International Journal of Pure and Applied Mathematics
Volume 92 No. 2 2014, 225-242
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v92i2.7

ijpam.eu

NEW CONVERGENCE METHOD WITH NONMONOTONE LINE SEARCH

Tahar Bouali[1 §], Yamina Laskri[2]

[1,2]Department of Mathematics
Badji Mokhtar University
Annaba 23000, ALGERIA

Abstract: In this paper, an efficient new nonlinear conjugate gradient method is proposed for the unconstrained optimization problems, which possesses the following property: the sufficient descent condition $g_k^T d_k = -\|g_k\|^2$ holds without any line search. Under the strong Wolf Non-monotone line search, we proved the global convergence of the FR method for strongly convex functions.

The numerical experiments show that the FR method is especially efficient.

AMS Subject Classification: 65K10, 90C30
Key Words: conjugate gradient, sufficient descent, non-monotone line search, global convergence,unconstrained optimization

1. Introduction

The first objective of this paper is to study the global convergence and practical computational performance of FR conjugate gradient method with strong wolf Nonmonotone line search for nonlinear unconstrained optimization .Consider the following unconstrained optimization problem

$$\min \{f(x) : x \in \mathbb{R}^n\} \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuosly differentiable. The general form of the conju-

Received: December 22, 2013

§Correspondence author

gate gradient method

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

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

where $g_k = \nabla f(x_k)$, α_k is a stepe-size obtained by some line search, and β_k [4, 5, 7, 12] is a scalar There are many ways to select β_k, and some well-known formulas are given by

$$\beta_k^{PRP} = \frac{g_k^T y_k}{\|g_{k-1}\|^2} \tag{1.4}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \tag{1.5}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \tag{1.6}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \tag{1.7}$$

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - g_{k-1})} \tag{1.8}$$

respectively, where $y_{k-1} = g_k - g_{k-1}$and $\|\|$ means the Euclidean norm.

In the convergence analysis and implementation of conjugate gradient method, the extended strong wolf Non-monotone line search, namely

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq m(k)} f(x_{k-j}) + \delta \alpha_k g_k^T d_k \tag{1.9}$$

$$\sigma_1 g_k^T d_k \leq g\left(x_{k+}\alpha_k d_k\right)^T d_k \leq -\sigma_2 g_k^T d_k, \tag{1.10}$$

which $0 < \delta \leq \sigma_1 \leq 1, 0 \leq \sigma_2 \leq 1, 0 \leq \lambda \leq 1$ and $M_0 \epsilon N$,

$$0 \leq m(k) \leq \min\{m(k-1) + 1, M_0\}, \quad m(0) = 0.$$

In addition, the sufficient descent condition, namely

$$g_k^T d_k \leq -c \|g_k\|^2 \tag{1.11}$$

When f is strongly convex, the stepsize α_k satisfies the following new non-monotone line search condition

$$f(x_k+\alpha_k d_k) - \lambda \max_{0\le j\le m(k)} f(x_{k-j}) - (1-\lambda) \min_{0\le j\le m(k)} f(x_{k-j}) \le \delta \alpha_k g_k^T d_k \quad (1.12)$$

The convergence of (1.5),(1.6)and (1.7) with some line search condition have been studied by many authors for many years. Al-Baali [1] proved the global convergence of FR method with the strong wolf line search.Liu and al [8] .and Dai and Yaun [3] extended this result to $\sigma = \frac{1}{2}$, the PRP method is generally believed to the most efficient conjugate gradient method. However, Powell [13] constructed a counter example and showed that the PRP method can circle infinitely without approaching the solution,which implies that this method is not globally convergent for general function.But the PRP + $\left(B_k^{PRP+} = \max\left\{0, B_k^{PRP}\right\}\right)$ method with the wolf line search is globally convergent when the sufficient descent condition

$$f(x_k+\alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k$$

is given, the HS method is very familiar with the PRP method,the DY method with the Wolf and strong Wolf line search is globally convergent without descent condition,but the descent property of the DY method depends on line search or convexity of the objective function.

In [6], the authors propose a nonmonotone Newton method, and analyze its convergence. Lucidi and Roma [10]present a nonmonotone algorithm of FR in 1995, see [9], Liu et al in the above-montioned nonmonotonic strong Wolf line search method to prove that two types of classical conjugate gradient algorithm of the PRP method and HS method for convex objective function in this nonmonotonic Wolf line search, global convergence. In 2002, Dai Yu Hong [2] the initial nonmonotone line search method on Grippo,Lampariello and Lucidi search and hybrid conjugate DY-CD method combined with the study. In 2006 [14] on this basis, given the amendments.

In this paper we presente new global convergence of FR method conjugate gradient with nonmonotone line search type strong Wolf.

Remark. The new nonmonotone line search can be viewed as some kind of convex combination of the extended strong Wofle line search and the extended nonmonotone line search, when

$\lambda = 0$ the new nonmonotone line search reduce to the extended strong Wofle line search, and when $\lambda =$ 1 the new nonmonotone line search reduce to the extended nonmonotone

line search.

This paper is organized as follows.We will present a new algorithm (Algorithm 2.8), and the sufficient descent property (1.11) of Algorithm 2.8 is also given in the next section. In Section 3 the global convergence results of the FR method are established.At last the preliminary numerical results are reported.

2. New Algorithm

Assumption 2.1. First given the general assumption of this section

(A_1) The level set $L_1 = \{x \setminus f(x) \leq f(x_1), x \epsilon R^n\}$ is bounded, where x_1 is the starting point

(A_2) f is strongly convex and differentiable in the level set L_1 and its gradient $g_k = \nabla f\left(x_k\right)$ lipschitz continuous . ie, there exist constants $L > 0$ and $\eta > 0$ making

$$\|g(x) - g(y)\| \leq L \|x - y\|, \ \forall x, y \in \mathcal{N} \tag{2.1}$$

and

$$(g(x) - g(y))^T (x - y) \geq \eta \|x - y\|^2 \tag{2.2}$$

If f satisfies Assumption (A_1) and (A_2), we can get that

$$\|g(x)\| \leq \gamma \tag{2.3}$$

for all $x \epsilon L_1$.

Now we give the following theorem, which illustrates that the formula (1.5) possesses the sufficient descent condition without any line searches.

Theorem 2.2. *Consider any method (1.2) and (1.3), where $\beta_k = \beta_k^{FR}$. Then for all $k \geq 1$*

$$g_k^T d_k = - \|g_k\|^2 . \tag{2.4}$$

Proof. Since $d_0 = -g_0$, we have $g_0^T d_0 = - \|g_0\|^2$

$$g_k^T d_k = g_k^T (-\theta_k g_k + \beta_k^{FR} d_{K-1})$$

$$= g_k^T (-\frac{d_{K-1}^T y_{k-1}}{\|g_{k-1}\|^2} g_k + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} d_{K-1})$$

□

$$= -\frac{d_{K-1}^T y_{k-1}}{\|g_{k-1}\|^2} \|g_k\|^2 + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_k^T d_{K-1}$$

$$= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} (-d_{K-1}^T y_{k-1} + g_k^T d_{K-1})$$

$$= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} (-d_{K-1}^T y_{k-1} + d_{K-1}^T g_k)$$

$$= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \left[-d_{K-1}^T (g_k - g_{k-1}) + d_{K-1}^T g_k\right]$$

$$= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} d_{K-1}^T g_{k-1}$$

Thus, we find a recurrence relation

$$g_k^T d_k = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} d_{K-1}^T g_{k-1}$$

so

$$\begin{aligned} g_k^T d_k &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} d_{K-1}^T g_{k-1} \\ &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \frac{\|g_{k-1}\|^2}{\|g_{k-2}\|^2} d_{K-2}^T g_{k-2} = \frac{\|g_k\|^2}{\|g_{k-2}\|^2} d_{K-2}^T g_{k-2} \\ &= \ldots \\ &= \frac{\|g_k\|^2}{\|g_0\|^2} d_0^T g_0 \\ &= \frac{\|g_k\|^2}{\|g_0\|^2} (-\|g_0\|^2) = -\|g_k\|^2 < 0 \end{aligned}$$

Lemma 2.3. *Suppose that assumption 2.1 holds and* α_k *is obtainded by the naw Nonmonotone line search(1.12) and (1.10). Then there exist* $c_1 > 0$*, such that*

$$\|s_k\| \geq \frac{c_1(-g_k^T d_k)}{\|d_k\|^2} \tag{2.5}$$

where $s_k = x_{k+1} - x_k$.

Proof. From assumption 2.1 (A_1)

$$\|g(x) - g(y)\| \leq L \|x - y\|$$

that is

$$\|y_k\| \leq L \|s_k\| \tag{2.6}$$

so we have □

$$y_k^T d_k \leq L \|s_k\| \|d_k\| \tag{2.7}$$

and from (1.10) we can get

$$y_k^T d_k = d_k^T (g_{k+1} - g_k) \geq (1 - \sigma_1)(- g_k^T d_k) \tag{2.8}$$

Thus, from (2.7) and (2.8) we obtain

$$\|s_k\| \geq \frac{(1 - \sigma_1)}{L} \times \frac{(- g_k^T d_k)}{\|d_k\|^2} \tag{2.9}$$

let $c_1 = \frac{(1-\sigma_1)}{L}$ and from (2.9), we obtain (2.5). Therefore, our proof is complete.

Lemma 2.4. *Suppose that assumption 2 holds, α_k is given (1.12) and (1.10), and $\beta_k = \beta_k^{FR}$.Then there exist a positive constant $M = \frac{L^2}{\eta}$ such that*

$$\frac{\|y_k\|^2}{y_k^T d_k} \leq M. \tag{2.10}$$

Proof. By the convexity, assumption we have

$$y_k^T d_k = d_k^T (g_{k+1} - g_k) \geq \eta \alpha_k \|d_k\|^2 \tag{2.11}$$

and from the lipschitz continuity (2.1), we can get that

$$\|y_k\| = \|g_{k+1} - g_K\| \leq L \|x_{k+1} - x_k\| = L\alpha_k \|d_k\| \tag{2.12}$$

Utilising (2.11) and (2.12), we can get that

$$\frac{\|y_k\|^2}{y_k^T d_k} \leq \frac{L^2 \alpha_k^2 \|d_k\|^2}{\eta \alpha_k^2 \|d_k\|^2} = \frac{L^2}{\eta} = M \tag{2.13}$$

which completes the proof. □

Lemma 2.5. *Assume that the following inequality holds for all k*

$$0 < m_1 \le \|g_k\| \le m_2 \tag{2.14}$$

and α_k is given by (1.12) and (1.10), then:

(1) there exist a positive constant $b > 1$ such that

$$|\beta_k| \le b \tag{2.15}$$

(2) there exist a positive constant λ, when $\|y_{k-1}\| \le \lambda$, we have $|\beta_k| \le \varepsilon$ for any $\varepsilon > 0$.

Proof. When$\beta_k = \beta_k^{FR}$, from (2.4) and (2.14) we can get

$$\begin{aligned}
|\beta_k| &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \frac{\|g_k - g_{k-1} + g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&= \frac{\|g_k - g_{k-1} + g_{k-1}\| \, \|g_k - g_{k-1} + g_{k-1}\|}{\|g_{k-1}\|^2} \\
&\le \frac{(\|g_k - g_{k-1}\| + \|g_{k-1}\|)(\|g_k - g_{k-1}\| + \|g_{k-1}\|)}{\|g_{k-1}\|^2} \\
&= \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\| \, \|g_{k-1}\| + \|g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&\le \frac{\|g_k\|^2 + \|g_{k-1}\|^2 + 2\|g_k\| \, \|g_{k-1}\| + 2\|g_k - g_{k-1}\| \, \|g_{k-1}\| + \|g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&\le \frac{\|g_k\|^2 + \|g_{k-1}\|^2 + 2\|g_k\| \, \|g_{k-1}\| + 2\{\|g_k\| + \|g_{k-1}\|\} \|g_{k-1}\| + \|g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&= \frac{\|g_k\|^2 + 2\|g_{k-1}\|^2 + 2\|g_k\| \, \|g_{k-1}\| + 2\|g_k\| \, \|g_{k-1}\| + 2\|g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&= \frac{\|g_k\|^2 + 4\|g_{k-1}\|^2 + 4\|g_k\| \, \|g_{k-1}\|}{\|g_{k-1}\|^2} \\
&\le \frac{9m_2^2}{m_1^2} = \left(\frac{3m_2}{m_1}\right)^2.
\end{aligned}$$

Let $b = \left(\frac{3m_2}{m_1}\right)^2$,obviously we have $b > 1$ because of $m_2 \ge m_1 > 0$.

Let $\lambda = \frac{2m_1^3}{9m_2^2}\varepsilon$

$$\begin{aligned}
|\beta_k| &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \leq \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\|\|g_{k-1}\| + \|g_{k-1}\|^2}{\|g_{k-1}\|^2} \\
&\leq \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\|\|g_{k-1}\| - g_{k-1}^T d_{k-1}}{\|g_{k-1}\|^2} \\
&\leq \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\|\|g_{k-1}\| + g_{k-1}^T d_{k-1}}{\|g_{k-1}\|^2} \\
&\leq \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\|\|g_{k-1}\|}{\|g_{k-1}\|^2} + \frac{g_{k-1}^T d_{k-1}}{\|g_{k-1}\|^2}
\end{aligned}$$

we have

$$g_{k-1}^T d_{k-1} < 0.$$

Therfore

$$\begin{aligned}
|\beta_k| &\leq \frac{\|g_k - g_{k-1}\|^2 + 2\|g_k - g_{k-1}\|\|g_{k-1}\|}{\|g_{k-1}\|^2} \\
&= \frac{\|g_k - g_{k-1}\|(\|g_k - g_{k-1}\| + 2\|g_{k-1}\|)}{\|g_{k-1}\|^2} \\
&\leq \frac{(\|g_k\| + 3\|g_{k-1}\|)\|y_{k-1}\|}{\|g_{k-1}\|^2} \\
&= \frac{4m_2}{m_1^2}\|y_{k-1}\| = \frac{8m_2^2}{2m_1^2 m_2}\|y_{k-1}\| \\
&\leq \frac{8m_2^2}{2m_1^2 m_1}\|y_{k-1}\| \leq \frac{9m_2^2}{2m_1^3}\|y_{k-1}\| \\
&\leq \frac{9m_2^2}{2m_1^3}\lambda \leq \varepsilon
\end{aligned}$$

This completes the proof. □

Lemma 2.6. *Suppose that assumption 2.1 (A_1) holds and α_k is obtained by the new Nonmonotone line search(1.12) and (1.10), denote $\zeta_k = \delta\alpha_k g_k^T d_k$, then $\{f(x_k)\}$ is nonincreasing and*

$$\lim_{k\to\infty} \zeta_{l(k+1)-1} = 0 \tag{2.16}$$

where

$$l(k) = \max\left\{ i \mid 0 \le k-i \le m(k), f(x_i) = \max_{0\le\ j\ \le m\langle k\rangle} f(x_k - j) \right\} \tag{2.17}$$

and $k - m(k) \le l(k) \le (k)$.

Proof. From (1.12) and (2.17) we can get that

$$f(x_k + \alpha_k d_k) \le \lambda \max_{0\le j\le m(k)} f(x_{k-j}) + (1-\lambda) \min_{0\le j\le m(k)} f(x_{k-j}) + \delta\alpha_k g_k^T$$

$$\le f(x_{l(k)}) + \zeta_k \tag{2.19}$$

□

because $\zeta_k < 0$,we can obtain

$$f(x_{(k+1)}) \le f(x_{l(k)}$$

from (2.17) and

$$m(k) \le m(k-1) + 1$$

for all k, we can get that

$$f(x_{l(k)}) \le \max_{0\le\ j\le m(k-1)+1} f(x_{k-j}) = \max\left\{ \max_{0\le\ j\le m(k-1)} f(x_{k-1-j}), f(x_k) \right\} \tag{2.20}$$

$$= \max f(x_{l(k-1)}), f(x_k) = f(x_{l(k-1),} \qquad k = 1, 2, \ldots$$

therefore $\{f(x_k)\}$ is nonincreasing and.Because
$l(k+1) - 1 \ge k + 1 - m(k+1) - 1 \ge k - M_0$, we have

$$f(x_{l(l(k+1)-1)}) \le f(x_{l(k-M_0)}) \tag{2.21}$$

from the above ineaquation and (2.19) we can obtain

$$f(x_{l(k+1)}) \le f(x_{l(l(k+1)-1)}) + \zeta_{l(k+1)-1}$$

$$\leq f(x_{l(k-M_0)}) + \zeta_{l(k+1)-1}$$

hence, we have

$$0 \leq -\zeta_{l(k+1)-1} \leq f(x_{l(k-M_0)}) - f(x_{l(k+1)}) \tag{2.22}$$

by (2.22) when assumption 2.1 (A_1) holds,we have

$$\lim_{k\to\infty} \zeta_{l(k+1)-1} = 0$$

Lemma 2.7. Suppose that α_k is given by (1.12) and (1.10) and $\beta_k = \beta_k^{FR}$. Then $\{l(k)\}$ is increasing

Proof. . Assume that

$$l(k+1) < l(k) \tag{2.23}$$

□

then we have

$$k+1 \geq k \geq l(k) > l(k+1) \geq k+1-m(k+1) \tag{2.24}$$

by the definition of $l(k+1)$ and (2.20), we can obtain

$$f(x_{l(k+1)}) \geq f(x_{l(k)}) \tag{2.25}$$

but from the lemma 2.6 we have

$$f(x_{l(k+1)}) \leq f(x_{l(k)}) \tag{2.26}$$

hence,by the definition of $l(k+1)$ and (2.25), we have

$$l(k+1) \geq l(k) \tag{2.27}$$

which is contradictory to (2.25) .Hence $l(k) \leq l(k+1)$, namely $\{l(k)\}$ is increasing

Now we can present a new descent conjugate gradient method as follows:

***FR* method with nonmonotone line search (algorithm)**

Stepe 1 : Given $x_1 \epsilon \mathbb{R}^n, \varepsilon \geq 0, 0 < \delta \leq \sigma_1 \leq 1, 0 \leq \sigma_2 \leq 1,\ M_0 \epsilon N\ , 0 \leq \lambda \leq 1$,

set $d_1 = -g_1, k = 1$, if $\|g_k\| \leq \varepsilon$, then stop.

Stepe 2 : Find $\alpha_k \geq 0$ by (1.12) and (1.10).

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

Stepe 4 : Compute $\beta_k = \beta_k^{FR}$ by the formula (1.5) and generate $d_{k+1}$by (1.3).

Stepe 5 : Set $k = k + 1$,go to step2.

3. Global Convergence

Theorem 3.1. Suppose that assumption holds and α_k is obtained by the new Nonmonotone line search (1.12) and (1.10), consider any iteration method of the form (1.2) and (1.3), where $\beta_k = \beta_k^{FR}$. Then

$$\lim_{k\to\infty} \inf \|g_k\| = 0 \tag{3.1}$$

Proof. . If assumption 2.1, then there exists a constantγ_1 such that

$$0 < \gamma_1 \leq \|g(x)\| \leq \gamma_2$$

for all k . From (1.3) and lemma 2.5 we have □

$$\begin{aligned} \left\|d_{l(k+1)-1}\right\| &\leq \left\|g_{l(k+1)-1}\right\| + \left|\beta_{l(k+1)-2}\right| \left\|d_{l(k+1)-2}\right\| \\ &\leq \gamma_2 + b \left\|d_{l(k+1)-2}\right\| \\ &\leq \cdots \\ &\leq \gamma_2 \sum_{j=0}^{l(k+1)-l(k)-3} (b)^j + (b)^{l(k+1)-l(k)-2} \left\|d_{l(k)+1}\right\| \\ &\leq \gamma_2 \sum_{j=0}^{l(k+1)-l(k)-2} (b)^j + (b)^{l(k+1)-l(k)-2} \left|\beta_{l(k)+1}\right| \left\|d_{l(k)}\right\| \end{aligned} \tag{3.2}$$

because

$$l(k+1) - l(k) \leq k + 1 - [k - m(k)] \leq M_0 + 1$$

we can get

$$\overline{\gamma} \sum_{j=0}^{l(k+1)-l(k)-2} (b)^j \leq \sum_{j=0}^{M_0} (b)^j, (b')^{l(k+1)-l(k)-2} \leq (b)^{M_0}$$

where

$$\gamma_2\sum_{j=0}^{M_0}(b)^j = h_1, (b)^{M_0} = h_2$$

Then

$$\left\|d_{l(k+1)-1}\right\| \leq h_1 + h_2\left|\beta_{l(k)}\right|\left\|d_{l(k)}\right\| \qquad (3.3)$$

from lemma 2.3and (1.10) we have

$$0 \leq \|y_k\| \leq \sqrt{My_k^T s_k} \leq \sqrt{M(1+\sigma_2)(-\alpha_k g_k^T d_k)}$$

from lemma 2.5 we not that $\varepsilon = \frac{1}{h_2 b^2}$, there exists an innegative integer k_0 when$k \geq k_0$ we have

$\left|\beta_{l(k)}\right| < \varepsilon$

Thus we can get

$$\begin{aligned}
\left\|d_{l(k+1)-1}\right\| &\leq h_1 + \frac{\left\|d_{l(k)}\right\|}{b^2} \\
&\leq h_1 + \frac{\left\|g_{l(k)}\right\| + \left|\beta_{l(k)}\right|\left\|d_{l(k)-1}\right\|}{b^2} \\
&\leq h_1 + \frac{\gamma_2 + b\left\|d_{l(k)-1}\right\|}{b^2} \\
&\leq h_3 + \frac{\left\|d_{l(k)-1}\right\|}{b}
\end{aligned}$$

where $h_3 = h_1 + \frac{\gamma_2}{b^2}$,we have a recursive equation which leads to

$$\begin{aligned}
\left\|d_{l(k+1)-1}\right\| &\leq h_3 + \frac{\left\|d_{l(k)-1}\right\|}{b} \\
&\leq h_3 + \frac{h_3 + \frac{\left\|d_{l(k-1)-1}\right\|}{b}}{b} \leq ... \\
&\leq h_3\sum_{j=0}^{k-k_0}(\frac{1}{b})^j + \frac{1^{k-k_0+1}}{b}\left\|d_{\imath(k_0)-1}\right\| \\
&\leq h_3\sum_{j=0}^{\infty}(\frac{1}{b})^j + \left\|d_{\imath(k_0)-1}\right\|
\end{aligned}$$

$$\leq h_3 \frac{b}{b-1} + \left\| d_{\imath(k_0)-1} \right\| \tag{3.4}$$

Applying $l(k) \geq k - m(k) \geq k - M_0$ and lemma 2.7 we can assume

$$l(i) - 1 \leq j < l(i+1) - 1, i \geq k_0 + 2$$

for all $j \geq l(k_0 + 2) - 1$, thus we have

$$\begin{aligned} \|d_j\| \leq & \quad \|g_j\| + |\beta_j| \, \|d_j - 1\| \leq \gamma_2 + b(\|g_{j-1}\| + |\beta_{j-1}| \, \|d_j - 1\|) \leq \dots \\ \leq & \quad \gamma_2 \sum_{t=0}^{j-l(i)} (b)^t + (b)^{j-l(i)+1} \left\| d_{l(i)-1} \right\| \end{aligned} \tag{3.5}$$

Therefore, from

$$\begin{aligned} j - l(i) + 1 & \leq [l(i+1) - 1] - l(i) + 1 \\ & \leq i + 1 - [i - m(i)] \\ & \leq M_0 + 1 \end{aligned}$$

and (3.5), we have

$$\|d_j\| \leq \gamma_2 \sum_{t=0}^{j-l(i)} (b)^t + (b)^{M_0+1} \left\| d_{l(i)-1} \right\| \tag{3.6}$$

From (3.4) and (3.6), we have

$$\|d_j\| \leq \gamma_2 \sum_{t=0}^{M_0} (b)^t + (b)^{M_0+1} \left[h_3 \frac{b}{b-1} + \left\| d_{\imath(k_0)-1} \right\| \right] \tag{3.7}$$

for all $i - 1 > k_0$. By using lemma 2.3 and lemma 2.4 we have

$$\begin{aligned} -\zeta_{l(k+1)-1} &= \frac{\rho \left\| s_{l(k+1)-1} \right\| (-g_{l(k+1)-1}, d_{l(k+1)-1})}{\left\| d_{l(k+1)-1} \right\|} \\ &\geq \rho c_1 \frac{(-g_{l(k+1)-1}, d_{l(k+1)-1})^2}{\left\| d_{l(k+1)-1} \right\|} \\ &\geq \rho c_1 \frac{\gamma_1^4}{\left\| d_{l(k+1)-1} \right\|} \geq 0 \end{aligned} \tag{3.8}$$

but from (2.16) and (3.8) we can get

Problems	Name
1	Roth function
2	Beale function
3	Helicalevalley function
4	Gulf research and development function
5	Powell singular function
6	Wood function
7	Kowalikand Osborne function
8	Brownand Dennis function
9	Watson function
10	Etended Rosenbrock function
11	Trigonometric function
12	Extended Powell singular function
13	Penalty functionI
14	Penalty functionII

Table 4.1: List of test problems

$$\lim_{k\to\infty} \frac{1}{\left\|d_{l(k+1)-1}\right\|} = 0$$

which is contradictory to (3.7)

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

4. Numerical Experiments

We performed some numerical experiments on conjugate methods with the new nonmonotone line search type strong Wolf. We tested the FR method,,DY method,HS method and PRP method.the methods were tested from [11], when $\varepsilon = 10^{-6}$, $\delta = 0.01, \sigma_1 = \sigma_2 = 0.1, \lambda = 0$, $M_0 = 150$ for when $\varepsilon = 10^{-6}$, $\delta = 0.01, \sigma_1 = \sigma_2 = 0.1, \lambda = \frac{1}{2}$, $M_0 = 150$.The tested problems are listed in Table 4.1, and the detail numerical results of our tests are reported in Tables 4.2, 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.

The star ($*$) denotes that this result is best one among these methods. In table 4.3, CPU times of the methods are given.Tables 4.2 and 4.3 show that the

Probl.	Dimen.	FR	DY	HS	PRP
1	2	11/70/56*	42/168/139	55/187/155	11/76/57
2	2	14/57/45	75/186/164	74/177/155	13/58/45*
3	3	40/148/123	37/118/98*	56/157/132	66/183/156
4	3	1/2/2*	1/2/2*	1/2/2*	1/2/2*
5	4	103/333/383	2286/4555/4547	425/1036/947	113/380/328
6	4	78/278/230*	100/291/240	184/438/399	118/357/304
7	4	68/249/220	536/1449/1271	254/723/633	93/269/240
8	5	41/178/136	39/158/121	44/171/133	37/156/123*
9	5	403/1210/1062	126/348/298	87/276/238*	133/374/338
	20	2287/7972/7045	1945/5658/4924	4375/12695/11223	3293/10458/9246
10	100	33/195/152	31/157/121	62/223/182	29/168/128*
	200	24/167/125*	26/160/121	25/159/117	25/175/132
11	100	56/137/120*	306/401/401	*FAIL*	59/120/114
	200	64/168/140	314/399/398	*FAIL*	64/135/127
12	500	219/490/465*	4796/6823/6822	5089/7049/7058	1645/2889/2889
	1000	38/69/64*	414/449/449	2406/3114/3115	147/251/252
13	500	6/13/7*	7/16/8	7/16/8	7/16/8
	1000	7/14/8	7/15/8	7/15/8	6/13/7*
14	1000	33/75/60*	52/107/102	52/114/109	36/78/75
	5000	32/75/60*	65/137/131	72/149/145	35/76/73

Table 4.2: The numerical results of the methods

FR method has the best performance with respect to the number of iterations and CPU time.All numerical results show that the efficiency of the FR method is encouraging

In order to rank the iterative numerical methods, one can compute the total nSumber of function and gradient evaluations by the formula

$$N_{total} = NF + 5 \times NG \tag{4.1}$$

Similarly, we compare PRP method, HS method, DY method with FR method as follows: for each problem i, compute the total numbers of function evaluations and gradient evaluations required by the evaluated methods and FR method by formula (4.1), and denote them by $N_{total,i}$ (EM) and $N_{total,i}$ (FR); then calculate the ratio

$$r_i\,[EM\ \ (J)] = \frac{N_{total,\ i}\ \ [EM\ \ (J)]}{N_{total,\ i}\ (FR)} \tag{4.2}$$

If $[EM\ \ (J_0)]$ method does not work for example i_0, but FR method can work, we replace the $r_{i_0}\,[EM\ \ (J_0)]$ by a positive constant τ_1 which define as follows:

$$\tau_1 = max\,\{r_i\ \ [EM\ \ (J_0)] : (i, j_0) \in S_1\} \tag{4.3}$$

where

$$S_1 = \{(i, j_0) : method\ \ j_0\ does\ not\ work\ for\ example\ i\} \tag{4.4}$$

If FR method does not work for example i_0, but $[EM\ \ (J_0)]$ method can work, we replace the $r_{i_0}\,[EM\ \ (J_0)]$ by a positive constant τ_2 which define as

Problems	dimension	FR	DY	HS	PRP
1	2	0.0451*	0.1775	0.1964	0.0516
2	2	0.0721	0.3100	0.2717	0.0713*
3	3	0.2530	0.1264*	0.1995	0.5374
4	3	0.0008*	0.0041	0.0044	0.0036
5	4	0.3892	8.9351	2.5667	0.7034
6	4	0.3788*	0.4389	0.7155	0.5266
7	4	0.3570	2.1098	0.6122	0.4177
8	5	0.2364	0.3521	0.1614*	0.2172
9	5	2.2057	0.5362	0.3141*	0.6104
	20	7.5890	6.100*	16.9904	10.9497
10	100	0.2136*	0.2332	0.5122	0.2718
	200	0.3657*	0.4020	0.3961	0.4309
11	100	0.3899*	1.9000	*FAIL*	0.4573
	200	1.9860	6.5000	*FAIL*	1.8896*
12	500	8.7861*	56.4843	57.1386	28.8092
	1000	1.8963*	15.0000	40.6372	6.1500
13	500	1.9425	2.1589	2.1428	1.8554*
	1000	8.3452	9.7060	8.5487	7.3760*
14	1000	0.5001*	0.6897	0.7284	0.5066
	5000	1.3960*	3.1576	3.6011	1.7479

Table 4.3: The corresponding CPU times of the methods

follows:

$$\tau_2 = \min \{r_i \ [EM \ \ (J_0)] : (i, j_0) \in S_1\} \tag{4.5}$$

Neither FR method nor $[EM \ \ (J_0)]$ method works, we define $r_{i_0} [EM \ \ (J_0)] = 1$.The geometric mean of these ratios for $[EM \ \ (J)]$ method over all the test problems is defined by

$$r\,[EM \ \ (J)] = \{\Pi_{i \in s} \ r_i\,[EM \ \ (J)]\}^{\frac{1}{|S|}} \tag{4.6}$$

where S denotes the set of the test problems and $|S|$ the number of elements in S.

1. The values of $r(FR)$, $r(PRP)$, $r(HS)$, and $r(DY)$ are listed in table 4.4

From Table 4.4 we can see that the new method is more efficient than FR method and PRP method.

FR	PRP	HS	DY
0.775	0.896	0.923	1.212

Table 4.4

References

[1] Al-Baali, M. Descent property and global convergence of the Fletcher-Reeves method with inexact line search. IMA J. Numer. Anal. 5 (1985), no. 1, 121–124. MR0777963 (86d:49043)

[2] Dai, Yuhong. A nonmonotone conjugate gradient algorithm for unconstrained optimization. J. Syst. Sci. Complex. 15 (2002), no. 2, 139–145. MR1900944 (2003b:90097)

[3] Dai, Y. H.; Yuan, Y. Convergence properties of the Fletcher-Reeves method. IMA J. Numer. Anal. 16 (1996), no. 2, 155–164. doi: 10.1093/imanum/16.2.155

[4] Dai, Y. H.; Yuan, Y. A nonlinear conjugate gradient method with a strong global convergence property. SIAM J. Optim. 10 (1999), no. 1, 177–182. doi 10.1137/S1052623497318992

[5] Fletcher, R.; Reeves, C. M. Function minimization by conjugate gradients. Comput. J. 7 1964 149–154. MR0187375 (32 #4827)

[6] Grippo, L.; Lampariello, F.; Lucidi, S. A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. 23 (1986), no. 4, 707–716. MR0849278 (87g:90105)

[7] Hestenes, M, R., and Stiefel, E, L.: Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standars Sect., 5 (49), pp. 409-436,* (1952).

[8] Liu, G, H., Han, J, Y., and Yin, H, X.: Global convergence of the Fletcher-Reeves algorithm with an inexact line search, *Appl. math. J. Chinese Univ. Ser. B, 10, pp. 75-82,* (1995).

[9] Liu, G. H.; Jing, L. L.; Han, L. X.; Han, D. A class of nonmonotone conjugate gradient methods for unconstrained optimization. J. Optim. Theory Appl. 101 (1999), no. 1, 127–140. MR1685579

[10] Lucidi, S.; Roma, M. Nonmonotone conjugate gradient methods for optimization. System modelling and optimization (Compiègne, 1993), 206–214, Lecture Notes in Control and Inform. Sci., 197, Springer, London, 1994. MR1294156

[11] Moré, J, J., Garbow, B, S., and Hillstrom, K, E.: Testing unconstrained optimization software, ACM Transactions on mathematical software., 17-41, (1981). ACM0098-3500/81/0300-0017

[12] Polyak, B, T.: The conjugate gradient method in extremal problems, *Comput. Math. Math. Phys., 9, pp. 94-112,* (1969). dx.doi.org/10.1016/0041-5553(69)90035-4

[13] Powell, M. J. D. Nonconvex minimization calculations and the conjugate gradient method. Numerical analysis (Dundee, 1983), 122–141, Lecture Notes in Math., 1066, Springer, Berlin, 1984. MR0760460.

[14] Zhang, Li; Zhou, Weijun; Li, Donghui. Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search. Numer. Math. 104 (2006), no. 4, 561–572. DOI 10.1007/s00211-006-0028-z