Vol.3, No.1, 85-90 (2011)
http://dx.doi.org/10.4236/ns.2011.31012

Natural Science

# A hybrid conjugate gradient method for optimization problems*

**Xiangrong Li, Xupei Zhao**

Department of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, China; xrli68@163.com

Received 14 October 2010; revised 18 November 2010; accepted 22 November 2010.

## ABSTRACT

**A hybrid method of the Polak-Ribière-Polyak (PRP) method and the Wei-Yao-Liu (WYL) method is proposed for unconstrained optimization problems, which possesses the following properties: i) This method inherits an important property of the well known PRP method: the tendency to turn towards the steepest descent direction if a small step is generated away from the solution, preventing a sequence of tiny steps from happening; ii) The scalar $\beta_k \geq 0$ holds automatically; iii) The global convergence with some line search rule is established for nonconvex functions. Numerical results show that the method is effective for the test problems.**

**Keywords:** Line Search; Unconstrained Optimization; Conjugate Gradient Method; Global Convergence

## 1. INTRODUCTION

We are interested to consider the unconstrained optimization problem

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

where $f: \Re^n \to \Re$ is continuously differentiable. It is well known that there are many methods for solving optimization problems (see [24,26,28-32,34] etc.), where the conjugate gradient(CG) method is a powerful line search method because of its simplicity and its very low memory requirement, especially for the large scale optimization problems [22,23,27], which can avoid, like steepest descent method, the computation and storage of some matrices associated with the Hessian of objective functions. The following iterative formula is often used by the nonlinear CG method

$$x_{k+1} = x_k + \alpha_k d_k, k = 0,1,2,\cdots \tag{1.2}$$

for (1.1), where $x_k$ is the current iterate point, $\alpha_k > 0$ is a steplength, and $d_k$ is the search direction designed by

$$d_k = \begin{cases} -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1 \\ -g_k, & \text{if } k = 0 \end{cases}, \tag{1.3}$$

where $\beta_k \in \Re$ is a scalar which determines the different conjugate gradient methods [4,5,8,9,12,13,15,16,18, 20,21,25,33] etc., and $g_k$ is the gradient of $f(x)$ at the point $x_k$. The well-known formula for $\beta_k$ from the computation point of view is the following PRP method

$$\beta_k^{PRP} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|^2}, \tag{1.4}$$

where $g_k$ and $g_{k+1}$ are the gradients $\nabla f(x_k)$ and $\nabla f(x_{k+1})$ of $f(x)$ at the point $x_k$ and $x_{k+1}$, respectively, and $\|\cdot\|$ denotes the Euclidian norm of vectors. Throughout this paper, we also denote $f(x_k)$ by $f_k$. Polak and Ribèire [18] proved that this method with the exact line search is globally convergent when the objective function is convex. Powell [19] gave a counter example to show that there exist nonconvex functions on which the **PRP** method does not converge globally even the exact line search is used. He suggested that $\beta_k$ should not be less than zero. Considering this suggestion, Gilbert and Nocedal [10] proved that the modified **PRP** method $\beta_k^+ = \max\{0, \beta_k^{PRP}\}$ is globally convergent with the weak Wolfe-Powell (**WWP**) line search technique and the assumption of sufficient descent condition. However, the global convergence of the **PRP** method is still open under the **WWP** line search rule.

Recently, Wei, Yao, and Liu(**WYL**) [21] propose a new conjugate gradient formula

$$\beta_k^{WYL} = \frac{g_{k+1}^T \left( g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k \right)}{\|g_k\|^2} \tag{1.5}$$

It is not difficult to deduce that

*This work is supported by China NSF grands 10761001 and Guangxi SF grands 0991028.

$$\beta_k^{WYL} = \frac{g_{k+1}^T\left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k\right)}{\|g_k\|^2} \geq \frac{\|g_{k+1}\|^2 - \|g_{k+1}\| \frac{\|g_{k+1}\|}{\|g_k\|}\|g_k\|}{\|g_k\|^2}$$
$$= 0$$

is true. The numerical results show that this method is competitive to the **PRP** method for the test problems of [17]. Under the sufficient descent condition, this method is globally convergent with the **WWP** line search.

These observations make us know that the sufficient descent condition

$$g_k^T d_k \leq -c\|g_k\|^2, c > 0 \textit{ is a constant holds for all } k \geq 0 \quad (1.6)$$

is very important to ensure the global convergence [1,2,10,14], and the scalar $\beta_k \geq 0$ also plays a very important role [10,19]. This motivates us to propose a hybrid method combining the **PRP** method and the **WYL** method. The hybrid method will possess some better properties of the **PRP** method and the **WYL** method: (i) the tendency to turn towards the steepest descent direction if a small step is generated away from the solution, preventing a sequence of tiny steps from happening; (ii) The scalar $\beta_k \geq 0$ holds automatically. The global convergence with the **WWP** line search of the presented method is established for nonconvex objective function. Numerical results show that this given method is competitive to the **PRP** method and the **WYL** method.

This paper is organized as follows. In the next section, the algorithm is stated. The global convergence is proved in Section 3, and the numerical results are reported in Section 4. The last section gives one conclusion.

## 2. ALGORITHM

Now we describe the given algorithm as follows. Here we call it Algorithm 1.

**Algorithm 1** (The hybrid algorithm of the **PRP** method and the **WYL** method)

Step 0: Choose an initial point $x_0 \in \Re^n, \varepsilon \in (0,1)$. Set $d_0 = -g_0 = -\nabla f(x_0)$, $k := 0$.

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

Step 2: Compute step size $\alpha_k$ by some line search rules.

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

Step 4: Calculate the search direction

$$d_{k+1} = -g_{k+1} + \beta_k^{P-W} d_k, \quad (2.1)$$

where $\beta_k^{P-W} = \max\left\{\beta_k^{PRP}, \beta_k^{WYL}\right\}$.

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

**Remark** i) If $x_{k+1} \approx x_k$, we have $g_{k+1} \approx g_k$ and $\|g_{k+1}\| \approx \|g_k\|$ which imply that $\beta_k^{PRP} \to 0$, and $\beta_k^{WYL} \to 0$, which means that $\beta_k^{P-W} \to 0$ if a small step is generated for all $k \geq 0$. Thus the given method inherits the better property of the **PRP** method: the directions will turn out to be the steepest descent directions if the tiny steps from happening.

ii) By the definition of the new formula $\beta_k^{P-W}$, we have

$$\beta_k^{P-W} = \max\left\{\beta_k^{PRP}, \beta_k^{WYL}\right\} \geq \beta_k^{WYL}$$
$$\geq \frac{\|g_{k+1}\|^2 - \|g_{k+1}\| \frac{\|g_{k+1}\|}{\|g_k\|}\|g_k\|}{\|g_k\|^2}$$
$$= 0$$

## 3. THE GLOBAL CONVERGENCE

The following assumptions are often needed to prove the convergence of the nonlinear conjugate gradient methods (see [5,9,10,20,21] etc.).

**Assumption 3.1** i) The function $f(x)$ has a lower bound on the level set $\Omega = \left\{x \in \Re^n \middle| f(x) \leq f(x_0)\right\}$, where $x_0$ is a given point and $\Omega$ is bounded.

ii) In an open convex set $\Omega_0$ that contains $\Omega$, *J* is differentiable and its gradient *g* is Lipschitz continuous, namely, there exists a constants $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \ \forall x, y \in \Omega_0. \quad (3.1)$$

### 3.1. The global Convergence with the Weak Wolfe-Powell Line Search

The weak Wolfe-Powell (**WWP**) search rule is to find a step length $\alpha_k$ such that

$$f(x_k + \alpha_k d_k) \leq f_k + \delta\alpha_k g_k^T d_k \quad (3.2)$$

and

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T, \quad (3.3)$$

where $\delta \in (0,1/2)$, $\delta \in (\delta,1)$. This line search technique is often used to study the convergence of conjugate gradient algorithms [6,27,34]. At present, the global convergence of the **PRP** method with the **WWP** line search is still open.

**Lemma 3.1** *Suppose that Assumption* 3.1 *holds. Let the sequence* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 1*,* $g_k^T d_k \leq 0$*, and the stepsize* $\alpha_k$ *be determined* *by the **WWP** line search* (3.2) *and* (3.3) *Then the zoutendijk condition* [34]

$$\sum_{k=0}^{\infty}\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2}<+\infty \tag{3.4}$$

*holds*.

**Proof.** By (3.3) and Assumption 3.1 ii), we have

$$-(1-\sigma)g_k^T d_k \le \left(g_{k+1}-g_k\right)^T d_k \le \alpha_k L\|d_k\|^2,$$

this means that $\alpha_k \ge -(1-\sigma)g_k^T d_k / L\|d_k\|^2$, which together with $g_k^T d_k \le 0$, and (3.2) implies that

$$\frac{(1-\sigma)\left(g_k^T d_k\right)^2}{L\|d_k\|^2}\le f_k - f_{k+1},$$

summing up this inequality from $k=0$ to $\infty$, and using Assumption 3.1 i), we can obtain this lemma. This completes the proof.

We will prove the global convergence of Algorithm 1 by contradiction. Then we assume that there exists a positive constant $\gamma>0$ such that

$$\|g_k\|\ge\gamma,\ \forall k\ge 0. \tag{3.5}$$

Using (3.5) deduces a contradiction to obtain our conclusion.

Similar to Lemma 3.3.1 in [6], based on Assumption 3.1, Lemma 3.1, the fact $\beta_k^{P-W}\ge 0$, and (3.5), we can get the following lemma.

**Lemma 3.2** *Let Assumption* 3.1 *hold and the sequences* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 1. *The sufficient descent condition* (1.6) *holds, and the stepsize* $\alpha_k$ *is determined by* (3.2) *and* (3.3). *Suppose that the inequalities* (3.5) *is true. Then we have* $d_k\ne 0$ *and*

$$\sum_{k=0}^{\infty}\|u_{k+1}-u_k\|^2<\infty,$$

*where* $u_k=\dfrac{d_k}{\|d_k\|}$.

**Proof.** These two inequalities (1.6) and (3.5) imply that $d_k\ne 0$ is true, for otherwise $g_k=0$, then $u_k=d_k/\|d_k\|$ is reasonable. Denote

$$r_{k+1}=-\frac{g_{k+1}}{d_{k+1}},\ \delta_k=\beta_k^{P-W}\frac{\|d_k\|}{\|d_{k+1}\|}$$

By (2.1), for $k\ge 0$, we have

$$u_{k+1}=r_{k+1}+\delta_k u_k,$$

this combining with $\|u_{k+1}\|=\|u_k\|=1$ shows that

$$\|r_{k+1}\|=\|u_{k+1}-\delta_k u_k\|=\|\delta_k u_{k+1}-u_k\| \tag{3.6}$$

The inequality $\beta_k^{P-W}\ge 0$ implies that $\delta_k\ge 0$ is true, then it follows that from (3.6) and triangular inequality

$$\begin{aligned}&\|u_{k+1}-u_k\|\\ &\le\|(1+\delta_k)u_{k+1}-(1+\delta_k)u_k\|\\ &\le\|u_{k+1}-\delta_k u_k\|+\|\delta_k u_{k+1}-u_k\|\\ &=2\|r_{k+1}\|.\end{aligned} \tag{3.7}$$

By (1.6) and (3.4), we get

$$\sum_{k\ge 0}\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}=\sum_{k\ge 1}\|r_{k+1}\|^2\|g_{k+1}\|^2<\infty$$

Which together with (3.5), we obtain

$$\sum_{k\ge 0}\|r_{k+1}\|^2<\infty$$

By the above inequality and (3.7), we get this lemma. The proof is complete.

The following property (*) was introduced by Gilbert and Nocedal [10], which pertains to the $\beta_k^{+}$ under the sufficient descent condition. The **WYL** formula also has this property. Now we show that this property (*) pertains to our method.

**Property** (*). Suppose that

$$0<r_1\le\|g_k\|\le r_2. \tag{3.8}$$

We say that the method has Property (*), if for all $k$, there exists constants $b>1$ and $\lambda>0$ such that $|\beta_k|\le b$ and

$$\|s_k\|\le\lambda\Rightarrow|\beta_k|\le\frac{1}{2b}$$

**Lemma 3.3** *Let Assumption* 3.1 *hold and the sequences* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 1. *Then the new formula* $\beta_k^{P-W}$ *possesses property* (*).

**Proof.** The result of this lemma is proved by the following two cases.

**Case i:** we consider $\beta_k^{PRP}$ By (3.1), we have

$$\left|\beta_k^{PRP}\right|=\frac{g_{k+1}^T\left(g_{k+1}-g_k\right)}{\|g_k\|^2}\le\frac{L\|g_{k+1}\|.\|s_k\|}{\|g_k\|^2}. \tag{3.9}$$

From Assumption 3.1 i), then there exists a constant $M_1>0$ such that

$$\|s_k\|\le M_1. \tag{3.10}$$

Let $b=\max\left\{2,\left(Lr_2/r_1^2\right)M_1\right\}>1$ and $\lambda=\gamma_1^2/2b\left(L\gamma_2\right)$, it follows that

$$\left|\beta_k^{PRP}\right|\le b$$

and

$$\left|\beta_k^{PRP}\right| \le \frac{L\|g_{k+1}\|\cdot\|s_k\|}{\|g_k\|^2} \le \frac{2L\gamma_2}{\gamma_1^2}\|s_k\| \le \frac{2L\gamma_2}{\gamma_1^2}\lambda = \frac{1}{2b}.$$

Then the $PRP$ formula $\beta_k^{PRP}$ has this property (*).

**Case ii:** let us consider $\beta_k^{WYL}$. Denote $Y_k = g_{k+1} - \|g_{k+1}\|/\|g_k\| g_k$, by (3.1), we get

$$\begin{aligned}\|Y_k\| &= \left\| g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k \right\| \\ &= \left\| g_{k+1} - g_k + g_k - \frac{\|g_{k+1}\|}{\|g_k\|} g_k \right\| \\ &\le \|g_{k+1} - g_k\| + \left| \|g_k\| - \|g_{k+1}\| \right| \quad (3.11) \\ &\le \|g_{k+1} - g_k\| + \|g_{k+1} - g_k\| \\ &\le 2L\|s_k\|,\end{aligned}$$

By (1.5), (3.11), (3.10) and (3.8) we have

$$\left|\beta_k^{WYL}\right| = \left|\frac{g_{k+1}^T Y_k}{\|g_k\|^2}\right| \le \frac{\|g_{k+1}\|\cdot\|Y_k\|}{\|g_k\|^2} \le \frac{2L\gamma_2\|s_k\|}{\gamma_1^2}, \quad (3.12)$$

let $b = \max\left\{2, \left(2L\gamma_2/\gamma_1^2\right)M_1\right\}$ and $\lambda = \gamma_1^2/2b\left(2L\gamma_2\right)$, it follows that (3.12) and the definition of $b$ and $\lambda$ that $b > 1$

$$\left|\beta_k^{WYL}\right| \le b, \text{ and } \left|\beta_k^{WYL}\right| \le \left(\frac{2L\gamma_2}{\gamma_1^2}\right)\|s_k\| \le \left(\frac{2L\gamma_2}{\gamma_1^4}\right)\lambda = \frac{1}{2b}$$

Thus, the formula $\beta_k^{WYL}$ also has the property (*).

Using the definition of the $\beta_k^{P-W} = \max\left\{\beta_k^{WYL}, \beta_k^{PRP}\right\}$, we conclude that the formula $\beta_k^{P-W}$ possesses the property (*). The proof is complete.

By Lemma 3.3, similar to Lemma 3.3.2 in [6], it is not difficult to prove the following result. Here we only state it as follows, but omit the proof.

**Lemma 3.4** (*Lemma* 3.3.2 *in* [6]) *Let the sequences* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm 1 and the conditions in Lemma* **3.3** *hold. If* $\beta_k^{P-W} > 0$ *and has property* (*), *then there exists a constant* $\lambda > 0$ *such that, for any* $\Delta \in N$ *and any index* $k_0$ *there is an index* $k > k_0$ *satisfying*

$$\left|k_{k,\Delta}^{\lambda}\right| > \frac{\lambda}{2},$$

*where* $k_{k,\Delta}^{\lambda} = \left\{i \in N : k \le i \le k + \Delta - 1, \|s_i\| > \lambda\right\}$, $N$ *denotes the set of positive integers, and* $\left|k_{k,\Delta}^{\lambda}\right|$ *denotes the numbers of elements in* $k_{k,\Delta}^{\lambda}$.

Finally, by Lemma 3.2 and Lemma 3.4, we present the global convergence theorem of Algorithm 1 with the **WWP** line search. Similar to Theorem 3.3.3 in [6], it is not difficult to prove the result, here we also give the process of the proof.

**Theorem 3.1** *Let the sequence* $\{g_k, d_k\}$ *be generated by Algorithm* 1 *with the weak Wolfe-Powell line search and the conditions in Lemma* 3.3 *hold. Then* $\lim_{k\to\infty}\inf\|g_k\| = 0$.

**Proof.** We will get this theorem by contradiction. Suppose that (3.5) is true, then the conditions in Lemma 3.2 and 3.3 hold. By Assumption 3.1 i), then there exists a constant $\xi_0 > 0$ such that

$$\|x\| \le \xi_0, \forall x \in \Omega \quad (3.13)$$

We also denote $u_i = d_i / \|d_i\|$, then for all integers $l, k\left(l \ge k\right)$, we have

$$\begin{aligned}x_l - x_{k-1} &= \sum_{i=k}^{l} \|s_{i-1}\| u_i \\ &= \sum_{i=k}^{l} \|s_{i-1}\| u_{k-1} + \sum_{i=k}^{l}\left(u_{i-1} - u_{k-1}\right).\end{aligned}$$

Taking the norm in both sides of the above equality, and using (3.13) we get

$$\sum_{i=k}^{l}\|s_{i-1}\| \le 2\xi_0 + \sum_{i=k}^{l}\|s_{i-1}\|\cdot\|u_{i-1} - u_{k-1}\|$$

Let $\Delta = \left[8\xi_0/\lambda\right]$ be the smallest integer where $\Delta$ does not less than $8\xi_0/\lambda$. By Lemma 3.2, there exists an index $k_0$ such that

$$\sum_{i\ge k}\|u_{i+1} - u_i\|^2 \le \frac{1}{4\Delta} \quad (3.14)$$

On the other hand, by Lemma 3.3, there exists $k \ge k_0$ satisfying

$$\left|k_{k,\Delta}^{\lambda}\right| > \frac{\Delta}{2} \quad (3.15)$$

For all $i \in \left[k, k+\Delta-1\right]$, by Cauchy-Schwarz inequality and (3.14), we obtain

$$\begin{aligned}\|u_{i-1} - u_{k-1}\| &\le \sum_{j=k}^{i-1}\|u_j - u_{j-1}\| \\ &\le \left(i-k\right)^{\frac{1}{2}}\left(\sum_{j=k}^{i-1}\|u_j - u_{j-1}\|^2\right)^{\frac{1}{2}} \\ &\le \Delta^{\frac{1}{2}}\left(\frac{1}{4\Delta}\right)^{\frac{1}{2}} = \frac{1}{2}.\end{aligned}$$

By the above inequality, (3.15) and (3.13), we have

$$2\xi_0 \ge \frac{1}{2}\sum_{i=k}^{k+\Delta-1}\|s_{i-1}\| > \frac{\lambda}{2}\left|k_{k,\Delta}^{\lambda}\right| > \frac{\lambda\Delta}{4},$$

Thus $\Delta < 8\xi_0/\lambda$, this contradicts with the definition of $\Delta$. Therefore, the conclusion of this theorem is right. This completes the proof.

## 4. NUMERICAL RESULTS

In this section, we report some numerical experiments.

The unconstrained optimization problems with the given initial points can be found at:

www.ici.ro/camo/neculai/SCALCG/testuo.pdf,

which were collected by Neculai Andrei. Since this new method is the hybrid method of the **PRP** method and the **WYL** method, we test Algorithm 1 with the **WWP** line search and compare its performance with those of the **WYL** [21] and the PRP [18] methods. The stop criterions are given below: we stop the program if the inequality $\|g(x_k)\| \le \varepsilon$ is satisfied or the inequality

$$\|g(x_k)\| \le \varepsilon\left(1+\left|f(x_k)\right|\right)$$

is satisfied, where $\varepsilon = 1.0\,D-5$. All the codes were written in Fortran and run on **PC** with 2.60 GHz **CPU** processor and **256 MB** memory and **Windows XP** operation system. **I**n the experiments, the parameters were chosen as $\delta = 1.0\,D-2$, $\sigma = 1.0\,D-1$. The dimension of the test problems is from 500 to 5000. The detailed numerical results are listed on the web site

http://210.36.18.9:8018/publication.asp?id=35392.

In **Figure 1**, "**WYL**", "**PRP**", and "**MPRP-WYL**" stand for the **WYL** method, the **PRP** method, and the new method, respectively.

**Figure 1** shows the performance of these methods relative to the iterative number of the function and gradient(**NFN**), which were evaluated using the profiles of Dolan and Moré [7]. It is easy to see that the ***MPRP-WYL*** is predominant among these three methods and the new method can solve about 99% of the test problems successfully. The ***PRP*** method is better than the ***WYL*** method for $1 \le t \le 1.2$ and the ***WYL*** method is better than the ***PRP*** method for $1.2 \le t \le 6$. Moreover, the ***PRP*** method solves about 98% of the test problems and the ***WYL*** method solve about 99% of the test problems successfully, respectively. In a word, the given method is competitive to the other two methods and the hybrid formula is notable.

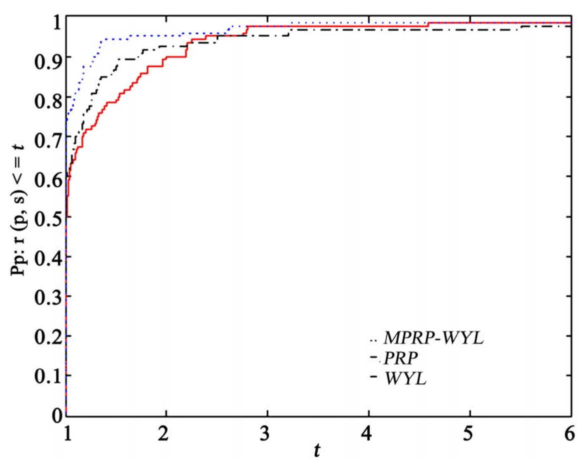


**Figure 1.** Performance profiles of conjugate gradient methods in Table 1 (NFN).

## 5. CONCLUSION

This paper gives a hybrid conjugate gradient method for solving unconstrained optimization. The global convergence for nonconvex functions with the **WWP** line search is established. The numerical results show that the given method is competitive to the other standard conjugate gradient methods for the test problems.

For further research, we should study the convergence of the new algorithm under other line search rules. Moreover, more numerical experiments and testing environments (such that [3]) for large practical problems should be done in the future.

## REFERENCES


[1] Ahmed, T. and Storey, D. (1990) Efficient hybrid conjugate gradient techniques. *Journal of Optimization Theory and Applications*, **64**, 379-394. doi:10.1007/BF00939455

[2] Al-Baali, A. (1985) Descent property and global convergence of the Flecher-Reeves method with inexact line search. *IMA Journal of Numerical Analysis*, **5**, 121-124. doi:10.1093/imanum/5.1.121

[3] Bongartz, K.E., Conn, A.R., Gould, N.I.M. and Toint, P.L. (1995) CUTE: Constrained and unconstrained testing environments. *ACM Transactions on Mathematical Software*, **21**, 123-160. doi:10.1145/200979.201043

[4] Dai, Y. (2002) A nonmonotone conjugate gradient algorithm for unconstrained optimization. *Journal of Systems Science and Complexity*, **15**, 139-145.

[5] Dai, Y. and Yuan, Y. (2000) A nonlinear conjugate gradient with a strong global convergence properties. *SIAM Journal on Optimization*, **10**, 177-182. doi:10.1137/S1052623497318992

[6] Dai, Y. and Yuan, Y. (1998) Nonlinear conjugate gradient methods. Shanghai Scientific and Technical Publishers, Shanghai.

[7] Dolan, E.D. and Moré, J.J. (2002) Benchmarking optimization software with performance profiles. *Mathematical Programming*, **91**, 201-213. doi:10.1007/s101070100263

[8] Fletcher, R. (1997) Practical Method of Optimization, Vol I: Unconstrained Optimization. 2nd Edition, Wiley, New York.

[9] Fletcher, R. and Reeves, C. (1964) Function minimization bu conjugate gradients. *Computer Journal*, **7**, 149-154. doi:10.1093/comjnl/7.2.149

[10] Gibert, J.C. and Nocedal, J. (1992) Global convergence properties of conugate gradient methods for optimization. *SIAM Journal on Optimization*, **2**, 21-42. doi:10.1137/0802003

[11] Grippo, L. and Lucidi, S. (1997) A globally convergent version of the Polak-Ribière gradient method. *Mathe-*

*matical Programming*, **78**, 375-391. doi:10.1007/BF02614362

[12] Hager, W.W. and Zhang, H.C. (2005) A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on Optimization*, **16**, 170-192. doi:10.1137/030601880

[13] Hestenes, M.R. and Stiefel, E. (1952) Method of conjugate gradient for solving linear equations. *Journal of Research National Bureau of Standards*, **49**, 409-436.

[14] Hu, Y.F. and Storey, C. (1991) Global convergence result for conjugate method. *Journal of Optimization Theory and Applications*, **71**, 399-405. doi:10.1007/BF00939927

[15] Li, G., Tang, C. and Wei, Z. (2007) New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. *Journal of Computational and Applied Mathematics*, **2**, 523-539. doi:10.1016/j.cam.2006.03.005

[16] Liu, Y. and Storey, C. (1992) Effcient generalized conjugate gradient algorithms, part 1: theory. *Journal of Optimization Theory and Applications*, **69**, 17-41.

[17] Moré, J.J., Garbow, B.S. and Hillstrome, K.E. (1981) Testing unconstrained optimization software. *ACM Transactions on Mathematical Software*, **7**, 17-41. doi:10.1145/355934.355936

[18] Polak, E. and Ribiere, G. (1969) Note sur la xonvergence de directions conjugees. *Rev. Francaise informat Recherche Operatinelle*, 3e Annee, **16**, 35-43.

[19] Powell, M.J.D. (1984) Nonconvex minimization calculations and the conjugate gradient method. *Lecture Notes in Mathematics*, **1066**, 122-141.

[20] Wei, Z., Li, G. and Qi, L. (2006) New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems. *Applied Mathematics and Computation*, **179**, 407-430. doi:10.1016/j.amc.2005.11.150

[21] Wei, Z., Yao, S. and Liu, L. (2006) The convergence properties of some new conjugate gradient methods. *Applied Mathematics and Computation*, **183**, 1341-1350. doi:10.1016/j.amc.2006.05.150

[22] Yu, G.H. (2007) Nonlinear self-scaling conjugate gradient methods for large-scale optimization problems. Thesis of Doctor's Degree, Sun Yat-Sen University.

[23] Yuan, G.L. (2009) Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. *Optimization Letters*, **3**, 11-21. doi:10.1007/s11590-008-0086-5

[24] Yuan, G.L. and Lu, X.W. (2008) A new line search method with trust region for unconstrained optimization. *Communications on Applied Nonlinear Analysis*, **15**, 35-49.

[25] Yuan, G.L. and Lu, X.W. (2009) A modified PRP conjugate gradient method. *Annals of Operations Research*, **166**, 73-90. doi:10.1007/s10479-008-0420-4

[26] Yuan, G.L., Lu, X.W. and Wei, Z.X. (2007) New two-point stepsize gradient methods for solving unconstrained optimization problems. *Natural Science Journal of Xiangtan University*, **29**, 13-15.

[27] Yuan, Y. and Sun, W. (1999) Theory and Methods of Optimization. Science Press of China, Beijing.

[28] Yuan, G.L. and Wei, Z.X. (2009) New line search methods for unconstrained optimization. *Journal of the Korean Statistical Society*, **38**, 29-39. doi:10.1016/j.jkss.2008.05.004

[29] Yuan, G.L. and Wei, Z.X. (2008) The superlinear convergence analysis of a nonmonotone BFGS algorithm on convex objective functions. *Acta Mathematica Sinica*, **24**, 35-42. doi:10.1007/s10114-007-1012-y

[30] Yuan, G.L. and Wei, Z.X. (2010) Convergence analysis of a modified BFGS method on convex minimizations. *Computational Optimization and Applications*, **47**, 237-255. doi:10.1007/s10589-008-9219-0

[31] Yuan, G.L. and Wei, Z.X. (2009) A Rank-One fitting method for unconstrained optimization problems. *Mathematica Applicata*, **22**, 118-122.

[32] Zhang, H.C. and Hager, W.W. (2004) A nonmonotone line search technique and its application to unconstrained optimization. *SIAM Journal on Optimization*, **14**, 1043-1056. doi:10.1137/S1052623403428208

[33] Zhang, L., Zhou, W. and Li, D. (2006) A descent modified Polak-Ribiére-Polyak conjugate method and its global convergence. *IMA Journal on Numerical Analysis*, **26**, 629-649. doi:10.1093/imanum/drl016

[34] Zoutendijk, G. (1970) Nonlinear programming computational methods. In: Abadie, J. Ed., *Integer and Nonlinear Programming*, NorthHolllad, Amsterdam, 37-86.

 