Int. J. Adv. Appl. Math. and Mech. 2(3) (2015) 51 - 59 (ISSN: 2347-2529)
Journal homepage: www.ijaamm.com

IJAAMM

International Journal of Advances in Applied Mathematics and Mechanics

A new spectral PRP conjugate gradient method with sufficient descent property

Research Article

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

1. Introduction

In this paper, we consider the following unconstrained optimization problem

$$\{\min f(x),\ \ x \in \mathscr{R}^n\}, \tag{1}$$

where $\mathscr{R}^n$ is the n-dimensional Euclidean space, $f : \mathscr{R}^n \to \mathscr{R}$ 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

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

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

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{3}$$

in which β_k is a scalar that characterizes the method. That is, different conjugate gradient methods correspond to different choices of the parameter β_k. There have been many famous formulas for parameter β_k, such as, Fletcher-Reeves (FR)[1], Polak-Ribière-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 β_k is given by

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

* 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 $\alpha_k = \max\{\rho^j | j = 0, 1, \dots\}$ satisfying

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^\top d_k, \tag{4}$$

where $\delta, \rho \in (0, 1)$ are constants. The weak Wolfe line search is to find a step-size α_k satisfying

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \rho \alpha_k g_k^\top d_k, \\ g(x_k + \alpha_k d_k)^\top d_k \ge \sigma g_k^\top d_k, \end{cases} \tag{5}$$

where $0 < \rho \le \sigma < 1$. The strong Wolfe line search is to compute α_k such that

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \rho \alpha_k g_k^\top d_k, \\ |g(x_k + \alpha_k d_k)^\top d_k| \le \sigma |g_k^\top d_k|, \end{cases} \tag{6}$$

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 Ribière [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 α_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.[10] with

$$\beta_k^{\mathrm{WYL}} = \frac{g_k^\top \left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right)}{\|g_{k-1}\|^2} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^\top g_{k-1}}{\|g_{k-1}\|^2}.$$

The β_k^{WYL} has the following nice properties: (1) $\beta_k^{\mathrm{WYL}} \ge 0$; (2) it can avoid jamming. However, the WYL method satisfies the sufficient descent condition

$$g_k^\top d_k \le -c\|g_k\|^2, c > 0 \tag{7}$$

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 β_ks are specified as follows

$$\beta_k^{\mathrm{ZPRP}} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}|}{\|g_{k-1}\|^2}, \beta_k^{\mathrm{HPRP}} = \frac{g_k^\top \left(g_k - \frac{g_k^\top g_{k-1}}{\|g_{k-1}\|^2} g_{k-1} \right)}{\|g_{k-1}\|^2},$$

$$\beta_k^{\mathrm{DPRP}} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}|}{\mu |g_k^\top d_{k-1}| + \|g_{k-1}\|^2}, \ (\mu > 1).$$

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 β_k^{DPRP} reduces to β_k^{ZPRP}, however, to ensure the sufficient descent property (7), μ must satisfy $\mu > 1$ in β_k^{DPRP}. Moreover, the numerical results show the conjugate gradient induced by β_k^{DPRP} is not very effective. Therefore, in this paper, we will give a modified version of β_k^{DPRP}, and investigate global convergence of the corresponding method when $0 \le \mu \le 1$. In fact, we give a new parameter defined as follows:

$$\beta_k^{\mathrm{MPRP}} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} \max\{g_k^\top g_{k-1}, 0\}}{\mu |g_k^\top d_{k-1}| + \|g_{k-1}\|^2}, \ (\mu \ge 0). \tag{8}$$

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:

$$d_k = -\theta_k^{\mathrm{WZ}} g_k + \beta_k^{\mathrm{PRP}} d_{k-1}, \quad \forall k \geq 1, \tag{9}$$

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

$$d_k = -g_k - \beta_k^{\mathrm{PRP}} \frac{g_k^\top d_{k-1}}{\|g_k\|^2} g_k + \beta_k^{\mathrm{PRP}} d_{k-1}. \tag{10}$$

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

$$d_k = -g_k - \beta_k^{\mathrm{MPRP}} \frac{g_k^\top d_{k-1}}{\|g_k\|^2} g_k + \beta_k^{\mathrm{MPRP}} d_{k-1}, d_0 = -g_0, \tag{11}$$

from which, we get the following method:
SPRP Method (Spectral PRP method)
Step 0. Given an initial point $x_0 \in \mathscr{R}^n$, $\mu \geq 0$, $0 < \delta < 1$, $0 < \rho < \sigma < 1$ 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 α_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 $g_k^\top g_{k-1} \geq 0$, then the SPRP method reduces to the classical Wei-Yao-Liu method in [10].
Remark 2.1. From (11), we can easily obtain

$$g_k^\top d_k = -\|g_k\|^2, \text{ and } \|g_k\| \leq \|d_k\|. \tag{12}$$

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 $L_0 = \{x | f(x) \leq f(x_0)\}$ 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 γ such that

$$\|g(x)\| \leq \gamma. \tag{13}$$

Lemma 3.1.
The parameter β_k^{MPRP} defined by (8) satisfies

$$0 \leq \beta_k^{\mathrm{MPRP}} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}. \tag{14}$$

Proof. From the definition of β_k^{MPRP}, we have

$$0 = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{\mu|g_k^\top d_{k-1}| + \|g_{k-1}\|^2} \leq \beta_k^{\mathrm{MPRP}} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}.$$

□

Lemma 3.2.
The sequence $\{d_k\}$ generated by the SPRP method satisfies

$$\|d_k\|^2 \le \|g_k\|^2 + (\beta_k^{\mathrm{MPRP}})^2 \|d_{k-1}\|^2, \quad \forall k \ge 1. \tag{15}$$

Proof. For all $k \ge 1$, set $\theta_k = 1 + \beta_k^{\mathrm{MPRP}} g_k^\top d_{k-1} / \|g_k\|^2$, then from (11), we have

$$g_k^\top d_k = -\theta_k \|g_k\|^2 + \beta_k^{\mathrm{MPRP}} g_k^\top d_{k-1}.$$

This together with (12) shows that

$$\beta_k^{\mathrm{MPRP}} g_k^\top d_{k-1} = (\theta_k - 1)\|g_k\|^2. \tag{16}$$

Squaring both sides of (11), we obtain

$$\|d_k\|^2 = \theta_k^2 \|g_k\|^2 - 2\theta_k \beta_k^{\mathrm{MPRP}} g_k^\top d_{k-1} + (\beta_k^{\mathrm{MPRP}})^2 \|d_{k-1}\|^2. \tag{17}$$

Substituting (16) into (17), we have

$$\begin{aligned}
&\|d_k\|^2 \\
&= \theta_k^2 \|g_k\|^2 - 2\theta_k(\theta_k - 1)\|g_k\|^2 + (\beta_k^{\mathrm{MPRP}})^2 \|d_{k-1}\|^2 \\
&= -(\theta_k - 1)^2 \|g_k\|^2 + \|g_k\|^2 + (\beta_k^{\mathrm{MPRP}})^2 \|d_{k-1}\|^2 \\
&\le \|g_k\|^2 + (\beta_k^{\mathrm{MPRP}})^2 \|d_{k-1}\|^2.
\end{aligned}$$

□

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 α_k is determined by the standard Armijo line search (4), then for all k,

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{18}$$

Proof. From (4), (12) and Assumption1 (I), we have

$$\sum_{k=0}^{\infty} \alpha_k \|g_k\|^2 < \infty. \tag{19}$$

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

$$\alpha_k \ge c_1 \frac{\|g_k\|^2}{\|d_k\|^2}. \tag{20}$$

The proof of (20) is divided into the following two cases.
Case (I): If $\alpha_k = 1$, then from (12), we have $\alpha_k = 1 \ge \frac{\|g_k\|^2}{\|d_k\|^2}$.
Case (II): If $\alpha_k < 1$, then by the Armijo line search condition, $\rho^{-1}\alpha_k$ does not satisfy inequality (4). That is,

$$f(x_k + \rho^{-1}\alpha_k d_k) > f(x_k) + \delta \rho^{-1} \alpha_k g_k^\top d_k. \tag{21}$$

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

$$\begin{aligned}
&f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) \\
&= \rho^{-1}\alpha_k g(x_k + t_k \rho^{-1}\alpha_k d_k)^\top d_k \\
&= \rho^{-1}\alpha_k g_k^\top d_k + \rho^{-1}\alpha_k (g(x_k + t_k\rho^{-1}\alpha_k d_k) - g_k)^\top d_k \\
&\le \rho^{-1}\alpha_k g_k^\top d_k + L\rho^{-2}\alpha_k^2 \|d_k\|^2.
\end{aligned}$$

Substituting the last inequality into (21), we have

$$\alpha_k \ge \frac{(1-\delta)\rho}{L} \frac{\|g_k\|^2}{\|d_k\|^2}.$$

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). □

Theorem 3.1.
Suppose that the conditions in Assumption A hold. Consider the SPRP method, and the step-size α_k is determined by the standard Armijo line search (4). Then we have

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{22}$$

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

$$\|g_k\| \geq \varepsilon, \ \ \forall k \geq 0.$$

From (14) and (15), we have

$$\begin{aligned}
&\frac{\|d_k\|^2}{\|g_k\|^4} \\
&\leq \frac{\|g_k\|^2 + (\beta_k^{\mathrm{MPRP}})^2\|d_{k-1}\|^2}{\|g_k\|^4} \\
&= (\beta_k^{\mathrm{MPRP}})^2 \frac{\|g_{k-1}\|^4}{\|g_k\|^4}\frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2} \\
&\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2} \\
&\leq \frac{1}{\|g_k\|^2} + \frac{1}{\|g_{k-1}\|^2} + \cdots + \frac{1}{\|g_0\|^2} \\
&\leq \frac{k+1}{\varepsilon^2}.
\end{aligned}$$

That is

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

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 α_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 α_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

$$d_k = -g_k - \beta_k^{\mathrm{MHS}} \frac{g_k^\top d_{k-1}}{\|g_k\|^2} g_k + \beta_k^{\mathrm{MHS}} d_{k-1}, d_0 = -g_0, \tag{23}$$

where

$$\beta_k^{\mathrm{MHS}} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\max\{g_k^\top g_{k-1}, 0\}}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}}, \ \ (\mu > 0). \tag{24}$$

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 $\varepsilon > 0$ such that $\|g_k\| \geq \varepsilon$ for all $k \geq 0$, then we have

$$\sum_{k=0}^{\infty} \|u_{k+1} - u_k\|^2 < +\infty, \tag{25}$$

where $u_k = d_k/\|d_k\|$.

Proof. From (5), we have

$$y_{k-1}^{\top} d_{k-1} \geq (1-\sigma)(-g_{k-1}^{\top} d_{k-1}) = (1-\sigma)\|g_{k-1}\|^2. \tag{26}$$

From (12) and $g_k \geq \varepsilon$ for all k, we have $\|d_k\| > 0$ for all k. Therefore, u_k is well-defined. Define

$$r_k = -\frac{\left(1 + \beta_k^{\mathrm{MHS}} \dfrac{g_k^{\top} d_{k-1}}{\|g_k\|^2}\right)}{\|d_k\|} g_k \ \text{ and } \ \delta_k = \beta_k^{\mathrm{MHS}} \frac{\|d_{k-1}\|}{\|d_k\|}.$$

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

$$\|u_k - u_{k-1}\| \leq \|(1+\delta_k)(u_k - u_{k-1})\| \leq \|u_k - \delta u_{k-1}\| + \|\delta u_k - u_{k-1}\| = 2\|r_k\|. \tag{27}$$

From (24) and (27), we have

$$\beta_k^{\mathrm{MHS}} \frac{|g_k^{\top} d_{k-1}|}{\|g_k\|^2} \leq \frac{\|g_k\|^2}{\mu |g_k^{\top} d_{k-1}|} \frac{|g_k^{\top} d_{k-1}|}{\|g_k\|^2} \leq \frac{1}{\mu}. \tag{28}$$

From (13) and (28), it follows that there exists a constant $M_1 \geq 0$ such that

$$\|-\left(1 + \beta_k^{\mathrm{MHS}} \frac{g_k^{\top} d_{k-1}}{\|g_k\|^2}\right) g_k\| \leq \|g_k\| + \frac{1}{\mu}\gamma \leq \gamma + \frac{1}{\mu}\gamma \doteq M_1. \tag{29}$$

Thus, from (18) and (29), we get

$$\sum_{k=0}^{\infty} \|r_k\|^2 \leq \sum_{k=0}^{\infty} \frac{M_1^2}{\|d_k\|^2} = \sum_{k=0}^{\infty} \frac{M_1^2}{\|g_k\|^4} \frac{\|g_k\|^4}{\|d_k\|^2} \leq \frac{M_1^2}{\varepsilon^4} \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty,$$

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

$$d_k = -g_k + \beta_k^{\mathrm{MHS}} d_{k-1}, d_0 = -g_0. \tag{30}$$

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.

Lemma 3.6.
If the step-size α_k is determined by the weak Wolfe line search (5), then the parameter β_k^{MHS} defined by (24) and the direction d_k determined by (30) satisfy

$$\beta_k^{\text{MHS}} \le \frac{g_k^\top d_k}{g_{k-1}^\top d_{k-1}}, \quad \text{and} \quad g_k^\top d_k < 0. \tag{31}$$

Proof. The proof is divided into the following two cases.
Case 1: If $g_k^\top g_{k-1} < 0$, then from (30), (24), we have

$$\begin{aligned} & g_k^\top d_k \\ &= \frac{1}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} \left(-\|g_k\|^2 y_k^\top d_{k-1} - \mu\|g_k\|^2 |g_k^\top d_{k-1}| + \|g_k\|^2 g_k^\top d_{k-1}\right) \\ &\le \frac{\|g_k\|^2}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} g_{k-1}^\top d_{k-1} \\ &= \beta_k^{\text{MHS}} g_{k-1}^\top d_{k-1}. \end{aligned}$$

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

$$\begin{aligned} & g_k^\top d_k \\ &= \frac{1}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} \left(-\|g_k\|^2 y_k^\top d_{k-1} - \mu\|g_k\|^2 |g_k^\top d_{k-1}| + \|g_k\|^2 g_k^\top d_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}| g_k^\top d_{k-1}\right) \\ &\le \frac{1}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} \left(\|g_k\|^2 g_{k-1}^\top d_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}| g_k^\top d_{k-1}\right) \\ &= \frac{1}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} \left(\left(\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}|\right) g_{k-1}^\top d_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}| y_{k-1}^\top d_{k-1}\right) \\ &\le \frac{1}{\mu|g_k^\top d_{k-1}| + y_{k-1}^\top d_{k-1}} \left(\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^\top g_{k-1}|\right) g_{k-1}^\top d_{k-1} \\ &= \beta_k^{\text{MHS}} g_{k-1}^\top d_{k-1}. \end{aligned}$$

From the inequalities in the above two cases and $g_0^\top d_0 = -\|g_0\|^2 < 0$, 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 α_k is determined by the weak Wolfe line search (5). Then, we have

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

4. Numerical Results

In this section, we present some numerical results to compare the performance of the SPRP method in this paper, 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 $\|g_k\| < 10^{-5}$. 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 τ is small, however, the CG_DESCENT method performs a little better than the other two methods when τ 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

FREUROTH	1000	68/939/117/0.1875	79/1135/151/0.1875	212/4697/748/1.0469
SROSENBR	5000	343 /4487 / 403 /0.7031	313 /3936 / 372 /0.5000	1001 /13100 /1159 /2.6875
BEALE	5000	41 /303 / 48 /0.1094	55 /409 / 65 /0.1406	68 /581 / 88 /0.1563
HIMMELBC	5000	28 /251 / 29 /0.0469	23 /195 / 26 /0.0469	31 /302 / 34 /0.0625
CLIFF	5000	93 /798 / 143 /0.2969	931 /6434 /1286 /2.0625	167 /1684 / 285 /0.5938
WOODS	5000	181 /2215 / 220 /0.3906	257 /2946 / 296 /0.4063	676 /8498 / 758 /1.0313
BDQRTIC	200	176 /2548 / 243 /0.1094	200 /3184 / 362 /0.1563	449 /6789 / 619 /0.3125
TRIDIA	100	320 /4168 / 388 /0.1406	322 /4024 / 386 /0.0625	335 /4435 / 390 /0.1406
NONDIA	5000	69 /1495 / 78 /0.2344	344 /7184 / 392 /1.0625	465 /10365 / 536 /1.0938
NONDQUAR	100	1310 /4484 /1312 /0.2500	1919 /6134 /1921 /0.2969	1966 /7778 /1969 /0.4688
EG2	100	1001 /48217 /6623 /7.8281	F	105 /1101 / 114 /0.0625
DIXMAANA	5001	16 / 68 / 17 /0.2344	19 / 71 / 20 /0.2031	18 / 87 / 19 /0.3906
DIXMAANB	5001	13 / 54 / 14 /0.1719	10 / 40 / 11 /0.1250	14 / 68 / 15 /0.2188
DIXMAANC	5001	16 / 69 / 17 /0.1875	20 / 78 / 21 /0.2500	18 / 90 / 19 /0.2188
DIXMAANE	1002	212 /624 / 213 /0.5156	193 /490 / 194 /0.3906	207 /636 / 208 /0.4219
EDENSCH	5000	33 /256 / 45 /0.3750	27 /276 / 53 /0.3594	41 /303 / 43 /0.4844
LIARWHD	5000	139 /2423 / 164 /0.4844	193 /3148 / 210 /0.5000	489 /8633 / 551 /1.2656
DIXON3DQ	100	440 /2132 / 441 /0.0781	539 /2380 / 540 /0.0625	601 /3003 / 625 /0.1406
ENGVAL1	5000	26 /274 / 55 /0.1563	37 /562 / 130 /0.1563	42 /619 / 157 /0.2969
DENSCHNA	5000	28 /134 / 29 /0.1406	28 /128 / 30 /0.1719	36 /199 / 37 /0.1563
DENSCHNB	5000	13 / 65 / 14 /0.1250	19 / 79 / 20 /0.1094	18 / 94 / 19 /0.0625
DENSCHNC	5000	37 /304 / 72 /0.3594	35 /577 / 140 /0.6875	61 /457 / 68 /0.3438
DENSCHNF	5000	31 /349 / 41 /0.1250	24 /256 / 30 /0.0781	36 /421 / 40 /0.0938
SINQUAD	1000	339 /4869 / 393 /0.6719	F	F
BIGGSB1	500	3649 /17531 /3650 /0.7031	2002 /8811 /2003 /0.3281	2599 /12908 /2683 /0.4688
EXTROSNB	5000	54 /728 / 67 /0.2031	37 /483 / 45 /0.1250	52 /719 / 63 /0.1250
ARGLINB	5000	F	329 /14192 / 517 /2.7656	583 /24599 / 652 /3.4844
FLETCHCR	5000	102 /1474 / 141 /0.3906	88 /1236 / 104 /0.2813	48 /711 / 58 /0.2031
HIMMELBH	5000	21 /111 / 22 /0.1406	22 /112 / 24 /0.1875	30 /230 / 57 /0.1094

References

[1] R. Fletcher, C. Reeves C. Function minimization by conjugate gradients. J. Comput. 7 (1964) 149-154 .

[2] B. Polak, Ribière, G., Note surla convergence des méthodes de directions conjuguées, Revue Francaise Imformat Recherche Opertionelle 16 (1969) 35-43.

[3] J. C. Gilber, J. Nocedal, Global convergence properties of conjugate gradient methods for optimization. SIAM J. Optim. 2 (1992) 21-42.

[4] Y. L. Liu, C. S. Storey, Efficient generalized conjugate gradient algorithms, part 1: Theory. J. Optim. Theory Appl. 69 (1991) 129-137.

[5] Y. H. Dai, Y. X. Yuan, A nonlinear conjugate gradient method with a strong global convergence property. SIAM J. Optim. 10 (2000) 177-182.

[6] M. R. Hestenes, E. L. Stiefel, Method of conjugate gradient for solving linear systems, J. Res. Nat. Bur. Stand. 49 (1952) 409-432.

[7] R. Fletcher, Practical methods of Optimization, Unconstrained Optimization, Vol. I, Wiley & Sons, New York, 1987.

[8] M. J. D. Powell, Nonconvex minimization calculations and the conjugate gradient method, in Numerical Analysis, D.F. Griffiths, ed., Lecture Notes in Mathematics, Vol. 1066, Springer-Verlag, Berlin, 1984.

[9] Y. H. Dai, Nonlinear conjugate gradient methods, Wiley Encyclopedia of Operations Research and Management Science, 2011.

[10] Z. X. Wei, S. W. Yao, L. Y. Liu, The convergence properties of some new conjugate gradient methods. Appl. Math. Comput. 183 (2006) 1341-1350.

[11] L. Zhang, An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization problems. Appl. Math. Comput. 215 (2009) 2269-2274.

[12] H. D. Huang, Y. J. Liu, Z. X. Wei, Global convergence of a modified PRP conjugate gradient method. Journal of Mathematical Research & Exposition 30(1) (2010) 141-148.

[13] Z. F. Dai, F. H. Wen, Another improved Wei-Yao-Liu nonlinear conjugate gradient method with sufficient descent property. Appl. Math. Comput. 218 (2012) 7421-7430.

[14] W. W. Hager, H. C. Zhang, A survey of nonlinear conjugate gradient methods, Pac. J. Optim. 2 (2006) 35-58.

[15] Z. Wan, Z. L. Yang, Y. L. Wang, New spectral PRP conjugate gradient method for unconstrained optimization.

Appl. Math. Lett. 24 (2012) 16-22.
[16] G. Zoutendijk, Nonlinear programming, computational methods, in Integer and Nonlinear Programming, J. Abadie, ed. North-Holland, Amsterdam, 1970.
[17] Y. H. Dai, L. Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods. Appl. Math. Optim. 43 (2001) 97-101.
[18] W. W. Hager, H. C. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. 16 (2005) 170-192.
[19] K. E. Bongartz, A. R. Conn, N. I. M. Gould, Toint P.L., CUTE: constrained and unconstrained testing environments. ACM Trans. Math. Softw. 21 (1995) 123-160 .
[20] X. Y. Zhou, Q. Sun, Stability analysis of a fractional-order HBV infection model, Int. J. Adv. Appl. Math. and Mech. 2(2) (2014) 1-6.