International Journal of Pure and Applied Mathematics

Volume 28 No. 2 2006, 169-176

A NEW HYBRID PROJECTION ALGORITHM FOR SOLVING SYSTEMS OF EQUATIONS

Chuanwei Wang[1 §], Jingzhao Liu[2]

[1]College of Informational Science and Engineering
Shandong Agricultural University
Tai'an Shandong, 271018, P.R. CHINA
e-mail: tw516@163.com

[2] Editorial Office
Journal of Qufu Normal University
Qufu Shandong, 273100, P.R. CHINA
e-mail: qfsdxb@mail.qfnu.edu.cn

Abstract: Based on the Newton-type method for solving nonlinear systems of equations developed by Solodov and Svaiter (1998), we propose an improved version of the method by adjusting the projection region in this paper. Under standard assumptions, we show the global convergence of the proposed algorithm.

AMS Subject Classification: 34A34
Key Words: nonlinear equations, global convergence, projection

1. Introduction

For a continuous mapping $F : R^n \to R^n$, consider the problem of finding solutions of the nonlinear system of equations

$$F(x) = 0. \tag{1.1}$$

It is well known that if the mapping F is monotone, i.e.,

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

then the solution set of (1.1) is convex (see [5]). Such problem has many ap-

Received: October 14, 2005

§Correspondence author

plications in the nonlinear variational inequality problem, the nonlinear complementarity problem, as well as the maximal monotone operator problem (see [3]).

There are numerous existing methods for solving problem (1.1), among which the well-known Newton method may be the simplest one which possesses local quadratic convergence rate. However, its efficiency requires not only that the initial point is sufficiently close to the solution point, but also requires the nonsingularity of the Jacibian matrix of the underlying mapping F at the solution point (see [1]). To overcome the drawbacks, some researchers developed many variants of the Newton methods having global convergence (see [2], [4], etc.). Different kinds of the variants, the hybrid method developed in [5] received much attention due to its nice global convergence property. To make the next iterate closer to the solution set, and thus to have fast convergence rate, we proposed an improved algorithm based on modifying the projection region for solving (1.1). Under standard assumptions, we show the global convergence of the proposed algorithm.

In this paper, we assume (see [5]):

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

(A2) the underlying mapping F is monotone.

Since the projection operator is involved in our proposed method, we give the definition of the projection operator and its some properties in the end of this section.

Let Ω be a nonempty closed convex subset of R^n, the projection from R^n to Ω is defined as:

$$P_\Omega[x] = \arg\min\{\|y - x\| \mid y \in \Omega\}.$$

The mapping $P_\Omega : R^n \to \Omega$ is called projection operator. The following properties about the projection operator are well known (see [6]).

Lemma 1.1. *Let Ω be a closed convex subset of R^n. For any $x, y \in R^n$, and $z \in \Omega$, it holds that:*

(1) $\|P_\Omega[x] - P_\Omega[y]\| \leq \|x - y\|$;

(2) $\langle P_\Omega[x] - x, z - P_\Omega[x]\rangle \geq 0$;

(3) $\|P_\Omega[x] - z\|^2 \leq \|x - z\|^2 - \|P_\Omega[x] - x\|^2$.

2. The Improved Algorithm

First, we give the description of the improved algorithm for solving (1.1).

Algorithm 2.1. *Step 0.* Choose an arbitrary initial point $x^0 \in R^n$, parameters $\sigma \in [0,1)$, $\lambda, \beta \in (0,1), \bar{\mu} \in [1/(\lambda(1-\sigma)), \infty)$, and set $k := 0$.

Step 1. If $F(x^k) = 0$, then stop.

Step 2. Choose $\mu_k \in [1/(\lambda(1-\sigma)), \bar{\mu}]$, and a positive semidefinite matrix $G_k \in R^{n\times n}$, then find $\bar{x}^k \in R^n$ by solving the following system of linear equations with respect to $x \in R^n$:

$$F(x^k) + (G_k + \mu_k I)(x - x^k) = r^k \tag{2.1}$$

with

$$\|r^k\| \le \sigma \|x^k - \bar{x}^k\|. \tag{2.2}$$

Step 3. Find $y^k = x^k + t_k(\bar{x}^k - x^k)$ with

$$\langle F(y^k), x^k - \bar{x}^k \rangle \ge \lambda(1-\sigma)\mu_k \|x^k - \bar{x}^k\|^2, \tag{2.3}$$

where $t_k = \beta^{m_k}$ and m_k is the smallest nonnegative integer such that (2.3) holds.

Step 4. Compute x^{k+1} via

$$x^{k+1} = P_{H_k^1 \cap H_k^2}[x^k], \tag{2.4}$$

where

$$H_k^1 := \{x \in R^n \mid \langle F(y^k), x - y^k \rangle \le 0\},$$

and

$$H_k^2 := \{x \in R^n \mid \langle x - x^k, x^{k-1} - x^k \rangle \le 0\}.$$

Set $k := k+1$ and go to Step 1.

From the choices of μ_k and G_k, we know that the coefficient matrix of the linear system of equations (2.1) is positive definite, so $\bar{x}^k$ satisfying the accuracy criterion (2.2) exists and it is easy to obtain. Second, the linesearch step (2.3) is well defined as shown from the following lemma. The proof is similar as that in [5] and hence is omitted.

Lemma 2.1. *The linesearch step (2.3) is well defined.*

Now, we consider the projection step in the algorithm.

Lemma 2.2. *Under assumptions (A1) and (A2), the intersection of H_k^1 and H_k^2 is nonempty. Moreover, it holds that $S \subseteq H_k^1 \cap H_k^2$.*

Proof. See the proof of the Lemma 3.1 in [6]. □

From the analysis above, we conclude that the algorithm is well defined.

Remark 2.1. In Algorithm 2.1, we choose a positive constant $\sigma \in [0, 1)$ instead of the altering ρ_k in [5], and let μ_k possess upper and lower bound and be bounded away from zero, which are independent of x^k. So Algorithm 2.1 is easier to implement in practice than that in [5].

Remark 2.2. Just as pointed out in [5], the hyperplane

$$H_k := \{x \in R^n \mid \langle F(y^k), x - y^k \rangle = 0\}$$

strictly separates the current iterate x^k from S, the solution set of (1.1). Let

$$z^k = x^k - \frac{\langle F(y^k), x^k - y^k \rangle}{\|F(y^k)\|^2} F(y^k)$$

be the new iterate obtained by the algorithm in [5], it is easy to see that $\langle F(y^k), z^k - y^k \rangle = 0$, which means that $z^k = P_{H_k^1}[x^k]$, and hence the new iterate obtained by Algorithm 2.1 is closer to the solution set of (1.1) than that by the algorithm in [5] at each step theoretically.

3. The Global Convergence

In the following analysis of the convergence of Algorithm 2.1, we assume that Algorithm 2.1 generates an infinite sequence $\{x^k\}$.

Lemma 3.1. *Suppose that assumptions (A1) and (A2) hold, the sequences $\{x^k\}$ and $\{y^k\}$ are generated by Algorithm 2.1, and that for $\forall\ k \geq 0, x^k \notin S$. Then $-F(y^k)$ is a decent direction of the merit function $\frac{1}{2}\|x - x^*\|^2$ at the point x^k, where $x^* \in S$.*

Proof. From $x^k - y^k = t_k(x^k - \bar{x}^k)$ and (2.3), we have

$$\begin{aligned}\langle F(y^k), x^k - y^k \rangle &= t_k \langle F(y^k), x^k - \bar{x}^k \rangle \geq t_k \lambda (1 - \sigma) \mu_k \|x^k - \bar{x}^k\|^2 \\ &= \frac{\lambda(1-\sigma)\mu_k}{t_k} \|x^k - y^k\|^2 \geq \|x^k - y^k\|^2, \quad (3.1)\end{aligned}$$

where the last inequality follows from $\mu_k \geq \frac{1}{\lambda(1-\sigma)}$ and $t_k \leq 1$.

By the monotonicity of F and $x^* \in S$, it holds that

$$\begin{aligned}\langle F(y^k), x^k - x^* \rangle &= \langle F(y^k), x^k - y^k \rangle + \langle F(y^k), y^k - x^* \rangle \\ &\geq \langle F(y^k), x^k - y^k \rangle + \langle F(x^*), y^k - x^* \rangle = \langle F(y^k), x^k - y^k \rangle. \quad (3.2)\end{aligned}$$

From $x^k \notin S$, (3.1) and (3.2) we know that the assertion holds. ∎

Lemma 3.2. *Suppose that assumptions (A1) and (A2) hold and that the sequences $\{x^k\}$ and $\{y^k\}$ are generated by Algorithm 2.1. Then the following statements hold.*

(1) $\{x^k\}$ *and* $\{y^k\}$ *are bounded;*

(2) $\lim_{k\to\infty}(x^k - x^{k+1}) = 0$;

(3) $\lim_{k\to\infty}(x^k - y^k) = 0$.

Proof. (1) For any $x^* \in S$, from (2.4) and Lemma 1.1, we have that

$$\|x^{k+1} - x^*\| \leq \|x^k - x^*\|, \quad \forall k \geq 0\,, \tag{3.3}$$

which means that the sequence $\{x^k\}$ is bounded.

From the monotonicity of F and (3.1), using the Cauchy-Schwartz inequality, it holds that

$$\|F(x^k)\| \geq \|x^k - y^k\|.$$

By the boundedness of $\{x^k\}$ and the continuity of F, the sequences $\{y^k\}$ is also bounded. The first assertion holds.

(2) Since $x^{k+1} \in H^2_{k+1}$, it follows from Lemma 1.1 that

$$\|P_{H^2_{k+1}}[x^{k+1}] - P_{H^2_{k+1}}[x^*]\|^2 \leq \|x^k - x^*\|^2 - \|P_{H^2_{k+1}}[x^{k+1}] - x^k\|^2,$$

that is,

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2,$$

which means that the sequence $\{\|x^k - x^*\|\}$ is non-increasing. By the boundedness of $\{x^k\}$, thus we have that

$$\lim_{k\to\infty}(x^k - x^{k+1}) = 0.$$

(3) Since $\{y^k\}$ is bounded and F is continuous, there exists a positive constant M, such that $\|F(y^k)\| \leq M$ for all $k \geq 0$.

Since

$$P_{H^1_k}[x^k] = x^k - \frac{\langle F(y^k), x^k - y^k\rangle}{\|F(y^k)\|^2} F(y^k),$$

and $x^{k+1} \in H^1_k$, it follows from the definition of the projection operator that

$$\|x^k - x^{k+1}\| \geq \|x^k - P_{H^1_k}[x^k]\| = \frac{\langle F(y^k), x^k - y^k\rangle}{\|F(y^k)\|} \geq \frac{\|x^k - y^k\|^2}{M}, \tag{3.4}$$

where the second inequality follows from (3.1). Thus the last assertion holds from (3,4) and the second assertion. □

We are now in the position to present our main result.

Theorem 3.1. *Suppose that assumptions (A1) and (A2) hold. Then the sequence* $\{x^k\}$ *generated by Algorithm 2.1 converges globally to a solution of (1.1).*

Proof. By Lemma 3.2, $\{x^k\}$ is bounded, so it has at least one cluster point. Let $\hat{x}$ be a cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ be the corresponding subsequence converging to $\hat{x}$.

Since $y^k = x^k + t_k(\bar{x}^k - x^k)$, from Lemma 2.3, it holds that

$$\lim_{k\to\infty} t_k\|\bar{x}^k - x^k\| = \lim_{k\to\infty} \|x^k - y^k\| = 0. \tag{3.5}$$

Now, we complete our proof in the following two different cases.

If

$$\limsup_{j\to\infty} t_{k_j} > 0,$$

then from (3.5) we have that

$$\liminf_{j\to\infty} \|\bar{x}^{k_j} - x^{k_j}\| = \lim_{j\to\infty} \|x^{k_j} - y^{k_j}\| = 0.$$

Thus from (2.2) it holds that

$$\liminf_{j\to\infty} r^{k_j} = 0.$$

Since F is continuous and $\bar{x}^{k_j}$ is the solution of (2.1), taking limit along the subsequence, we have $F(\hat{x}) = 0$ from (2.1), which implies that $\hat{x}$ is a solution of (1.1). Replacing x^* with $\hat{x}$ because of the arbitrariness of $x^* \in S$ in (3.4), we have that

$$\|x^{k+1} - \hat{x}\| \leq \|x^k - \hat{x}\|,$$

which shows that the whole sequence $\{x^k\}$ converges to $\hat{x}$, a solution of (1.1).

Second, we consider the case that $\lim_{j\to\infty} t_{k_j} = 0$. In this case, we will also show that $\lim_{j\to\infty} \|\bar{x}^{k_j} - x^{k_j}\| = 0$. Suppose that $\|\bar{x}^{k_j} - x^{k_j}\| \geq d > 0$. Without loss of generality, we assume that $\{\bar{x}^{k_j}\}$ is also a convergent subsequence, and denote $\lim_{j\to\infty}(\bar{x}^{k_j} - x^{k_j}) = \hat{z}$, then $\|\hat{z}\| \geq d$.

By the choice of t_{k_j}, we know that

$$\langle F(x^{k_j} + \beta^{m_{k_j}-1}(\bar{x}^{k_j} - x^{k_j})), x^{k_j} - \bar{x}^{k_j}\rangle < \lambda(1-\sigma)\mu_{k_j}\|\bar{x}^{k_j} - x^{k_j}\|^2.$$

Since $\bar{x}^k$ is a solution of (2.1), thus for any $k \geq 0$, we have that

$$\begin{aligned}\langle F(x^k), x^k - \bar{x}^k\rangle &= (x^k - \bar{x}^k)^\top (G_k + \mu_k I)(x^k - \bar{x}^k) + (x^k - \bar{x}^k)^\top r^k \\ &\geq \mu_k \|x^k - \bar{x}^k\|^2 - \|r^k\| \cdot \|x^k - \bar{x}^k\| \geq (1-\sigma)\mu_k \|x^k - \bar{x}^k\|^2,\end{aligned}$$

where the first inequality follows from the Cauchy-Schartz inequality, the second one from (2.2) and $\mu_k \geq 1/(\lambda(1-\sigma)) > 1$.

The above two inequalities mean that (for sufficiently large j, $m_{k_j} > 1$)

$$\begin{aligned}(1-\sigma)\mu_{k_j}\|\bar{x}^{k_j} - x^{k_j}\|^2 + \langle F(x^{k_j} + \beta^{m_{k_j}-1}(\bar{x}^{k_j} - x^{k_j})) \\ - F(x^{k_j}), x^{k_j} - \bar{x}^{k_j}\rangle < \lambda(1-\sigma)\mu_{k_j}\|\bar{x}^{k_j} - x^{k_j}\|^2.\end{aligned}$$

Letting $j \to \infty$, and using the continuity of F again, we have

$$(1-\sigma)\|\hat{z}\|^2 \leq \lambda(1-\sigma)\|\hat{z}\|^2.$$

Since $\|\hat{z}\| \geq d > 0$, hence $\lambda \geq 1$, which contradicts with the fact that $\lambda \in (0,1)$. Therefore,

$$\lim_{j\to\infty} \|\bar{x}^{k_j} - x^{k_j}\| = 0.$$

Using the similar arguments as that for the first case, we conclude that the whole sequence $\{x^k\}$ converges to a solution of (1.1). □

Acknowledgments

This work was supported by the Natural Science Foundation of Shandong Province China (Y2003A02) and the National Natural Science Foundation of China (10231060).

References

[1] J.E. Dennis, R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, N.J. (1983).

[2] J.L. Nazareth, L. Qi, Globalization of Newton's method for solving nonlinear equations, *J. Numerical Algebra Appl.*, **90** (1996), 653-673.

[3] J.S. Pang, L. Qi, Nonsmooth equations: Motivation an algorithms, *SIAM J. Optimization*, **3** (1993), 443-465.

[4] L. Qi, J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming*, **58** (1993), 353-367.

[5] M.V. Solodov, B.F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, In: *Reformulation: Piecewise Smooth, Semismooth and Smoothing Methods* (Ed-s: M. Fukushima, L. Qi), Kluwer Academic Publishers (1998).

[6] Y.J. Wang, A new projection and contraction method for variational inequalities, *Journal of Pure Mathematics and Applications*, **13** (2002), 483-493.