Hindawi
Abstract and Applied Analysis
Volume 2017, Article ID 7238134, 6 pages
https://doi.org/10.1155/2017/7238134

Research Article

Modification of Nonlinear Conjugate Gradient Method with Weak Wolfe-Powell Line Search

Ahmad Alhawarat[1] and Zabidin Salleh[2]

[1]*Department of Mathematics, College of Science, Isra University, Amman, Jordan*
[2]*School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia*

Correspondence should be addressed to Zabidin Salleh; zabidin@umt.edu.my

Received 21 November 2016; Revised 5 February 2017; Accepted 12 February 2017; Published 5 March 2017

Academic Editor: Patricia J. Y. Wong

Conjugate gradient (CG) method is used to find the optimum solution for the large scale unconstrained optimization problems. Based on its simple algorithm, low memory requirement, and the speed of obtaining the solution, this method is widely used in many fields, such as engineering, computer science, and medical science. In this paper, we modified CG method to achieve the global convergence with various line searches. In addition, it passes the sufficient descent condition without any line search. The numerical computations under weak Wolfe-Powell line search shows that the efficiency of the new method is superior to other conventional methods.

1. Introduction

The nonlinear CG method is a useful tool to find the minimum value of function for unconstrained optimization problems. Let us consider the following form

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and its gradient is denoted by $g(x) = \nabla f(x)$. The method to find a sequence of points $\{x_k\}$ starting from initial point $x_0 \in \mathbb{R}^n$ is given by the iterative formula:

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

where x_k is the current iteration point and $\alpha_k > 0$ is the step size obtained by some line search. The search direction d_k is defined by

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

where $g_k = g(x_k)$ and β_k is known as the conjugate gradient coefficient.

Strong Wolfe-Powell (SWP) line search is the most popular inexact line search, which is depending on a reduction in function and decreasing the search area to find step length. In addition, it forces the step length to be closed to stationary point or local minimum of function, so it is useful method to find the step size.

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

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq \sigma \left| g_k^T d_k \right|, \tag{5}$$

where $0 < \delta < \sigma < 1$. In fact, SWP line search is modified from weak Wolfe-Powell (WWP), so we find that the step length satisfies (4) and

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k. \tag{6}$$

However, WWP line search may accept the step length far from stationary or local minimum of function. Dai [1] proposed two Armijo type line searches: the first one matches the global convergence for any $\beta_k \geq 0$ using methods (2) and (3). By this line search, the global convergence for FR, nonnegative PRP, and CD methods have been established. To match the global convergence of original PRP method, he designed another line search proposed as follows.

Given a constant $\lambda \in (0,1)$, $\delta > 0$ and $\sigma \in (0,1)$ determine the smallest integer $m \geq 0$, if it defines $\alpha_k = \lambda^m$, then the vectors x_{k+1} and d_{k+1} given by (2) and (3) satisfy (4) with

$$0 \neq g(x_k + \alpha_k d_k)^T d_{k+1} \leq -\sigma \|d_{k+1}\|^2, \quad \sigma \in (0,1), \tag{7}$$

where $\delta \in (0, 1/2)$ and $\sigma \in (\delta, 1)$ are two constants.

The most popular formulas for β_k are as follows: Hestenes-Stiefel (HS) [2], Fletcher-Reeves (FR) [3], Polak-Ribière-Polyak (PRP) [4], Conjugate Descent (CD) [5], Liu-Storey (LS) [6], Dai-Yuan (DY) [7], Wei et al. (WYL) [8], and Hager and Zhang (HZ) [9].

$$\begin{aligned}
\beta_k^{HS} &= \frac{g_k^T(y_{k-1})}{d_{k-1}^T(g_k - g_{k-1})}, \\
\beta_k^{FR} &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \\
\beta_k^{PRP} &= \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \\
\beta_k^{CD} &= -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \\
\beta_k^{LS} &= -\frac{g_k^T(y_{k-1})}{d_{k-1}^T g_{k-1}}, \\
\beta_k^{DY} &= \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})}, \\
\beta_k^{WYL} &= \frac{g_k^T(g_k - (\|g_k\| / \|g_{k-1}\|) g_{k-1})}{\|g_{k-1}\|^2}, \\
\beta_k^{HZ} &= \max\{\beta_k^N, \eta_k\},
\end{aligned} \tag{8}$$

where $\beta_k^N = (1/d_k^T y_k)(y_k - 2d_k(\|y_k\|^2/d_k^T y_k))^T g_k$, $\eta_k = -1/\|d_k\| \min\{\eta, \|g_k\|\}$ with $y_k = g_{k+1} - g_k$ and $\eta > 0$ being a constant.

The global convergence of FR method with exact line search was achieved by Zoutendijk [10], Al-Baali [11] proved that FR method is globally convergent under strong Wolfe condition when $\sigma < 1/2$, and later Liu et al. [12] extended the result to $\sigma \leq 1/2$. Its behavior on numerical computation is unpredictable. In few cases, it is as efficient as PRP method. However, generally, it is very slow. In addition, DY and CD have the same performance as FR method under exact line search with strong global convergence. Global convergence of PRP method for convex objective function under exact line search was proved by Polak and Ribière in 1969 [4]. Later, Powell gave out a counterexample showing that there exists nonconvex function, which PRP method does not converge globally, although the exact line search is used. Powell suggested the importance of achieving the global convergence of PRP method, and it should not be negative. Gilbert and Nocedal [13] proved that nonnegative PRP method is globally convergent with the Wolfe-Powel line search. HS method and LS method have the same performance as PRP with exact line search. Therefore, PRP method is the most efficient method when it is compared to the other conjugate gradient methods. For more, the reader can see the following references [14–19].

In 2006, Wei et al. [8] gave a new positive CG method, and it seems like original PRP method which has been studied in both exact line search and inexact line search, and many modifications have appeared, such as the following [20–23], respectively.

A little modification from β_k^{WYL}, Zhang [21] presented the following CG method:

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - (\|g_k\| / \|g_{k-1}\|)\left|g_k^T g_{k-1}\right|}{\|g_{k-1}\|^2}. \tag{9}$$

In the same manner, β_k^{DPRP} construct the following CG by using the denominator of β_k^{NPRP}:

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - (\|g_k\| / \|g_{k-1}\|)\left|g_k^T g_{k-1}\right|}{w\left|g_k^T d_{k-1}\right| + \|g_{k-1}\|^2}. \tag{10}$$

In addition, $\beta_k^{MLS^*}$ is constructed by using the numerator of β_k^{WYL}:

$$\beta_k^{MLS^*} = \frac{g_k^T(g_k - (\|g_k\| / \|g_{k-1}\|) g_{k-1})}{-g_{k-1}^T d_{k-1} + m\left|g_k^T d_{k-1}\right|}, \tag{11}$$

where $m \geq 0$ and $w \geq 1$.

The descent condition plays important rule in CG method given by

$$g_k^T d_k < 0, \quad k \geq 0. \tag{12}$$

If we extend (12) to the following form,

$$g_k^T d_k \leq -c\|g_k\|^2, \quad k \geq 0, \ c > 0, \tag{13}$$

then the search direction satisfies the sufficient descent condition.

In this paper, we will present the new formula and the algorithm in Section 2. Furthermore, we will establish the global convergence of our method with several line searches in Section 3. Numerical results with conclusion will be presented in Sections 4 and 5, respectively.

2. The Modified Formula

In this section, $\beta_k^{HZ^*}$ is presented which is extended to $\beta_k^{MLS^*}$ and β_k^{NPRP} method; that is,

$$\beta_k^{HZ^*} = \frac{\|g_k\|^2 - (\|g_k\| / \|g_{k-1}\|)\left|g_k^T g_{k-1}\right|}{-g_{k-1}^T d_{k-1} + \theta\left|g_k^T d_{k-1}\right|}, \tag{14}$$

where $\|\cdot\|$ means the Euclidean norm, and $\theta > 1$.

Algorithm 1.

Step 1 (initialization). Given x_0, set $k = 0$.

Step 2. Compute β_k based on (14).

Step 3. Compute d_k based on (3). If $\|g_k\| = 0$, then stop.

Step 4. Compute α_k based on some line search; we use in numerical section WWP line search with $\sigma = 0.1$ and $\delta = 0.001$.

Step 5. Update new point based on (2).

Step 6. Convergent test and stopping criteria: if $f(x_k) < f(x_{k+1})$ and $\|g_k\| \leq 10^{-6}$ then stop; otherwise, go to Step 1 with $k = k + 1$.

3. The Global Convergence Analysis for $\beta_k^{HZ^*}$ Method

The following assumption is needed to be used in following theorems.

Assumption 2. (I) $f(x)$ is bounded from below on the level set $\Omega = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$, where x_1 is the starting point.

(II) In some neighborhood N of Ω, f is continuous and differentiable, and its gradient is Lipschitz continuous; that is, for any $x, y \in N$, there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \leq L\|x - y\|$.

Lemma 3. *Let Assumption 2 hold. Consider any method in form (2), (3), and α_k satisfies the WWP line search (4) and (6), in which the search direction is descent. Then, the following condition holds:*

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < \infty. \quad (15)$$

Substituting (13) into (15), it follows that

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (16)$$

3.1. The Sufficient Descent Condition with Convergence Properties for SWP Line Search

Theorem 4. *Let sequences g_k and d_k be generated by methods (2), (3), and (14); then (13) holds, where $c \in (0,1)$.*

Proof. We use proof by induction. From (3), we know that for $k = 0$ it is hold. Suppose that it is true until $k - 1$; that is,

$$g_{k-1}^T d_{k-1} \leq -c\|g_{k-1}\|^2; \quad (17)$$

then

$$\frac{1}{-g_{k-1}^T d_{k-1}} \leq \frac{1}{c\|g_{k-1}\|^2}. \quad (18)$$

Now multiply (3) by g_k^T:

$$\begin{aligned} g_k^T d_k &= g_k^T \left(-g_k + \beta_k d_{k-1}\right) = -\|g_k\|^2 + \beta_k g_k^T d_{k-1} \\ &\leq -\|g_k\|^2 + \left|g_k^T d_{k-1}\right| \frac{\|g_k\|^2}{\theta\left|g_k^T d_{k-1}\right|} \\ &= -\left(1 - \frac{1}{\theta}\right)\|g_k\|^2, \end{aligned} \quad (19)$$

where $\theta > 1$. Take $c = 1 - 1/\theta$ and complete the proof. □

3.2. Global Convergence under WWP Line Search. Gilbert and Nocedal [13] present an important theorem to find the global convergence for a nonnegative part of PRP method; it is summarized by Theorem 5. In addition, [13] presents a nice property called Property ∗, which plays strong roles in studies of CG methods.

Property ∗. Consider a method of form (1) and (2), and suppose $0 < \gamma \leq \|g_k\| \leq \overline{\gamma}$; we say that the method possesses Property ∗ if there exists constant $b > 1$ and $\lambda > 0$, where for all $k \geq 1$, and we get $|\beta_k| \leq b$, and if $\|x_k - x_{k-1}\| \leq \lambda$, then $|\beta_k| \leq 1/2b$.

Theorem 5 (see [13]). *Consider that any CG method of form (2) and (3) achieves the following conditions that hold:*

(I) $\beta_k \geq 0$

(II) *The sufficient descent condition (13)*

(III) *Zoutendijk condition*

(IV) *Property ∗*

(V) *Assumption 2*

Then the iterates are globally convergent.

Lemma 6. *Suppose that Assumption 2 holds with Algorithm 1; then $\beta_k^{HZ^*}$ satisfy Property ∗.*

Proof. Since $\beta_k^{HZ^*} \leq \beta_k^{MLS^*}$ and since $\beta_k^{MLS^*}$ satisfies Property ∗, $\beta_k^{HZ^*}$ also achieves Property ∗; for more we suggest that the reader reads Lemma 3.6 [24]. The proof is completed. □

The following corollary is a result from Theorem 5 and Lemma 3.

Corollary 7. *Let sequences x_k be generated by Algorithm 1. If Assumption 2 holds true, then any line search satisfies Zoutendijk condition; we have* $\liminf_{k\to\infty}\|g_k\| = 0$.

3.3. Global Convergence Properties for Armijo Type Line Search

Theorem 8. *Suppose Assumption 2 is true. Consider the methods of form (2) and (3) with $\beta_k^{HZ^*}$, and α_k is obtained by (4) and (7). Then we have* $\liminf_{k\to\infty}\|g_k\| = 0$.

Proof. By using Lemma 2.8 in [1], we achieve

$$\alpha_k > c, \quad c \in (0,1). \tag{20}$$

Using (2) and (7), then

$$\|d_k\| \le \sigma^{-1} \|g_k\|. \tag{21}$$

From (2), (4), (7), and (20), we have

$$\lim_{k\to\infty} \|d_k\| = 0. \tag{22}$$

From Assumption 2 and (21), we obtain

$$\|g_{k+1}\| \le \left(1 + \frac{L}{\sigma}\right) \|g_k\|. \tag{23}$$

From (3),

$$\|g_{k+1}\| \le \|d_{k+1}\| + \beta_{k+1} \|d_k\|. \tag{24}$$

Using (23), (13), (14), and (24), then

$$\|g_{k+1}\| \le \|d_{k+1}\| + \frac{(1+L/\sigma)^2}{c} \|d_k\|, \tag{25}$$

where $c \in (0,1)$. Take the limit and use (22), and then we have $\liminf_{k\to\infty}\|g_k\| = 0$. The proof is completed. □

4. Numerical Results and Discussions

To analyze the efficiency of the new method, we selected some of the test functions in Table 1 from CUTEr [25], Andrei [26], and Adorio and Diliman [24]. We performed a comparison with other CG methods, including NPRP and DPRP methods using weak Wolfe-Powell line search with $\delta = 0.001$. The tolerance ε is selected to 10^{-6} for all algorithms to investigate the rapidity of the iteration methods towards the optimal. The gradient value is taken as the stopping criteria. Here, the stopping criteria considered $\|g_k\| \le 10^{-6}$. Since the parameters NPRP and DPRP are tested based on weak Wolfe-Powell line search, the modified parameters HZ* are tested based on weak Wolfe line search with values of $\sigma = 0.1$ and $\delta = 0.001$. In addition, the values of $\theta = 2$ and $w = 2$ are for HZ* and DPRP parameters, respectively.

We used Matlab 7.9 subroutine program, with CPU processor Intel (R) Core (TM), i3 CPU, and 2 GB DDR2 RAM under strong Wolfe line search. The performance results are shown in Figures 1 and 2, respectively, using a performance profile introduced by Dolan and Moré [27]. This performance measure was introduced to compare a set of solvers S on a set of problems ρ. Assuming n_s solvers and n_p problems in S and ρ, respectively, the measure $t_{p,s}$ is defined as the computation time (e.g., the number of iterations or the CPU time) required for solver s to solve problem p.

To create a baseline for comparison, the performance of solver s on problem p is scaled by the best performance of any solver in S on the problem using the ratio:

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s}: s \in S\}}. \tag{26}$$

Table 1: The test functions.

Number	Function	Dimension/s
1	EXTENDED WHITE & HOlST	500, 1000, 5000, 10000
2	EXTENDED ROENBROCK	500, 1000, 5000, 10000
3	EXTENDED BEALE	500, 1000, 5000, 10000
4	EXTENDED HIMMELBLAU	500, 1000, 5000, 10000
5	EXTENDED DENSCHNB	500, 1000, 5000, 10000
6	SIX HUMP	2
7	THREE HUMP	2
8	BOOTH	2
9	SHALLOW	500, 1000, 5000, 10000
10	DIXMAANA	1500, 3000, 6000, 9000
11	DIXMAANB	1500, 3000, 6000, 9000
12	NONDIA (Shanno-78)	500, 1000, 5000, 1000
13	DQDRTIC	500, 1000, 5000, 10000
14	RAYDAN 1	500, 1000, 5000, 10000
15	EXTENDED TRIDIAGONAL 1	500, 1000, 5000, 1000
16	GENERALIZED QUARTIC GQ1	500, 1000, 5000, 10000
17	DIAGONAL4	500, 1000, 5000, 10000
18	EXTENDED POWELL	4
19	PERTURBED QUADRATIC	500, 1000, 5000
20	EXTENDED CLIFF	10, 20, 30, 40
21	A QUADRATIC FUNCTION QF2	500, 1000, 5000, 10000
22	DIAGONAL 2	500, 1000, 5000, 10000
23	SUM SQUARES	500, 1000, 5000, 10000
24	ZETTL	2
25	DIXMAANC	1500, 3000, 6000, 9000
26	NONDIA	500, 1000, 5000, 10000

Let the parameter $r_M \ge r_{p,s}$ for all p, s be selected, and further assume that $r_{p,s} = r_M$ if and only if the solver s does not solve problem p. As we would like to obtain an overall assessment of the performance of a solver, we defined the measure:

$$P_s(t) = \frac{1}{n_p} \text{size}\left\{p \in \rho: r_{p,s} \le t\right\}. \tag{27}$$

Thus, $P_s(t)$ is the probability for solver $s \in S$ that the performance ratio $r_{p,s}$ is within a factor $t \in \mathbb{R}$ of the best possible ratio. If we define the function p_s as the cumulative distribution function for the performance ratio, then the performance measure $p_s : \mathbb{R} \to [0,1]$ for a solver is nondecreasing and piecewise continuous function from the right. The value of $p_s(1)$ is the probability that the solver achieves the best performance of all of the solvers. In general, a solver with high values of $P_s(t)$, which would appear in the upper right corner of the figure, is preferable.

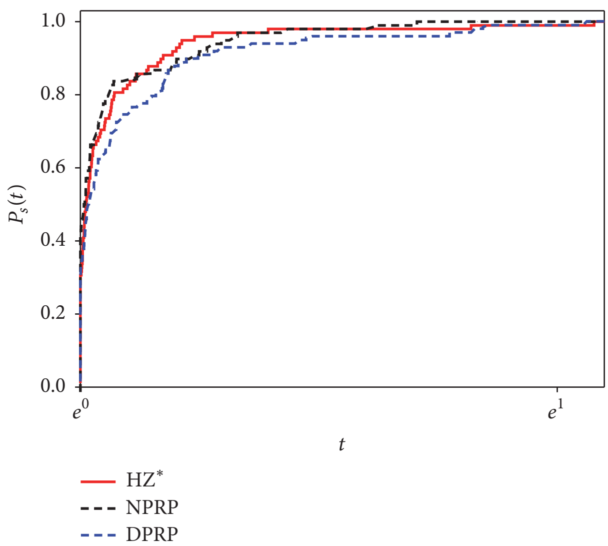

FIGURE 1: Performance profile based on the CPU time with weak Wolfe-Powell line search.

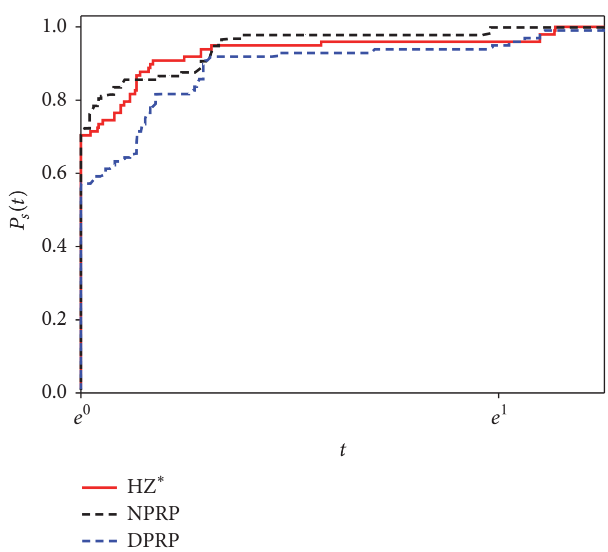

FIGURE 2: Performance profile based on the number of iterations with the weak Wolfe-Powell line search.

It is clear that HZ^* parameter is strong competitive with NPRP parameter and slightly better in some cases for all graphs in Figures 1, 2, 3, and 4 which include the number of iterations, CPU times, gradient evaluations, and function evaluations. On the other hand, it is clear that HZ^* parameter outperforms DPRP parameter in all performance profiles.

5. Conclusion

In this paper, we proposed a new modification of conjugate gradient method extended from NPRP methods. Our numerical results had shown that the new coefficient is comparable compared to other conventional CG methods. This method converges globally with several line searches with descent

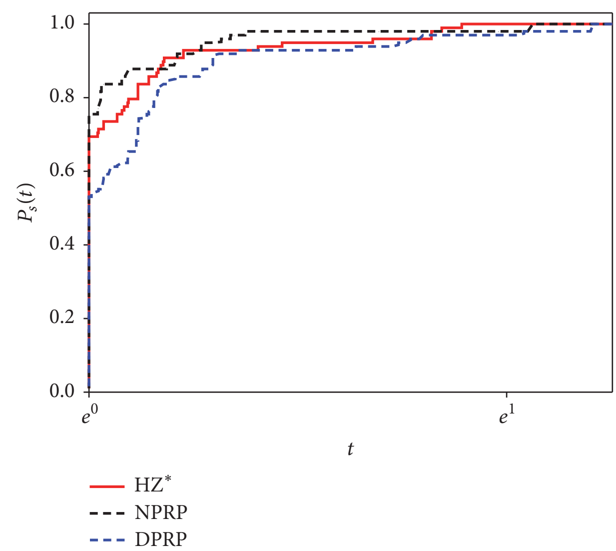

FIGURE 3: Performance profile based on the number of gradient evaluations with weak Wolfe-Powell line search.

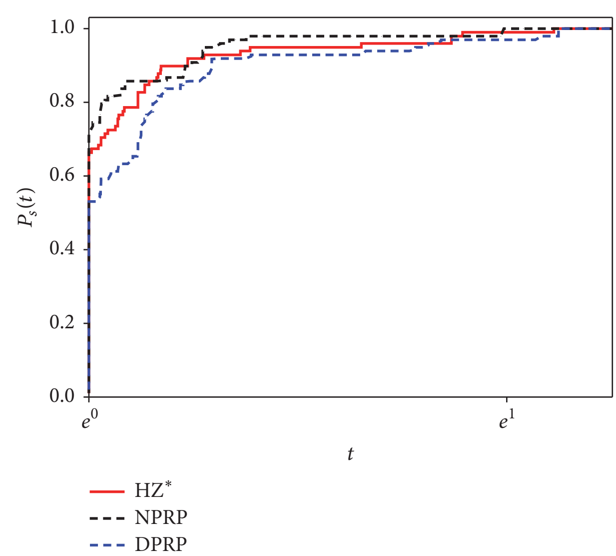

FIGURE 4: Performance profile based on the function evolutions with weak Wolfe-Powell line search.

direction. However, in future, we will focus on speed using hybrid methods. Additionally, we will try to compare several line searches with modern CG method.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Y.-H. Dai, "Conjugate gradient methods with Armijo-type line searches," *Acta Mathematicae Applicatae Sinica*, vol. 18, no. 1, pp. 123–130, 2002.

[2] M. R. Hestenes and E. Stiefel, *Methods of Conjugate Gradients for Solving Linear Systems*, vol. 49, National Bureau of Standards, Washington, DC, USA, 1952.

[3] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

[4] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," *Revue Française d'Automatique, Informatique, Recherche Opérationnelle*, vol. 3, no. 16, pp. 35–43, 1969.

[5] R. Fletcher, *Practical methods of optimization*, Wiley-Interscience John Wiley & Sons, New York, NY, USA, 2nd edition, 2001.

[6] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms. I. Theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.

[7] Y. H. Dai and Y. Yuan, *Nonlinear Conjugate Gradient Methods*, Shanghai Science and Technology Publisher, Shanghai, China, 2000.

[8] Z. Wei, S. Yao, and L. Liu, "The convergence properties of some new conjugate gradient methods," *Applied Mathematics and Computation*, vol. 183, no. 2, pp. 1341–1350, 2006.

[9] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.

[10] G. Zoutendijk, "Nonlinear programming, computational methods," *Integer and Nonlinear Programming*, pp. 37–86, 1970.

[11] M. Al-Baali, "Descent property and global convergence of the Fletcher-Reeves method with inexact line search," *IMA Journal of Numerical Analysis*, vol. 5, no. 1, pp. 121–124, 1985.

[12] G. H. Liu, J. Y. Han, and H. X. Yin, "Global convergence of the Fletcher-Reeves algorithm with inexact linesearch," *Applied Mathematics*, vol. 10, no. 1, pp. 75–82, 1995.

[13] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, 1992.

[14] A. Alhawarat, M. Mamat, M. Rivaie, and I. Mohd, "A new modification of nonlinear conjugate gradient coefficients with global convergence properties, World Academy of Science, Engineering and Technology, International Science Index 85," *International Journal of Mathematical, Computational, Physical and Quantum Engineering*, vol. 8, no. 1, pp. 54–60, 2014.

[15] A. Alhawarat, M. Mamat, M. Rivaie, and Z. Salleh, "An efficient hybrid conjugate gradient method with the strong Wolfe-Powell line search," *Mathematical Problems in Engineering*, vol. 2015, Article ID 103517, 7 pages, 2015.

[16] Z. Salleh and A. Alhawarat, "An efficient modification of the Hestenes-Stiefel nonlinear conjugate gradient method with restart property," *Journal of Inequalities and Applications*, vol. 2016, no. 1, article no. 110, 2016.

[17] A. Alhawarat, Z. Salleh, M. Mamat, and M. Rivaie, "An efficient modified Polak–Ribière–Polyak conjugate gradient method with global convergence properties," *Optimization Methods and Software*, pp. 1–14, 2016.

[18] M. Al-Baali, Y. Narushima, and H. Yabe, "A family of three-term conjugate gradient methods with sufficient descent property for unconstrained optimization," *Computational Optimization and Applications*, vol. 60, no. 1, pp. 89–110, 2015.

[19] W. W. Hager and H. Zhang, "The limited memory conjugate gradient method," *SIAM Journal on Optimization*, vol. 23, no. 4, pp. 2150–2168, 2013.

[20] Y. Shengwei, Z. Wei, and H. Huang, "A note about WYL's conjugate gradient method and its applications," *Applied Mathematics and Computation*, vol. 191, no. 2, pp. 381–388, 2007.

[21] L. Zhang, "An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation," *Applied Mathematics and Computation*, vol. 215, no. 6, pp. 2269–2274, 2009.

[22] Z. Dai and F. Wen, "Another improved Wei-Yao-Liu nonlinear conjugate gradient method with sufficient descent property," *Applied Mathematics and Computation*, vol. 218, no. 14, pp. 7421–7430, 2012.

[23] H. Huang and S. Lin, "A modified Wei–Yao–Liu conjugate gradient method for unconstrained optimization," *Applied Mathematics and Computation*, vol. 231, pp. 179–186, 2014.

[24] E. P. Adorio and U. P. Diliman, "Mvf-multivariate test functions library in C for unconstrained global optimization," 2005.

[25] I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, "CUTE: constrained and unconstrained testing environment," *ACM Transactions on Mathematical Software*, vol. 21, no. 1, pp. 123–160, 1995.

[26] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 147–161, 2008.

[27] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.