A new spectral PRP conjugate gradient method with sufficient descent property

Xiaoli Sun *
School of Mathematics and Statistics, Zaozhuang University, Beian Street, 277160, Shandong, China

Received 29 December 2014; accepted (in revised version) 08 March 2015

Abstract: In this paper, a new spectral PRP conjugate gradient method is proposed, which can always generate sufficient descent direction without any line search. Under some standard conditions, global convergence of the proposed method is established when the standard Armijo or weak Wolfe line search is used. Moreover, we extend these results to the HS method. Numerical comparisons are reported with some existing modified PRP methods by utilizing test problems in the CUTEr library.

MSC: 90C33 • 49J40

Keywords: Conjugate gradient method • Sufficient descent property • Global convergence

© 2015 IJAAMM all rights reserved.

1. Introduction

In this paper, we consider the following unconstrained optimization problem

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

(1)

where $\mathbb{R}^n$ is the $n$-dimensional Euclidean space, $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, and its gradient $g(x)$ is available. Here, we abbreviate $f(x_k)$ and $g(x_k)$ by $f_k$ and $g_k$, respectively. Nonlinear conjugate gradient methods for solving (1) generate sequence of iterates recurrently by

$$x_{k+1} = x_k + \alpha_k d_k, \ k = 0, 1, 2, \ldots,$$

(2)

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

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

(3)

in which $\beta_k$ is a scalar that characterizes the method. That is, different conjugate gradient methods correspond to different choices of the parameter $\beta_k$. There have been many famous formulas for parameter $\beta_k$, such as, Fletcher-Reeves (FR)[1], Polak-Ribiére-Polyak (PRP)[2,3], Liu-Storey (LS) [4], Dai-Yuan (DY) [5], Hestenes-Stiefel (HS) [6] and the Conjugate Descent (CD) [7]. In this paper, we focus our attention on the PRP method in which the parameter $\beta_k$ is given by

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

* Corresponding author.
E-mail address: xiaolisun1981@163.com
where $y_k = g_k - g_{k-1}$ and $\| \cdot \|$ means the Euclidean norm. In the convergence analysis and implementations of conjugate gradient methods, one often requires the line search to be the inexact line search such as the Armijo line search or the weak Wolfe line search, or the strong Wolfe line search. The Armijo line search is to find a step-size $\alpha_k = \max \{ \rho^j | j = 0, 1, \ldots \}$ satisfying
\[
f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k,
\]
where $\delta, \rho \in (0, 1)$ are constants. The weak Wolfe line search is to find a step-size $\alpha_k$ satisfying
\[
\begin{align*}
&f(x_k + \alpha_k d_k) \leq f(x_k) + \rho \alpha_k g_k^T d_k, \\
g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k,
\end{align*}
\]
where $0 < \rho \leq \sigma < 1$. The strong Wolfe line search is to compute $\alpha_k$ such that
\[
\begin{align*}
&f(x_k + \alpha_k d_k) \leq f(x_k) + \rho \alpha_k g_k^T d_k, \\
&\|g(x_k + \alpha_k d_k)^T d_k\| \leq \sigma \|g_k^T d_k\|,
\end{align*}
\]
where $0 < \rho < 1/2$ and $\sigma \in (\rho, 1)$. The PRP method has been regarded as one of the most efficient conjugate gradient methods and has been studied by many researchers. Polak and Ribi`ere [2] proved that the PRP method with the exact line search is globally convergent under a strong convexity assumption for the objective function $f$. However, for general functions, Powell [8] showed that the PRP method can cycle infinitely without approaching a solution even if the step-size $\alpha_k$ is chosen to the least positive minimizer of the line search function. In addition, the PRP method may generate an uphill search direction even for strongly convex functions when the strong Wolfe line search is used[9]. Therefore, great attentions are given to finding modified PRP methods which not only have nice numerical performance but also have global convergence, see [10-12], or the recent survey paper [14] for modified PRP type methods. One well-known modified PRP method (denoted by WYL method) was proposed by Wei et al.

\[
\beta_k^{WYL} = \frac{g_k^T \left( g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right) - \|g_k\|^2 - \frac{3g_k^T g_k}{\|g_k\|^2} g_{k-1}^T g_{k-1}}{\|g_{k-1}\|^2}.
\]

The $\beta_k^{WYL}$ has the following nice properties: (1) $\beta_k^{WYL} \geq 0$; (2) it can avoid jamming. However, the WYL method satisfies the sufficient descent condition
\[
g_k^T d_k \leq -c\|g_k\|^2, \quad c > 0
\]
only when $\sigma < 1/4$ in the strong Wolfe line search, which is a little restricted condition. To obtain nice convergence property, this parameter was further revised by Zhang [11], Huang et al.[12] and Dai et al.[13], in which the parameters of $\beta_k$s are specified as follows
\[
\begin{align*}
&\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_k^T g_{k-1}}{\|g_{k-1}\|^2}, \beta_k^{HPRP} = \frac{g_k^T \left( g_k - \frac{g_k^T g_{k-1}}{\|g_k\|^2} g_{k-1} \right)}{\|g_{k-1}\|^2}, \\
&\beta_k^{ZPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_k^T g_{k-1}}{\mu \|g_k^T d_{k-1}\| + \|g_k\|^2}, \quad (\mu > 1).
\end{align*}
\]

The conjugate gradient methods induced by the above three parameters are globally convergent under some Armijo-type line search or the weak Wolfe line search. Obviously, if $\mu = 0$, then $\beta_k^{DPRP}$ reduces to $\beta_k^{ZPRP}$, however, to ensure the sufficient descent property (7), $\mu$ must satisfy $\mu > 1$ in $\beta_k^{DPRP}$. Moreover, the numerical results show the conjugate gradient induced by $\beta_k^{DPRP}$ is not very effective. Therefore, in this paper, we will give a modified version of $\beta_k^{DPRP}$, and investigate global convergence of the corresponding method when $0 \leq \mu \leq 1$. In fact, we give a new parameter defined as follows:
\[
\begin{align*}
&\beta_k^{MPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \max \{g_k^T g_{k-1}, 0\}}{\mu \|g_k^T d_{k-1}\| + \|g_k\|^2}, \quad (\mu \geq 0).
\end{align*}
\]

To do so, the remainder of the paper is organized as follows: In Section 2, motivation and new spectral PRP method (denoted by SPRP method) are given. In Section 3, we prove the global convergence of the SPRP method with the standard Armijo or weak Wolfe line search. We also extend these results to the HS method in this section. In Section 4, we report some numerical results to test the proposed method.
2. Motivation and the SPRP method

More recently, Wan et al.[15] proposed a spectral PRP method, called WPRP method, whose direction \( d_k \) is defined by:

\[
d_k = -\theta_k^{WZ} g_k + \beta_k^{PRP} d_{k-1}, \quad \forall k \geq 1,
\]

where \( \theta_k^{WZ} = d_{k-1}^T y_{k-1}/\|g_{k-1}\|^2 - d_{k-1}^T g_{k-1} g_{k-1}^T / \|g_k\|^2 \|g_{k-1}\|^2 \). An important property of the WPRP method is that its generated direction satisfies the sufficient condition \( g_k^T d_k = -\|g_k\|^2 \). Wan et al.[15] also established the global convergence of the WPRP method with a modified Wolfe line search. Combining with \( g_k^T d_k = -\|g_k\|^2 \), (9) can be rewritten as

\[
d_k = -g_k - \beta_k^{PRP} \frac{g_k^T d_{k-1}}{\|g_k\|^2} g_k + \beta_k^{PRP} d_{k-1}.
\]

Therefore, replacing \( \beta_k^{PRP} \) by \( \beta_k^{MPRP} \) in (10), we get a new iteration of the form

\[
d_k = -g_k - \beta_k^{MPRP} \frac{g_k^T d_{k-1}}{\|g_k\|^2} g_k + \beta_k^{MPRP} d_{k-1}, \quad d_0 = -g_0.
\]

from which, we get the following method:

**SPRP Method** (Spectral PRP method)

**Step 0.** Given an initial point \( x_0 \in \mathbb{R}^n, \mu \geq 0, 0 < \delta < 1, 0 < \rho < \sigma < 1 \) and set \( d_0 = -g_0, k := 0 \).

**Step 1.** If \( \|g_k\| = 0 \) then stop; otherwise go to step 2.

**Step 2.** Compute \( d_k \) by (11) and determine the step-size \( \alpha_k \) by the standard Armijo line search (4) or the weak Wolfe line search (5).

**Step 3.** Set \( x_{k+1} = x_k + \alpha_k d_k \), and \( k := k + 1 \); go to Step 1.

For convenience, we call the above method as SPRP method. Obviously, if the line search is exact and \( g_k^T g_{k-1} \geq 0 \), then the SPRP method reduces to the classical Wei-Yao-Liu method in [10].

**Remark 2.1.** From (11), we can easily obtain

\[
g_k^T d_k = -\|g_k\|^2, \quad \text{and} \quad \|g_k\| \leq \|d_k\|.
\]

This indicates that the SPRP method satisfy the sufficient descent property.

3. Global convergence

The following assumptions are often used in the global convergence analysis of conjugate gradient methods.

**Assumption A**

(I) The level set \( L_0 = \{ x \in \mathbb{R}^n | f(x) \leq f(x_0) \} \) is bounded.

(II) In some neighborhood \( N \) of \( L_0 \), the function \( f \) is continuously differentiable and its gradient \( g(x) \) is Lipschitz continuous, i.e., there exists a constant \( L > 0 \) such that

\[
\|g(x) - g(y)\| \leq L \|x - y\|, \quad \forall x, y \in N.
\]

Assumptions (I) and (II) imply that there exist a positive constant \( \gamma \) such that

\[
\|g(x)\| \leq \gamma.
\]

**Lemma 3.1.**

The parameter \( \beta_k^{MPRP} \) defined by (8) satisfies

\[
0 \leq \beta_k^{MPRP} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}.
\]

**Proof.** From the definition of \( \beta_k^{MPRP} \), we have

\[
0 = \frac{\|g_k\|^2 - \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \|g_k\| \|g_{k-1}\|}{\mu |g_k^T d_{k-1}| + \|g_{k-1}\|^2} \leq \beta_k^{MPRP} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}.
\]
Consider the SPRP method. If the step-size \( \alpha \) was originally given by Zoutendijk in lemma, called the Zoutendijk condition, is often used to prove global convergence of conjugate gradient method. It is determined by the standard Armijo line search condition, \( \rho \), and Assumption 1 (I), we have

\[ g_k^\top d_k = -\theta_k\|g_k\|^2 + \beta_k^\text{MPRP} g_k^\top d_k. \]

This together with (12) shows that

\[ \beta_k^\text{MPRP} g_k^\top d_k = (\theta_k - 1)\|g_k\|^2. \]

Squaring both sides of (11), we obtain

\[ \|d_k\|^2 = \theta_k^2\|g_k\|^2 - 2\theta_k\|g_k\|^2 + (\beta_k^\text{MPRP})^2\|d_k\|^2. \]

Substituting (16) into (17), we have

\[
\begin{align*}
\|d_k\|^2 &= \theta_k^2\|g_k\|^2 - 2\theta_k(\theta_k - 1)\|g_k\|^2 + (\beta_k^\text{MPRP})^2\|d_k\|^2 \\
&= -\theta_k(\theta_k - 1)^2\|g_k\|^2 + \|g_k\|^2 + (\beta_k^\text{MPRP})^2\|d_k\|^2 \\
&\leq \|g_k\|^2 + (\beta_k^\text{MPRP})^2\|d_k\|^2.
\end{align*}
\]

Now, we prove the global convergence of the SPRP method under the standard Armijo line search (4). The following lemma, called the Zoutendijk condition, is often used to prove global convergence of conjugate gradient method. It was originally given by Zoutendijk in [16].

**Lemma 3.3.**
Consider the SPRP method. If the step-size \( \alpha_k \) is determined by the standard Armijo line search (4), then for all \( k \),

\[
\sum_{k=0}^\infty \frac{\|g_k\|^4}{\|d_k\|^2} < \infty.
\]

**Proof.** From (4), (12) and Assumption 1 (I), we have

\[
\sum_{k=0}^\infty \alpha_k\|g_k\|^2 < \infty.
\]

First, we prove there exists a constant \( c_1 > 0 \) such that the following inequality holds for all \( k \),

\[
\alpha_k \geq c_1 \frac{\|g_k\|^2}{\|d_k\|^2}.
\]

The proof of (20) is divided into the following two cases.

Case (I): If \( \alpha_k = 1 \), then from (12), we have \( \alpha_k = 1 \geq \frac{\|g_k\|^2}{\|g_k\|^2} \).

Case (II): If \( \alpha_k < 1 \), then by the Armijo line search condition, \( \rho \) does not satisfy inequality (4). That is,

\[ f(x_k + \rho^{-1} \alpha_k d_k) > f(x_k) + \delta \rho^{-1} \alpha_k g_k^\top d_k. \]

By the mean-value theorem and the above inequality, there is a \( t_k \in (0, 1) \) such that \( x_k + t_k \rho^{-1} \alpha_k d_k \in N \) and

\[
\begin{align*}
f(x_k + \rho^{-1} \alpha_k d_k) - f(x_k) &= \rho^{-1} \alpha_k g(x_k + t_k \rho^{-1} \alpha_k d_k)^\top d_k \\
&= \rho^{-1} \alpha_k g_k^\top d_k + \rho^{-1} \alpha_k (g(x_k + t_k \rho^{-1} \alpha_k d_k) - g_k)^\top d_k \\
&\leq \rho^{-1} \alpha_k g_k^\top d_k + L \rho^{-1} \alpha_k^2 \|d_k\|^2.
\end{align*}
\]

Substituting the last inequality into (21), we have

\[
\alpha_k \geq \frac{(1 - \delta)\rho}{L} \frac{\|g_k\|^2}{\|d_k\|^2}.
\]

Letting \( c_1 = \min\{1, \frac{(1 - \delta)\rho}{L}\} \), from the above two cases, we get (20). Substituting (20) into (19), we have (18). \( \square \)
Theorem 3.1.
Suppose that the conditions in Assumption A hold. Consider the SPRP method, and the step-size $\alpha_k$ is determined by the standard Armijo line search (4). Then we have

$$\liminf_{k \to \infty} \|g_k\| = 0.$$ \hspace{1cm} (22)

Proof. Assume that the conclusion (22) is not true. Then there exists a constant $\epsilon > 0$ such that

$$\|g_k\| \geq \epsilon, \quad \forall k \geq 0.$$ 

From (14) and (15), we have

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{\|g_k\|^2 + (\beta_k^\text{MPRP})^2 \|d_{k-1}\|^2}{\|g_k\|^4} = (\beta_k^\text{MPRP})^2 \frac{\|g_{k-1}\|^4 \|d_{k-1}\|^2}{\|g_k\|^4 \|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2} \|g_k\|^2 \leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2} \|g_k\|^2 \leq \frac{1}{\|g_{k-1}\|^2} + \frac{1}{\|g_k\|^2} + \cdots + \frac{1}{\|g_0\|^2} \leq \frac{k + 1}{\epsilon^2}.$$ 

That is

$$\sum_{k=0}^{n} \|g_k\|^4 \leq \sum_{k=0}^{n} \frac{\epsilon^2}{k + 1} = \infty.$$ 

This contradicts (18). Therefore, (22) holds.

In what follows, we shall establish the global convergence of the SPRP method with the weak Wolfe line search (5). First, we also can get the Zoutendijk condition (18).

Lemma 3.4.
Consider any method in the form of (2), where $d_k$ is a descent direction and $\alpha_k$ satisfies weak Wolfe condition (5) or strong Wolfe condition (6). Then the conclusion in Lemma 3.3 also holds.

Following the same argument of Theorem 3.1, we have the following globally convergent result.

Theorem 3.2.
Suppose that the conditions in Assumption A hold. Consider the SPRP method, and the step-size $\alpha_k$ is determined by the weak Wolfe line search (5). Then we have

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

Similar to the SPRP method, we can define a spectral HS method (denoted by SHS1 method), in which the direction is determined by

$$d_k = -g_k - \beta_k^\text{MHS} \frac{g_k^T d_{k-1}}{\|g_k\|^2} g_k + \beta_k^\text{MHS} d_{k-1}, \quad d_0 = -g_0,$$ \hspace{1cm} (23)

where

$$\beta_k^\text{MHS} = \frac{\|g_k\|^2 - \|g_{k-1}\|^2}{\|d_{k-1}\|^2 \|g_k\|^2 + \mu |g_k^T d_{k-1} + y_{k-1}^T d_{k-1}|}, \quad (\mu > 0).$$ \hspace{1cm} (24)

Obviously, the above direction also satisfies the sufficient descent property (12). Moreover, it also satisfies the Zoutendijk condition (18) if the SHS1 method utilizes the weak Wolfe line search. The next lemma corresponds to Lemma 3.4 in [17] and Lemma 3.1 in [18].
Lemma 3.5.
Suppose that Assumption A holds. Let \( \{x_k\} \) be the sequence generated by SHS1 with the weak Wolfe line search (5). If there exists a constant \( \epsilon > 0 \) such that \( \|g_k\| \geq \epsilon \) for all \( k \geq 0 \), then we have
\[
\sum_{k=0}^{\infty} \|u_{k+1} - u_k\|^2 < +\infty,
\]
(25)
where \( u_k = d_k/\|d_k\| \).

Proof. From (5), we have
\[
y_k^T d_{k-1} \geq (1 - \sigma) (-g_k^T d_{k-1}) = (1 - \sigma) \|g_k\|^2.
\]
(26)
From (12) and \( g_k \geq \epsilon \) for all \( k \), we have \( \|d_k\| > 0 \) for all \( k \). Therefore, \( u_k \) is well-defined. Define
\[
r_k = - \left( 1 + \beta_k^MHS \frac{\|g_k\|^2}{\|d_k\|^2} \right) \frac{g_k}{\|d_k\|} \quad \text{and} \quad \delta_k = \beta_k^MHS \frac{\|d_{k-1}\|}{\|d_k\|}.
\]
Then, we have
\[
u_k = r_k + \delta u_{k-1}.
\]
Since \( u_{k-1} \) and \( u_k \) are unit vectors, we can write
\[
\|r_k\| = \|u_k - \delta u_{k-1}\| = \|\delta u_k - u_{k-1}\|.
\]
Noting that \( \delta_k \geq 0 \), we get
\[
\|u_k - u_{k-1}\| \leq (1 + \delta_k) \|u_k - u_{k-1}\| \leq \|u_k - \delta u_{k-1}\| + \|\delta u_k - u_{k-1}\| = 2 \|r_k\|.
\]
(27)
From (24) and (27), we have
\[
\beta_k^MHS \|g_k\|^2 \frac{\|g_{k-1}\|^2}{\|d_k\|^2} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \frac{\|g_{k-1}\|^2}{\|d_k\|^2} \leq \frac{1}{\mu}.
\]
(28)
From (13) and (28), it follows that there exists a constant \( M_1 \geq 0 \) such that
\[
\| - \left( 1 + \beta_k^MHS \frac{g_k^T d_{k-1}}{\|g_k\|^2} \right) g_k \| \leq \|g_k\| + \|g_k\| \gamma + \frac{1}{\mu} \gamma \leq \frac{1}{\mu} \gamma + \frac{1}{\mu} \gamma \leq M_1.
\]
(29)
Thus, from (18) and (29), we get
\[
\sum_{k=0}^{\infty} \|r_k\|^2 \leq \sum_{k=0}^{\infty} \|d_{k-1}\|^2 \leq \sum_{k=0}^{\infty} \frac{M_1^2}{\mu} \|g_k\|^4 \|d_{k-1}\|^2 \leq \frac{M_1^2}{\mu^4} \sum_{k=0}^{\infty} \|d_{k-1}\|^2 < +\infty,
\]
which together with (27) completes the proof.

The following theorem establishes the global convergence of the SHS1 method with Wolfe line search (5). The proof is analogous to that of Theorem 3.2 in [18].

Theorem 3.3.
Let the Assumption A holds. Then the sequence \( \{x_k\} \) generated by the SHS1 method satisfies
\[
\liminf_{k \to \infty} \|g_k\| = 0.
\]

Proof. The proof is omitted.

Similar to the NHS method in [11], we can define another SHS method (denoted by SHS2 method), in which the direction is determined by
\[
d_k = -g_k + \beta_k^{MHS} d_{k-1}, \quad d_0 = -g_0.
\]
(30)
In the following, we shall establish the global convergence of the SHS2 method with the weak Wolfe line search (5). First, we have the following property.
In this section, we present some numerical results to compare the performance of the SPRP method in this paper, search

SHS2 method.

Case 2: If $g_k^T g_{k-1} < 0$, then from (30), (24), we have

\[
g_k^T d_k = \frac{1}{\mu |g_k^T d_{k-1}| + y_{k-1}^T d_{k-1}} \left( -\|g_k\|^2 y_k^T d_{k-1} - \mu \|g_k\|^2 g_k^T d_{k-1} + \|g_k\|^2 g_k^T g_{k-1}^T d_{k-1} \right)
\]

\[
\leq \frac{1}{\mu |g_k^T d_{k-1}| + y_{k-1}^T d_{k-1}} \left( \|g_k\|^2 g_k^T g_{k-1}^T d_{k-1} \right)
\]

\[
= \beta_k \mathbf{MHS} g_{k-1}^T d_{k-1}.
\]

Case 2: If $g_k^T g_{k-1} \geq 0$, then from (30), (24), we get

\[
g_k^T d_k = \frac{1}{\mu |g_k^T d_{k-1}| + y_{k-1}^T d_{k-1}} \left( -\|g_k\|^2 y_k^T d_{k-1} - \mu \|g_k\|^2 g_k^T d_{k-1} + \|g_k\|^2 g_k^T g_{k-1}^T d_{k-1} \right)
\]

\[
\leq \frac{1}{\mu |g_k^T d_{k-1}| + y_{k-1}^T d_{k-1}} \left( \|g_k\|^2 g_k^T g_{k-1}^T d_{k-1} \right)
\]

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

\[
= \beta_k \mathbf{MHS} g_{k-1}^T d_{k-1}.
\]

From the inequalities in the above two cases and $g_0^T d_0 = -\|g_0\|^2 < 0$, we have (31) by induction.

Using Lemma 3.6 and the same argument of Theorem 3.3 in [5], we have the following global convergence of the

SHS2 method.

**Theorem 3.4.**

Let Assumption A hold. \{x_k\} be generated by the SHS2 method. If step-size $\alpha_k$ is determined by the weak Wolfe line search (5), Then, we have

\[
\lim \inf_k \|g_k\| = 0.
\]

## 4. Numerical Results

In this section, we present some numerical results to compare the performance of the SPRP method in this paper, search

the DTPRP method in [13] and the CG..DESCENT method in [18].

- SPRP: the SPRP method with the weak Wolfe line search (5), with $\mu = 10^{-4}, \rho = 0.1, \sigma = 0.5$.
- DTPRP: the DTPRP method with the weak Wolfe line search (5), with $\mu = 1.2, \rho = 0.1, \sigma = 0.5$.
- CG..DESCENT: the CG..DESCENT method with the weak Wolfe line search (5), with $\rho = 0.1, \sigma = 0.5$.

All codes were written in Matlab 7.1, and run on a portable computer with 2.10GHz CPU processor, 2.0GB RAM memory, and Windows XP operating system. We stopped the iteration if the number of iteration exceed 5000 or $\|g_k\| < 10^{-3}$. Here, we use some test problems in [19] with different dimension. Our numerical results are listed in the form NI/NF/NG/CPU, where the symbols NI, NF, NG and CPU mean the number of iterations, the number of function evaluations, the number of gradient evaluations and the CPU time in seconds, respectively. 'F' means the method failed. Table 1 list the numerical results of the SPRP method and the DTPRP method and the CG..DESCENT method.

Table 1 reveals that all the test methods are efficient, and the SPRP method perform better than the DTPRP method and the CG..DESCENT method when $\tau$ is small, however, the CG..DESCENT method performs a little better than the other two methods when $\tau$ is large, that is it can ultimately solve almost 100% of the test problems. Therefore, the proposed SPRP method is comparable with the famous CG..DESCENT method.
Table 1. The results for the methods on the tested problems

<table>
<thead>
<tr>
<th>Method</th>
<th>Problem Size</th>
<th>Iterations</th>
<th>CPU Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREUROTH</td>
<td>1000</td>
<td>68/939/117</td>
<td>0.1875</td>
</tr>
<tr>
<td>SROSENBR</td>
<td>5000</td>
<td>343/4487/403</td>
<td>0.7031</td>
</tr>
<tr>
<td>BEALE</td>
<td>5000</td>
<td>41/303/48</td>
<td>0.1094</td>
</tr>
<tr>
<td>HIMMELBC</td>
<td>5000</td>
<td>28/251/29</td>
<td>0.0469</td>
</tr>
<tr>
<td>CLIFF</td>
<td>5000</td>
<td>93/738/143</td>
<td>0.2969</td>
</tr>
<tr>
<td>WOODS</td>
<td>5000</td>
<td>181/2213/220</td>
<td>0.3906</td>
</tr>
<tr>
<td>BDQRTIC</td>
<td>200</td>
<td>176/2548/243</td>
<td>0.1094</td>
</tr>
<tr>
<td>TRIDIA</td>
<td>100</td>
<td>320/4168/388</td>
<td>0.1406</td>
</tr>
<tr>
<td>NONDIA</td>
<td>5000</td>
<td>69/1495/78</td>
<td>0.2344</td>
</tr>
<tr>
<td>NONDQVAR</td>
<td>100</td>
<td>1310/4484/1312</td>
<td>0.2500</td>
</tr>
<tr>
<td>EG2</td>
<td>1000</td>
<td>1001/4821/6623</td>
<td>0.7828</td>
</tr>
<tr>
<td>DIXMAANA</td>
<td>5001</td>
<td>16/68/17</td>
<td>0.2344</td>
</tr>
<tr>
<td>DIXMAANB</td>
<td>5001</td>
<td>13/54/14</td>
<td>0.1719</td>
</tr>
<tr>
<td>DIXMACN</td>
<td>5001</td>
<td>16/69/17</td>
<td>0.1875</td>
</tr>
<tr>
<td>DIXMAANE</td>
<td>1002</td>
<td>212/624/213</td>
<td>0.5156</td>
</tr>
<tr>
<td>EDENSC</td>
<td>5000</td>
<td>33/256/45</td>
<td>0.3750</td>
</tr>
<tr>
<td>LIARWHD</td>
<td>5000</td>
<td>139/2423/164</td>
<td>0.4844</td>
</tr>
<tr>
<td>DIXON3DQ</td>
<td>100</td>
<td>440/2132/441</td>
<td>0.0781</td>
</tr>
<tr>
<td>ENGVALI</td>
<td>5000</td>
<td>26/274/55</td>
<td>0.1563</td>
</tr>
<tr>
<td>DENSCHNA</td>
<td>5000</td>
<td>28/134/29</td>
<td>0.1406</td>
</tr>
<tr>
<td>DENSCHNB</td>
<td>5000</td>
<td>13/65/14</td>
<td>0.1250</td>
</tr>
<tr>
<td>DENSCHNC</td>
<td>5000</td>
<td>37/304/72</td>
<td>0.3594</td>
</tr>
<tr>
<td>DENSCHNF</td>
<td>5000</td>
<td>31/349/41</td>
<td>0.1250</td>
</tr>
<tr>
<td>SINQUAD</td>
<td>5000</td>
<td>339/4869/393</td>
<td>0.6719</td>
</tr>
<tr>
<td>BIGGSB1</td>
<td>5000</td>
<td>3649/1753/3650</td>
<td>0.7031</td>
</tr>
<tr>
<td>EXTROSBN</td>
<td>5000</td>
<td>54/728/67</td>
<td>0.2031</td>
</tr>
<tr>
<td>ARGLINB</td>
<td>5000</td>
<td>54/3291/57</td>
<td>2.7656</td>
</tr>
<tr>
<td>FLETCHCR</td>
<td>5000</td>
<td>102/1474/141</td>
<td>0.3906</td>
</tr>
<tr>
<td>HIMMELBH</td>
<td>5000</td>
<td>21/111/22</td>
<td>0.1406</td>
</tr>
</tbody>
</table>

References


