0.6719 |

where $\beta_{\min} = 10^{-10}$ and $\beta_{\max} = 10^{10}$. This parabolic model is the same as the one described in [15]. We stop the iteration if the iteration number exceeds 1000 or the inequality $\|F(x_k)\| \leq 10^{-6}$ is satisfied. The method in [7] (denoted by WPM) is implemented with the following parameters: $G_k \equiv 0$, $\sigma = 0$, $\lambda = 0.95$, $\beta = 0.5$, and $\mu_k \equiv 2.5$. The method in [9] (denoted by YSGP) is implemented with the following parameters: $\beta = 0.5$, $\sigma = 0.01$, and $r = 0.001$.

For Problem 11, the initial point is set as $x_0 = (1, 1, \ldots, 1)$, and Table 1 gives the numerical results by Algorithm 1, WPM, and YSGP with different dimensions, where Iter. denotes the iteration number, Fn denotes the number of function evaluations, and CPU denotes the CPU time in seconds when the algorithm terminates. Table 2 lists the numerical results of Problem 12 with different initial points. The numerical results given in Tables 1 and 2 show that Algorithm 1 performs a little better than YSGP in [9] and obviously better than WPM in [7], since it requires much lower number of iterations or less CPU time than WPM in [7] and a little lower number of iterations or less CPU time than YSGP in [9]. So the proposed method is promising.

## 5. Conclusions

A globally convergent matrix-free method to solve constrained equations has been developed, which is not only derivative-free but also completely matrix-free. Consequently, it can be applied to solve large-scale nonsmooth constrained equations. We established the global convergence without the requirement of differentiability of the equations and presented the linear convergence rate under standard conditions. We also report some numerical results to show the efficiency of the proposed method.

Numerical results indicate that the parameters $r$ and $\gamma$ influence the performance of the method, so the choice of the positive constants $r$ and $\lambda$ is our future work.

## Conflict of Interests

All the authors of the paper declare that they do not have any conflict of interests, and there are no financial or personal relationships with other people or organizations that can inappropriately influence our work in this paper.

## Acknowledgments

The authors would like to thank the referees for giving them many valuable suggestions and comments, which improve this paper greatly. This work is supported by the Nature Science Foundation of Shandong Province (ZR2012AL08).

## References

[1] E. Zeidler, *Nonlinear Functional Analysis and Its Applications*, Springer, 1990.

[2] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, John Wiley and Sons, New York, NY, USA, 1996.

[3] K. Meintjes and A. P. Morgan, "A methodology for solving chemical equilibrium systems," *Applied Mathematics and Computation*, vol. 22, no. 4, pp. 333–361, 1987.

[4] L. Qi, X. J. Tong, and D. H. Li, "Active-set projected trust-region algorithm for box-constrained nonsmooth equations," *Journal of Optimization Theory and Applications*, vol. 120, no. 3, pp. 601–625, 2004.

[5] M. Ulbrich, "Nonmonotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems," *SIAM Journal on Optimization*, vol. 11, no. 4, pp. 889–917, 2001.

[6] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY, USA, 1970.

[7] C. Wang, Y. Wang, and C. Xu, "A projection method for a system of nonlinear monotone equations with convex constraints," *Mathematical Methods of Operations Research*, vol. 66, no. 1, pp. 33–46, 2007.

[8] F. Ma and C. Wang, "Modified projection method for solving a system of monotone equations with convex constraints," *Journal of Applied Mathematics and Computing*, vol. 34, no. 1-2, pp. 47–56, 2010.

[9] Z. Yu, J. Lin, J. Sun, Y. Xiao, L. Liu, and Z. Li, "Spectral gradient projection method for monotone nonlinear equations with convex constraints," *Applied Numerical Mathematics*, vol. 59, no. 10, pp. 2416–2423, 2009.

[10] L. Zheng, "A new projection algorithm for solving a system of nonlinear equations with convex constraints," *Bulletin of the Korean Mathematical Society*, vol. 50, no. 3, pp. 823–832, 2013.

[11] W. La Cruz and M. Raydan, "Nonmonotone spectral methods for large-scale nonlinear systems," *Optimization Methods & Software*, vol. 18, no. 5, pp. 583–599, 2003.

[12] L. Zhang and W. Zhou, "Spectral gradient projection method for solving nonlinear monotone equations," *Journal of Computational and Applied Mathematics*, vol. 196, no. 2, pp. 478–484, 2006.

[13] W. Cheng, "A PRP type method for systems of monotone equations," *Mathematical and Computer Modelling*, vol. 50, no. 1-2, pp. 15–20, 2009.

[14] L. Zhang, W. Zhou, and D.-H. Li, "A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence," *IMA Journal of Numerical Analysis*, vol. 26, no. 4, pp. 629–640, 2006.

[15] W. La Cruz, J. M. Martínez, and M. Raydan, "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations," *Mathematics of Computation*, vol. 75, no. 255, pp. 1429–1448, 2006.