International Journal of Pure and Applied Mathematics
Volume 79 No. 1 2012, 87-96
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

A MODIFIED PRP PROJECTION METHOD FOR NONLINEAR EQUATIONS WITH CONVEX CONSTRAINTS

Lian Zheng
Department of Mathematics and Computer Science
Yangtze Normal University
Fuling, Chongqing, 408100, P.R. CHINA

Abstract: We present a modification of PRP projection method for nonlinear equations with convex constraints, which combines elements of the modified PRP conjugate gradient method and the halfspace projection method. The main modification is to use a different formula for PRP parameter to obtain a new search direction and linearsearch procedure, and hence to construct a new class of hyperplanes which strictly separate the current iterate from the solution set. Our method is proved to be globally convergent under very mild assumptions.

AMS Subject Classification: 90C25, 90C30
Key Words: nonlinear equations, PRP conjugate gradient method, halfspace projection algorithm, global convergence

1. Introduction

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping and $C \subset \mathbb{R}^n$ be a nonempty closed convex set. The inner product and norm are denoted by $\langle \cdot, \cdot \rangle$ and $\|.\|$, respectively. Consider the problem of finding

$$x^* \in C \quad \text{such that} \quad F(x^*) = 0. \tag{1}$$

Received: May 14, 2012

Let S denote the solution set of problem (1). Throughout this paper, we assume that S is nonempty and F has the property that

$$\langle F(y), y - x^* \rangle \geq 0, \quad \text{for all } y \in C \text{ and all } x^* \in S. \tag{2}$$

The property (2) holds if F is monotone or more generally pseudomontone on C in the sense of Karamardian [1].

Nonlinear equations have wide applications in reality. In recent years, many numerical methods for problem (1) with smooth mapping F have been proposed. These methods include the Newton method, quasi-Newton method, Levenberg-Marquardt method, trust region methods and their variants, see [2-7]. They are popular because of their locally fast convergent rate. Recently, great efforts have been made to find a solution of nonsmooth systems of equations, see [8-11]. But the above-mentioned methods need to solve a linear system using the Jacobian matrix or an approximation of it, which is not suitable for large scale nonlinear equations.

Based on this fact that the well-known spectral gradient method and the conjugate gradient method are very efficient approaches for large-scale unconstrained optimization problems due to their simplicity and low storage, see [12-14]. Recently, they have been extended to solve large-scale nonlinear systems of equations, see [15-19]. They are proved to be globally convergent if F is monotone and Lipschitz continuous without any differentiability requirement. However, the monotonicity and Lipschitz continuity of the mapping F seems to be too stringent a requirement for the purpose of ensuring global convergence property of the proposed methods.

The purpose of this paper is to relax the assumptions of the monotonicity and Lipschitz continuity of F. We present a modification of PRP type method proposed by Cheng [17] for solving systems of equations. The main modification is to use a different PRP parameter β_k to obtain a new search direction and a new linearsearch procedure, and hence to construct a new class of hyperplanes. Our method possesses the following properties: (a) the proof of the global convergence need only F to be continuous and to satisfy condition (2); (b) the sequence of the distances from the iterates to the solution set of problem (1) is decreasing; (c) our hyperplane covers those of [17-19].

The remaining part of this paper is distributed as follows. In the next section, we give some preliminaries and the details of the algorithm. The global convergence analysis of the method is proved in Section 3.

2. Preliminaries and Algorithms

For a nonempty closed convex set $\Omega \subset \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, the projection of x onto Ω is defined as:

$$\Pi_\Omega(x) = \arg\min\{\|y - x\| \,|y \in \Omega\}.$$

We have the following properties on the projection operator, see [20].

Lemma 2.1. *Let $\Omega \subset \mathbb{R}^n$ be a closed convex set. Then it holds that*

$$\|x - \Pi_\Omega(y)\|^2 \leq \|x - y\|^2 - \|y - \Pi_\Omega(y)\|^2, \forall\, x \in \Omega,\, y \in \mathbb{R}^n.$$

Algorithm 2.1. Choose an intial point $x_0 \in C$, parameters $\beta, \xi \in (0, 1)$, $\sigma > 0$, $a_1, a_2 \geq 0$ and $a_1 \neq a_2$. Set $d_0 = -F(x_0)$ and $k := 0$.

Step 1. Find the trial point $z_k = x_k + \alpha_k d_k$, where $\alpha_k = \beta^{m_k}$ with m_k being the smallest nonnegative integer m such that

$$-\langle F(x_k + \beta^m d_k), d_k\rangle \geq t_2^{-1}\sigma\beta^m\|d_k\|^2\|t_1F(x_k) + t_2F(x_k + \beta^m d_k)\|, \quad (3)$$

where

$$t_1 = \frac{a_1}{\|F(x_k)\|},\ t_2 = \frac{a_2}{\|F(x_k + \beta^m d_k)\|}, \quad (4)$$

Stop if $\|F(z_k)\| = 0$.

Step 2. Compute the projection of x_k on C_k by

$$x_{k+1} = \Pi_{C_k}(x_k), \quad (5)$$

where $C_k := \{x \in C : h_k(x) \leq 0\}$ and

$$h_k(x) := \langle t_1F(x_k) + t_2F(z_k), x - z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle. \quad (6)$$

Stop if $\|F(x_{k+1})\| = 0$.

Step 3. Compute d_{k+1} by

$$d_{k+1} = -F(x_{k+1}) + \beta_k d_k, \quad (7)$$

where

$$\beta_k = \frac{\langle F(x_{k+1}), x_{k+1} - x_k\rangle}{\|F(x_k)\|^2}. \quad (8)$$

If $\langle d_{k+1}, F(x_{k+1})\rangle > -\xi\|F(x_{k+1})\|^2$, set $d_{k+1} = -F(x_{k+1})$.

Let k:=k+1, Go to Step 1.

Remark 2.1. Compared with the algorithm in [17]: The parameter β_k^{PRP} in [17] is gived by

$$\beta_k^{PRP} = \frac{\langle F(x_{k+1}), F(x_{k+1}) - F(x_k)\rangle}{\|F(x_k)\|^2},$$

which is different from β_k of Algorithm 2.1. In addition, our linearsearch procedure and hyperplane are also different from ones of [17]. When taking $a_1 = 0$, our hyperplane degrade into

$$\{x \in \mathbb{R}^n | \langle F(z_k), x - z_k\rangle = 0\},$$

which is used by [17-19].

Remark 2.2. Now we analyze the feasibility of Algorithm 2.1. It is obvious from Step 3 of Algorithm 2.1 that

$$-\langle F(x_k), d_k\rangle \geq \xi \|F(x_k)\|^2, \text{ for all } k. \tag{9}$$

Therefore the line search procedure, i.e., Step 1 of Algorithm 2.1, is well defined. In fact, we suppose that the conclusion doesn't hold, then there exists a nonnegative integer k_0 such that (3) is not satisfied for any nonnegative integer m, i.e.,

$$-\langle F(x_{k_0}+\beta^m d_{k_0}), d_{k_0}\rangle < t_2^{-1}\sigma\beta^m \|d_{k_0}\|^2 \|t_1 F(x_{k_0}) + t_2 F(x_{k_0}+\beta^m d_{k_0})\|, \text{ for all } m.$$

Letting $m \to \infty$, by the continuity of F and $\beta \in (0, 1)$, we obtain

$$-\langle F(x_{k_0}), d_{k_0}\rangle \leq 0.$$

Which contradict inequality (9). We obtain the desired conclusion. In addition, it is also easy to see from equation (4) that $t_1 F(x_k) + t_2 F(z_k) \neq 0$ because of $a_1 \neq a_2$, which means that Step 2 is also well defined. Consequently, our Algorithm 2.1 is well defined.

3. Convergence Analysis

In this section, we first prove two lemmas, and then analyze the global convergence of Algorithm 2.1.

Lemma 3.1. *Let $\bar{x}$ be a solution of problem (1), condition (2) holds and the function h_k be defined by equation (6). Then*

$$h_k(x_k) \geq \sigma\alpha_k^2\|d_k\|^2\|t_1F(x_k)+t_2F(z_k)\| \text{ and } h_k(\bar{x}) \leq 0. \tag{10}$$

In particular, if $d_k \neq 0$, then $h_k(x_k) > 0$.

Proof.

$$\begin{aligned} h_k(x_k) &= \langle t_1F(x_k)+t_2F(z_k), x_k - z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle \\ &= t_2\langle F(z_k), x_k - z_k\rangle \\ &= -\alpha_k t_2\langle F(z_k), d_k\rangle \\ &\geq \sigma\alpha_k^2\|d_k\|^2\|t_1F(x_k)+t_2F(z_k)\|, \end{aligned}$$

where the inequality follows from inequality (3).

$$\begin{aligned} h_k(\bar{x}) &= \langle t_1F(x_k)+t_2F(z_k), \bar{x} - z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle \\ &\leq t_1\langle F(x_k), \bar{x} - z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle \\ &= t_1\langle F(x_k), \bar{x} - x_k\rangle \\ &\leq 0, \end{aligned}$$

where the two inequalities follow from condition (2).

If $d_k \neq 0$, then $h_k(x_k) > 0$. The proof is completed.

Remark 3.1. Lemma 3.1 means that the hyperplane

$$H_k := \{x \in \mathbb{R}^n | \langle t_1F(x_k)+t_2F(z_k), x - z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle = 0\}$$

strictly separates the current iterate from the solution set of problem (1). Therefore, it is reasonable to let the next iterate x_{k+1} be the projection of x_k onto the halfspace C_k.

Remark 3.2. If Algorithm 2.1 generates an infinite sequence, then $F(x_k) \neq 0$ and $F(z_k) \neq 0, \forall k$. From inequality (9) and the Cauchy-Schwartz inequality, we obtain

$$\|d_k\| \geq \xi\|F(x_k)\| > 0. \tag{11}$$

Let $x^* \in S$. Since

$$\begin{aligned} \langle t_1F(x_k)+t_2F(z_k), x_k - x^*\rangle &= t_1\langle F(x_k), x_k - x^*\rangle \\ &\quad + t_2\langle F(z_k), x_k - x^*\rangle \\ &= t_1\langle F(x_k), x_k - x^*\rangle \end{aligned}$$

$$
\begin{aligned}
&+ t_2\langle F(z_k), x_k - z_k\rangle \\
&+ t_2\langle F(z_k), z_k - x^*\rangle \\
&\geq t_2\langle F(z_k), x_k - z_k\rangle \\
&\geq \sigma\alpha_k^2\|d^k\|^2\|t_1F(x_k) + t_2F(z_k)\| \\
&> 0,
\end{aligned}
$$

where the first inequality follows from condition (2), the second one follows from (3), and the last one follows inequality (11).

Which shows that $-(t_1F(x_k)+t_2F(y_k))$ is a descent direction of the function $\frac{1}{2}\|x - x^*\|^2$ at the point x_k.

Lemma 3.2. *Suppose that F is continuous and the sequence $\{x_k\}$ generated by Algorithm 2.1 is bounded. If there exists a constant $\varepsilon > 0$, such that $\|F(x_k)\| \geq \varepsilon$ for all k and*

$$lim_{k\to\infty}\alpha_k\|d_k\| = 0. \tag{12}$$

Then $\{d_k\}$ is also bounded.

Proof. Since F is continuous and the sequence $\{x_k\}$ is bounded, there exists a positive constant r such that

$$\|F(x_k)\| \leq r, \text{ for all } k. \tag{13}$$

From the definition of d_k, we obtain

$$\|d_k\| \leq \|F(x_k)\| + |\beta_{k-1}|\|d_{k-1}\| \tag{14}$$

By inequality (6) and inequality (9), we have

$$
\begin{aligned}
h_{k-1}(z_{k-1}) &= \langle t_1F(x_{k-1}) + t_2F(z_{k-1}), z_{k-1} - z_{k-1}\rangle + \alpha_{k-1}t_1\langle F(x_{k-1}), d_{k-1}\rangle \\
&= \alpha_{k-1}t_1\langle F(x_{k-1}), d_{k-1}\rangle \\
&\leq -\alpha_{k-1}t_1\xi\|F(x_{k-1})\|^2 < 0.
\end{aligned}
$$

It follows that $z_{k-1} \in C_{k-1}$. This together with (5) and Lemma 2.1 means

$$\|x_k - x_{k-1}\| \leq \|x_{k-1} - z_{k-1}\|, \text{ for all } k \tag{15}$$

From (12), (13), (15) and the condition $\|F(x_k)\| \geq \varepsilon$, we obtain

$$|\beta_{k-1}| \leq \frac{\|F(x_k)\|\|x_k - x_{k-1}\|}{\|F(x_{k-1})\|^2}$$

$$\begin{aligned} &\leq \frac{\|F(x_k)\|\|x_{k-1}-z_{k-1}\|}{\|F(x_{k-1})\|^2} \\ &= \frac{\|F(x_k)\|\alpha_{k-1}\|d_{k-1}\|}{\|F(x_{k-1})\|^2} \\ &\leq \frac{r}{\varepsilon^2}\alpha_{k-1}\|d_{k-1}\| \to 0, \text{ as } k\to\infty, \end{aligned}$$

which implies that there exists a constant $\varepsilon_0 \in (0,1)$ and an integer k_0 such that

$$|\beta_{k-1}| \leq \varepsilon_0, \text{ for all } k > k_0. \tag{16}$$

By (13),(14) and (16), we obtain

$$\begin{aligned} \|d_k\| &\leq r+\varepsilon_0\|d_{k-1}\| \leq r+\varepsilon_0 r+\varepsilon_0^2\|d_{k-2}\| \leq ... \\ &\leq r(1+\varepsilon_0+\varepsilon_0^2+...+\varepsilon_0^{k-k_0-1})+\varepsilon_0^{k-k_0}\|d_{k_0}\| \\ &\leq \frac{r}{1-\varepsilon_0}+\|d_{k_0}\|, \text{ for all } k>k_0. \end{aligned}$$

Taking $M = \max\{\|d_0\|, \|d_1\|, ..., \|d_{k_0}\|, \frac{r}{1-\varepsilon_0} + \|d_{k_0}\|\}$. We obtain the desired conclusion.

Now, we turn to consider the convergence of Algorithm 2.1. Certainly, if Algorithm 2.1 terminates at Step k, then $\|F(z_k)\| = 0$ or $\|F(x_k)\| = 0$. This means that x^k or z_k is a solution of problem (1). Therefore, in the following analysis, we assume that Algorithm 2.1 always generates an infinite sequence. That is to say, $\|F(z_k)\| \neq 0$ and $\|F(x_k)\| \neq 0$, for all k.

Theorem 3.1. *If F is continuous on C, condition (2) holds and the solution set of problem (1) is not empty, then the sequence $\{x_k\} \subset \mathbb{R}^n$ generated by Algorithm 2.1 globally converges to a solution of problem (1).*

Proof. Let $\bar{x}$ be a solution of problem (1). It follows from Lemma 3.1 that $\bar{x} \in C_k$. Since $x_{k+1} = \Pi_{C_k}(x_k)$, we obtain

$$x_{k+1} = x_k - \frac{\langle t_1F(x_k)+t_2F(z_k), x_k-z_k\rangle + \alpha_k t_1\langle F(x_k), d_k\rangle}{\|t_1F(x_k)+t_2F(z_k)\|^2}(t_1F(x_k)+t_2F(z_k)), \tag{17}$$

it follows from Lemma 2.1 and inequality (3) that

$$\begin{aligned} \|x_{k+1}-\bar{x}\|^2 &\leq \|x_k-\bar{x}\|^2 - \|x_{k+1}-x_k\|^2 \\ &= \|x_k-\bar{x}\|^2 - \frac{t_2^2|\langle F(z_k), x_k-z_k\rangle|^2}{\|t_1F(x_k)+t_2F(z_k)\|^2} \\ &\leq \|x_k-\bar{x}\|^2 - \sigma^2\alpha_k^4\|d_k\|^4. \end{aligned}$$

Which implies that the sequence $\{\|x_{k+1} - \bar{x}\|^2\}$ is nonincreasing and hence is a convergent sequence. Therefore, $\{x_k\}$ is bounded and

$$\lim_{k\to\infty}\alpha_k\|d_k\| = 0. \tag{18}$$

We claim that $\lim_{k\to\infty}\inf\|F(x_k)\| = 0$. In fact, if

$$\lim_{k\to\infty}\inf\|F(x_k)\| > 0 \tag{19}$$

holds, from (18) and Lemma 3.2, we obtain that $\{d_k\}$ is bounded. By inequalities (11) and (19), we have

$$\lim_{k\to\infty}\inf\|d_k\| \geq \xi\lim_{k\to\infty}\inf\|F(x_k)\| > 0.$$

This together with (18) implies

$$\lim_{k\to\infty}\alpha_k = 0.$$

By the choice of α_k, (3) implies that

$$-\langle F(x_k + \beta^{m_k-1}d_k), d_k\rangle < t_2^{-1}\sigma\beta^{m_{k-1}}\|d_k\|^2\|t_1F(x_k) + t_2F(x_k + \beta^{m_k-1}d_k)\|, \tag{20}$$

Since $\{x_k\}$ and $\{d_k\}$ are bounded, there exists some infinite index set K such that

$$\lim_{k\in K}x_k = \hat{x}, \text{ and } \lim_{k\in K}d_k = \hat{d},$$

where $\hat{x}$ and $\hat{d}$ are accumulation points of $\{x_k\}$ and $\{d_k\}$, respectively. We obtain by taking the limit in (20) for $k \in K$ that $-\langle F(\hat{x}), \hat{d}\rangle \leq 0$. Which contradicts (9). Therefore, $\lim_{k\to\infty}\inf\|F(x_k)\| = 0$. Since F is continuous and $\{x_k\}$ is a bounded sequence, there exists an accumulation point $\breve{x}$ of $\{x_k\}$ such that $F(\breve{x}) = 0$. This implies that $\breve{x}$ solves problem (1). Replacing $\bar{x}$ by $\breve{x}$ in the preceding argument, we obtain that the sequence $\{\|x_k - \breve{x}\|\}$ is nonincreasing and hence converges. Since $\breve{x}$ is an accumulation point of $\{x_k\}$, some subsequence of $\{\|x_k - \breve{x}\|\}$ converges to zero. This shows that the whole sequence $\{\|x_k - \breve{x}\|\}$ converges to zero, and hence $\lim_{k\to\infty}x_k = \breve{x}$. This completes the proof.

Remark 3.3. Compared with the proof of the global convergence in literatures [15-19], our conditions is weaker.

Acknowledgments

This work was supported by the Educational Science Foundation of Chongqing, Chongqing of China (grant KJ111309).

References

[1] S. Karamardian, Complementarity problems over cones with monotone and pseudomonotone maps. J. Optim. Theory Appl., 18(1976), 445-454.

[2] J. E. Dennis, J. J. Moré, Quasi-Newton method, motivation and theory, SIAM Rev., 19(1977), 46-89.

[3] J. E. Dennis, R. B. Schnable, Numerical methods for unconstrained optimization and nonlinear equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[4] X. J. Tong, L. Qi, On the convergence of a trust-region method for solving constrained nonlinear equations with degenerate solution. J. Optim. Theory Appl., 123(2004), 187-211.

[5] J. Y. Fan, Y. X. Yuan, On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. Computing, 74(2005), 23-39.

[6] J. Y. Fan, J. Y. Pan, An improved trust region algorithm for nonlinear equations. Comput. Optim. Appl., 48(2011), 59-70.

[7] D. H. Li, W. Y. Cheng, Recent progress in the globally convergence of quasi-Newton methods for nonlinear equations, Hokkaido Math. J., 36(2007), 729-743.

[8] M. V. Solodov, B. F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations. In: Fukushima M, Qi L (eds) Reformulation: piecewise smooth,semismooth and smoothing methods. Kluwer, Holanda, (1998), 355-369.

[9] C. W. Wang, Y. J. Wang, A superlinearly convergent projection method for constrained systems of nonlinear equations. J. Glob. Optim., 44(2009), 283-296.

[10] C. W. Wang, Y. J. Wang, C. L. Xu, A projection method for a system of nonlinear monotone equations with convex constraints, Math.Method Oper. Res., 66(2007), 33-46.

[11] F. M. Ma, C. W. Wang, Modified projection method for solving a system of monotone equations with convex constraints, J. Appl. Math. Comput., 34(2010), 47-56.

[12] S. Q. Du, Y. Y. Chen, Global convergence of a modified spectral FR conjugate gradient method, Appl. Math. Comput., 202(2008), 766-770.

[13] M. Li, H. Y. Feng, A sufficient descent LS conjugate gradient method for unconstrained optimization problems. Appl. Math. Comput.,218(2011), 1577-1586.

[14] L. Zhang, W. J. Zhou, D. H. Li, A descent modified Polak-Ribiére-Polyak conjugate gradient method and its global convergence, IMAJ. Numer. Anal., 26(2006), 629-640.

[15] Z. Wan, Z. L. Yang, Y. L. Wang, New spectral PRP conjugate gradient method for unconstrained optimization, Appl. Math. Lett., 24(2011), 16-22.

[16] Z. S. Yu, J. Lin, J. Sun, etc., Spectral gradient projection method for monotone nonlinear equations with convex constraints, Appl. Numer. Math., 59(2009), 2416-2423.

[17] W. Y. Cheng, A PRP type method for systems of monotone equations, Math. Comp. Model., 50(2009), 15-20.

[18] L. Zhang, W. J. Zhou, Spectral gradient projection method for solving nonlinear monotone equtions, J. Comput. Appl. Math., 196(2006), 478-484.

[19] Q. R. Yan, X. Z. Peng, D. H. Li, A globally convergent derivative-free method for solving large-scale nonlinear monotone equations, J. Comput. Appl. Math., 234(2010), 649-657.

[20] E. H. Zarantonello, Projections on convex sets in Hilbert spaces and spectral theory. Academic Press, New York(1971).