Bull. Korean Math. Soc. **50** (2013), No. 1, pp. 73–81
http://dx.doi.org/10.4134/BKMS.2013.50.1.073

GLOBAL CONVERGENCE OF AN EFFICIENT HYBRID CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION

Jinkui Liu and Xianglin Du

Abstract. In this paper, an efficient hybrid nonlinear conjugate gradient method is proposed to solve general unconstrained optimization problems on the basis of CD method [2] and DY method [5], which possess the following property: the sufficient descent property holds without any line search. Under the Wolfe line search conditions, we proved the global convergence of the hybrid method for general nonconvex functions. The numerical results show that the hybrid method is especially efficient for the given test problems, and it can be widely used in scientific and engineering computation.

1. Introduction

The primary objective of this paper is to study the global convergence property and practical computational performance of a new hybrid conjugate gradient method with the Wolfe line search for nonlinear unconstrained optimization.

Consider the following unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1.1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ is smooth and its gradient is available. The conjugate gradient method is very useful for solving (1.1) especially when n is large, and has the following iterative formulas:

$$x_{k+1}=x_k+\alpha_k d_k, \tag{1.2}$$

$$d_k=\begin{cases}-g_k, & k=1,\\ -g_k+\beta_k d_{k-1}, & k\ge 2,\end{cases} \tag{1.3}$$

where x_k is the current iteration point, g_k is the gradient of f at x_k, α_k is a positive scalar and called the steplength which is determined by some line

Received April 1, 2011; Revised April 19, 2012.
2010 *Mathematics Subject Classification.* 46N10.
Key words and phrases. unconstrained optimization, conjugate gradient method, the Wolfe line search, descent property, global convergence.
This work was supported by the Nature Science Foundation of Chongqing Education Committee (KJ121112).

search, d_k is the search direction, and β_k is a scalar. There are many ways to select β_k, and some well-known formulas are given by

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} \text{ (Polak-Ribiere-Polak [9], [8]),} \tag{1.4}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \text{ (Dai-Yuan [2]),} \tag{1.5}$$

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} \text{ (Fletcher [5]),} \tag{1.6}$$

respectively, where $\|\cdot\|$ is the Euclidean norm and $y_{k-1} = g_k - g_{k-1}$. The corresponding methods are called the PRP method, DY method and CD method, respectively.

In the convergence analysis and implementations of conjugate gradient methods, one often requires the line search to satisfy the strong Wolfe line search conditions, namely

$$f(x_k + \alpha_k d_k) - f(x_k) \le \delta \alpha_k g_k^T d_k, \tag{1.7}$$

$$|g(x_k + \alpha_k d_k)^T d_k| \le -\sigma g_k^T d_k, \tag{1.8}$$

where $0 < \delta < \sigma < 1$. The PRP method is regarded as the best one in practical computation. However, the PRP method has no global convergence in some situations. So some people have studied modified PRP methods. For example, Gilbert and Nocedal [6] proved that the conjugate gradient method with $\beta_k = \max\{\beta_k^{PRP}, 0\}$ converged globally (The corresponding method is the famous PRP$^+$ method), where the strong Wolfe line search and the sufficient descent condition were satisfied. Dai and Yuan [2] proved that the DY method could produce a descent search direction at every iteration and converge globally, where the line search satisfied the Wolfe line search conditions, namely, (1.7) and

$$g(x_k + \alpha_k d_k)^T d_k > \sigma g_k^T d_k, \tag{1.9}$$

where $0 < \delta < \sigma < 1$. Dai and Yuan [4] proved the CD method could ensure all search direction downhill, as long as the strong Wolfe line search conditions were satisfied. Dai [1] proposed a new conjugate gradient method in which

$$\beta_k = \frac{\|g_k\|^2}{\max\{d_{k-1}^T y_{k-1}, -d_{k-1}^T g_{k-1}\}},$$

and proved the global convergence under the nonmonotone line search. Wei, Yao and Liu [10] gave a new conjugate gradient method where the parameter β_k satisfies the following formula:

$$\beta_k^{VPRP} = \frac{g_k (g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1})}{\|g_{k-1}\|^2}.$$

They discussed the global convergence of the VPRP under the exact line search, the Wolfe line search and the Grippo Lucidi line search, respectively. The corresponding method is called the VPRP method in this paper.

The aim of this paper is to choose β_k to ensure that d_k is a descent direction and, at the same time, ensure the global convergence.

Under the inexact line searches, based on the effectively global convergence of the DY method and the well descent property of the CD method, we choose β_k to satisfy:

$$\beta_k = \begin{cases} 0, & \text{if } d_{k-1}^T g_k \leq \sigma d_{k-1}^T g_{k-1}, \\ \beta_k^{CD}, & \text{if } \sigma d_{k-1}^T g_{k-1} < d_{k-1}^T g_k \leq 0, \\ \beta_k^{DY}, & \text{if } 0 < d_{k-1}^T g_k < \mu d_{k-1}^T (g_k - g_{k-1}), \\ \mu \|g_k\|^2 / d_{k-1}^T g_k, & \text{if } d_{k-1}^T g_k \geq \mu d_{k-1}^T (g_k - g_{k-1}), \end{cases} \tag{1.10}$$

where $0 < \mu \leq \sigma$. In this paper, the corresponding descent method is called as CDY method, and we can prove that the CDY method has the sufficient descent property and the global convergence property.

2. The descent property

In order to prove the global convergence of the CDY method, the objective function $f(x)$ satisfies the following assumption.

Assumption (H):

(1) $f(x)$ is bounded from below on the level set $\Omega = \{x \mid f(x) \leq f(x_1)\}$.

(2) The objective function $f(x)$ is continuously differentiable, and its gradient $g(x)$ is Lipschitz continuous on the open set Γ containing Ω, i.e., there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\| \quad \text{for all} \quad x, y \in \Gamma. \tag{2.1}$$

CDY method:

Data: $x_1 \in \mathbb{R}^n$, $\varepsilon \geq 0$.

Step 1: Set $d_1 = -g_1$, if $||g_1|| \leq \varepsilon$, then stop.

Step 2: Compute α_k by some inexact line search.

Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = g(x_{k+1})$, if $||g_{k+1}|| \leq \varepsilon$, then stop.

Step 4: Compute β_{k+1} by (1.10), and generate d_{k+1} by (1.3).

Step 5: Set $k = k + 1$, go to Step 2.

In most references, we can see that the sufficient descent condition

$$g_k^T d_k \leq -c\|g_k\|^2, \ c > 0$$

is always given which plays a vital role in guaranteeing the global convergence properties of conjugate gradient methods. Furthermore, we have the following lemma which illustrates that the CDY method has the sufficient property for any line search.

Lemma 1. *Consider the iteration* (1.2)-(1.3), *where* α_k *is computed by any inexact line search, and* β_k *satisfies* (1.10). *Then if* $g_k \neq 0$ *for* $k \geq 1$, *we have that*

$$g_k^T d_k \leq -\|g_k\|^2 \quad \textit{for} \quad \forall \geq 1. \tag{2.2}$$

Proof. To obtain this result, we divide the proof into four aspects as follows.

(i) If $d_{k-1}^T g_k \leq \sigma g_{k-1}^T d_{k-1}$, we get $\beta_k = 0$. By multiplying (1.3) with g_k, we have

$$g_k^T d_k = -\|g_k\|^2 + \beta_k g_k^T d_{k-1}. \tag{2.3}$$

Obviously, $g_k^T d_k = -\|g_k\|^2 < 0$.

(ii) If $\sigma g_{k-1}^T d_{k-1} < g_k^T d_{k-1} \leq 0$, then $\beta_k = \beta_k^{CD}$. From (1.6), we have $\beta_k = \beta_k^{CD} > 0$. From (2.3), we easily have $g_k^T d_k \leq -\|g_k\|^2 < 0$.

(iii) If $0 < g_k^T d_{k-1} < \mu d_{k-1}^T (g_k - g_{k-1})$, we have $\beta_k = \beta_k^{DY}$. Then from (2.3) and (1.5), we have

$$g_k^T d_k = -\|g_k\|^2 + \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})} \cdot d_{k-1}^T g_k. \tag{2.4}$$

From (2.4) and $0 < g_k^T d_{k-1} < \mu d_{k-1}^T (g_k - g_{k-1})$, we have

$$g_k^T d_k < -\|g_k\|^2 + \frac{\|g_k\|^2}{\frac{1}{\mu} \cdot d_{k-1}^T g_k} \cdot d_{k-1}^T g_k = -(1-\mu)\|g_k\|^2 < 0.$$

(iv) If $g_k^T d_{k-1} \geq \mu d_{k-1}^T (g_k - g_{k-1})$, we have $\beta_k = \mu \|g_k\|^2 / d_{k-1}^T g_k$. Then from (2.3), we have

$$g_k^T d_k = -\|g_k\|^2 + \mu \|g_k\|^2 = -(1-\mu)\|g_k\|^2 \leq -(1-\sigma)\|g_k\|^2 < 0.$$

This completes the proof of Lemma 1. □

3. Global convergence

In this section, we will study the global convergence of the CDY method with the Wolfe line search. The following lemma, often called the Zoutendijk condition, is used to prove the global convergence of nonlinear conjugate gradient methods. It was originally given by Zoutendijk [11].

Lemma 2. *Suppose Assumption* (H) *holds. Consider any iteration of the form* (1.2)-(1.3), *where* d_k *satisfies* $d_k^T g_k < 0$ *for* $k \in \mathbb{N}$ *and* α_k *satisfies the Wolfe line search* (1.7) *and* (1.9). *Then*

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{3.1}$$

Theorem 3. *Suppose that Assumption* (H) *holds. Consider any iteration of the form* (1.2)-(1.3), *where* β_k *is computed by* (1.10) *and* α_k *satisfies the Wolfe line search* (1.7) *and* (1.9). *Then we have either* $g_k = 0$ *for some* k, *or*

$$\liminf_{k \to +\infty} \|g_k\| = 0. \tag{3.2}$$

Proof. If $g_k = 0$ holds for some finite k, then we know that x_k is a stationary point. Otherwise, we prove the conclusion (3.2) by contradiction.

Suppose that (3.2) dose not hold. This means that the gradients remain bounded away from zero, and hence there exists $\gamma > 0$ such that

$$\|g_k\| \geq \gamma \quad \text{for} \quad \forall\ k \geq 1. \tag{3.3}$$

From (1.9) and (2.2), we have

$$d_{k-1}^T(g_k - g_{k-1}) \geq -(1-\sigma)d_{k-1}^T g_{k-1} > 0. \tag{3.4}$$

In the following, we will prove that $\beta_k \leq \beta_k^{DY}$ holds, for $\forall\ k \geq 1$.

First, if $\sigma g_{k-1}^T d_{k-1} \leq g_k^T d_{k-1} \leq 0$, then we have $\beta_k = \beta_k^{CD}$ and

$$\begin{aligned} -g_{k-1}^T d_{k-1} &\geq -g_{k-1}^T d_{k-1} + g_k^T d_{k-1} \\ &\geq -(1-\sigma)d_{k-1}^T g_{k-1}. \end{aligned} \tag{3.5}$$

From (1.5), (1.6) and (3.5), we have

$$\beta_k = \beta_k^{CD} \leq \beta_k^{DY}. \tag{3.6}$$

Second, if $0 < g_k^T d_{k-1} < \mu d_{k-1}^T(g_k - g_{k-1})$, then form (1.10), we have

$$\beta_k = \beta_k^{DY}. \tag{3.7}$$

Third, if $g_k^T d_{k-1} \geq \mu d_{k-1}^T(g_k - g_{k-1})$, from (3.4), we have $\mu d_{k-1}^T(g_k - g_{k-1}) > 0$. Then

$$\beta_k = \mu\|g_k\|^2/d_{k-1}^T g_k \leq \beta_k^{DY}. \tag{3.8}$$

Y. H. Dai and Y. Yuan [3] gives an equivalent formula to (1.5):

$$\beta_k^{DY} = \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}. \tag{3.9}$$

What's more, if $d_{k-1}^T g_k < \sigma g_{k-1}^T d_{k-1}$, we have $\beta_k = 0$. So from (3.6)-(3.9), we can get

$$\beta_k \leq \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} \quad \text{for} \quad \forall\ k \geq 1. \tag{3.10}$$

On the one hand, writing (1.3) as $d_k + g_k = \beta_k d_{k-1}$ and squaring it, we get

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + (\beta_k)^2\|d_{k-1}\|^2.$$

From (3.10), we have

$$\|d_k\|^2 \leq -\|g_k\|^2 - 2g_k^T d_k + \left(\frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}\right)^2 \|d_{k-1}\|^2.$$

Dividing above inequality by $(g_k^T d_k)^2$, we have

$$\begin{aligned}\frac{\|d_k\|^2}{(g_k^T d_k)^2} &\le \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \left(\frac{\|g_k\|}{g_k^T d_k} + \frac{1}{\|g_k\|}\right)^2 + \frac{1}{\|g_k\|^2} \\ &\le \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2}.\end{aligned} \tag{3.11}$$

Using (3.11) recursively and noting that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, we get

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \sum_{i=1}^{k} \frac{1}{\|g_k\|^2}. \tag{3.12}$$

Then we get from this and (3.3) that

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \frac{\gamma^2}{k}, \tag{3.13}$$

which indicates

$$\sum_{k\ge 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = +\infty. \tag{3.14}$$

This contradicts the Zoutendijk condition (3.1). Therefore the conclusion (3.2) holds. □

Corollary. *Suppose that Assumption* (H) *holds. Consider any iteration of the form* (1.2)-(1.3)*, where* β_k *is computed by* (1.10) *and* α_k *satisfies the strong Wolfe line search* (1.7)-(1.8)*. Then we have either* $g_k = 0$ *for some* k*, or*

$$\liminf_{k\to+\infty} \|g_k\| = 0. \tag{3.15}$$

Proof. If α_k satisfies the strong Wolfe line search (1.7)-(1.8), then α_k must satisfy the Wolfe line search (1.7) and (1.9). Therefore the statement follows Theorem 3. □

4. Numerical results

The purpose of this section is to present computational supports for the CD method, DY method, PRP$^+$ method, VPRP method and CDY($\mu = 10^{-6}$) method. Under the strong Wolfe line searches, the methods were tested for a set of standard unconstrained minimization test problems from [7], where $\delta = 0.01$, $\sigma = 0.1$. The termination condition of the experiments is $\|g_k\| \le \varepsilon$, where $\varepsilon = 10^{-6}$.

The test problems are listed in Table 1, and the detail numerical results of our tests are reported in Tables 2-4. In Tables, "Number" and "Name" denote the problem number and problem name, respectively. "Dim" denotes the dimension of the test problems. "——" means the method fails. In Table 2, the detailed numerical results are listed in the form NI/NF/NG, where NI, NF, NG denote the number of iterations, function evaluations, and gradient

Table 1. List of test problems

Number	Name
1	Freudenstein and Roth function
2	Beale function
3	Helical valley function
4	Gulf research and development function
5	Powell singular function
6	Wood function
7	Kowalik and Osborne function
8	Brown and Dennis function
9	Watson function
10	Penalty function I
11	Trigonometric function
12	Extended Powell singular function
13	Discrete boundary value function
14	Discrete integral equation function
15	Broyden tridiagonal function

Table 2. The numerical results of the methods

Number	Dim	CD	DY	PRP^+	VPRP	CDY
1	2	51/187/152	42/168/138	11/72/56*	15/90/70	11/76/57
2	2	73/177/155	75/186/164	13/58/45*	21/65/47	15/57/44
3	3	56/157/132	37/118/98*	65/181/156	57/174/146	45/147/123
4	3	1/2/2*	1/2/2*	1/2/2*	1/2/2*	1/2/2*
5	4	421/1036/947	2286/4555/4545	113/379/328	87/300/253*	102/383/333
6	4	184/438/399	100/291/240	118/357/304	236/608/543	78/278/230*
7	4	254/723/633	536/1449/1271	93/269/240	66/228/197*	68/249/220
8	4	44/171/133	39/158/121	37/156/123*	——	41/178/136
9	5	87/277/239*	127/348/299	133/374/330	130/377/324	402/1208/1062
	15	4377/12694/11221	1845/5658/4924	3290/10457/9244	1800/5631/4954*	2288/7973/7044
10	100	62/223/182	31/157/121	29/168/128*	34/224/170	33/195/152
	200	25/159/117	26/160/121	25/175/132	34/239/181	24/167/125*
11	100	——	306/401/400	58/120/113	56/145/119	56/137/121*
	200	——	315/399/398	64/135/128	58/126/122*	64/160/146
12	500	748/1780/1689	2778/5384/5374	105/342/297	178/612/526	100/379/324*
	1000	578/1419/1292	4329/8191/18168	198/693/595	106/355/305*	154/623/552
13	500	5089/7049/7048	4796/6823/6822	1645/2889/2888	1461/2534/2533	219/491/466*
	1000	2406/3114/3113	414/449/448	147/251/250	157/275/274	38/69/64*
14	500	7/15/8	7/15/8	6/13/7*	7/15/8	7/15/8
	1000	7/15/8	7/15/8	6/13/7*	7/15/8	7/15/8
15	500	52/112/107	49/106/101	35/78/74	35/78/73	33/75/60*
	1000	70/149/145	64/137/133	34/76/72	36/81/77	32/75/61*

evaluations, respectively. The star "$*$" denotes that this result is best one among these methods. In Table 3, CPU times of the methods are given. The final values and standard values are reported in Table 4. "Standard" means standard value of the test problem.

Tables 2-3 show that the CDY method has the best performance with respect to the number of iterations and the CPU time. What is more, Table 4 also show that the CDY method relative to the final values of the test problems is comparable with that of the PRP+ method and VPRP method. All numerical results show that the efficiency of the CDY method is encouraging.

TABLE 3. The corresponding CPU times of the methods

Number	Dim	CD	DY	PRP$^+$	VPRP	CDY
1	2	0.1863	0.1675	0.0510	0.0494	0.0452*
2	2	0.2517	0.3000	0.0714*	0.0735	0.0720
3	3	0.1985	0.1263*	0.4376	0.1982	0.2531
4	3	0.0036	0.0037	0.0033	0.0034	0.0009*
5	4	2.5067	8.9351	0.7034	0.3132*	0.3892
6	4	0.7156	0.4189	0.5263	1.2076	0.3787*
7	4	0.6112	2.1098	0.4172	0.3289*	0.3579
8	4	0.1614*	0.3521	0.2172	——	0.2364
9	5	0.3142*	0.5312	0.6102	0.5762	2.2067
	15	16.9905	6.0000*	10.9487	6.8933	7.5893
10	100	0.5118	0.2325	0.2712	0.3102	0.2136*
	200	0.3952	0.4015	0.4306	0.6212	0.3657*
11	100	——	1.9000	0.4573	0.4384	0.3899*
	200	——	6.5000	1.8896*	1.9796	1.9860
12	500	12.9834	32.8723	2.4000	3.5617	2.2670*
	1000	28.7000	43.2000	7.3000	6.7891*	11.2701
13	500	58.1386	55.4843	27.8092	21.1132	8.7861*
	1000	39.6372	13.0000	6.1000	7.2000	1.8963*
14	500	2.1428	2.1554	1.8556*	2.1450	1.9325
	1000	8.5477	9.7059	7.3761*	8.5380	8.3450
15	500	0.7284	0.6892	0.5063	0.5092	0.5011*
	1000	3.6011	3.1574	1.6478	1.7671	1.4968*

TABLE 4. The final values and standard values of the methods

Number	Dim	CD	DY	PRP$^+$	VPRP	CDY	Standard
1	2	48.9843*	48.9843*	48.9843*	48.9843*	48.9843*	48.9842
2	2	2.0949e-013	7.5278e-014	5.7209e-016	3.1303e-019*	3.4197e-014	0
3	3	1.7392e-015*	1.2045e-014	7.5337e-015	5.2175e-015	4.5381e-015	0
4	3	0.0385*	0.0385*	0.0385*	0.0385*	0.0385*	0
5	4	1.7630e-010	2.8606e-011	1.4371e-012*	5.2126e-010	2.0535e-011	0
6	4	4.6149e-014	2.6023e-014*	3.0109e-013	2.6463e-014	8.6269e-014	0
7	4	3.0751e-004*	3.0751e-004*	3.0751e-004*	3.0751e-004*	3.0751e-004*	3.07505e-004
8	4	8.5822e+004*	8.5822e+004*	8.5822e+004*	——	8.5822e+004*	8.5822e+004
9	5	0.0172*	0.0172*	0.0172*	0.0172*	0.0172*	0.0172
	15	1.6303e-009	2.8710e-010	3.4585e-008	3.7954e-008	4.0916e-010*	4.7224e-010
10	100	9.0249e-004*	9.0249e-004*	9.0249e-004*	9.0249e-004*	9.0249e-004*	9.0249e-004
	200	0.0019*	0.0019*	0.0019*	0.0019*	0.0019*	0.0019
11	100	——	1.8410e-006*	1.8410e-006*	1.8410e-006*	1.8410e-006*	0
	200	——	1.0051e-006*	1.1542e-006	1.1542e-006	1.1542e-006	0
12	500	4.4691e-010	3.1934e-010	4.3989e-011*	7.3238e-010	7.8245e-010	0
	1000	9.9418e-010	4.5246e-010	1.2283e-010	5.9312e-011*	6.2169e-011	0
13	500	9.2306e-009	7.4621e-009	8.9750e-009	9.0694e-009	1.0275e-009*	0
	1000	1.2344e-009*	1.2373e-009	1.2644e-009	1.2616e-009	1.2926e-009	0
14	500	1.0452e-013	1.1936e-013	1.4385e-015	1.2338e-015*	1.1623e-014	0
	1000	2.0882e-013	5.7282e-015	2.8738e-015	2.4651e-015*	2.3219e-014	0
15	500	1.3526e-014	1.1159e-014	6.6559e-015*	7.4750e-015	7.1464e-015	0
	1000	9.6485e-015	1.0529e-014	1.1090e-014	4.4833e-015	3.9079e-015*	0

References

[1] Y. H. Dai, *A nonmonotone conjugate gradient algorithm for unconstrained optimization*, J. Syst. Sci. Complex. **15** (2002), no. 2, 139–145.
[2] Y. H. Dai and Y. X. Yuan, *A Nonlinear conjugate gradient with a strong global convergence property*, SIAM. J. Optimization **10** (2000), 177–182.
[3] ______, *Nonlinear Conjugate Gradient Method*, Shanghai Scientific and Technical, Shanghai, China, 2000.
[4] ______, *Convergence properties of the conjugate descent method*, Adv. in Math. (China) **25** (1996), no. 6, 552–562.
[5] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Ltd., 1987.
[6] J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim. (1992), no. 1, 21–42.
[7] J. J. More, B. S. Garbow, and K. E. Hillstrome, *Testing unconstrained optimization software*, ACM Trans. Math. Software **7** (1981), no. 1, 17–41.
[8] B. T. Polak, *The conjugate gradient method in extreme problems*, USSR Comput. Math. Math. Phys. **9** (1969), 94–112.
[9] E. Polak and G. Ribiére, *Note sur la convergence de méthodes de directions conjuguées*, Rev. Française Informat. Recherche Opérationnelle **3** (1969) no. 16, 35–43.
[10] Z. Wei, S. Yao, and L. Liu, *The convergence properties of some new conjugate gradient methods*, Appl. Math. Comput. **183** (2006), no. 2, 1341–1350.
[11] G. Zoutendijk, *Nonlinear Programming Computational Methods*, Integer and nonlinear programming, pp. 37–86. North-Holland, Amsterdam, 1970.

Jinkui Liu
School of Mathematics and Statistics
Chongqing Three Gorges University
Chongqing, P. R. China
E-mail address: liujinkui2006@126.com

Xianglin Du
School of Mathematics and Statistics
Chongqing Three Gorges University
Chongqing, P. R. China
E-mail address: duxianglin2002@126.com