


Article

A Modified Sufficient Descent Polak–Ribière–Polyak Type Conjugate Gradient Method for Unconstrained Optimization Problems

Xiuyun Zheng * and Jiarong Shi 

School of Science, Xi'an University of Architecture and Technology, Xi'an 710055, China; shijiarong@xauat.edu.cn

* Correspondence: xyzheng@xauat.edu.cn; Tel.: +86-029-8220-5670

Received: 4 July 2018; Accepted: 3 September 2018; Published: 6 September 2018

Abstract: In this paper, a modification to the Polak–Ribière–Polyak (PRP) nonlinear conjugate gradient method is presented. The proposed method always generates a sufficient descent direction independent of the accuracy of the line search and the convexity of the objective function. Under appropriate conditions, the modified method is proved to possess global convergence under the Wolfe or Armijo-type line search. Moreover, the proposed methodology is adopted in the Hestenes–Stiefel (HS) and Liu–Storey (LS) methods. Extensive preliminary numerical experiments are used to illustrate the efficiency of the proposed method.

Keywords: unconstrained optimization; conjugate gradient method; sufficient descent; global convergence

1. Introduction

Conjugate gradient methods are among the most popular methods for solving optimization problem, especially for large-scale problems due to the simplicity and low storage of their iterative form [1].

Consider the following unconstrained optimization problem:

$$\min\{f(x) : x \in R^n\}, \quad (1)$$

where $f : R^n \rightarrow R$ is continuously differentiable. Let x_0 be any initial point of the solution of the problem (1), then the conjugate gradient method generates an iteration sequence as follows:

$$x_{k+1} = x_k + \alpha_k d_k, k = 0, 1, 2, \dots \quad (2)$$

where x_k is the k th iterative point, $\alpha_k > 0$ is a steplength which is obtained by carrying out some line search, and d_k is a search direction defined by

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

where $g_k = g(x_k)$ denotes the gradient of the function $f(x)$ at x_k and β_k is a scalar that determines different conjugate gradient methods [2–6]. In this paper, we focus our attention on well-known methods such as Polak–Ribière–Polyak (PRP) [4,5], Hestenes–Stiefel (HS) [3] and Liu–Storey (LS) [6] methods which share the same numerator $g_k^T y_{k-1}$ in β_k . The update parameters of these methods are, respectively, given by

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \quad (4)$$

where $y_{k-1} = g_k - g_{k-1}$, $\|\cdot\|$ denotes the Euclidean norm of vectors. Other nonlinear conjugate gradient methods and their global convergence can be found in [1,7].

It is well-known that the PRP, HS and LS methods are generally regarded to be the most efficient methods in practical computation. This can be attributed to the property (*), which was derived in Gilbert and Nocedal [8]. Polak–Ribière [4] obtained the global convergence of the PRP method for the strongly convex functions with exact line search. Yuan [9] also obtained the global convergence of the PRP method under the assumption that the search direction satisfies a descent condition:

$$g_k^T d_k < 0, \forall k \geq 0. \tag{5}$$

and the following standard Wolfe line search

$$\begin{cases} f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \end{cases} \tag{6}$$

where $0 < \delta \leq \sigma < 1$.

Their convergence properties are not so good in many situations. Powell [10] gave a counter example which showed that there exist nonconvex functions on which the PRP method does not converge globally even if the exact line search is used. Inspired by Powell’s work, Gilbert and Nocedal [8] proved that the modified PRP method is globally convergent in which β_k is given by $\beta_k^{PRP+} = \max\{\beta_k^{PRP}, 0\}$. The search direction prevents effectively jamming phenomena from occurring and satisfies the descent property (5) or the following sufficient descent condition:

$$g_k^T d_k \leq -c \|g_k\|^2, \forall k \geq 0, c > 0, \tag{7}$$

which is very important for establishing the global convergence of the proposed method. In [11], Hager and Zhang proposed a modified HS formula for β_k defined by

$$\beta_k^{HZ} = \beta_k^{HS} - 2 \frac{\|y_k\|^2 g_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2}.$$

More specifically, in their proposed method, called CG-DESCENT. They showed that the method possesses the sufficient descent property with $c = 7/8$. Afterwards, they also presented the following extension of β_k^{HZ} :

$$\beta_k^{HZ} = \beta_k^{HS} - \theta_k \frac{\|y_k\|^2 g_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2},$$

where θ_k is a nonnegative parameter. If $\theta_k \geq \theta > 1/4$, then the method possesses the sufficient descent property with $c = 1 - 1/(4\theta)$. Cheng [12] developed a two term PRP-based descent method satisfying (7) by use of the projection technique for unconstrained optimization problem. Yu et al. [13] proposed a modified form of β_k^{PRP} as follows:

$$\beta_k^{DPRP} = \beta_k^{PRP} - \mu \frac{\|y_k\|^2}{\|g_k\|^4} g_{k+1}^T d_k.$$

It is important that if $\mu > \frac{1}{4}$, then the condition (7) is achieved with $c = 1 - \frac{1}{4\mu}$. Yuan [14] present a new PRP formula defined by

$$\beta_k^{MPRP} = \beta_k^{PRP} - \min\{\beta_k^{PRP}, \mu \frac{\|y_k\|^2}{\|g_k\|^4} g_{k+1}^T d_k\},$$

where $\mu > \frac{1}{4}$ guaranteeing the descent property (7) and $\beta_k^{MPRP} \geq 0$. Livieris and Pintelas [15] proposed a new class of spectral conjugate gradient methods which ensures sufficient descent independent of the accuracy of the line search.

Wei et al. [16] gave a variant of the PRP method called the VPRP method. The parameter β_k in the VPRP method is given by

$$\beta_k^{VPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2}.$$

Based on the VPRP method, Zhang [17] made a little modification and obtained the NPRP method as follows,

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2},$$

and established the sufficient descent property (7) of the NPRP method. Recently, Zhang [18] proposed a three-term conjugate gradient method called MPRP method in which the direction d_k takes the following form:

$$d_k = -g_k + \beta_k^{PRP} d_{k-1} - \frac{g_k^T d_{k-1}}{\|g_{k-1}\|^2} y_{k-1}, \tag{8}$$

leading to the MPRP method with the sufficient descent property. This property always holds independent of any line search and the convexity of the objective function. Under the following line search

$$f(x_k + \alpha_k d_k) \leq f(x_k) - \delta \alpha_k^2 \|d_k\|^2, \tag{9}$$

where $\alpha_k = \max\{\rho^i, i = 1, 2, \dots\}$ with $0 < \rho, \delta < 1$, the global convergence of the MPRP method is established. Note that the MPRP method in [18] will reduce to the standard PRP method if exact line search is used and converges globally under the line search (9). However, it fails to converge under the weak Wolfe line search (6). The main reason lies in the trust region property (Lemma 1 in Section 2) that is not satisfied by the MPRP method. Based on the method in [12,18], Dong et al. [19] propose a three-term PRP-type conjugate gradient method which always satisfies the sufficient descent condition independently of line searches employed.

Motivated by the above observations, we propose a modified three-term PRP formula based on (8), which possesses not only the sufficient descent property but also the trust region feature. In the following, we first reformulate the search direction (8) into a new form, which can be written as follows:

$$d_k = -g_k + \beta_k^{PRP} d_{k-1} - \beta_k^{PRP} \frac{g_k^T d_{k-1}}{g_k^T y_{k-1}} y_{k-1}. \tag{10}$$

Then, we can consider the following general iteration form:

$$d_k = -g_k + \beta_k d_{k-1} - \beta_k \frac{g_k^T d_{k-1}}{g_k^T y_{k-1}} y_{k-1}. \tag{11}$$

with any β_k , it is not difficult to deduce that the direction defined by (11) satisfies

$$d_k^T g_k = -\|g_k\|^2, \tag{12}$$

which is independent of any line search and the convexity of the objective function.

In this paper, we further study the PRP method and suggest a new three term PRP method to improve the numerical performance and obtain better property of the PRP method. The remaining of this paper is organized as follows. In Section 2, we present a modified PRP method by using a new technique and establish its global convergence. In Section 3, the new technique is extended to the HS and LS method. In the last section, some numerical results are reported to show the modified methods are efficient.

2. The Modified PRP Method and Its Properties

In order to have the sufficient descent condition and keep simple structure and good properties, we take a modification to the denominator of the PRP formula, namely,

$$\beta_k^{ZPRP} = \frac{g_k^T y_{k-1}}{\max\{\mu \|d_{k-1}\| \|y_{k-1}\|, \|g_{k-1}\|^2\}}, \tag{13}$$

where $\mu > 0$. For convenience, we call the iterative form by (2), (11) and (13) a ZPRP method. It is obvious that the ZPRP method reduces to the PRP method if $\mu \|d_{k-1}\| \|y_{k-1}\| \leq \|g_{k-1}\|^2$.

Then, we give the modified PRP type conjugate gradient method below (Algorithm 1).

Algorithm 1 Modified PRP-type Conjugate Gradient Method

- Step 0: Given an initial point $x_0 \in R^n$, $\epsilon > 0$, $d_0 = -g_0$, set $k := 0$.
 - Step 1: If $\|g_k\| \leq \epsilon$, then stop. Otherwise, go to Step 2.
 - Step 2: Find the step size α_k satisfying a suitable line search.
 - Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$.
 - Step 4: Compute the search direction d_{k+1} by (11) where $\beta_{k+1} = \beta_{k+1}^{ZPRP}$ is given by (13).
 - Step 5: Set $k := k + 1$ and go to Step 1.
-

The following lemma shows that the direction d_k determined by (11) satisfies a trust region property.

Lemma 1. Let d_k be defined by (11) with β_k^{ZPRP} , then we have

$$\|d_k\| \leq \left(1 + \frac{2}{\mu}\right) \|g_k\|. \tag{14}$$

Proof of Lemma 1. By (13), for all $k \geq 1$, we have

$$\begin{aligned} |\beta_k^{ZPRP}| &\leq \frac{\|g_k\| \|y_{k-1}\|}{\max\{\mu \|d_{k-1}\| \|y_{k-1}\|, \|g_{k-1}\|^2\}} \\ &\leq \frac{\|g_k\|}{\mu \|d_{k-1}\|} \end{aligned} \tag{15}$$

From (11), (13) and (15), we obtain

$$\begin{aligned} \|d_k\| &\leq \|g_k\| + |\beta_k^{ZPRP}| \|d_{k-1}\| + |\beta_k^{ZPRP}| \frac{\|g_k\| \|d_{k-1}\|}{|g_k^T y_{k-1}|} \|y_{k-1}\| \\ &\leq \|g_k\| + \frac{1}{\mu} \|g_k\| + \frac{|g_k^T y_{k-1}|}{\mu \|d_{k-1}\| \|y_{k-1}\|} \frac{\|g_k\| \|d_{k-1}\|}{|g_k^T y_{k-1}|} \|y_{k-1}\| \\ &= \left(1 + \frac{2}{\mu}\right) \|g_k\|. \end{aligned} \tag{16}$$

The proof is completed. \square

3. Global Convergence of the ZPRP Method

In this section, we come to show the global convergence of our proposed method. The following assumptions are often used in the literature to analyze the global convergence of conjugate gradient methods with inexact line searches.

Assumption 1

- (i) The level set $\Omega = \{x \in R^n | f(x) \leq f(x_0)\}$ is bounded.
- (ii) In some neighborhood N of Ω , f is continuously differentiable and its gradient is Lipschitz continuous, that is, there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \leq L \|x - y\|, \forall x, y \in N$.

We first prove the ZPRP method is globally convergent with Wolfe line search (6). Under Assumption 1, we give a useful Zoutendijk condition [20].

Lemma 2. Suppose that Assumption 1 holds. Consider the method in the form of (2) and (3) where d_k is a descent direction and α_k satisfies the Wolfe line search conditions (6). Then we have

$$\sum_{k \geq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{17}$$

Obviously, the Zoutendijk condition (17) and (12) imply that

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

Theorem 1. Suppose that Assumption 1 holds. Consider the ZPRP method, and α_k is obtained by the Wolfe conditions (6). Then we have

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \tag{19}$$

Proof of Theorem 1. By Lemma 1, we have $\|d_k\| \leq (1 + \frac{2}{\mu})\|g_k\|$. Let $C = 1 + \frac{2}{\mu}$, then we get

$$\|d_k\|^2 \leq C^2 \|g_k\|^2, \tag{20}$$

which implies

$$\sum_{k=0}^{\infty} \|g_k\|^2 \leq C^2 \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty. \tag{21}$$

Hence, (19) holds. The proof is completed. \square

Next, we prove the global convergence of the ZPRP method under the condition (9).

Theorem 2. Suppose that Assumption 1 holds. Consider the ZPRP method and α_k satisfies the Armijo line search (9). Then we have

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

Proof of Theorem 2. Suppose that the conclusion is not true. Then there exists a constant $\epsilon > 0$ such that $\forall k \geq 0$,

$$\|g_k\| \geq \epsilon. \tag{23}$$

From (9) and Assumption 1 (i), we have

$$\lim_{k \rightarrow \infty} \alpha_k^2 \|d_k\|^2 = 0. \tag{24}$$

If $\liminf_{k \rightarrow \infty} \alpha_k > 0$, we get from (24) that $\liminf_{k \rightarrow \infty} \|d_k\| = 0$. From (12), we get $\liminf_{k \rightarrow \infty} \|g_k\| = 0$, which contradicts (23).

Suppose $\liminf_{k \rightarrow \infty} \alpha_k = 0$, then there is an infinite index set K such that

$$\lim_{k \in K, k \rightarrow \infty} \alpha_k = 0. \tag{25}$$

From (9), it follows that when $k \in K$ is sufficiently large, $\rho^{-1}\alpha_k$ satisfies the following inequality,

$$f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) > -\delta \rho^{-2} \alpha_k^2 \|d_k\|^2. \tag{26}$$

By Assumption 1 (ii) and the mean value theorem, there is a $\eta_k \in (0, 1)$ such that

$$\begin{aligned} f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) &= \rho^{-1}\alpha_k g(x_k + \eta_k \rho^{-1}\alpha_k d_k)^T d_k \\ &= \rho^{-1}\alpha_k (g(x_k + \eta_k \rho^{-1}\alpha_k d_k) - g(x_k))^T d_k + \rho^{-1}\alpha_k g(x_k)^T d_k \\ &\leq L \rho^{-2} \alpha_k^2 \|d_k\|^2 - \rho^{-1}\alpha_k g_k^T d_k \\ &= L \rho^{-2} \alpha_k^2 \|d_k\|^2 - \rho^{-1}\alpha_k \|g_k\|^2 \end{aligned} \tag{27}$$

By (20), (26) and (27), we can get that

$$\|g_k\|^2 \leq \rho^{-1}(\delta + L)\alpha_k C^2. \tag{28}$$

Together with (25), (28) implies $\lim_{k \in K, k \rightarrow \infty} \|g_k\| = 0$. This also yields a contradiction. The proof is completed. \square

4. Extension to the HS and LS Method

In this section, we extend the idea above to the HS and LS method. The corresponding method is called the ZHS method and the ZLS method in which β_k is respectively defined by

$$\beta_k^{ZHS} = \frac{g_k^T y_{k-1}}{\max\{\mu \|d_{k-1}\| \|y_{k-1}\|, d_{k-1}^T y_{k-1}\}}, \beta_k^{ZLS} = \frac{g_k^T y_{k-1}}{\max\{\mu \|d_{k-1}\| \|y_{k-1}\|, -g_{k-1}^T d_{k-1}\}} \quad (29)$$

where $\mu > 0$. It is obvious that $\beta_k^{ZLS} = \beta_k^{ZPRP}$ since $g_k^T d_k = -\|g_k\|^2$. Hence, we now only need to discuss the global convergence of the ZHS method.

The following theorem shows that the ZHS method converges globally with the Wolfe line search (6).

Theorem 3. *Let Assumption 1 hold. Consider the ZHS method and α_k is obtained by the Wolfe line search (6), then*

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (30)$$

Proof of Theorem 3. Suppose by contradiction that the conclusion is not true. Then there exists a constant $\epsilon > 0$ such that $\|g_k\| > \epsilon, \forall k \geq 1$. From (29), it follows that

$$|\beta_k^{ZHS}| \leq \frac{\|g_k\| \|y_{k-1}\|}{\mu \|d_{k-1}\| \|y_{k-1}\|} = \frac{\|g_k\|}{\mu \|d_{k-1}\|}. \quad (31)$$

By (11) with $\beta_k = \beta_k^{ZHS}$, we can get that

$$\begin{aligned} \|d_k\| &\leq \|g_k\| + |\beta_k^{ZHS}| \|d_{k-1}\| + |\beta_k^{ZHS}| \frac{\|g_k\| \|d_{k-1}\|}{|g_k^T y_{k-1}|} \|y_{k-1}\| \\ &\leq \|g_k\| + \frac{\|g_k\|}{\mu} + \frac{\|g_k\|}{\mu} \\ &= (1 + \frac{2}{\mu}) \|g_k\|. \end{aligned} \quad (32)$$

Hence, combing with (17),

$$\sum_{k=0}^{\infty} \|g_k\|^2 \leq (1 + \frac{2}{\mu})^2 \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (33)$$

which leads to a contradiction. The proof is completed. \square

The following result shows that the ZHS method with the Armijo line search (9) possesses global convergence.

Theorem 4. *Let Assumption 1 hold. Consider the ZHS method and α_k is obtained by the line search (9), then*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (34)$$

Proof of Theorem 4. The proof is similar to the proof of the global convergent property of the ZPRP method given in Theorem 2 in this paper. We omit it here. \square

5. Numerical Experiments

In this section, we report some numerical results on some of the unconstrained optimization problems in the CUTE [21] test problem library. We test the ZPRP method and ZHS method, and compare the performance of the these two methods with the MPRP method in [18]. The parameters

$\delta = 10^{-4}, \rho = 0.3$ and $\mu = 0.001$. All codes were written in MATLAB R2012a and run on PC with 3.00 GHz CPU processor and Win 7 operation system. We use the stopping iteration $\|g_k\|_\infty \leq 10^{-6}$. The detailed numerical results are listed on the web site: <http://mathxiuxiu.blog.sohu.com/326066259.html>.

We first evaluate the performance of the ZPRP method with that of CG-DESCENT proposed by Hager and Zhang (2005) and all methods with the Wolfe line search (6). Figures 1–3 show the numerical performance of the above methods related to the total number of iterations, the number of function and gradient evaluations, CPU time, respectively, which are evaluated using the profiles of Dolan and Moré [22]. For each method, we plot the fraction P of problems for which the method is within a factor t of the smallest number of iterations, or the smallest number of function evaluations or least CPU time, respectively. The left side of the figure gives the percentage of the test problems for which a method is the fastest; the right side gives the percentage of the test problems that are successfully solved by each of the methods. The top curve is the method that solved the most problems in a time that was within a factor t of the best time. Clearly, the ZPRP method has the better performance since it illustrates the best probability of being the optimal solver, outperforming CG-DESCENT. From Figure 1, we can obtain the ZPRP method solves about 59.5% of the test problems with the least number of function evaluations while CG-DESCENT solve about 56.5% of the test problems. For the total number of function and gradient evaluations, in Figure 2 illustrates that the ZPRP method solves 55.2% of the test problems with the least number of function and gradient evaluations while CG-DESCENT solve about 52.2% of the test problems. Therefore, the ZPRP method outperforms CG-DESCENT.

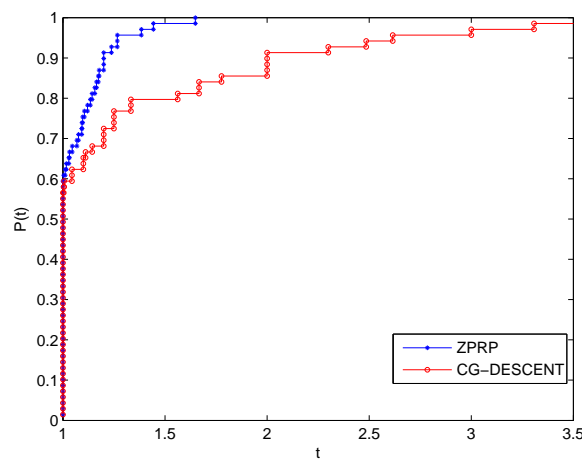


Figure 1. The number of iteration.

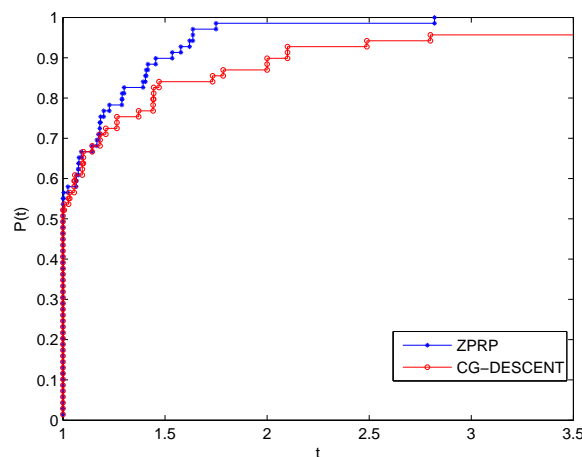


Figure 2. The total number of function and gradient evaluations.

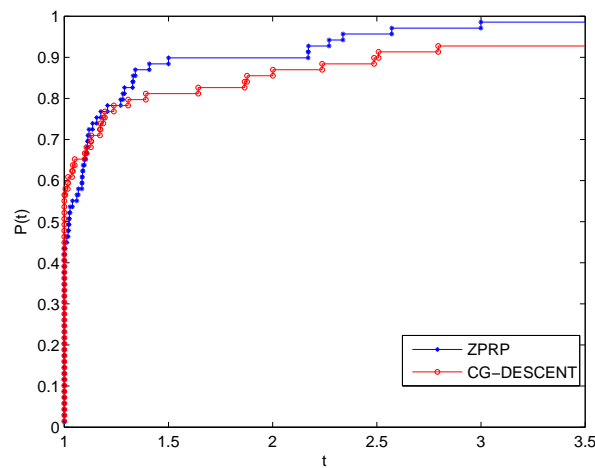


Figure 3. The total CPU time.

In the sequel, we compare the performance of the ZPRP method with that of the ZHS method and the MPRP method in [18] and all methods with the line search (9). Figures 4–6 list the performance of the above methods relative to CPU time, the number of function evaluations and the number of gradient evaluations, respectively. From Figure 4, we can observe that the ZPRP method outperforms the MPRP and ZHS method. More analytically, the performance profile for the number of iteration shows that ZPRP can solve 61% of the test problems with the least number while MPRP and ZHS solve about 47.5% and 45.2% of the test problems, respectively. As regards the number of function and gradient evaluations, Figure 5 shows that the ZPRP solves 80% of the test problems with the least number. Hence, the performance of the ZPRP method slightly better than that of the MPRP and ZHS methods.

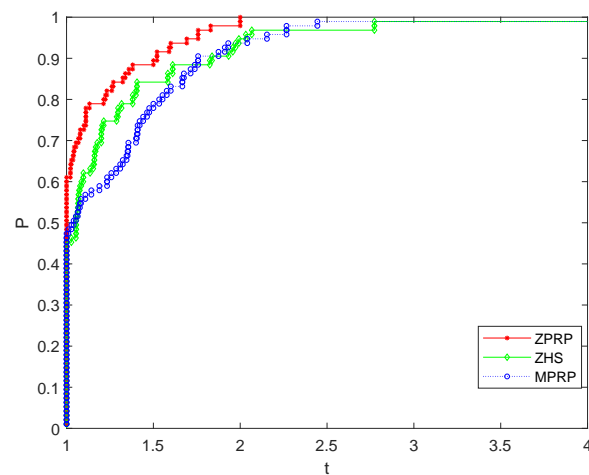


Figure 4. The number of iteration.

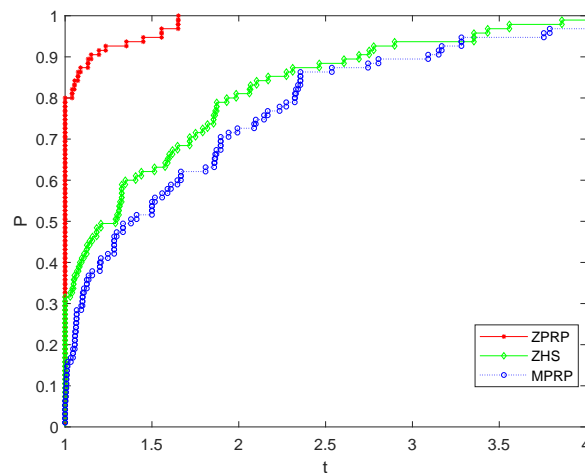


Figure 5. The total number of function and gradient evaluations.

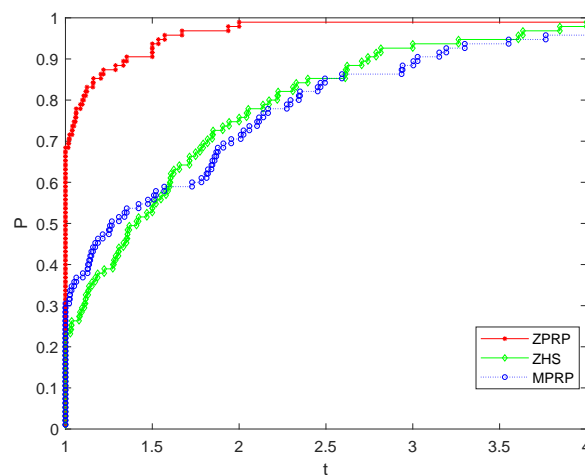


Figure 6. The total CPU time.

6. Conclusions

In this paper, we first proposed a modified PRP formula which provides sufficient descent directions for the objective function independent of any line search. Then we applied the technique to HS and LS conjugate gradient methods which also ensure the sufficient descent property. The global convergence of modified methods are established under the standard Wolfe line search or Armijo line search. Moreover, numerical experiments show that the proposed methods are promising. Our future work is concentrated on applying our coefficient β_k with spectral conjugate gradient method [15] which ensures sufficient descent independent of the accuracy of the line search and studying the convergence properties of a spectral conjugate gradient method.

Author Contributions: J.S. made the numerical experiments, X.Z. was in charge of the overall research of the paper.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the editor and the reviewers for their very careful reading and constructive comments on this paper. This work is supported by the National Natural Science Foundation of China (No. 11702206), the Natural Science Basic Research Plan of Shaanxi Province (No. 2018JQ1043) and Shaanxi Provincial Education Department (No. 16JK1435).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dai, Y.; Yuan, Y. *Nonlinear Conjugate Gradient Methods*; Shanghai Scientific and Technical Publishers: Shanghai, China, 2000.
2. Fletcher, R.; Reeves, C. Function minimization by conjugate gradients. *Comput. J.* **1964**, *7*, 149–154. [[CrossRef](#)]
3. Hestenes, M.; Stiefel, E. Methods of conjugate gradient for solving linear systems. *J. Res. Natl. Bur. Stand.* **1952**, *49*, 409–436. [[CrossRef](#)]
4. Polak, B.; Ribière, G. Note sur la convergence des méthodes de directions conjuguées. *Rev. Fr. Inform. Rech. Oper.* **1969**, *3*, 35–43.
5. Polyak, B. The conjugate gradient method in extreme problems. *USSR Comput. Math. Math. Phys.* **1969**, *9*, 94–112. [[CrossRef](#)]
6. Liu, Y.; Sorey, C. Efficient generalized conjugate gradient algorithms. Part 1 Theory. *J. Optim. Theory Appl.* **1991**, *69*, 177–182. [[CrossRef](#)]
7. Hager, W.W.; Zhang, H. A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.* **2006**, *2*, 35–58.
8. Gibert, J.; Nocedal, J. Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.* **1992**, *2*, 21–42. [[CrossRef](#)]
9. Yuan, Y. Analysis on the conjugate gradient method. *Optim. Method Softw.* **1993**, *2*, 19–29. [[CrossRef](#)]
10. Powell, M. Nonconvex minimization calculations and the conjugate gradient method. *Numer. Anal. Lect. Notes Math.* **1984**, *1066*, 122–141. [[CrossRef](#)]
11. Hager, W.W.; Zhang, H. A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J. Optim.* **2005**, *16*, 170–192. [[CrossRef](#)]
12. Cheng, W. A Two-Term PRP-Based Descent Method. *Numer. Func. Anal. Opt.* **2007**, *28*, 1217–1230. [[CrossRef](#)]
13. Yu, G.; Guan, L.; Li, G. Global convergence of modified Polak-Ribière-Polyak conjugate gradient methods with sufficient descent property. *J. Ind. Manag. Optim.* **2008**, *4*, 565–579. [[CrossRef](#)]
14. Yuan, G. Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. *Optim. Lett.* **2009**, *3*, 11–21. [[CrossRef](#)]
15. Livieris, I.; Pintelas, P. A new class of spectral conjugate gradient methods based on a modified secant equation for unconstrained optimization. *J. Comput. Appl. Math.* **2013**, *239*, 396–405. [[CrossRef](#)]
16. Wei, Z.; Yao, S.; Liu, L. The convergence properties of some new conjugate gradient methods. *Appl. Math. Comput.* **2006**, *183*, 1341–1350. [[CrossRef](#)]
17. Zhang, L. An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation. *Appl. Math. Comput.* **2009**, *215*, 2269–2274. [[CrossRef](#)]
18. Zhang, L.; Zhou, W.; Li, D. A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence. *IMA J. Numer. Anal.* **2006**, *26*, 629–640. [[CrossRef](#)]
19. Dong, X.; Liu, H.; He, Y.; Babaie-Kafaki, S.; Ghanbari, R. A New Three-Term Conjugate Gradient Method with Descent Direction for Unconstrained Optimization. *Math. Model. Anal.* **2016**, *21*, 399–411. [[CrossRef](#)]
20. Zoutendijk, G. Nonlinear Programming, Computational Methods. In *Ineger and Nonlinear Programming*; Abadie, J., Ed.; North-Holland: Amsterdam, The Netherlands, 1970; pp. 37–86.
21. Bongartz, K.; Conn, A.; Gould, N.; Toint, P. CUTE: Constrained and unconstrained testing environments. *ACM T. Math. Softw.* **1995**, *21*, 123–16. [[CrossRef](#)]
22. Dolan, E.; Morè, J. Benchmarking optimization software with performance profiles. *Math. Program.* **2002**, *91*, 201–213. [[CrossRef](#)]

