Volume 36, Number 3, 2011, pp. 611–622

A LINEARLY CONVERGENT CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION PROBLEMS

SUN MIN AND LIU JING

Abstract. In this paper, a new conjugate gradient method with a simple formula β_k is given for unconstrained optimization problems. It possesses sufficient descent property for any line search. Global convergence results of the new method with the Goldstein line search and the Armijo line search are discussed. For uniformly convex functions, the method has linear convergence rate. Preliminary computational experiments are included to illustrate the efficiency of the proposed method in minimizing large-scale non-convex optimization problems.

Consider the unconstrained nonlinear optimization problem

$$\min f(x),\ x \in R^n, \tag{0.1}$$

where $f : R^n \to R$ is a nonlinear function and its gradient $g(x)$ is available.

Conjugate gradient method is very effective for solving large-scale unconstrained optimization problems (0.1) due to its low memory requirements, and its iterative formula is given by

$$x_{k+1} = x_k + \alpha_k d_k, \tag{0.2}$$

with

$$d_k = \begin{cases} -g_k,\ k = 1, \\ -g_k + \beta_k d_{k-1},\ k \geq 2, \end{cases} \tag{0.3}$$

where x_1 is a given initial point, α_k is a step-length along d_k which is computed by carrying out some line search, g_k denotes $g(x_k)$ and β_k is a suitable scalar given by different formulae which result in distinct conjugate gradient methods. Well-known conjugate gradient methods include Fletcher-Reeves (FR) method [1], Polak-Ribiere-Polyak (PRP) method [2,3], Dai-Yuan (DY) method [4], Conjugate Descent (CD) method [5], and Hestenes-Stiefel (HS) method [5]. The parameters β_k of these methods are given by

$$\beta_k^{\mathrm{FR}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2},\ \beta_k^{\mathrm{PRP}} = \frac{g_k^\top (g_k - g_{k-1})}{d_{k-1}^\top (g_k - g_{k-1})},$$

Received October 7, 2009; in revised form September 8, 2010.

2000 *Mathematics Subject Classification.* Primary: 90C30; Secondary 90C33.

Key words and phrases. Conjugate gradient method, unconstrained optimization problems, global convergence, linear convergence rate.

$$\beta_k^{\mathrm{DY}} = \frac{\|g_k\|^2}{d_{k-1}^\top(g_k - g_{k-1})},\ \beta_k^{\mathrm{CD}} = -\frac{\|g_k\|^2}{g_{k-1}^\top d_{k-1}},\ \beta_k^{\mathrm{HS}} = \frac{g_k^\top(g_k - g_{k-1})}{d_{k-1}^\top(g_k - g_{k-1})}.$$

It is generally believed that the FR method has nice global convergence properties. The HS and PRP methods have been regarded as two of the most efficient conjugate gradient methods in practical computation and are preferred over the FR method. Nevertheless, Powell [5] showed that the PRP method with exact line search can cycle without approaching a solution point. Under the sufficient descent condition, Gilbert and Nocedal [6] showed that the modified PRP method $\beta_k = \max\{0, \beta_k^{\mathrm{PRP}}\}$ is globally convergent with the Wolfe-Powell line search. For the unmodified PRP method, Grippo and Lucidi [7] proved its convergence with a new line search.

The following step-length line searches are often used in the convergence analysis of the conjugate gradient methods:

(1) Armijo line search. Let m_k be the minimum nonnegative integer m such that $\alpha_k = \beta^{m_k}$ satisfies

$$f(x_k + \alpha_k d_k) - f(x_k) \le \delta\beta^m g_k^\top d_k,\ \beta, \delta \in (0,1). \tag{0.4}$$

(2) Goldstein line search. Find an $\alpha_k > 0$ such that

$$\alpha_k \mu_2 g_k^\top d_k \le f(x_k + \alpha_k d_k) - f(x_k) \le \alpha_k \mu_1 g_k^\top d_k \tag{0.5}$$

where $0 < \mu_1 < \mu_2 < 1$.

(3) Weak Wolfe-Powell line search (WWP). Find an $\alpha_k > 0$ satisfying the right-hand side inequality of (5) and

$$g(x_k + \alpha_k d_k)^\top d_k \ge \sigma g_k^\top d_k,\ \sigma \in (\mu_1, 1). \tag{0.6}$$

(4) Strong Wolfe-Powell line search (SWP). Find an $\alpha_k > 0$ satisfying the right-hand side inequality of (5) and

$$|g(x_k + \alpha_k d_k)^\top d_k| \le \sigma |g_k^\top d_k|,\ \sigma \in (\mu_1, 1). \tag{0.7}$$

The Wolfe-Powell line search and the Armijo type line search can guarantee the global convergence in many conjugate gradient methods [6-9], however the conjugate gradient method with Goldstein line search is very fewer except for the paper[11]. Can the Goldstein line search guarantee the global convergence and linear convergence?

In this paper, we present a new conjugate gradient method with the Goldstein line search for unconstrained optimization problems and prove its global convergence under some mild conditions. In addition, we show that the new method with the Armijo line search is also convergent. When the objective function is uniformly convex, we investigate the linear convergence rate of this new method.

The remainder of the paper is organized as follows. We describe the algorithm and analyze its simple properties in Section 1. In Section 2, we prove its global convergence under some mild conditions. Linear convergence rate is analyzed in Section 3, and preliminary computational results are given in Section 4. The conclusion is given in Section 5.

1. Algorithm

Throughout the paper, we always suppose the following assumptions hold.

(H1): The objective function f has lower bound on the level set $L_0 = \{x \in R^n | f(x) \leq f(x_1)\}$.

(H2): In some neighborhood N of L_0, g is uniformly continuous on an open convex set B that contains L_0.

(H2′) The gradient g is Lipschitz continuous on the open convex set B, i.e. there exists an $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|. \tag{1.1}$$

It is obvious that (H2′) implies (H2).

Now we consider the following algorithm.

Algorithm 1.1.

Step 1: Given $x_1 \in R^n$, $\epsilon > 0$, $t > 1$, set $k = 1$.

Step 2: If $\|g_k\| < \epsilon$, then stop; compute α_k by the Armijo line search (0.4) or the Goldstein line search (0.5), and compute d_k by (0.3), where

$$\beta_k = \frac{\|g_k\|}{t \cdot \|d_{k-1}\|} \text{ if } k \geq 2. \tag{1.2}$$

Step 3: Let the next iterative be $x_{k+1} = x_k + \alpha_k d_k$, and $k := k + 1$. Go to Step 2.

The following lemmas are very useful for studying conjugate gradient methods.

Lemma 1.1. *For all $k \geq 1$,*

$$g_k^\top d_k \leq -\frac{t-1}{t}\|g_k\|^2. \tag{1.3}$$

Proof. If $k = 1$, then

$$g_k^\top d_k = -\|g_k\|^2 \leq -\frac{t-1}{t}\|g_k\|^2.$$

If $k \geq 2$, by (0.3) and (1.2) we have

$$g_k^\top d_k = -\|g_k\|^2 + \beta_k g_k^\top d_{k-1} \leq -\|g_k\|^2 + \beta_k\|g_k\| \cdot \|d_{k-1}\| \leq -\frac{t-1}{t}\|g_k\|^2.$$

This completes the proof. □

Remark 1.1. From Lemma 1.1 and $t > 1$, the new conjugate gradient method can guarantee the sufficient descent condition for any line search.

Remark 1.2. [11] If $f(x)$ has lower bound and $g_k^\top d_k < 0$, then there exists α_k satisfying the Goldstein line search, i.e. the line search (0.5) is well-defined.

Remark 1.3. If x_k is not a stationary point of $f(x)$, i.e. $\|g_k\| \neq 0$, then by Lemma 1.1 we have $\|d_k\| \neq 0$, so β_k is well-defined.

Lemma 1.2. *For all $k \geq 1$,*

$$\|d_k\| \leq \frac{1+t}{t}\|g_k\|. \tag{1.4}$$

Proof. If $k=1$, then

$$\|d_k\| = \|-g_k\| \le \frac{1+t}{t}\|g_k\|.$$

If $k \ge 2$, by (0.3) and (1.2), we have

$$\|d_k\| = \|-g_k + \beta_k d_{k-1}\| \le \|g_k\| + \beta_k\|d_{k-1}\| \le \frac{1+t}{t}\|g_k\|.$$

This completes the proof. □

2. Global convergence

In this section, the convergence property of the new method with the Goldstein line search will be studied. Then, some results about the convergence properties of the new method with the Armijo line search are given.

2.1. **The convergence properties with the Goldstein line search.** In this subsection we assume that the step size α_k is computed by the Goldstein line search (0.5). The following lemma, which is the well-known Zoutendijk condition, is very useful in the proof of global convergence of the conjugate gradient methods.

Theorem 2.1. *If* (H1) *and* (H2′) *hold and Algorithm 1.1 generates an infinite sequence* $\{x_k\}$*, then*

$$\sum_{k=1}^{\infty} \frac{(g_k^\top d_k)^2}{\|d_k\|^2} < +\infty. \tag{2.1}$$

Proof. From the right-hand side inequality of (0.5) we have

$$\sum_{k=1}^{\infty} |\alpha_k g_k^\top d_k| < +\infty. \tag{2.2}$$

By the mean value theorem applied to the left-hand side of (0.5), there exists $\theta_k \in (0,1)$ such that

$$\alpha_k g(x_k + \theta_k\alpha_k d_k)^\top d_k = f(x_k + \alpha_k d_k) - f(x_k) \ge \alpha_k\mu_2 g_k^\top d_k.$$

By (H2′) and the Cauchy-Schwarz inequality we have

$$\alpha_k L\|d_k\|^2 \ge \|g(x_k+\theta_k\alpha_k d_k) - g_k\|\cdot\|d_k\| \ge (g(x_k+\theta_k\alpha_k d_k)-g_k)^\top d_k \ge -(1-\mu_2)g_k^\top d_k,$$

i.e.

$$\alpha_k \ge \frac{-(1-\mu_2)g_k^\top d_k}{L\|d_k\|^2}. \tag{2.3}$$

From the above inequality and (2.2), (2.1) holds. The proof is complete. □

Remark 2.1. By (1.3), (1.4) and (2.3) there exists a constant parameter $c>0$ such that

$$\alpha_k \ge c \doteq \frac{(1-\mu_2)t(t-1)}{L(1+t)^2}. \tag{2.4}$$

Theorem 2.2. *If the conditions in Theorem 2.1 hold then*

$$\lim_{k\to\infty} \|g_k\| = 0.$$

Proof. By (1.3) and (2.1) we have

$$\lim_{k\to\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = 0.$$

Therefore, by the above equality and (1.4) it follows that

$$\lim_{k\to\infty} \frac{t^2}{(1+t)^2}\|g_k\|^2 = \lim_{k\to\infty} \frac{t^2\|g_k\|^4}{(1+t)^2\|g_k\|^2} \leq \lim_{k\to\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = 0.$$

This completes the proof. □

Theorem 2.3. *If* (H1) *and* (H2) *hold and Algorithm 1.1 generates an infinite sequence* $\{x_k\}$*, then the sequence* $\{\|g_k\|\}$ *has an upper bound.*

Proof. By Lemma 1.1 and the Cauchy-Schwarz inequality we have

$$\|d_k\| \geq \frac{t-1}{t}\|g_k\|. \tag{2.5}$$

If we assume that $\{\|g_k\|\}$ has no bound, then there exists an infinite subset K such that

$$\lim_{k\in K, k\to\infty} \|g_k\| = +\infty. \tag{2.6}$$

By the right-hand side inequality of (0.5) and (1.3) we have

$$\begin{aligned} +\infty &> \sum_{k=1}^{\infty}(f_k - f_{k+1}) \\ &> \sum_{k\in K}(f_k - f_{k+1}) \\ &\geq -\mu_1 \sum_{k\in K} \alpha_k g_k^\top d_k \geq \mu_1 \frac{t-1}{t} \sum_{k\in K} \alpha_k\|g_k\|^2 \geq \mu_1 \frac{t-1}{t+1} \sum_{k\in K} \alpha_k\|d_k\|^2. \end{aligned}$$

Thus,

$$\lim_{k\in K, k\to\infty} \alpha_k\|d_k\|^2 = 0. \tag{2.7}$$

By (2.5) we have

$$\lim_{k\in K, k\to\infty} \|d_k\| = +\infty. \tag{2.8}$$

By (2.7) and (2.8) we have

$$\lim_{k\in K, k\to\infty} \alpha_k\|d_k\| = 0. \tag{2.9}$$

By the mean value theorem applied to the left-hand side of (0.5) and Lemma 1.1 we have

$$\begin{aligned} & \|g(x_k+\theta_k\alpha_k d_k)-g_k\|\cdot\|d_k\| \\ \geq\ & (g(x_k+\theta_k\alpha_k d_k)-g_k)^\top d_k \\ \geq\ & -(1-\mu_2)g_k^\top d_k \\ \geq\ & (1-\mu_2)(t-1)\|g_k\|^2/t \\ \geq\ & (1-\mu_2)(t-1)\|g_k\|\cdot\|d_k\|/(1+t), \end{aligned}$$

where $\theta_k \in (0,1)$. Using (2.9), we get

$$\lim_{k\in K, k\to\infty}\|g_k\|=0,$$

which contradicts (2.6). This shows that $\{\|g_k\|\}$ has an upper bound. This completes the proof. □

Theorem 2.4. *If* (H1) *and* (H2) *hold and Algorithm 1.1 generates an infinite sequence* $\{x_k\}$*, then*

$$\lim_{k\to\infty}\|g_k\|=0.$$

Proof. Assume that there is an infinite subset $K\subset N$ such that

$$\|g_k\|>\epsilon,\ \forall k\in K.$$

In the proof of Theorem 2.3, we have seen that

$$\mu_1\frac{t-1}{t}\sum_{k\in K}\alpha_k\|g_k\|^2<+\infty.$$

Thus,

$$\lim_{k\in K, k\to\infty}\alpha_k=0. \tag{2.10}$$

By Lemma 1.2, Theorem 2.3 and (2.4) we have

$$\lim_{k\in K, k\to\infty}\alpha_k\|d_k\|=0. \tag{2.11}$$

In the proof of Theorem 2.3 we have seen that

$$\begin{aligned} & \|g(x_k+\theta_k\alpha_k d_k)-g_k\|\cdot\|d_k\| \\ \geq\ & (1-\mu_2)(t-1)\|g_k\|\cdot\|d_k\|/(1+t) \end{aligned}$$

Combing this with (2.11) we get

$$\lim_{k\in K, k\to\infty}\|g_k\|=0,$$

which contradicts the assumption. This completes the proof. □

2.2. **The convergence properties with the Armijo line search.** In this subsection we assume that the step size α_k is computed by the Armijo line search.

Theorem 2.5. *If* (H1) *and* (H2′) *hold and Algorithm 1.1 generates an infinite sequence* $\{x_k\}$ *then*

$$\lim_{k\to\infty} \|g_k\| = 0. \tag{2.12}$$

Proof. Let

$$K_1 = \{k|\alpha_k = 1\},\ K_2 = \{k|\alpha_k < 1\}.$$

For $k \in K_1$ we have

$$f(x_{k+1}) - f(x_k) \le \delta g_k^\top d_k \le -\frac{\delta(t-1)}{t}\|g_k\|^2, \tag{2.13}$$

where the first inequality follows from (0.4) and the second one follows from (1.3).

If $k \in K_2$ then $\alpha_k < 1$, and this shows that $\alpha = \alpha_k/\beta$ cannot satisfy (0.4) and thus

$$f(x_k + \alpha d_k) - f(x_k) > \delta\alpha g_k^\top d_k.$$

Applying the mean value theorem to the left-hand side of the above inequality we see that there exists $\theta_k \in [0,1]$ such that

$$\alpha g(x_k + \theta_k \alpha d_k)^\top d_k > \delta\alpha g_k^\top d_k,$$

and by (H2') and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} & L\alpha\|d_k\|^2 \\ \ge\ & \|g(x_k + \theta_k\alpha d_k) - g_k\| \cdot \|d_k\| \\ \ge\ & (g(x_k + \theta_k\alpha d_k) - g_k)^\top d_k \\ \ge\ & -(1-\delta) g_k^\top d_k. \end{aligned}$$

Therefore we have

$$\alpha_k \ge \frac{-\beta(1-\delta)g_k^\top d_k}{L\|d_k\|^2},$$

and from (1.2) and(1.3) we have

$$\alpha_k \ge \frac{-\beta(1-\delta)(t-1)\|g_k\|^2}{tL\|d_k\|^2} \ge \frac{\beta t^2(1-\delta)(t-1)\|g_k\|^2}{tL(1+t)^2\|g_k\|^2} = \frac{\beta t(1-\delta)(t-1)}{L(1+t)^2} \doteq \widetilde{c}.$$

By (0.4) we have

$$f(x_k + \alpha_k d_k) - f(x_k) \le \delta\widetilde{c} g_k^\top d_k \le -\frac{\delta\widetilde{c}(t-1)}{t}\|g_k\|^2. \tag{2.14}$$

By (2.13) and (2.14) we have

$$\widehat{c}\|g_k\|^2 \le f(x_k) - f(x_{k+1}), \tag{2.15}$$

where

$$\widehat{c} = \min\{\frac{\delta\widetilde{c}(t-1)}{t}, \frac{\delta(t-1)}{t}\}.$$

From (2.15) we can obtain that $\{f_k\}$ is a decreasing sequence and has a bound from below. This shows that $\{f_k\}$ has a limit. Summing both side of (2.15) from $k = 0$ to ∞, we have

$$\sum_{k=0}^{\infty} \widehat{c}\|g_k\|^2 < +\infty.$$

We assert that (2.12) holds. The proof is complete. □

3. Linear convergence rate

In this section we will discuss the convergence rate of the new method. First, we give the definition of a uniformly convex function.

Definition 3.1. f is uniformly convex on R^n if there exists a constant $\mu > 0$ such that

$$(g(x) - g(y))^\top (x - y) \geq \mu\|x - y\|^2, \ \forall\ x, y \in R^n,$$

where $g(x) = \nabla f(x)$.

Now, we further assume that

(H3): f is uniformly convex and twice continuously differentiable.

Obviously, assumption (H3) implies (H1),(H2) and (H2′).

Lemma 3.1 ([8]). *If* (H3) *holds, then f has the following properties:*

(1) *f has a unique minimizer on R^n, say x^*.*

(2) *There exist $m > 0$, $M > 0$ and $\epsilon > 0$ such that*

$$\frac{1}{2}m\|x - x^*\|^2 \leq f(x) - f(x^*) \leq \frac{1}{2}M\|x - x^*\|^2, \ \forall x \in N(x^*, \epsilon),$$

$$m\|x - x^*\| \leq \|g(x)\| \leq M\|x - x^*\|, \ \forall x \in N(x^*, \epsilon).$$

The following theorem is inspired by Theorem 4.1 in [8].

Theorem 3.2. *If* (H3) *holds and $\mu_1 < 1/2$ then $\{x_k\}$ converges to x^* at least R-linearly.*

Proof. If (H3) holds then there exists k' such that $x_k \in N(x^*, \epsilon_0) \forall k \geq k$. Without loss of generality we can assume that $x_1 \in N(x^*, \epsilon_0)$. By the proof of Theorem 2.3 and Remark 2.1 we have

$$f_k - f_{k+1} \geq \frac{\mu_1(t-1)c}{t}\|g_k\|^2. \tag{3.1}$$

By Lemma 3.1(2) and (3.1) we obtain

$$\begin{aligned} f_k - f_{k+1} &\geq \frac{\mu_1(t-1)cm^2}{t}\|x_k - x^*\|^2 \\ &\geq \frac{2\mu_1(t-1)cm^2}{Mt}(f_k - f^*). \end{aligned}$$

Setting

$$\theta = m\sqrt{\frac{2\mu_1(t-1)c}{Mt}}$$

we have

$$f_k - f_{k+1} \geq \theta^2 (f_k - f^*). \tag{3.2}$$

Now we prove that $\theta < 1$. In fact, by the definition of c, and noting that $m \leq M \leq L$, we have

$$\begin{aligned}\theta^2 &= \frac{2\mu_1 (t-1) c m^2}{Mt} = \frac{2\mu_1 (t-1)^2 m^2 (1-\mu_2)}{ML(1+t)^2} \\ &\leq \frac{2\mu_1 m^2}{ML} \leq 2\mu_1 < 1.\end{aligned}$$

Setting

$$\omega = \sqrt{1-\theta^2}$$

it obviously hold $\omega < 1$, and (3.2) implies that

$$\begin{aligned} f_{k+1} - f^* &\leq (1-\theta^2)(f_k - f^*) \\ &= \omega^2 (f_k - f^*) \\ &\leq \cdots \\ &\leq \omega^{2(k-k')} (f_{k'+1} - f^*). \end{aligned}$$

By Lemma 3.1 and the above inequality we have

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \frac{2}{m}(f_{k+1} - f^*) \\ &\leq \omega^{2(k-k')} \frac{2(f_{k'+1} - f^*)}{m}, \end{aligned}$$

thus

$$\|x_k - x^*\| \leq \omega^k \sqrt{\frac{2(f_{k'+1} - f^*)}{m\omega^{2(k'+1)}}}$$

and

$$\lim_{k\to\infty} \|x_k - x^*\|^{1/k} \leq \omega < 1.$$

This shows that $\{x_k\}$ converges to x^* at least R-linearly. The proof is complete. □

4. Numerical reports

In this section we give some preliminary computational results to test the ability of the proposed Algorithm 1.1 in MATLAB 7.1. We take $\mu_1 = 0.38, \mu_2 = 0.75$ in the Goldstein line search, and $t = 2$ in the algorithm. The stopping criterion is

$$\|g_k\| \leq 10^{-6}.$$

We test the following conjugate gradient methods.

NCG1: Algorithm 1.1 with the Goldstein line search.

NCG2: Algorithm 1.1 with the Armijo line search.

PRP1: the PRP method with the Goldstein line search.

PRP2: the PRP method with the Armijo line search.

PRP$^+$: the PRP formula with nonnegative values $\beta_k = \max\{0, \beta_k^{\text{PRP}}\}$ and the Goldstein line search conditions.

FR: the FR method with the Armijo line search.

DY: the DY method with the Armijo line search.

HS: the HS method with the Armijo line search.

The numerical results are listed in Tables 1 and 2, where the items in each column have the following meanings:

N: the dimension of the problem;

NI/NF/CPU: the number of iterations, function evaluations, and CPU time;

F means that the number of the iteration exceeds 100.

Problem 4.1.

$$f(x) = \sum_{i=1}^{n}(x_i - 1)^2 + [\sum_{i=1}^{n}\frac{1}{i}(x_i - 1)]^2 + [\sum_{i=1}^{n}\frac{1}{i}(x_i - 1)]^4.$$

The initial point:

$$x_i = 1 - \frac{i}{n}, i = 1, 2, \cdots, n.$$

Table 1. Numerical results for Problem 4.1

N	NCG1	PRP1	PRP$^+$
2	12/57	11/44	13/53
10	10/48	17/74	18/79
100	14/69	17/72	17/73
1000	15/82	19/86	17/77
5000	14/74	21/99	20/95
10000	15/85	20/94	20/95

Problem 4.2.

$$f(x) = \sum_{i=1}^{n}(e^{x_i} - x_i).$$

The initial point:

$$x_i = \frac{n}{n-1}, i = 1, 2, \cdots, n.$$

Table 2. Numerical results for Problem 4.2

N	NCG2	PRP2	FR	DY
50	6/13/0.0100	7/14/0.0100	9/18/0.0100	9/18/0.0100
100	7/14/0.0100	7/14/0.0100	9/18/0.0100	9/18/0.0100
500	8/16/0.0200	7/14/0.0200	10/20/0.0100	10/20/0.0100
1000	8/16/0.0200	7/14/0.0200	F	13/26/1.4621
5000	9/18/0.2604	7/14/0.2504	F	13/26/7.3105

Problem 4.3 (Beale function).

$$f(x) = \sum_{i=1}^{3} f_i(x)^2, \;\; f_i(x) = y_i - x_1(1 - x_2^i), i = 1, 2, 3$$
$$y_1 = 1.5, y_2 = 2.25, y_3 = 2.625, x_0 = (1, 1)^\top.$$

Problem 4.4 (Rosenbrock function).

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, x_0 = (-3.635, 5.621)^\top.$$

Problem 4.5 (Cube function).

$$f(x) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2, x_0 = (1.2, 1)^\top.$$

Problem 4.6.

$$f(x) = (-x_1 + x_2 + x_3)^2 + (x_1 - x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2, x_0 = (100, -1, 2, 5)^\top.$$

Table 3. Numerical results for Problems 4.3-4.6

P	NCG2	PRP2	FR	HS
P3	54/108/3.3849	71/142/7.7111	35/70/1.9128	34/68/1.6624
P4	40/79/2.9643	90/179/11.9171	F	123/245/8.9429
P5	60/119/4.5766	42/83/3.9357	59/117/3.9056	F
P6	24/47/1.7325	7/13/0.3705	21/41/1.0615	40/79/2.2132

The above numerical results shows that the new method is efficient in practice. First, the new method avoids the evaluation of second derivatives of f. Second, it also avoids the storage of any matrix associated with quasi Newton type method. This comparison shows that the search direction of the new method is a good descent direction at each iteration.

However, how to choose a suitable parameter t for different problems deserves further research.

5. Conclusions

In this paper we developed a new conjugate gradient method for unconstrained optimization problems. Under mild conditions, we obtained global convergence with the Armijo line search and the Goldstein line search of the new method, and linear convergence rate was also studied. Numerical tests showed our method is encouraging in practical computation, compared with other similar methods.

Acknowledgments

The author would like to thank the anonymous referees for constructive comments and suggestions that greatly improved the paper.

References

[1] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, *Computer Journal* **7** (1964), 149-154.

[2] B. Polay and G. Ribiere, Note surla convergence des methods de directions conjuguees, *Reue Francaise de Recherche Operationnele* **16** (1969), 35-43.

[3] B. T. Polaykm, The conjugate gradient method in extreme problems. *USSR Computational Mathematics and Mathematical Physics* **9** (1969), 94-112.

[4] Y. H. Dai and Y. X. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Journal on optimmization* **10** (2000), 177-182.

[5] R.Flryvhrt, *Practical Methods of Optimization*, Vol. 1: Unconstrained Optimization (New York: Wiley and Sons), 1987.

[6] J. C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM Journal on Optimization* **2** (1992), 21-42.

[7] L. Grippo and S. Lucidi, A global convergence version of Polak-Ribiere gradient method, *Math. Program.* **78** (1997), 375-391.

[8] Shi Zhenjun and Shen Jie, Convergence of the Polak-Ribiere-Polyak conjugate method, *Nonlinear Analysis* **66** (2007), 1428-1441.

[9] Zhang Li, Zhou Weijun, and Li Donghui, Global convergence of a modified Fletch-Reeves conjugate gradient method with Armijo-type line search, *Numer. Math.* **104** (2006), 561-572.

[10] J. J. More, B. S. Garbow and K. E. Hillstrom, Testing unconstrained optimization software, *ACM Trans. Math. Softw.* **7** (1981), 17-41.

[11] B. C. Jiao, L. P. Chen, and C. Y. Pan, Convergence properties of a hybrid conjugate gradient mehtod with Goldstein line search, *Mathematica Numerica Sinca.* **29**(2) (2007), 137-146.

[12] M. Sun and J. Liu, A New Family of Conjugate Gradient Method with Armijo Line search, *Int. Journal of Math. Analysis* **4**(20) (2010), 987-993.

Department of Mathematics and Information Science,
Zaozhuang University, Shandong 277160, China.
E-mail address: sunmin_2008@yahoo.com.cn

School of Mathematics and Statistics,
Zhejiang University of Finance and Economics,
Hangzhou, 310018, China.
E-mail address: lifj8899@163.com