Scientific
Research

Three New Hybrid Conjugate Gradient Methods for Optimization*

Anwa Zhou, Zhibin Zhu, Hao Fan, Qian Qing
*School of Mathematics and Computational Science, Guilin University of Electronic Technology,
Guilin, China*
E-mail: *zhouanwa@126.com*
Received November 10, 2010; *revised January* 9, 2011; *accepted January* 14, 2011

Abstract

In this paper, three new hybrid nonlinear conjugate gradient methods are presented, which produce sufficient descent search direction at every iteration. This property is independent of any line search or the convexity of the objective function used. Under suitable conditions, we prove that the proposed methods converge globally for general nonconvex functions. The numerical results show that all these three new hybrid methods are efficient for the given test problems.

Keywords: Conjugate Gradient Method, Descent Direction, Global Convergence

1. Introduction

In this paper, we consider the unconstrained optimization problem:

$$\min f(x), x \in R^n \qquad (1.1)$$

where $f : R^n \to R$ is continuously differentiable and the gradient of f at x is denoted by $g(x)$.

Due to its simplicity and its very low memory requirement, the conjugate gradient (CG) method plays a very important role for solving (1.1). Especially, when the scale is large, the CG method is very efficient. Let $x_0 \in R^n$ be the initial guess of the solution of problem (1.1). A nonlinear conjugate gradient method is usually designed by the iterative form

$$x_{k+1} = x_k + \alpha_k d_k, k = 0, 1, \cdots, \qquad (1.2)$$

where x_k is the current iterate point, $\alpha_k > 0$ is a step-length which is determined by some line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if } k > 0, \end{cases} \qquad (1.3)$$

where g_k denotes $g(x_k)$, and β_k is a scalar. There are some well-known formulas for β_k, which are given as follows:

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad [1];$$

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \quad [2];$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \quad [3];$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad [4];$$

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad [5,6];$$

$$\beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \quad [7];$$

where $y_{k-1} = g_k - g_{k-1}$ and $\|\cdot\|$ stands for the Euclidean norm of vectors.

Although the methods above are equivalent [8,9] when f is a strictly convex quadratic function and α_k is calculated by the exact line search, their behaviors for general objective functions may be far different. Generally, in the convergence analysis of conjugate gradient methods, one hopes the inexact line search such as the Wolfe conditions, the strong Wolfe conditions or the strong *Wolfe conditions, which are showed respectively as follows:

1) The Wolfe line search is to find α_k such that

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k, \\ d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k, \end{cases} \quad (1.4)$$

with $0 < \delta < \dfrac{1}{2}$, and $\delta < \sigma < 1$.

2) The strong Wolfe line search is to find α_k such that

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k, \\ \left| d_k^T g(x_k + \alpha_k d_k) \right| \le -\sigma d_k^T g_k, \end{cases} \quad (1.5)$$

with $0 < \delta < \dfrac{1}{2}$, and $\delta < \sigma < 1$.

3) The strong *Wolfe line search is to find α_k such that

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k, \\ \sigma d_k^T g_k \le d_k^T g(x_k + \alpha_k d_k) \le 0, \end{cases} \quad (1.6)$$

with $0 < \delta < \dfrac{1}{2}$, and $\delta < \sigma < 1$.

For general functions, Zoutendjk [10] and Al-Baali [11] had proved the global convergence of the FR method with different line searches. And Powell [12] gave a counter example which showed that there exist nonconvex functions such that the PRP method may cycle and does not approach any stationary point even with exact line search. Although one would be satisfied with its global convergence, the FR method performs much worse than the PRP (HS, LS) method in real computations. In other words, in practical computation, the PRP method, the HS method, and the LS method are generally believed to be the most efficient conjugate gradient methods since these methods essentially perform a restart if a bad direction occurs. A similar case happens to the DY method and the CD method. That is to say the convergences of the CD , DY and FR methods are established [1-3], however their numerical results are not so well. Resently, some good results on the nonlinear conjugate gradient method are given. Combining the good numerical performance of the PRP and HS methods and the nice global convergence properties of the FR and DY methods, recently, [13] and [14] proposed some hybrid methods which we call the H1 method and the H2 method, respectively, that is,

$$\beta_k^{H1} = \max\left\{0, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}, \quad (1.7)$$

$$\beta_k^{H2} = \max\left\{0, \min\left\{\beta_k^{HS}, \beta_k^{DY}\right\}\right\}. \quad (1.8)$$

Gilbert and Nocedal [15] extended H1 to the case that

$$\beta_k = \max\left\{-\beta_k^{FR}, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}.$$

Numerical performances show that the H1 and the H2 methods are better than the PRP method [13,14,16].

As we all know, the FR , DY and CD methods are descent methods, but their descent properties depend on the line search such as the strong Wolfe line search (1.5). Similar to the descent three terms PRP method in [17], Zhang *et al.* [18,19] proposed a modified FR method which we call the MFR method, that is,

$$\begin{aligned} MFR: d_k &= -g_k + \beta_k^{FR} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k \\ &= -\left(1 + \beta_k^{FR} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{FR} d_{k-1}. \end{aligned} \quad (1.9)$$

And [18] also gave an equivalent form to the MFR method. Similarly, Zhang [19] also proposed a modified DY method called the MDY method, that is,

$$\begin{aligned} MDY: d_k &= -g_k + \beta_k^{DY} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k \\ &= -\left(1 + \beta_k^{DY} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{DY} d_{k-1} \end{aligned} \quad (1.10)$$

It is easy to see that the MFR and MDY methods have an important property that the search directions satisfy $g_k^T d_k = -\|g_k\|^2$, which depends neither on the line search used nor on the convexity of the objective function; moreover these two methods reduce to the FR method and the DY method respectively with exact line search.

[18] has explored the convergence and efficiency of the MFR method for nonconvex functions with the Wolfe line search or Armijo line search. Based on the idea of the H1 and the H2 methods, recently, Zhang-Zhou [20] replaced β_k^{FR} in (1.9) and β_k^{DY} in (1.10) with β_k^{H1} in (1.7) and β_k^{H2} in (1.8), respectively, and proposed two new hybrid PRP–FR and HS–DY methods called the NH1 method and the NH2 method, respectively, that is,

$$NH1: d_k = -\left(1 + \beta_k^{H1} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{H1} d_{k-1} \quad (1.11)$$

$$NH2: d_k = -\left(1 + \beta_k^{H2} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{H2} d_{k-1} \quad (1.12)$$

Obviously, these two new hybrid methods still satisfy

$$g_k^T d_k = -\|g_k\|^2, \quad (1.13)$$

which shows that they are descent and independent of any line search used. [20] proved the global convergence of these two methods and also showed their efficiency in real computations.

Similarly, based on the idea of the methods all above, we consider the CD method, and propose three new hy-

brid conjugate gradient(CG) methods which we call the H3 method, the MCD method and the NH3 method, respectively, that is

$$H3 : \beta_k^{H3} = \max\left\{0, \min\left\{\beta_k^{LS}, \beta_k^{CD}\right\}\right\} \qquad (1.14)$$

$$MCD : d_k = -g_k + \beta_k^{CD} d_{k-1} + \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k$$

$$= -\left(1 + \beta_k^{CD} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{CD} d_{k-1} \qquad (1.15)$$

$$NH3 : d_k = -\left(1 + \beta_k^{H3} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right) g_k + \beta_k^{H3} d_{k-1} \qquad (1.16)$$

From these three methods above, it is not difficult to see that the MCD and the NH3 methods also satisfy $g_k^T d_k = -\|g_k\|^2$, which shows that they are sufficient descent methods. In the next section, the new algorithms are given. The global convergence of the proposed methods are proved in Section 3. We give the numerical experiments in Section 4, and in Section 5, the conclusion is presented.

2. Algorithm

2.1. Algorithm 1 (The H3 Algorithm)

Step 0: Choose an initial point $x_0 \in R^n$, $0 < \varepsilon \ll 1$,

$0 < \delta < \frac{1}{2}$, and $\delta < \sigma < 1$. Set $d_0 = -g_0, k := 0$.

Step 1: If $\|g_k\| < \varepsilon$, then stop; Otherwise go to the next step.

Step 2: Compute step size α_k by strong *Wolfe line search rule (1.6).

Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$. If $\|g_{k+1}\| < \varepsilon$, then stop.

Step 4: Calculate the search direction
$d_{k+1} = -g_{k+1} + \beta_{k+1}^{H3} d_k$.

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

2.2. Algorithm 2 (The MCD (Or the NH3) Algorithm)

Step 0: Choose an initial point $x_0 \in R^n$, $0 < \varepsilon \ll 1$,

$0 < \delta < \frac{1}{2}$, and $\delta < \sigma < 1$. Set $d_0 = -g_0, k := 0$.

Step 1: If $\|g_k\| < \varepsilon$, then stop; otherwise go to the next step.

Step 2: Compute step size α_k by Wolfe line search rule (1.4).

Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$. If $\|g_{k+1}\| < \varepsilon,,$ then stop.

Step 4: Calculate the search direction d_{k+1} by (1.15) (or (1.16)).

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

3. The Global Convergence

Assumption A

1) The level set $\Omega = \left\{x \in R^n \,\big|\, f(x) \le f(x_0)\right\}$ is bounded, where $x_0 \in R^n$, is a given point.

2) In an open convex set N that contains Ω, f is continuously differentiable and its gradient g is Lipschitz continuous, namely, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in N. \qquad (3.1)$$

Since $\{f(x_k)\}$ is decreasing, it is clear that the sequence $\{x_k\}$ generated by Algorithm 1 and Algorithm 2 is contained in Ω. In addition, we can get from Assumption A that there exists a constant B and $\gamma_1 > 0$, such that

$$\|x\| \le B, \|g(x)\| \le \gamma_1, \forall x \in \Omega. \qquad (3.2)$$

In the latter part of the paper, we always suppose that the conditions in Assumption A hold. Then there is an useful lemma, which was originally given in [10,21].

Lemma 3.1 Let $\{x_k\}$ be generated by (1.2) and d_k is a descent direction. If α_k is determined by the Wolfe line search (1.4), then we have

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty. \qquad (3.3)$$

From Lemma 3.1 and (1.13) that for the MCD and NH3 methods with the Wolfe line search, we can easily obtain the following condition

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \qquad (3.4)$$

We now establish the global convergence theorem for Algorithm 1 and Algorithm 2 .

3.1. The Global Convergence of the H3 Method

For simlpicity, here, we list the Theorem 2.3 in [14] as the following Lemma 3.2 without proof.

Lemma 3.2 Suppose that x_0 is an initial point, Consider the method (1.2) and (1.3), where α_k is computed by the Wolfe line search(1.4), and β_k is such that

$r_k \in [-c, 1]$, where $r_k = \frac{\beta_k}{\beta_k^{DY}}, c = \frac{1-\sigma}{1+\sigma} > 0$.

Then if $g_k \neq 0$ for all $k \ge 0$, we have that

$$g_k^T d_k < 0, \forall k \geq 0.$$

Further, the method converges in the sense that

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

From the Lemma 3.2 above, similar to Corollary 2.4 in [14], we give the global convergence of the H3 method(Algorithm 1).

Theorem 3.3 Suppose that x_0 is an initial point, Consider the Algorithm1, then we have either $g_k = 0$ for some $k \geq 0$, or $\liminf_{k \to \infty} \|g_k\| = 0$.

Proof From the second inequality of (1.6) and the definitions of $\beta_k^{H3}, \beta_k^{CD}$ and β_k^{DY}, it follows that

$$0 \leq \beta_k^{H3} \leq \beta_k^{CD} \leq \beta_k^{DY}.$$

Therefore the statement follows lemma 3.2. □

3.2. The Global Convergence of the MCD Method

Now, we establish the global convergence theorem for the MCD method.

Theorem 3.4 Let $\{x_k\}$ be generated by the MCD method (Algorithm 2, where d_k satisfies (1.15)), then we have

$$\liminf_{k \to \infty} \|g_k\| = 0. \tag{3.5}$$

Proof Suppose by contradiction that the desired conclusion is not true, that is to say, there exists a constant $\varepsilon > 0$ such that

$$\|g_k\| \geq \varepsilon, \forall k \geq 0. \tag{3.6}$$

Set $h_k = 1 + \beta_k^{CD} \dfrac{g_k^T d_{k-1}}{\|g_k\|^2}$, and then we have

$\beta_k^{CD} g_k^T d_{k-1} = h_k \|g_k\|^2 - \|g_k\|^2$. From (1.15) and (1.13), it follows that

$$\begin{aligned}
\|d_k\|^2 &= \left\|\beta_k^{CD} d_{k-1} - h_k g_k\right\|^2 \\
&= \left(\beta_k^{CD}\right)^2 \|d_{k-1}\|^2 - 2h_k \beta_k^{CD} d_{k-1}^T g_k + h_k^2 \|g_k\|^2 \\
&= \left(\beta_k^{CD}\right)^2 \|d_{k-1}\|^2 - 2h_k \left(h_k \|g_k\|^2 - \|g_k\|^2\right) d_{k-1}^T g_k + h_k^2 \|g_k\|^2 \\
&= \left(\beta_k^{CD}\right)^2 \|d_{k-1}\|^2 + 2h_k \|g_k\|^2 - h_k^2 \|g_k\|^2 \\
&= \left(\beta_k^{CD}\right)^2 \|d_{k-1}\|^2 - 2h_k d_k^T g_k - h_k^2 \|g_k\|^2 \tag{3.7}
\end{aligned}$$

Dividing both sides of (3.7) by $\left(g_k^T d_k\right)^2$, we get, from (1.13) and the definition of β_k^{CD}, that

$$\frac{\|d_k\|^2}{\|g_k\|^4} = \frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2}$$

$$\begin{aligned}
&= \left(\beta_k^{CD}\right)^2 \frac{\|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \|g_k\|^2}{\left(g_k^T d_k\right)^2} \\
&= \left(\frac{\|g_k\|^2}{\|g_{k-1}\|^2}\right)^2 \frac{\|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \|g_k\|^2}{\left(g_k^T d_k\right)^2} \\
&= \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} - \frac{1}{\|g_k\|^2}\left(h_k^2 + 2h_k + 1 - 1\right) \\
&= \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} - \frac{(h_k+1)^2}{\|g_k\|^2} + \frac{1}{\|g_k\|^2} \\
&\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2} \\
&\leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2} \leq \frac{k}{\varepsilon^2}.
\end{aligned}$$

The last inequality implies

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \varepsilon^2 \sum_{k=0}^{\infty} \frac{1}{k} = \infty,$$

which contradicts (3.4). The proof is then completed. □

3.3. The Global Convergence of the NH3 Method

The same as the Theorem 3.4, we can establish the following global convergence theorem for the NH3 method.

Theorem 3.5 Let $\{x_k\}$ be generated by the NH3 method (Algorithm 2, where d_k satisfies (1.16)), then we have

$$\liminf_{x \to \infty} \|g_k\| = 0. \tag{3.8}$$

Proof Suppose by contradiction that the desired conclusion is false, that is to say, there exists a constant $\varepsilon > 0$ such that

$$\|g_k\| \geq \varepsilon, \forall k \geq 0. \tag{3.9}$$

Similar to (3.7), we get from (1.16) that

$$\|d_k\|^2 = \left(\beta_k^{H3}\right)^2 \|d_{k-1}\|^2 - 2h_k d_k^T g_k - h_k^2 \|g_k\|^2, \tag{3.10}$$

where

$$h_k = 1 + \beta_k^{H3} \frac{g_k^T d_{k-1}}{\|g_k\|^2}.$$

Notice that

$$\beta_k^{H3} \leq \left|\beta_k^{CD}\right|, \forall k \geq 0. \tag{3.11}$$

Dividing both sides of (3.10) by $\left(g_k^T d_k\right)^2$, we get, from (3.11), (1.13) and the definition of β_k^{H3}, that

$$\frac{\left\| d_k \right\|^2}{\left\| g_k \right\|^4} = \frac{\left\| d_k \right\|^2}{\left(g_k^T d_k \right)^2}$$

$$= \left(\beta_k^{H3} \right)^2 \frac{\left\| d_{k-1} \right\|^2}{\left(g_k^T d_k \right)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \left\| g_k \right\|^2}{\left(g_k^T d_k \right)^2}$$

$$\leq \left(\beta_k^{CD} \right)^2 \frac{\left\| d_{k-1} \right\|^2}{\left(g_k^T d_k \right)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \left\| g_k \right\|^2}{\left(g_k^T d_k \right)^2}$$

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

$$\leq \sum_{i=0}^{k} \frac{1}{\left\| g_i \right\|^2} \leq \frac{k}{\varepsilon^2},$$

which contradicts (3.4). This finishes the proof. □

4. Numerical Experiments

In this section, we carry out some numerical experiments. These three algorithms have been tested on some problems from [22]. The results are summarized in the following three tables: **Table 1-3**. For each test problem, No. is the number of the test problem in [22], x_0 is the initial point, x_k the final point, k the number of times of iteration for each problem.

These three tables show the performance of these three methods relative to the iterations, It is easily to see that, for each algorithm, are all very efficient, especially for the problems such as s201, s207, s240, s311. The results for each problem are accurate, and with less number of times of iteration.

5. Conclusions

We have proposed three new hybrid conjugate gradient (CG) methods, that are, the H3 method, the MCD me-

Table 1. The detail information of numerical experiments for H3 algorithm.

No.	x_0	x_k	$\left\| g_k \right\|$	k
S201	(8,9)	(5.0000000, 6.0000000)	7.08791050e-007	25
S205	(1,1)	(2.9999973, 0.4999993)	8.17619783e-007	188
S207	(−1.2,1)	(0.9999993, 0.9999983)	7.08905387e-007	61
S240	(100,−1, 2.5)	(1.3367494e-007, −1.3367494e-009, 3.3418736e-009)	8.81057842e-007	29
S311	(1, 1)	(2.9999999, 2.0000000)	3.63497147e-007	20
S314	(2,2)	(1.8064954, 1.3839575)	9.83714228e-007	339

Table 2. The detail information of numerical experiments for MCD algorithm.

No.	x_0	x_k	$\left\| g_k \right\|$	k
S201	(8,9)	(5.0000001,5.9999999)	9.53649845e-007	34
S205	(1,1)	(2.9999968,0.4999992)	9.97179740e-007	253
S207	(−1.2,1)	(0.9999992,0.9999979)	8.41893782e-007	151
S240	(100,−1,2.5)	(−9.909208e-008, 3.1120991e-008, 2.660865e-008)	8.37563602e-007	41
S311	(1,1)	(2.9999999,2.0000000)	7.05200476e-007	24
S314	(2,2)	(1.8064954,1.3839575)	9.94928488e-007	130

Table 3. The detail information of numerical experiments for NH3 algorithm.

No.	x_0	x_k	$\left\| g_k \right\|$	k
S201	(8, 9)	(5.0000001,5.9999999)	9.53649845e-007	34
S205	(1,1)	(2.9999972,0.4999993)	9.89188900e-007	418
S207	(−1.2,1)	(0.9999990,0.99999751)	9.59168006e-007	168
S240	(100,−1,2.5)	(-9.9092086e-008, 3.1120991e-008, 2.6608656e-008)	8.37563602e-007	41
S311	(1,1)	(2.9999999,2.0000000)	5.87169265e-007	25
S314	(2,2)	(1.8064954,1.3839575)	9.82136064e-007	339

thod and the NH3 method, where the last two methods produce sufficient descent search direction at every iteration. This property depends neither on the line search used nor on the convexity of the objective function. Under suitable conditions, we proposed the global convergence of these three new methods even for nonconvex minimization. And numerical experiments in section 4 showed that the new three algorithms are all efficient for the given test problems.

6. References

[1] R. Fletcher and C. Reeves, "Function Minimization by Conjugate Gradients," *The Computer Journal*, Vol. 7, No. 2, 1964, pp. 149-154. doi:10.1093/comjnl/7.2.149

[2] R. Fletcher, "Practical Methods of Optimization, Unconstrained Optimization," Wiley, New York, 1987.

[3] Y. H. Dai and Y. Yuan, "A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property," *SIAM Journal on Optimization*, Vol. 10, No. 1, 1999, pp. 177-182. doi:10.1137/S1052623497318992

[4] M. R. Hestenes and E. L. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, Vol. 49, No. 6, 1952, pp. 409-432.

[5] B. Polak and G. Ribiere, "Note Surla Convergence des Méthodes de Directions Conjuguées," *Revue Francaise d'Informatique et de Recherche Opérationnelle*, Vol. 16,

No. 1, 1969, pp. 35-43.

[6] B. T. Polyak, "The Conjugate Gradient Method in Extreme Problems," *USSR Computational Mathematics and Mathematical Physics*, Vol. 9, No. 4, 1969, pp. 94-112. doi:10.1016/0041-5553(69)90035-4

[7] Y. L. Liu and C. S. Storey, "Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory," *Journal of Optimization Theory and Applications*, Vol. 69, No. 1, 1991, pp. 129-137. doi:10.1007/BF00940464

[8] Y. Yuan and W. Sun, "Theory and Methods of Optimization," Science Press of China, Beijing, 1999.

[9] Y. H. Dai and Y. Yuan, "Nonlinear Conjugate Gradient Methods," Shanghai Scientific and Technical Publishers, Shanghai, 1998.

[10] G. Zoutendijk, "Nonlinear Programming, Computational Methods," In: J. Abadie Ed., *Integer and Nonlinear Programming,* North-Holland, Amsterdam, 1970, pp. 37-86.

[11] M. Al-Baali, "Descent Property and Global Convergence of the Fletcher–Reeves Method with Inexact Line Search," *IMA Journal of Numerical Analysis*, Vol. 5, No. 1, 1985, pp.121-124. doi:10.1093/imanum/5.1.121

[12] M. J. D. Powell, "Nonconvex Minimization Calculations and the Conjugate Gradient Method," *Lecture Notes in Mathematics*, Vol. 1066, No. 122, 1984, pp. 121-141.

[13] D. Touati-Ahmed and C. Storey, "Efficient Hybrid Conjugate Gradient Techniques," *Journal of Optimization Theory and Applications*, Vol. 64, No. 2, 1990, pp. 379–397. doi:10.1007/BF00939455

[14] Y. H. Dai and Y. Yuan, "An Efficient Hybrid Conjugate Gradient Method for Unconstrained Optimization," *Annals of Operations Research*, Vol. 103, No. 1-4, 2001,

pp. 33-47. doi:10.1023/A:1012930416777

[15] J. C. Gilbert and J. Nocedal, "Global Convergence Properties of Conjugate Gradient Methods for Optimization," *SIAM Journal Optimization*, Vol. 2, No. 1, 1992, pp. 21-42. doi:10.1137/0802003

[16] W. W. Hager and H. Zhang, "A New Conjugate Gradient Method with Guaranteed Descent and an Efficient Line Search," *SIAM Journal Optimization*, Vol. 16, No. 1, 2005, pp. 170-192. doi:10.1137/030601880

[17] L. Zhang, W. J. Zhou and D. H. Li, "A Descent Modified Polak-Ribière-Polyak Conjugate Gradient Method and Its Global Convergence," *IMA Journal of Numerical Analysis*, Vol. 26, No. 4, 2006, pp. 629-640. doi:10.1093/imanum/drl016

[18] L. Zhang, W. J. Zhou and D. H. Li, "Global Convergence of a Modified Fletcher-Reeves Conjugate Method with Armijo-Type Line Search," *Numerische Mathematik*, Vol. 104, No. 4, 2006, pp. 561-572. doi:10.1007/s00211-006-0028-z

[19] L. Zhang, "Nonlinear Conjugate Gradient Methods for Optimization Problems," Ph.D. Thesis, Hunan University, 2006.

[20] L. Zhang and W. J. Zhou, "Two Descent Hybrid Conjugate Gradient Methods for Optimization," *Journal of Computational and Applied Mathematics*, Vol. 216, No. 1, 2008, pp. 251-264. doi:10.1016/j.cam.2007.04.028

[21] P. Wolfe, "Convergence Conditions for Ascent Methods," *SIAM Review*, Vol. 11, No. 2, 1969, pp. 226-235. doi:10.1137/1011036

[22] W. Hock and K. Schittkowski, "Test Examples for Nonlinear Programming Codes," *Journal of Optimization Theory and Applications*, Vol. 30, No. 1, 1981, pp. 127-129. doi:10.1007/BF00934594