Asian-European Journal of Mathematics
(2021) 2150187 (10 pages)

DOI: 10.1142/S1793557121501874

New hyrid conjugate gradient method as a convex combination of HZ and CD methods

Amira Hamdi*, Badreddine Sellami† and Mohammed Belloufi‡

Laboratory Informatics and Mathematics (LiM)
Mohamed Cherif Messaadia University
Souk-Ahras, Algeria
**smalika335@gmail.com*
†basellami@yahoo.fr
‡m.belloufi@univ-soukahras.dz

Communicated by B. K. Dass
Received November 18, 2020
Revised December 27, 2020
Accepted January 1, 2021
Published March 6, 2021

In this paper, a new hybrid conjugate gradient algorithm is proposed for solving unconstrained optimization problems, the conjugate gradient parameter β_k is computed as a convex combination of β_k^{HZ} and β_k^{CD}. Under the wolfe line search, we prove the sufficient descent and the global convergence. Numerical results are reported to show the effectiveness of our procedure.

Keywords: Unconstrained optimization; hybrid conjugate gradient method; global convergence; numerical results.

1. Introduction

Consider the following nonlinear unconstrained optimization problem:

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

where the smooth nonlinear function and its gradient are available. There are many different method for solving the problem (1.1). Here, we are interested in conjugate gradient methods which have low memory requirements and strong local and global convergence properties. To solve the problem (1.1) starting from an initial guess $x_0 \in \mathbb{R}^n$, a nonlinear conjugate gradient method generates a sequence x_k as

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

where $\alpha_k > 0$ is a step size, received from the line search and d_k is the search directions defined by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad d_0 = -g_0, \tag{1.3}$$

where $g_k = \nabla f(x_k)$, and β_k is an important parameter.The different choices for the parameter β_k correspond to different CG methods. Consider$\| . \|$the Euclidean norm and define $y_k = g_{k+1} - g_k$. The strong Wolfe line search conditions [16] frequently used in the conjugate gradient methods are as follows:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{1.4}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \tag{1.5}$$

where d_k is a descent direction and $0 < \delta < \sigma < 1$.

There are many conjugate gradient methods; a great contribution in this sphere is given by Hagar and Zhang [13]. Different conjugate methods correspond to different values of the scalar parameter β_k. Hybrid conjugate gradient methods combine different conjugate gradient methods to improve the behavior of these methods, which have been widely studied by many authors (see [1–3, 9, 15]).

The methods of Fletcher and Reeves (FR) [7], of Dai and Yuan (DY) [5] and the Conjugate Descent (CD) proposed by Fletcher [6] are as follows: $\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}$, $\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{y_k^T d_k}$, $\beta_k^{\mathrm{CD}} = -\frac{\|g_{k+1}\|^2}{g_k^T d_k}$ have strong convergence properties and, in the same time, they may have modest practical performance due to jamming [1, 2]. On the other hand, the following methods of Polak–Ribire [12] and Polyak (PRP) [14], of Hestenes and Stiefel (HS) [8] or of Liu and Storey (LS) [11]: $\beta_k^{\mathrm{PRP}} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$, $\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T d_k}$, $\beta_k^{LS} = -\frac{g_{k+1}^T y_k}{g_k^T d_k}$ may not always be convergent, but they often have better computational performances.One of the conjugate gradient method which is strong in theory is suggested by Hager and Zhang [13]

$$\beta_k^{\mathrm{HZ}} = \frac{g_k^T (y_{k-1} - 2\frac{\|y_{k-1}\|^2}{s_{k-1}^T y_{k-1}} s_{k-1})}{d_{k-1}^T y_{k-1}}. \tag{1.6}$$

To achieve a method which possesses a good performance and strong convergence, we suggest a hybridization of CD and HZ methods as a convex combination to exploit the interesting features of each method.

The structure of the paper is as follows. Section 2 introduces our hybrid conjugate gradient algorithm, HYBRID,and proves that it generates descent directions satisfying the sufficient descent condition under certain circumstances and algorithm. In Sec. 3, its convergence analysis is shown. In Sec. 4, some numerical experiments and performance profiles of Dolan–Moré [18] corresponding to this new hybrid conjugate gradient algorithm are presented. Finally, we make conclusions.

2. A New Hybrid Conjugate Gradient Method

We use the following conjugate gradient parameter where β_k presents here the convex combination of CD and HZ:

$$\beta_k^{\mathrm{New}} = (1 - \theta_k)\beta_k^{\mathrm{HZ}} + \theta_k \beta_k^{\mathrm{CD}} \tag{2.1}$$

and θ_k is a scalar parameter satisfying $0 \le \theta_k \le 1$ which is to be determined. Hence, the direction d_k^{New} is given by

$$d_0^{\text{New}} = -g_0, \quad d_{k+1}^{\text{New}} = -g_{k+1} + \beta_k^{\text{New}} d_k. \tag{2.2}$$

Note that; if $\theta_k = 0$, then $\beta_k^{\text{New}} = \beta_k^{\text{HZ}}$ and if $\theta_k = 1$, $\beta_k^{\text{New}} = \beta_k^{\text{CD}}$, from the other side, if $0 < \theta_k < 1$, then the parameter θ_k is selected in such a way that at every iteration, the conjugacy condition $(d_{k+1} y_k = 0)$. is satisfied independently of the line search.

Clearly,

$$d_{k+1} = -g_{k+1} + (1-\theta_k)\frac{1}{d_k^T y_k}\left(y_k^T g_{k+1} - 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}\right) d_k + \theta_k \frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k \tag{2.3}$$

after some algebra, we have

$$\theta_k = \frac{2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}{\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}. \tag{2.4}$$

It is possible that θ_k, calculated as in (1.3), has the values outside the interval [0, 1]. So, we fix it

$$\theta_k = \begin{cases} 0 & \text{if } \dfrac{2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}{\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}} \le 0, \\ \dfrac{2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}{\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}} & \text{if } 0 < \dfrac{2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}{\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}} < 1, \\ 1 & \text{if } \dfrac{2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}}{\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1}} \ge 1. \end{cases} \tag{2.5}$$

Theorem 2.1. *If the relations* (2.1), (2.2) *hold, then*

$$d_{k+1}^{\text{New}} = (1-\theta_k) d_{k+1}^{\text{HZ}} + \theta_k d_{k+1}^{\text{CD}}. \tag{2.6}$$

Proof. We have $d_{k+1}^{\text{New}} = -g_{k+1} + \beta_k^{\text{New}} d_k$. After adding and subtracting$(\theta_k g_{k+1})$, we obtain

$$d_{k+1}^{\text{New}} = (1-\theta_k)(-g_{k+1} + \beta_k^{\text{HZ}} d_k) + \theta_k(-g_{k+1} + \beta_k^{\text{CD}} d_k) \tag{2.7}$$

from (2.7), implying

$$d_{k+1}^{\text{New}} = (1-\theta_k) d_{k+1}^{\text{HZ}} + \theta_k d_{k+1.}^{\text{CD}}$$ □

The following assumptions are generally utilized in global convergence of CG algorithms.

Assumption 1. The level set $\mathrm{S} = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ is bounded, i.e. there exists a constant $B > 0$, such that

$$||x|| \leq B, \quad \text{for all } x \in \mathrm{S}. \tag{2.8}$$

Assumption 2. In a neighborhood $\mathbf{N}$ of S, the function f is continuously differentiable and its gradient $\nabla f(x)$ is lipschitz continuous, i.e. there exists a constant $0 < L < \infty$ such that

$$||\nabla f(x) - \nabla f(y)|| \leq L||x - y|| \quad \text{for all } x, y \in \mathbf{N}. \tag{2.9}$$

Under these assumptions, there exists a constant $\Gamma \geq 0$, such that

$$||\nabla f(x)|| \leq \Gamma, \tag{2.10}$$

for all $x \in \mathbf{S}$ ([2]).

2.1. *Algorithm and sufficient descent condition*

(HCDHZ Algorithm)

Initialization: Choose an initial point $x_0 \in \mathbb{R}^n$, $\epsilon > 0$. Compute $f(x_0)$ and $g_0 = f(x_0)$. Set $d_0 = -g_0$, the initial guess $\alpha_0 = \frac{1}{||g_0||^2}$ and $k = 0$.

Step 1: If $||g_k|| < \epsilon$ then Stop, else go to step 2.

Step 2: Compute α_k by the strong Wolfe line search (1.4), (1.5).

Step 3: Generate the next iterate by $x_{k+1} = x_k + \alpha_k d_k$.
Compute $g_{k+1} = f(x_{k+1})$ and $y_k = g_{k+1} - g_k$.

Step 4: If $\frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k^T y_k - y_k^T g_{k+1} + 2\frac{||y_k||^2}{d_k^T y_k} d_k^T g_{k+1} = 0$, then $\theta_k = 0$, else compute θ_k as in (2.6).Compute β_k as in (1.2).

Step 5: Compute $d = -g_{k+1} + \beta_k^{\mathrm{New}} d_k$. If the restart criterion of Powell condition

$$|g_{k+1}^T g_k| \geq 0.2||g_{k+1}||^2 \tag{2.11}$$

is satisfied, then $d_{k+1} = -g_{k+1}$, else define $d_{k+1} = d$.

Step 6: Compute the initial guess $\alpha_k = \alpha_{k-1} \frac{\|d_{k-1}\|}{\|d_k\|}$. Put $k = k + 1$ and go to step 1.

First, from Theorem 2.1, we have if $\theta_k = 0$ then $d_{k+1}^{\mathrm{New}} = d_{k+1}^{\mathrm{HZ}}$ sufficient descent condition holds for the hybrid method, if it holds for HZ method. Hager and Zhang prove in ([13]) that d_{k+1}^{HZ} satisfies the sufficient descent condition for all k.

Theorem 2.2. *If* $d_k^T y_k \neq 0$ *and*

$$d_0 = -g_0, \quad d_{k+1} = -g_{k+1} + \tau d_k, \tag{2.12}$$

for any $\tau \in [\beta_k^{\mathrm{HZ}}, \max\{\beta_k^{\mathrm{HZ}}, 0\}]$, *then*

$$g_{k+1}^T d_{k+1} \leq -\frac{7}{8}\|g_{k+1}\|^2, \tag{2.13}$$

then the search direction d_k *satisfies the sufficient descent condition.*

Proof. Suppose $\tau = \beta_k^{\mathrm{HZ}}$. Multiplying (2.4) by g_{k+1}^T, we have

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \beta_k^{\mathrm{HZ}} g_{k+1}^T d_k \\ &= \frac{y_k^T g_{k+1}(d_k^T y_k)(g_{k+1}^T d_k) - \|g_{k+1}\|^2 (d_k^T y_k)^2 - 2||y_k||^2 (g_{k+1}^T d_k)^2}{(d_k^T y_k)^2} \end{aligned}$$

then by the inequality $(u^T v \leq \frac{1}{2}(||u||^2 + ||v||^2))$, we have

$$\begin{aligned} &g_{k+1}^T d_{k+1} \\ &\quad \leq \frac{\frac{1}{2}(\frac{1}{4}(d_k^T y_k)^2 \|g_{k+1}\|^2 + 4(g_{k+1}^T d_k)^2 \|y_k\|^2) - \|g_{k+1}\|^2 (d_k^T y_k)^2 - 2\|y_k\|^2 (g_{k+1}^T d_k)^2}{(d_k^T y_k)^2} \\ &\quad = -\frac{7}{8}\|g_{k+1}\|^2. \end{aligned}$$

If $\tau \neq \beta_k^{\mathrm{HZ}}$, then $\tau \in [\beta_k^{\mathrm{HZ}}, 0]$. After multiplying (2.6) by g_{k+1}^T, then we obtain

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \tau g_{k+1}^T d_k.$$

If $g_{k+1}^T d_k \geq 0$, then (2.7) follows immediately, since $\tau \leq 0$.

If $g_{k+1}^T d_k \geq 0$, then

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \tau g_{k+1}^T d_k \leq -\|g_{k+1}\|^2 + \beta_k^{\mathrm{HZ}} g_{k+1}^T d_k. \qquad \square$$

Second, from Theorem 2.1 if $\theta_k = 1$, we have $d_{k+1}^{\mathrm{New}} = d_{k+1}^{\mathrm{CD}}$.

So, if the sufficient descent holds for CD method, it holds for HCDHZ method. The following theorem prove the sufficient descent for CD method.

Theorem 2.3. *Assume that Assumptions* 1 *and* 2 *hold, and the conditions* (1.4), (1.5) *hold, then the search direction* d_k *satisfies the following sufficient descent condition*:

$$g_{k+1}^T d_{k+1} \leq -K\|g_{k+1}\|^2, \quad \forall\, k \geq 0, \tag{2.14}$$

where $K = 1 - \sigma, \forall\ \sigma < 0.5$.

Proof.

$$d_k^{\mathrm{New}} = d_k^{\mathrm{CD}} = -g_{k+1} + \frac{\|g_{k+1}\|^2}{-g_k^T d_k} d_k. \tag{2.15}$$

Multiplying by $gk+1^T$, we get

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

$$\leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{-g_k^T d_k}(-\sigma g_k^T d_k).$$

Now, we have

$$g_{k+1}^T d_{k+1}^{\text{CD}} \leq -(1-\sigma)\|g_{k+1}\|^2,$$

where $K = 1-\sigma > 0$. □

Finally, $(0 < \theta_k < 1)$: There exist two real numbers μ_1, μ_2 such that $0 < \mu_1 \leq \theta_k \leq \mu_2 < 1$, Then
$g_{k+1}^T d_{k+1}^{\text{New}} \leq \mu_1 g_{k+1}^T d_{k+1}^{\text{CD}} + (1-\mu_2) g_{k+1}^T d_{k+1}^{\text{HZ}}$. We evidently can achieve that there exists a number $K > 0$, such that

$$g_{k+1}^T d_{k+1}^{\text{New}} \leq -K\|g_{k+1}\|^2. \tag{2.16}$$

3. Convergence Analysis

Lemma 3.1 ([17]). *Assume that Assumptions* 1 *and* 2 *hold. Consider any method of the form* (1.1), *where* d_k *is a descent direction and* α_k *satisfies the standard Wolfe conditions* (1.4) *and* (1.5). *Then we have that*

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

We see that, if α_k satisfies the strong Wolfe line searches, then α_k must satisfy the Wolfe line searches. Thus, lemma 3.1 also holds under the strong Wolfe line searches.

Lemma 3.2 ([10]). *Suppose that Assumptions* 1 *and* 2 *hold. If* d_k *is a descent direction and* α_k *satisfies*

$$g_{k+1} d_k \geq \sigma g_k d_k, \quad \sigma < 1 \tag{3.1}$$

then

$$\alpha_k \geq \frac{(1-\sigma)}{L} \frac{|d_k^T g_k|}{\|d_k\|^2}. \tag{3.2}$$

Proof. We use (3.1), the Cauchy Schwarz inequality, it holds that

$$-(1-\sigma) g_k^T d_k \leq d_k^T (g_{k+1} - g_k) \leq L\alpha_k \|d_k\|^2.$$

Since d_k is a descent direction and $\sigma < 1$, then it is not difficult to get the assertion. Obviously, from (1.5) and (2.13), the stepsize α_k obtained in HCDHZ algorithm satisfies (3.2). According to the Assumptions 1 and (2.13), it is easy to obtain that $g_k^T d_k \neq 0$ for $\forall\ k \geq 0$. Thus, $\alpha_k = 0$ does not satisfy (1.5). This indicates that α_k obtained in HCDHZ algorithm is not equal to zero, i.e. there exists a constant $\lambda > 0$ such that

$$\alpha_k \geq \lambda, \quad \text{for all } k \geq 0. \tag{3.3}$$

The following theorem establishes the global convergence of HCDHZ algorithm with the strong Wolfe line searches. □

Theorem 3.1. *Consider the iterative method of the form* (1.1), (1.6), *Let all conditions of Theorem* 2.2 *hold, then either* $g_k = 0$ *for some* k, *or*

$$\lim_{k\to+\infty} \inf \|g_k\| = 0. \tag{3.4}$$

Proof. Suppose that (3.4) does not hold. then there exists $r > 0$ such that

$$\|g_k\| \geq r. \tag{3.5}$$

From Theorem 2.3, we have

$$g_k^T d_k \leq -K\|g_k\|^2, \quad \text{for all } K.$$

The fort wolfe condition gives

$$d_k^T y_k \geq -(1-\sigma) g_k^T d_k \geq K(1-\sigma)\|g_k\|^2. \tag{3.6}$$

It follows, from Assumptions 1 and 2, that

$$\|y_k\| = \|g_{k+1} - g_k\| \leq \|x_{k+1} - x_k\| \leq L.D, \tag{3.7}$$

where D is the diameter of $\{\|s_k\| : k = 0, 1, \ldots\}$.

We have

$$\begin{aligned} |\beta_k^{\text{New}}| &\leq |\beta_k^{\text{HZ}}| + |\beta_k^{\text{CD}}| \\ &\leq \frac{\|y_k\|\|g_{k+1}\|}{K(1-\sigma)\|g_k\|^2} + 2\frac{\|y_k\|^2\|d_k\|\|g_{k+1}\|}{K^2(1-\sigma)^2\|g_k\|^4} + \frac{\|g_{k+1}\|^2}{-K(1-\sigma)\|g_k\|^2} \\ &\leq \frac{\gamma LD}{K(1-\sigma)r^2}\left(1 + 2\frac{LD^2}{\lambda K(1-\sigma)r^2}\right) + \frac{\gamma^2}{-K(1-\sigma)r^2} = M \end{aligned}$$

and

$$\begin{aligned} \|d_{k+1}\| &\leq \|g_{k+1}\| + |\beta_k^{\text{New}}| \frac{\|s_k\|}{\alpha_k} \\ &\leq \gamma + M\frac{D}{\lambda} = E \end{aligned}$$

which implies that

$$\sum_{k\geq 0} \frac{1}{\|d_k\|^2} = +\infty. \tag{3.8}$$

On the other hand, we have that

$$\sum_{k\geq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = +\infty,$$

which contradicts with (3.8), Hence, (3.5) does not hold, and the claim (3.4) is proved. □

4. Numerical Example

In this section, we present the computational performance of a Mathematical implementation of HCDHZ algorithm on a set of unconstrained optimization test problems. The comparisons with other conjugate gradient algorithms are presented using the performance profiles of Dolan and Moré [18] and the parameters in the strong Wolfe line searches are chosen to be $\delta = 10^{-4}$ and $\sigma = 0.1$. We stop the iterations if the criterion $\|g_k\|_\infty \leq \epsilon = 10^{-6}$ is satisfied. On the one hand, for the ith problem, let f_i^{M1} and f_i^{M2} be the optimal values obtained by M1 method and M2 method, respectively. We say that, for the ith particular problem, the performance of M1 method is better than the performance of M2 method if

$$|f_i^{M1} - f_i^{M2}| < 10^{-3} \tag{4.1}$$

and number of iterations, or CPU time of M1 method is less than that of M2 method, respectively. We compare the performance of HCDHZ to the HZ and CD conjugate gradient algorithms. Figures 1 and 2 represent the performance profiles of HCDHZ versus HZ and CD based on the CPU time and number of iterations,

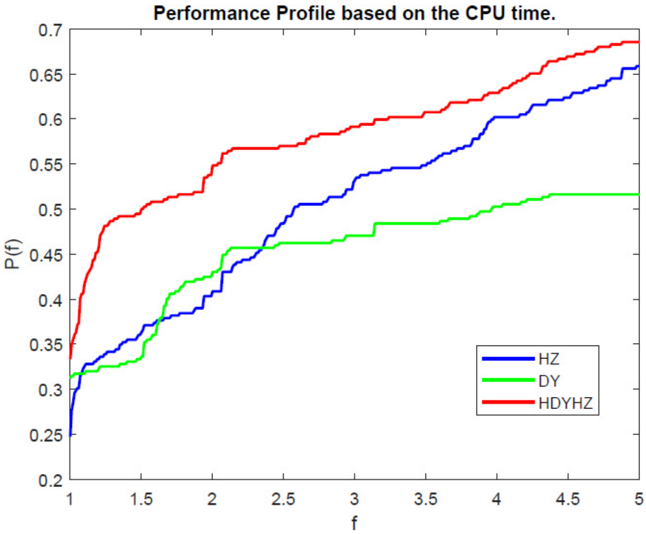

Fig. 1.

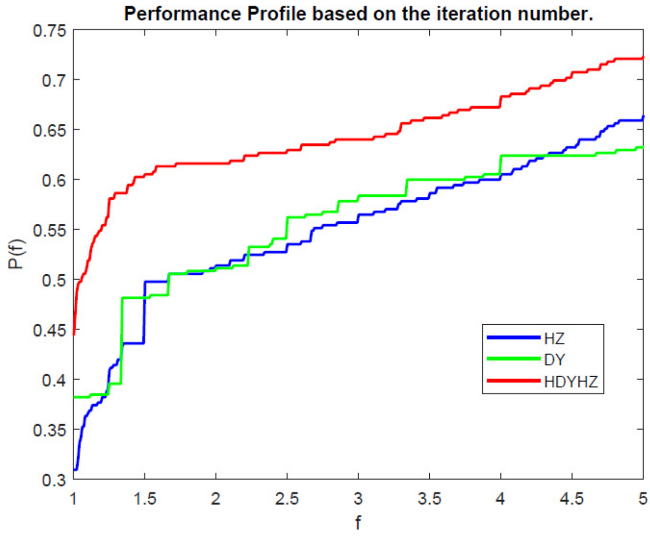

Fig. 2.

respectively, and they show that our procedure HCDHZ is better in terms of effectiveness and robustness.

5. Conclusion

In this paper, a new hybrid conjugate gradient algorithm is proposed and analyzed.The parameter θ_k is computed as a convex combination of HZ and CD, in such that the conjugacy condition is satisfied. The sufficient descent and global convergence was proved and the numerical performance supports the effectiveness and robustness of our procedure.

References

1. N. Andrei, A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of hestenes-stiefel and dai-yuan, *Stud. Inform. Control* **17**(1) (2008) 57.
2. N. Andrei, *New Hybrid Conjugate Gradient Algorithms for Unconstrained Optimization*, Encyclopedia of Optimization (Springer, Boston, MA, 2009), pp. 2560–2571.
3. S. S. Djordjevic, New hybrid conjugate gradient method as a convex combination of FR and PRP methods, *Filomat* **30**(11) (2016) 3083–3100.
4. A. Y. Al-Bayati and N. H. Al-Assady, Conjugate Gradient Method, Technical Report, No. (1/86), School of Computer Studies, Leeds University, UK, 1986.
5. Y.-H. Dai and Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim.* **10**(1) (1999) 177–182.
6. R. Fletcher, *Practical Methods Of Optimization*, Unconstrained Optimization, Vol. 1 (John Wiley& Son, New York, 1980).
7. R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, *Comput. J.* **7**(2) (1964) 149–154.

8. M. R. Hestenes and E. Stiefel, *Methods of Conjugate Gradients for Solving Linear System*, Vol. 49 (NBS, Washington, DC, 1952).
9. J. Hou, Z. C. Wen and Q. Chang, An unconstrained optimization reformulation for the Nash game, *J. Interdiscip. Math.* **21**(5) (2018) 1303–1307.
10. J. K. Liu and S. J. Li, New hybrid conjugate gradient method for unconstrained optimization, *Appl. Math. Comput.* **245** (2014) 36–43.
11. Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, part 1: Theory, *J. Optim. Theory Appl.* **69**(1) (1991) 129–137.
12. E. Polak and G. Ribiere, Note sur la convergence de mthodes de directions conjugues, *ESAIM: Mathematical Modelling and Numerical Analysis-Modlisation Mathmatique et Analyse Numriqueh* **3**(R1) (1969) 35–43.
13. W. W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM J. Optim.* **16**(1) (2005) 170–192.
14. B. T. Polyak, The conjugate gradient method in extremal problems, *USSR Comput. Math. Math. Phys.* **9**(4) (1969) 94–112.
15. P. Mtagulwa and P. Kaelo, A convergent modified HS-DY hybrid conjugate gradient method for unconstrained optimization problems, *J. Inform. Optim. Sci.* **40**(1) (2019) 97–113.
16. P. Wolfe, Convergence conditions for ascent methods, *SIAM Rev.* **11**(2) (1969) 226–235.
17. G. Zoutendijk, Nonlinear programming, computational methods, in *Integer and Nonlinear Programming*, (North-Holland, Amsterdam, 1970), pp. 37–86.
18. E. D. Dolan and J. J. Mor, Benchmarking optimization software with performance profiles, *Math. Program Ser. A* **91** (2002) 201–213.