Research Article

A Scaled Conjugate Gradient Method for Solving Monotone Nonlinear Equations with Convex Constraints

Sheng Wang and Hongbo Guan

Hunan Institute of Technology, Hengyang, Hunan 421008, China

Correspondence should be addressed to Sheng Wang; wangshengmath@163.com

Based on the Scaled conjugate gradient (SCALCG) method presented by Andrei (2007) and the projection method presented by Solodov and Svaiter, we propose a SCALCG method for solving monotone nonlinear equations with convex constraints. SCALCG method can be regarded as a combination of conjugate gradient method and Newton-type method for solving unconstrained optimization problems. So, it has the advantages of the both methods. It is suitable for solving large-scale problems. So, it can be applied to solving large-scale monotone nonlinear equations with convex constraints. Under reasonable conditions, we prove its global convergence. We also do some numerical experiments show that the proposed method is efficient and promising.

1. Introduction

In this paper, we consider the following convex constrained monotone equations:

$$F(x) = 0, \quad x \in \Omega, \tag{1}$$

where $F : R^n \to R^n$ is a continuous and monotone function. The feasible region Ω is a nonempty closed convex set. Monotone means that

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

The algorithms of solving monotone nonlinear equations $F(x) = 0$ have strong relationship to algorithms of solving optimization problems. It's known that the function $f(x)$ is strictly function is equivalent to that the vector function ∇f is strictly monotone which means $(\nabla f(x) - \nabla f(y))^T (x - y) > 0$, and the definition of monotone nonlinear equations is same to this. The strictly convex function must exists unique minimum point, so the minimum point is a stable point of the convex functions, namely, the point which the gradient vector $\nabla f(x) = 0$. The monotone vector function $F(x) = 0$ can be seen as a gradient vector of some strictly convex function. There exists strictly convex function $f(x)$, satisfying $\nabla f(x) = F(x)$, Therefore, solving $\min f(x)$ is equivalent to solving $F(x) = 0$.

Nonlinear monotone equations arise in wide variety of applications, such as subproblems in the generalized proximal algorithms with Bergman distances [1]. In power engineering, the operations of a power system are described by a system of nonlinear equations, called the power flow equations, which are constrained by some operating constraints.

It has received much attention for the unconstrained nonlinear monotone equations [2–5]. Solodov and Svaiter [2] proposed a Newton-type method and a good property of the method is that the whole sequence of iterates converges to a solution of the system without any regularity assumptions. Under some weaker conditions, Zhou and Toh [4] showed that the Solodov and Svaiter's method is super linear convergence. Zhou and Li [5, 6] extended Solodov and Svaiter's projection method to the BFGS method and limited memory BFGS method. Zhang and Zhou [3] combined the spectral gradient method and the projection method of Solodov and Svaiter, proposed a spectral gradient projection method. Wang et al. [7] extended Solodov and Svaiter's projection method to solve monotone equations with convex constraints. Yu et al. [8] proposed a spectral gradient projection algorithm for monotone nonlinear equations with convex constraints by combining a modified spectral gradient method and the projection method. A good property of the method is that the linear system is not necessary at

each iteration. Xiao and Zhu [9] extended CG_DESCENT to solve large-scale nonlinear convex constrained monotone equations in compressive sensing by combining with the projection method of Solodov and Svaiter. At each iteration, the proposed method is not necessary to compute the Jacobian information or store any matrix.

This paper is organized as follows. In Section 2, we propose a SCALCG method for solving monotone nonlinear equation with convex constraints. Under reasonable conditions, we prove its global convergence in Section 3. In Section 4, we do some numerical experiments show that our method are efficient and promising.

2. The Method

In this section, we propose our method. At first, we simply review the SCALCG method presented by Andrei [10] for the following unconstrained optimization problems.

$$\min f(x), \quad x \in R^n, \tag{3}$$

where $F : R^n \rightarrow R$ is a continuously differentiable function, g_k is its gradient at point x_k.

The method of Andrei generate a sequence $\{x_k\}$ of approximations to the minimum x^* of f, in which

$$x_{k+1} = x_k + \alpha_k d_k, \tag{4}$$

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \theta_{k+1} \left(\frac{g_{k+1}^T s_k}{y_k^T s_k} \right) y_k$$
$$- \left[\left(1 + \theta_{k+1} \frac{y_k^T y_k}{y_k^T s_k} \right) \frac{g_{k+1}^T s_k}{y_k^T s_k} - \theta_{k+1} \frac{g_{k+1}^T y_k}{y_k^T s_k} \right] s_k, \tag{5}$$

where $\theta_{k+1} = s_k^T s_k / y_k^T s_k$.

Based on the SCALCG method, we now introduce our method for solving (1). Inspired by (5), we define d_k as

$$d_k = \begin{cases} -F_0, & k = 0, \\ -\theta_k F_k + \theta_k \left(\dfrac{F_k^T s_{k-1}}{y_{k-1}^T s_{k-1}} \right) y_{k-1} & \\ \quad - \left[\left(1 + \theta_k \dfrac{y_{k-1}^T y_{k-1}}{y_{k-1}^T s_{k-1}} \right) \dfrac{F_k^T s_{k-1}}{y_{k-1}^T s_{k-1}} \right. & \\ \quad \left. -\theta_k \dfrac{F_k^T y_{k-1}}{y_{k-1}^T s_{k-1}} \right] s_{k-1}, & k \geq 1, \end{cases} \tag{6}$$

where $y_k = \gamma_k + \lambda_k t_k \|F_k\| d_k$, $\gamma_k = F_{k+1} - F_k$, $\lambda_k = 1 + \|F_k\|^{-1} \max\{0, -\langle \gamma_k, t_k d_k \rangle / \|t_k d_k\|^2\}$, $s_k = z_k - x_k = t_k d_k$, t_k is a step length which will be defined later. The definition of y_k is similar to the one in [9].

Lemma 1. *Let $\{d_k\}$ be generated by (6), then for any k, we have*

$$F_k^T d_k < 0. \tag{7}$$

Proof. If $k = 0$, we have $F_0^T d_0 = -\|F_0\|^2 < 0$.

If $k \geq 1$, we obtain

$$F_k^T d_k = -\theta_k \|F_k\|^2 + \theta_k \left(\frac{F_k^T s_{k-1}}{y_{k-1}^T s_{k-1}} \right) F_k^T y_{k-1}$$
$$- \left[\left(1 + \theta_k \frac{\|y_{k-1}\|^2}{y_{k-1}^T s_{k-1}} \right) \frac{F_k^T s_{k-1}}{y_{k-1}^T s_{k-1}} \right.$$
$$\left. -\theta_k \frac{F_k^T y_{k-1}}{y_{k-1}^T s_{k-1}} \right] F_k^T s_{k-1}$$
$$= \frac{1}{\left(y_{k-1}^T s_{k-1} \right)^2} \left[-\theta_k \|F_k\|^2 \left(y_{k-1}^T s_{k-1} \right)^2 \right.$$
$$+ \theta_k F_k^T s_{k-1} F_k^T y_{k-1} y_{k-1}^T s_{k-1}$$
$$- F_k^T s_{k-1} F_k^T s_{k-1} y_{k-1}^T s_{k-1}$$
$$- \theta_k \|y_{k-1}\|^2 F_k^T s_{k-1} F_k^T s_{k-1}$$
$$\left. + \theta_k F_k^T y_{k-1} F_k^T s_{k-1} y_{k-1}^T s_{k-1} \right]$$
$$= \frac{1}{\left(y_{k-1}^T s_{k-1} \right)^2} \left[-\theta_k \|F_k\|^2 \left(y_{k-1}^T s_{k-1} \right)^2 \right.$$
$$- \theta_k \|y_{k-1}\|^2 \left(F_k^T s_{k-1} \right)^2$$
$$+ 2\theta_k F_k^T s_{k-1} F_k^T y_{k-1} y_{k-1}^T s_{k-1}$$
$$\left. - \left(F_k^T s_{k-1} \right)^2 y_{k-1}^T s_{k-1} \right], \tag{8}$$

where

$$2 F_k^T s_{k-1} F_k^T y_{k-1} y_{k-1}^T s_{k-1}$$
$$= 2 \left(y_{k-1}^T s_{k-1} \right) F_k^T \cdot \left(F_k^T s_{k-1} \right) y_{k-1}$$
$$\leq \left\| y_{k-1}^T s_{k-1} F_k \right\|^2 + \left\| F_k^T s_{k-1} y_{k-1} \right\|^2$$
$$= \left(y_{k-1}^T s_{k-1} \right)^2 \|F_k\|^2 + \left(F_k^T s_{k-1} \right)^2 \|y_{k-1}\|^2. \tag{9}$$

So, we have

$$F_k^T d_k \leq -\frac{\left(F_k^T s_{k-1} \right)^2}{y_{k-1}^T s_{k-1}}. \tag{10}$$

By the definition of λ_k, the following inequality holds

$$\lambda_k \geq 1 - \|F_k\|^{-1} \frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2}. \tag{11}$$

So, we obtain

$$y_k^T s_k = \langle \gamma_k + \lambda_k \|F_k\| s_k, s_k \rangle = \langle \gamma_k, s_k \rangle + \lambda_k \|F_k\| \|s_k\|^2$$
$$\geq \langle \gamma_k, s_k \rangle + \|F_k\| \|s_k\|^2 - \langle \gamma_k, s_k \rangle = \|F_k\| \|s_k\|^2 > 0. \tag{12}$$

It can be seen that

$$F_k^T d_k \leq -\frac{\left(F_k^T s_{k-1}\right)^2}{y_{k-1}^T s_{k-1}} < 0. \tag{13}$$

□

The steps of our method are stated as follows.

Algorithm 2. Consider the following steps.

Step 0. Choose an initial point $x_0 \in \Omega$, and constants $\varepsilon > 0$, $\sigma \in (0,1)$, $\rho \in (0,1)$, $\xi > 0$ Set $k := 0$.

Step 1. Stop if $\|F_k\| \leq \varepsilon$. Otherwise, compute d_k by (6).

Step 2. Let $t_k = \max\{\xi\rho^i : i = 0, 1, 2, \ldots\}$ which satisfies

$$-\langle F(x_k + t_k d_k), d_k \rangle \geq \sigma t_k \|d_k\|^2. \tag{14}$$

Let $z_k = x_k + t_k d_k$.

Step 3. Compute

$$x_{k+1} = P_\Omega [x_k - \alpha_k F(z_k)], \tag{15}$$

where

$$\alpha_k = \frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2}. \tag{16}$$

Step 4. Let $k := k + 1$. Go to Step 1.

3. Convergence Analysis

In this section, we establish the global convergence of Algorithm 2. For our purpose, we assume that F satisfies the following assumptions.

Condition A. Consider the following.

(1) The mapping F is Lipchitz continuous, it means that it satisfies

$$\|F(x) - F(y)\| \leq L \|x - y\|, \quad \forall x, y \in \Omega. \tag{17}$$

(2) The solution set of (1), denoted by S, is nonempty.

Lemma 3. *Algorithm 2 is well defined.*

Proof. We just need prove that Step 2 is well defined in Algorithm 2. We take the limit of the both sides of (14), we have

$$\lim_{t_k \to 0} -\langle F(x_k + t_k d_k), d_k \rangle = \lim_{t_k \to 0} -F_k^T d_k > 0,$$
$$\lim_{t_k \to 0} \sigma t_k \|d_k\|^2 = 0. \tag{18}$$

So Algorithm 2 is well defined. □

Lemma 4. *Suppose Condition A hold, the step length t_k satisfies*

$$t_k \geq \min\left\{\xi, \frac{\rho\delta}{(L+\sigma)} \frac{\|F_k\|^2}{\|d_k\|^2}\right\}. \tag{19}$$

Proof. If the algorithm stops at some iteration k then $\|F_k\| = 0$, so that x_k is a solution of (1). From now on, we assume that $F_k \neq 0$ for any k. It is easy to see that $d_k \neq 0$ from (7).

If $t_k \neq \xi$, by the line search process, we know that $t_k' = \rho^{-1} t_k$ does not satisfies (14), that is

$$-\langle F(z_k'), d_k \rangle \leq \sigma t_k' \|d_k\|^2, \tag{20}$$

where $z_k' = x_k + t_k' d_k$.

From (7), we know

$$F_k^T d_k < 0, \tag{21}$$

So, for any $k \geq 0$, there exists a positive number $\delta > 0$, such that $F_k^T d_k \leq -\delta\|F_k\|^2$.

From (7) and condition (1), we have

$$\begin{aligned}
\delta\|F_k\|^2 &\leq -F_k^T d_k = \langle -F(x_k), d_k \rangle \\
&= \langle F(z_k') - F(x_k), d_k \rangle - \langle F(z_k'), d_k \rangle \\
&\leq L \|z_k' - x_k\| \|d_k\| + \sigma t_k' \|d_k\|^2 \\
&= L t_k' \|d_k\|^2 + \sigma t_k' \|d_k\|^2 \\
&= \rho^{-1}(L + \sigma) t_k \|d_k\|^2.
\end{aligned} \tag{22}$$

So we get

$$t_k \geq \frac{\rho\delta}{(L+\sigma)} \frac{\|F_k\|^2}{\|d_k\|^2},$$
$$t_k \geq \min\left\{\xi, \frac{\rho\delta}{(L+\sigma)} \frac{\|F_k\|^2}{\|d_k\|^2}\right\}. \tag{23}$$

□

Lemma 5. *Suppose Condition A hold and $\overline{x} \in S$, the sequence $\{x_k\}$ is generated by Algorithm 2. Then the sequence $\{\|F_k\|\}$ is bounded. That means for all $k \geq 0$, there exists a positive $M > 0$, such that*

$$\|F(x_k)\| \leq M. \tag{24}$$

Proof. From (2), we have

$$\begin{aligned}
\langle F(z_k), x_k - \overline{x} \rangle &= \langle F(z_k), x_k - z_k + z_k - \overline{x} \rangle \\
&= \langle F(z_k), x_k - z_k \rangle + \langle F(z_k), z_k - \overline{x} \rangle \\
&= \langle F(z_k), x_k - z_k \rangle \\
&\quad + \langle F(z_k) - F(\overline{x}), z_k - \overline{x} \rangle \\
&> \langle F(z_k), x_k - z_k \rangle.
\end{aligned} \tag{25}$$

From the non-expansiveness of the projection operator, it holds

$$
\begin{aligned}
\|x_{k+1} - \overline{x}\|^2 &= \|P_\Omega \left[x_k - \alpha_k F(z_k) \right] - \overline{x}\|^2 \\
&= \|P_\Omega \left[x_k - \alpha_k F(z_k) \right] - P_\Omega(\overline{x})\|^2 \\
&\le \|x_k - \alpha_k F(z_k) - \overline{x}\|^2 \\
&= \|x_k - \overline{x}\|^2 - 2\alpha_k \langle F(z_k), x_k - \overline{x} \rangle \\
&\quad + \alpha_k^2 \|F(z_k)\|^2 \\
&\le \|x_k - \overline{x}\|^2 - 2\alpha_k \langle F(z_k), x_k - z_k \rangle \\
&\quad + \alpha_k^2 \|F(z_k)\|^2 \\
&= \|x_k - \overline{x}\|^2 - 2\frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2} \\
&\quad \times \langle F(z_k), x_k - z_k \rangle \\
&\quad + \frac{\langle F(z_k), x_k - z_k \rangle^2}{\|F(z_k)\|^4} \|F(z_k)\|^2 \\
&= \|x_k - \overline{x}\|^2 - \frac{\langle F(z_k), x_k - z_k \rangle^2}{\|F(z_k)\|^2} \\
&\le \|x_k - \overline{x}\|^2.
\end{aligned}
\tag{26}
$$

It is easy to see

$$
\begin{aligned}
\|x_{k+1} - \overline{x}\|^2 &\le \|x_k - \overline{x}\|^2 \le \|x_{k-1} - \overline{x}\|^2 \\
&\le \|x_{k-2} - \overline{x}\|^2 \le \cdots \le \|x_0 - \overline{x}\|^2.
\end{aligned}
\tag{27}
$$

Since $F(x)$ is Lipchitz continuous, we get

$$
\|F(x_k)\| = \|F(x_k) - F(\overline{x})\| \le L\|x_k - \overline{x}\| \le L\|x_0 - \overline{x}\|.
\tag{28}
$$

Let $M = L\|x_0 - \overline{x}\|$, then (46) is established. □

Lemma 6. *Suppose Condition A hold, and the sequence $\{x_k\}$ and $\{z_k\}$ are generated by Algorithm 2. Then, $-F(z_k)$ is a decent direction of the function $(1/2)\|x - \overline{x}\|^2$ at the point x_k, where $\overline{x} \in S$.*

Proof. The gradient of the function $(1/2)\|x - \overline{x}\|^2$ is $g_k = x_k - \overline{x}$.

From (2), it can be seen that

$$
\begin{aligned}
\langle F(z_k), x_k - \overline{x} \rangle &= \langle F(z_k), x_k - z_k + z_k - \overline{x} \rangle \\
&= \langle F(z_k), x_k - z_k \rangle + \langle F(z_k), z_k - \overline{x} \rangle \\
&= \langle F(z_k), x_k - z_k \rangle \\
&\quad + \langle F(z_k) - F(\overline{x}), z_k - \overline{x} \rangle \\
&> 0.
\end{aligned}
\tag{29}
$$

So, we obtain

$$
\langle -F(z_k), x_k - \overline{x} \rangle < 0.
\tag{30}
$$

□

Lemma 7. *Suppose Condition A hold, and the sequence $\{x_k\}$ and $\{z_k\}$ are generated by Algorithm 2. Then we have the following:*

(1) $\{x_k\}$ *and* $\{z_k\}$ *are bounded.*

(2) $\lim_{k \to \infty}(x_k - z_k) = 0$.

Particularly, we have

$$
\lim_{k \to \infty} t_k \|d_k\| = 0.
\tag{31}
$$

(3) $\lim_{k \to \infty}(x_k - x_{k+1}) = 0$.

Proof. (1) From (26), we have

$$
\begin{aligned}
\|x_{k+1} - \overline{x}\|^2 &\le \|x_k - \overline{x}\|^2 \le \|x_{k-1} - \overline{x}\|^2 \\
&\le \|x_{k-2} - \overline{x}\|^2 \le \cdots \le \|x_0 - \overline{x}\|^2.
\end{aligned}
\tag{32}
$$

So the sequence $\{x_k\}$ is bounded.

From (2), (14), and (24), we get

$$
\begin{aligned}
\langle F(z_k), x_k - z_k \rangle &= \langle F(z_k), -t_k d_k \rangle = -t_k \langle F(z_k), d_k \rangle \\
&\ge \sigma t_k^2 \|d_k\|^2 = \sigma \|x_k - z_k\|^2, \\
\langle F(z_k), x_k - z_k \rangle &= \langle F(z_k) - F(x_k), x_k - z_k \rangle \\
&\quad + \langle F(x_k), x_k - z_k \rangle \\
&\le \|F(x_k)\| \|x_k - z_k\| \le M \|x_k - z_k\|.
\end{aligned}
\tag{33}
$$

So, the following inequality holds

$$
\sigma \|x_k - z_k\|^2 \le \|F(x_k)\| \|x_k - z_k\|.
\tag{34}
$$

That is,

$$
\|x_k - z_k\| \le \frac{M}{\sigma}.
\tag{35}
$$

So, the sequence $\{z_k\}$ is bounded. □

(2) From (26), we obtain

$$
\begin{aligned}
\|x_{k+1} - \overline{x}\|^2 &\le \|x_k - \overline{x}\|^2 - \frac{\langle F(z_k), x_k - z_k \rangle^2}{\|F(z_k)\|^2} \\
&\le \|x_k - \overline{x}\|^2 - \frac{\sigma^2 \|x_k - z_k\|^4}{\|F(z_k)\|^2}.
\end{aligned}
\tag{36}
$$

Since the function $F(x)$ is continuous, and the sequence $\{z_k\}$ is bounded, so the sequence $\|F(z_k)\|$ is bounded, that is

for all $k \geq 0$, that exists a positive $M_1 > 0$, such that $\|F(z_k)\| \leq M_1$. Then, we get

$$\|x_k - z_k\|^4 \leq \frac{M_1^2}{\sigma^2} \left(\|x_k - \overline{x}\|^2 - \|x_{k+1} - \overline{x}\|^2 \right),$$

$$\sum_{k=0}^{\infty} \|x_k - z_k\|^4 \leq \sum_{k=0}^{\infty} \frac{M_1^2}{\sigma^2} \left(\|x_k - \overline{x}\|^2 - \|x_{k+1} - \overline{x}\|^2 \right) < +\infty. \tag{37}$$

So, we have

$$\lim_{k \to \infty} (x_k - z_k) = 0. \tag{38}$$

Particularly, we obtain

$$\lim_{k \to \infty} t_k \|d_k\| = \lim_{k \to \infty} \|x_k - z_k\| = 0. \tag{39}$$

(3) From the non-expansiveness of the projection operator, it holds

$$\begin{aligned} \|x_k - x_{k+1}\| &= \|x_k - P_{\Omega} (x_k - \alpha_k F(z_k))\| \\ &= \|P_{\Omega} (x_k) - P_{\Omega} (x_k - \alpha_k F(z_k))\| \\ &\leq \|x_k - (x_k - \alpha_k F(z_k))\| = \|\alpha_k F(z_k)\| \\ &= \left\| \frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2} F(z_k) \right\| \leq \|x_k - z_k\|. \end{aligned} \tag{40}$$

So, we obtain

$$\lim_{k \to \infty} \|x_k - x_{k+1}\| = 0. \tag{41}$$

Theorem 8. *Suppose Condition A hold, and the sequence $\{x_k\}$ is generated by Algorithm 2. Then, we have*

$$\lim_{k \to \infty} \inf \|F_k\| = 0. \tag{42}$$

Proof. If (42) does not hold, for any $k \geq 0$, there exist $\varepsilon > 0$, such that

$$\|F_k\| \geq \varepsilon. \tag{43}$$

From the nonexpansiveness of the projection operator, it holds

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|P_{\Omega} [x_k - \alpha_k F(z_k)] - x_k\| \\ &\leq \|(x_k - \alpha_k F(z_k)) - x_k\| \\ &= \alpha_k \|F(z_k)\|. \end{aligned} \tag{44}$$

By the definition of α_k and Cauchy-Schwartz inequality, we have

$$\begin{aligned} \|x_{k+1} - x_k\| &\leq \frac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2} \|F(z_k)\| \\ &\leq \frac{\|F(z_k)\| \|x_k - z_k\|}{\|F(z_k)\|^2} \|F(z_k)\| \\ &= \|x_k - z_k\| = t_k \|d_k\|. \end{aligned} \tag{45}$$

By the definition of y_k, assumption (1) and (45), we obtain

$$\begin{aligned} \|y_k\| &= \|\gamma_k + \lambda_k t_k \|F_k\| d_k\| \leq \|\gamma_k\| + \lambda_k \|F_k\| \|t_k d_k\| \\ &= \|\gamma_k\| \\ &\quad + \left(1 + \|F_k\|^{-1} \max \left\{ 0, -\frac{\langle \gamma_k, t_k d_k \rangle}{\|t_k d_k\|^2} \right\} \right) \|F_k\| \|t_k d_k\| \\ &\leq \|\gamma_k\| + \left(1 + \|F_k\|^{-1} \frac{|\langle \gamma_k, t_k d_k \rangle|}{\|t_k d_k\|^2} \right) \|F_k\| \|t_k d_k\| \\ &\leq \|\gamma_k\| + \left(1 + \|F_k\|^{-1} \frac{\|\gamma_k\| \|t_k d_k\|}{\|t_k d_k\|^2} \right) \|F_k\| \|t_k d_k\| \\ &= 2 \|\gamma_k\| + \|F_k\| \|t_k d_k\| \\ &= 2 \|F(x_{k+1}) - F(x_k)\| + \|F_k\| \|t_k d_k\| \\ &\leq 2L \|x_{k+1} - x_k\| + \|F_k\| \|t_k d_k\| \\ &\leq 2L t_k \|d_k\| + M t_k \|d_k\| = (2L + M) t_k \|d_k\|. \end{aligned} \tag{46}$$

From (12), we get

$$y_{k-1}^T s_{k-1} \geq \|F_{k-1}\| \|s_{k-1}\|^2 \geq \varepsilon \|s_{k-1}\|^2. \tag{47}$$

From (46), we have

$$\begin{aligned} \|y_{k-1}\| &\leq (2L + M) \|s_{k-1}\|, \\ \theta_k &= \frac{s_{k-1}^T s_{k-1}}{y_{k-1}^T s_{k-1}} \leq \frac{\|s_{k-1}\|^2}{\varepsilon \|s_{k-1}\|^2} = \frac{1}{\varepsilon}. \end{aligned} \tag{48}$$

From (7), we get

$$F_k d_k \leq -\delta \|F_k\|^2. \tag{49}$$

So, we obtain

$$\delta \|F_k\|^2 \leq -F_k d_k \leq \|F_k\| \|d_k\|. \tag{50}$$

That is,

$$\|d_k\| \geq \delta \|F_k\| \geq \delta \varepsilon. \tag{51}$$

From (6), (24), (43), and (47), we have

$$\begin{aligned} \|d_k\| &\leq \|\theta_k F_k\| + \|\theta_k\| \frac{\|F_k\| \|s_{k-1}\| \|y_{k-1}\|}{y_{k-1}^T s_{k-1}} \\ &\quad + \left[\left(1 + \theta_k \frac{\|y_{k-1}\|^2}{y_{k-1}^T s_{k-1}} \right) \frac{\|F_k\| \|s_{k-1}\|^2}{y_{k-1}^T s_{k-1}} \right. \\ &\quad \left. + \theta_k \frac{\|F_k\| \|y_{k-1}\| \|s_{k-1}\|}{y_{k-1}^T s_{k-1}} \right] \\ &\leq \frac{1}{\varepsilon} \|F_k\| + \frac{1}{\varepsilon} \frac{\|F_k\| \|s_{k-1}\| (2L + M) \|s_{k-1}\|}{\varepsilon \|s_{k-1}\|^2} \end{aligned}$$

TABLE 1: Test results for Problem 9 with given initial points.

Init	$(1, 1, \ldots, 1)^{\mathrm{T}}$			$(2, 2, \ldots, 2)^{\mathrm{T}}$		
Dim	Iter	Time	Fn	Iter	Time	Fn
100	53	0.041288	$9.490010e-6$	60	0.046265	$9.294530e-6$
500	117	0.168041	$9.291395e-6$	130	0.253245	$9.861726e-6$
1000	166	0.709354	$9.629278e-6$	184	0.663908	$9.984013e-6$
2000	237	4.222668	$9.743963e-6$	262	2.370548	$9.975847e-6$
5000	382	42.01209	$9.623651e-6$	421	21.02392	$9.696351e-6$
Init	$(10, 10, \ldots, 10)^{\mathrm{T}}$			$(1, 0, 1, 0, \ldots, 1, 0)^{\mathrm{T}}$		
Dim	Iter	Time	Fn	Iter	Time	Fn
100	77	0.060589	$7.784485e-6$	55	0.042016	$7.946500e-6$
500	155	0.279761	$9.519559e-6$	120	0.149927	$9.872650e-6$
1000	216	0.813048	$9.260490e-6$	171	0.594531	$9.730717e-6$
2000	303	2.733373	$9.558623e-6$	244	2.273943	$9.991830e-6$
5000	479	23.90998	$9.962504e-6$	393	19.72807	$9.983895e-6$

TABLE 2: Test results for Problem 10 with given initial points.

Init	$(1, 1, \ldots, 1)^{\mathrm{T}}$			$(2, 2, \ldots, 2)^{\mathrm{T}}$		
Dim	Iter	Time	Fn	Iter	Time	Fn
100	26	0.026912	$6.603881e-6$	32	0.025124	$6.915301e-6$
500	51	0.088748	$7.539315e-6$	63	0.149986	$8.376139e-6$
1000	69	0.2554104	$9.920450e-6$	86	0.356989	$9.752235e-6$
2000	96	0.986898	$9.891781e-6$	120	1.278378	$8.932150e-6$
5000	151	13.81959	$9.365707e-6$	187	20.48943	$9.698312e-6$
Init	$(10, 10, \ldots, 10)^{\mathrm{T}}$			$(1, 0, 1, 0, \ldots, 1, 0)^{\mathrm{T}}$		
Dim	Iter	Time	Fn	Iter	Time	Fn
100	68	0.054123	$8.040204e-6$	24	0.018994	$7.446192e-6$
500	140	0.241559	$9.126295e-6$	47	0.128392	$9.168601e-6$
1000	194	0.712503	$9.667015e-6$	65	0.282163	$8.583682e-6$
2000	271	2.927190	$9.540461e-6$	90	0.889466	$9.452969e-6$
5000	425	46.31353	$9.117053e-6$	141	7.183196	$9.881571e-6$

$$+ \left(1 + \frac{1}{\varepsilon} \frac{(2L+M)^2 \|s_{k-1}\|^2}{\varepsilon \|s_{k-1}\|^2} \right) \frac{\|F_k\| \|s_{k-1}\|^2}{\varepsilon \|s_{k-1}\|^2}$$

$$+ \frac{1}{\varepsilon} \frac{\|F_k\| (2L+M) \|s_{k-1}\|^2}{\varepsilon \|s_{k-1}\|^2}$$

$$= \frac{1}{\varepsilon} \|F_k\| + \frac{1}{\varepsilon^2} (2L+M) \|F_k\|$$

$$+ \left(1 + \frac{(2L+M)^2}{\varepsilon^2} \right) \frac{1}{\varepsilon} \|F_k\| + \frac{(2L+M)}{\varepsilon^2} \|F_k\|$$

$$\leq \left(\frac{2}{\varepsilon} + \frac{2(2L+M)}{\varepsilon^2} + \frac{(2L+M)^2}{\varepsilon^3} \right) M. \tag{52}$$

Let $(2/\varepsilon + 2(2L+M)/\varepsilon^2 + (2L+M)^2/\varepsilon^3)M = C$, then for all $k \geq 0$, we have

$$\|d_k\| \leq C. \tag{53}$$

From (19), (43), and (53), it can be seen that

$$t_k \|d_k\| \geq \min \left\{ \xi, \frac{\rho \delta}{L+\sigma} \frac{\|F_k\|^2}{\|d_k\|^2} \right\} \|d_k\|$$

$$= \min \left\{ \xi \|d_k\|, \frac{\rho \delta}{L+\sigma} \frac{\|F_k\|^2}{\|d_k\|} \right\} \tag{54}$$

$$\geq \min \left\{ \xi \delta \varepsilon, \frac{\rho \delta \varepsilon^2}{C(L+\sigma)} \right\}.$$

The last inequality yields a contradiction with (31), so (42) holds. □

4. Numerical Experiments

In this section, we do some numerical experiments to test the performance of Algorithm 2 on the following two problems. The algorithm was coded in Matlab and run on a personal

computer with a 2.3 GHZ CPU and 2 GB memory and Windows XP operating system.

For each test problem, the termination condition is

$$F\left(x_k\right) \leq 10^{-5}. \tag{55}$$

We set $\xi = 1$, $\rho = 0.1$, $\sigma = 0.0001$. We test both problems with the number of variables $n = 100, 500, 1000, 2000$, and 5000 and start form different initial points. The meaning of the columns in Tables 1 and 2 is stated as follows. "Dim" means the dimension of the problem, "Init" means the initials points, "Iter" means the number of iterations, "Time" stands for CPU time in seconds, and "Fn" stands for the final norm of equations.

Problem 9. The F is taken as $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^{\mathrm{T}}$, where

$$f_i\left(x\right) = e^{x_i} - 2, \quad i = 1, 2, \ldots, n, \ \Omega = R_+^n. \tag{56}$$

Problem 10. The F is taken as $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^{\mathrm{T}}$, where

$$f_i\left(x\right) = 2x_i - \sin\left(\left|x_i - 1\right|\right), \quad i = 1, 2, \ldots, n, \ \Omega = R_+^n. \tag{57}$$

Tables 1 and 2 show that our method is efficient. It is suitable for solving large-scale monotone equations with convex constraints.

5. Conclusions

In this paper, we have proposed a SCALCG method for solving nonlinear monotone equations with convex constraints. Under some wild conditions, we proved its global convergence.

Preliminary numerical experiments have illustrated that the proposed method works well for Problems 9 and 10.

Acknowledgment

References

[1] M. V. Solodov and A. N. Iusem, "Newton-type methods with generalized distances for constrained optimization," *Optimization*, vol. 41, no. 3, pp. 257–278, 1997.

[2] M. V. Solodov and B. F. Svaiter, "A globally convergent inexact Newton method for system of monotone equations," in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, pp. 355–369, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[3] 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.

[4] G. Zhou and K. C. Toh, "Superlinear convergence of a Newton-type algorithm for monotone equations," *Journal of Optimization Theory and Applications*, vol. 125, no. 1, pp. 205–221, 2005.

[5] W. Zhou and D. Li, "A globally convergent BFGS method for nonlinear monotone equations without any merit functions," *Mathematics of Computation*, vol. 77, no. 264, pp. 2231–2240, 2008.

[6] W. Zhou and D. Li, "Limited memory BFGS method for nonlinear monotone equations," *Journal of Computational Mathematics*, vol. 25, no. 1, pp. 89–96, 2007.

[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] 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.

[9] Y. H. Xiao and H. Zhu, "A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing," *Journal of Mathematical Analysis and Applications*, vol. 405, no. 1, pp. 310–319, 2013.

[10] N. Andrei, "A scaled BFGS preconditioned conjugate gradient algorithm for unconstrained optimization," *Applied Mathematics Letters*, vol. 20, no. 6, pp. 645–650, 2007.