Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 3, pp. 699–715
DOI : 10.47974/JIOS-1354

Efficient one-parameter family of conjugate gradient methods

Samia Khelladi *
Djamel Benterki †
Laboratory of Fundamental and Numerical Mathematics LMFN
Faculty of Sciences
University of Setif-1
Ferhat Abbas 19000
Algeria

Abstract

We present an efficient one-parameter family of conjugate gradient methods for unconstrained optimization problems. These methods are defined using a combination of the Polak-Ribiere-Polyak method and the Rivaie-Mustafa-Ismail-Leong method. We prove the global convergence based on the Wolfe line search for nonlinear objective functions. Finally, we give some numerical experiments, proving the efficiency of the proposed approach.

Subject Classification: *(2010) 65K05, 90C26, 90C30.*

Keywords: *Unconstrained optimization, Global convergence, Conjugate gradient methods, Line search.*

1. Introduction

Consider the following unconstrained optimization problem

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

where $f: \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function.

Conjugate gradient methods are significant ones for solving unconstrained optimization problems (1) especially for large scale problems, which have the following form:

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

* *E-mail:* samia.boukaroura@univ-setif.dz (Corresponding Author)
† *E-mail:* djbenterki@univ-setif.dz

© Taru Publications

where x_k is the current iterate point, $\alpha_k > 0$ is the step size found by one of the line search methods, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_1, & k = 1, \\ -g_k + \beta_k d_{k-1}, & k \geq 2, \end{cases} \tag{3}$$

where $g_k = \nabla f(x_k)$ is the gradient of f at x_k, and β_k is an appropriately defined real scalar, known as the conjugate gradient parameter.

After the introduction of the nonlinear conjugate gradient method of Fletcher and Reeves (FR) in 1964 [8], many parameters β_k have been proposed which give a different conjugate gradient direction d_k. The known parameters β_k are those of Polak-Ribiere-Polyak (PRP) [14, 15], Conjugate Descent (CD) [7], Liu-Storey (LS) [12], Dai-Yuan (DY) [5], Dai and Liao [2] and Rivaie-Mustafa-Ismail-Leong (RMIL) [17]. Later, many combinations and new families of conjugate gradient methods were proposed, like the different hybridization methods proposed by Yang et al. [20], Dalladji et al. [6] and Mtagulwa and Kaelo [13] and the new families of conjugate methods proposed by Sellami and Chaib [18, 19], also in [21], we find some efficient methods of the type Dai-Liao, named DHSDL and DLSDL.

Several researchers have used quasi-Newton methods, known for their efficiency to compute the search direction of the conjugate gradient method [10].

The formulas of the β_k mentioned above are:

$$\beta_k^{FR} = \frac{\| g_k \|^2}{\| g_{k-1} \|^2}, \beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\| g_{k-1} \|^2}, \beta_k^{CD} = -\frac{\| g_k \|^2}{g_{k-1}^T d_{k-1}},$$

$$\beta_k^{LS} = -\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}}, \beta_k^{DY} = -\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}}, \beta_k^{RMIL} = \frac{g_k^T y_{k-1}}{\| d_{k-1} \|^2},$$

$$\beta_k^{DHSDL} = \frac{\| g_k \|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} | g_k^T g_{k-1} |}{\mu | g_k^T d_{k-1} | + d_{k-1}^T g_{k-1}} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \mu > 1, t > 0,$$

$$\beta_k^{DLSDL} = \frac{\| g_k \|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} | g_k^T g_{k-1} |}{\mu | g_k^T d_{k-1} | - d_{k-1}^T g_{k-1}} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \mu > 1, t > 0,$$

$$\beta_k^{SC} = \frac{(1-\lambda)\|g_k\|^2 + \lambda(-g_k^T d_{k-1})}{-g_{k-1}^T d_{k-1}}, 0 < \lambda < 1,$$

$$\beta_k^{SC*} = \frac{\lambda_k \left\| g_k \right\|^2 + (1-\lambda_k) g_k^T y_{k-1})}{(1-\lambda_k-\mu_k)\left\| g_{k-1} \right\|^2 + (\lambda_k+\mu_k)(y_{k-1}^T d_{k-1})},$$

with $0 < \lambda_k < 1$ and $0 < \mu_k < 1-\lambda_k$,

where $y_{k-1} = g_k - g_{k-1}$, $s_{k-1} = x_k - x_{k-1}$ and $\|.\|$ denotes the Euclidean vector norm.

We see that the formula of β_k^{PRP} and β_k^{RMIL} are only different in the denominators.

In this paper, we propose a combination of the two denominators to obtain the following new family of conjugate gradient methods, where our proposed β_k is:

$$\beta_k^{KB} = \frac{g_k^T y_{k-1}}{\lambda \| g_{k-1} \|^2 + (1-\lambda) \| d_{k-1} \|^2}, \text{ where } 0 \le \lambda \le 1. \tag{4}$$

We see that the above formula of β_k^{KB} is a special form of

$$\beta_k^{KB} = \frac{\phi_k}{\phi'_{k-1}}, \tag{5}$$

where ϕ_k satisfies that

$$\phi_k = g_k^T y_{k-1}, \tag{6}$$

and

$$\phi'_{k-1} = \lambda \| g_{k-1} \|^2 + (1-\lambda) \| d_{k-1} \|^2 \text{ where } 0 \le \lambda \le 1. \tag{7}$$

The step size α_k is determined using the following Wolfe line search conditions

$$f(x_k + \alpha_k d_k) \le f(x_k) + \rho \alpha_k g_k^T d_k, \tag{8}$$

$$g_{k+1}^T d_k \ge \sigma g_k^T d_k, 0 < \rho < \sigma < 1. \tag{9}$$

The rest of this paper is organized as follows. We give some preliminaries in the section 2. In section 3, we give convergence theorems for general method (2) – (3) with β_k^{KB} defined by (4). Section 4 includes the main convergence properties of the new method with the Wolfe line search. In section 5, we present some compared numerical tests to show the performance of the different considered methods. Finally, some conclusions end the paper.

2. Preliminaries

Throughout this paper, we assume that $g_k \neq 0$ for all k, else a stationary point has been found. Firstly, we give the following assumption concerning the objective function.

Assumption 2.1:

(i) The function f is bounded on the level set $\mathcal{L} = \{x \in \mathbb{R}^n / f(x) \leq f(x_1)\}$.

(ii) In some neighborhood $\mathcal{N}$ of $\mathcal{L}$, f is differentiable, and its gradient g is Lipschitz continuous. Namely, there exists a constant $L > 0$ such that

$$\| g(x) - g(y) \| \leq L \| x - y \|, \text{ for all } x, y \in \mathcal{N}. \tag{10}$$

Assumption 2.2: The level set $\mathcal{L} = \{x \in \mathbb{R}^n / f(x) \leq f(x_1)\}$ is bounded.

It is well known that if f satisfies Assumption 2.1 and 2.2, then there exists a positive constant γ, such that

$$\| g(x) \| \leq \gamma, \forall x \in \mathcal{L}. \tag{11}$$

To prove the global convergence of the nonlinear conjugate gradient methods, we use the following lemma, called the Zoutendijk condition [22].

Lemma 2.1: *Suppose Assumption 2.1 holds. Let the sequence* $\{x_k\}$ *be generated by* (2) *and* d_k *satisfies* $g_k^T d_k < 0$. *If* α_k *is determined by the Wolfe line search conditions* (8) – (9), *then we have*

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\| d_k \|^2} < +\infty \tag{12}$$

In the later sections, we need the following lemmas, the first of which is derived from [16], whereas the second is self-evident and will be used for many times.

Lemma 2.2: [16] *Let* $\{a_i\}$ *and* $\{b_i\}$ *be a positive number sequences. If*

$$\sum_{k \geq 1} a_k = \infty, \tag{13}$$

and there exist two constants c_1 *and* c_2 *such that for all* $k \geq 1$

$$b_k \leq c_1 + c_2 \sum_{i=1}^{k} a_i, \tag{14}$$

then we have

$$\sum_{k\geq 1}\frac{a_k}{b_k}=\infty. \tag{15}$$

Lemma 2.3: *Consider the following 1-dimensional function,*

$$\rho(t)=\frac{a+bt}{c+dt}, t\in\mathbb{R}, \tag{16}$$

where a,b,c *and* $d\neq 0$ *are given real numbers. If*

$$bc-ad>0, \tag{17}$$

then $\rho(t)$ *is strictly monotonically increasing for* $t\neq\frac{-c}{d}$, *otherwise, if*

$$bc-ad<0, \tag{18}$$

then $\rho(t)$ *is strictly monotonically decreasing for* $t\neq\frac{-c}{d}$.

3. Algorithm and convergence properties

Now, we can present a new descent conjugate gradient method, using our conjugate parameter

$$\beta_k^{KB}=\frac{g_k^T y_{k-1}}{\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2}, \text{ where } 0\leq\lambda\leq 1.$$

3.1 *Algorithm*

- Step 0: Given a starting point x_1 and a parameter $0<\varepsilon<1$.
- Step 1: Set $k=1$ and compute $d_1=-g_1$.
- Step 2: If $\|g_k\|\leq\varepsilon$, Stop; else go to Step3.
- Step 3: Calculate the step size $\alpha_k\in]0,1]$ via Wolfe line search.
- Step 4: Take $x_{k+1}=x_k+\alpha_k d_k$.
- Step 5: Take $y_k=g_{k+1}-g_k$ and go to Step6.
- Step 6: Calculate

$$\beta_{k+1}^{KB}=\frac{g_{k+1}^T y_k}{\lambda\|g_k\|^2+(1-\lambda)\|d_k\|^2}, \text{ where } 0\leq\lambda\leq 1.$$

- Step7: Calculate the search direction $d_{k+1} = -g_{k+1} + \beta_{k+1}^{KB} d_k$.
- Step8: Let $k := k+1$ and go back to Step2.

To prove the global convergence of Algorithm 3.1, we need the following lemma.

For simplicity, we define

$$t_k = \frac{\| d_k \|^2}{\phi_k^2}, \tag{19}$$

and

$$r_k = -\frac{g_k^T d_k}{\phi_k}. \tag{20}$$

Lemma 3.1: *For the method* (2)–(3) *with* β_k^{KB} *defined by* (5), *the equality:*

$$t_k = -2\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=1}^{k} \frac{\| g_i \|^2}{\phi_i^2} = 2\sum_{i=1}^{k} \frac{r_i}{\phi_i} - \sum_{i=1}^{k} \frac{\| g_i \|^2}{\phi_i^2}, \tag{21}$$

holds for all $k \geq 1$.

Proof: Since $d_1 = -g_1$, (21) holds for $k = 1$. For $i \geq 2$, it follows from (3) that

$$d_i + g_i = \beta_i^{KB} d_{i-1}.$$

Squaring both sides of the above equation, we get that

$$\| d_i \|^2 = -\| g_i \|^2 - 2 g_i^T d_i + (\beta_i^{KB})^2 \| d_{i-1} \|^2 . \tag{22}$$

Dividing (22) by ϕ_i^2 and applying (5) and (19),

$$t_i = \frac{\| d_{i-1} \|^2}{(\phi'_{i-1})^2} - 2\frac{g_i^T d_i}{\phi_i^2} - \frac{\| g_i \|^2}{\phi_i^2}. \tag{23}$$

Using (6)–(7) and since $d_1 = -g_1$, we get that

$$\frac{\| d_1 \|^2}{\phi_1'^2} = \frac{\| g_1 \|^2}{\| g_1 \|^4} = \frac{\| g_1 \|^2}{\phi_1^2}.$$

Summing the expression (23) over *i*, we obtain

$$t_k = t_1 - 2\sum_{i=2}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=2}^{k} \frac{\| g_i \|^2}{\phi_i^2}.$$

Since $d_1 = -g_1$ and $t_1 = \frac{\|g_1\|^2}{\phi_1^2}$, the above relation is equivalent to (21). So (21) holds for all $k \geq 1$.

Theorem 3.1: *Suppose that* x_1 *is a starting point for which Assumption 2.1 holds. Consider the method* (2)–(3) *and* (5), *if for all* k, d_k *satisfy* $g_k^T d_k < 0$ *and* α_k *is determined by the Wolfe line search conditions* (8)–(9), *and if*

$$\sum_{k \geq 1} r_k^2 = \infty, \tag{24}$$

then we have

$$\liminf_{k \to \infty} \| g_k \| = 0. \tag{25}$$

Proof: We have

$$d_i = -g_i + \beta_i^{KB} d_{i-1},$$

then

$$g_i^T d_i + \| g_i \|^2 = \beta_i^{KB} g_i^T d_{i-1}.$$

Squaring both sides of the above equation, we get that

$$-2 g_i^T d_i - \| g_i \|^2 \leq \frac{(g_i^T d_i)^2}{\| g_i \|^2} \tag{26}$$

Dividing (26) by ϕ_i^2 and applying (21)

$$t_k \leq \sum_{i=1}^{k} \frac{(g_i^T d_i)^2}{\| g_i \|^2 \, \phi_i^2}, \tag{27}$$

or equivalently,

$$t_k \leq \sum_{i=1}^{k} \frac{r_i^2}{\| g_i \|^2}. \tag{28}$$

We proceed by contradiction. Assume that (25) is not true, namely

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

Then, there exist a positive constante γ such that

$$\| g_k \| \geq \gamma, \text{ for all } k.$$

In this case, it follows by (28) that

$$t_k \leq \frac{1}{\gamma^2} \sum_{i=1}^{k} r_i^2. \tag{29}$$

The above relation (24) and Lemma 2.2 yield

$$\sum_{i \geq 1} \frac{r_i^2}{t_i} = \infty. \tag{30}$$

By the definition of t_k and r_k, we know that (30) contradicts (12). Therefore (25) is true.

Theorem 3.2: *Suppose that x_1 is a starting point for which Assumption 2.1 holds. Consider the method (2)–(3) and (5), if for all k, d_k satisfy $g_k^T d_k < 0$ and α_k is determined by the Wolfe line search conditions (8)–(9), and if*

$$\sum_{k \geq 1} \frac{\| g_k \|^2}{\phi_k^2} = \infty, \tag{31}$$

then we have

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

Proof: Noting that

$$t_k \geq 0 \text{ for all } k,$$

we can get from (21) that

$$-2 \sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} \geq \sum_{i=1}^{k} \frac{\| g_i \|^2}{\phi_i^2}, \tag{32}$$

which yields that

$$4 \sum_{i=1}^{k} \frac{(g_i^T d_i)^2}{\| g_i \|^2 \phi_i^2} \geq -4 \sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=1}^{k} \frac{\| g_i \|^2}{\phi_i^2} \geq \sum_{i=1}^{k} \frac{\| g_i \|^2}{\phi_i^2}. \tag{33}$$

Thus if (31) holds, we also have that

$$\sum_{k\geq 1}\frac{(g_k^T d_k)^2}{\|g_k\|^2\,\phi_k^2}=+\infty.$$

Because (27) still holds, it follows from (33) and Lemma 2.2 that

$$\sum_{k\geq 1}\frac{(g_k^T d_k)^2}{\|g_k\|^2\|d_k\|^2}=+\infty.$$

The above relation and Lemma 2.1 clearly give (25). This completes the proof.

Thus, we have proved two convergence theorems for the general method (2)–(3) with β_k^{KB} defined by (5).

4. Global convergence of new conjugate gradient method

In this section, we establish some global convergence of the new method (4) under certain line searches conditions.

We consider the method (2)–(3) with β_k^{KB} defined by (5) where

$$\phi_k=g_k^T y_{k-1} \tag{34}$$

$$\phi'_{k-1}=\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2,\ \text{where } 0\leq\lambda\leq 1. \tag{35}$$

(34)–(35) and (3) show that

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2+\beta_k^{KB}g_k^T d_{k-1}\\ &= -\|g_k\|^2+\frac{g_k^T y_{k-1}}{\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2}g_k^T d_{k-1}. \end{aligned} \tag{36}$$

The above relation implies that

$$g_k^T d_k=\frac{y_{k-1}^T d_{k-1}-(\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2)}{\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2}\|g_k\|^2. \tag{37}$$

Thus by (36), we deduce an equivalent form of β_k^{KB},

$$\beta_k^{KB}=\frac{y_{k-1}^T d_{k-1}}{\lambda\|g_{k-1}\|^2+(1-\lambda)\|d_{k-1}\|^2}\frac{\|g_k\|^2}{g_k^T d_{k-1}}. \tag{38}$$

Substituting (37) into (34), we obtain that

$$\phi_k=\frac{y_{k-1}^T d_{k-1}}{g_k^T d_{k-1}}\|g_k\|^2. \tag{39}$$

By this relation, we can show an important property of ϕ_k under Wolfe line search and hence obtain the global convergence of the new method (38) under some assumptions.

Theorem 4.1: *Let x_1 be a starting point for which Assumption 2.1 and Assumption 2.2 hold. Consider the method (2)−(3) and (5), if for all k, d_k is a descent direction and α_k is determined by the Wolfe line search conditions (8)−(9), then we have*

$$\frac{\phi_k}{\| g_k \|^2} \leq \frac{1}{1-\sigma}. \tag{40}$$

Further, the method converges in the sense that

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

Proof: Since (9), we have that

$$g_k^T d_{k-1} \geq \sigma g_{k-1}^T d_{k-1}$$

By direct calculation, we show that

$$y_{k-1}^T d_{k-1} \geq (1-\sigma)(-g_k^T d_{k-1}) \tag{41}$$

Dividing (39) by $\| g_k \|^2$ and applying (41), we obtain the inequality (40).

It follows from (11) and (41) that

$$\sum_{k\geq 1} \frac{\| g_k \|^2}{\phi_k^2} = +\infty.$$

Thus, (25) follows from Theorem 3.2.

The following theorem shows the descent property condition of each search direction and global convergence of the new method.

Theorem 4.2: *Let x_1 be a starting point for which Assumption 2.1 and Assumption 2.2 hold. Consider the method (2)−(3) and (5), and α_k determined by the Wolfe line search conditions (8)−(9), then we have for all k, d_k is a descent direction, i.e., d_k satisfies $g_k^T d_k < 0$.*

Further, the method converges in the sense that

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

Proof: If $k=1$ we have $d_1=-g_1$, then $g_1^T d_1=-\|g_1\|^2<0$.

Suppose that $k\geq 1$.

First we have from (10)

$$\|y_k\|\leq L\|s_k\| \text{ and } \|s_k\|=\|x_{k+1}-x_k\|=\alpha_k\|d_k\|.$$

Using the Cauchy-Schwartz inequality, we get

$$d_k^T y_k\leq \|d_k\|\|y_k\|\leq L\alpha_k\|d_k\|^2.$$

In the other hand, we have

$$\begin{aligned}
g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2+\beta_{k+1}^{KB} g_{k+1}^T d_k\\
&= -\|g_{k+1}\|^2+\frac{g_{k+1}^T y_k}{\lambda\|g_k\|^2+(1-\lambda)\|d_k\|^2} g_{k+1}^T d_k\\
&\leq -\|g_{k+1}\|^2+\frac{\|g_{k+1}\|^2\|y_k\|\|d_k\|}{\lambda\|g_k\|^2+(1-\lambda)\|d_k\|^2}\\
&\leq -\|g_{k+1}\|^2+\frac{L\alpha_k\|d_k\|^2}{\lambda\|g_k\|^2+(1-\lambda)\|d_k\|^2}\|g_{k+1}\|^2\\
&= -(1+\frac{L\alpha_k\|d_k\|^2}{\lambda\|g_k\|^2+(1-\lambda)\|d_k\|^2})\|g_{k+1}\|^2\\
&< 0
\end{aligned}$$

Then d_k is a descent direction.

Finaly, from Theorem 4.1, we get that

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

5. Numerical tests

Some numerical tests are reported to illustrate the behavior of the new proposed conjugate gradient method. The step size α_k is determined via the Wolfe line search. The test functions used in the examples are taken from [1] collection. For each test function, we have taken three values of λ and different values for the number of variables n.

We use the Matlab Language with a precision $\varepsilon=10^{-6}$.

We designate by:

- k : The number of iterations required to obtain the solution.

- *Time* : The execution time in second.
- n : The size of the problem (the number of variables).

We have take different values of $\lambda \in]0,1[$ $(\lambda = 0.25,\ \lambda = 0.5,\ \lambda = 0.75)$.

Example 1: We consider the Raydan 2 function

$$f(x) = \sum_{i=1}^{n} (\exp(x_i) - x_i).$$

The initial point is $x_1 = (1,1,\ldots,1)^T$.
The obtained results are summarized in the following tables:
For $\lambda = 0.25$, we have

n	k	*Time*	$\| g_k \|$
10	20	0.061145	7.3697e – 07
100	22	0.053863	5.8263e – 07
200	22	0.081342	8.2396e – 07
500	24	0.238925	9.9746e-07

For $\lambda = 0.5$, we have

n	k	*Time*	$\| g_k \|$
10	20	0.021559	7.3697e – 07
100	22	0.045234	5.8263e – 07
200	22	0.064795	8.2396e – 07
500	24	0.221002	9.9746e – 07

For $\lambda = 0.75$, we have

n	k	*Time*	$\| g_k \|$
10	20	0.013427	7.3697e – 07
100	22	0.042645	5.8263e – 07
200	22	0.067794	8.2396e – 07
500	24	0.258386	9.9746e – 07

Example 2: We consider the Diagonal 5 function

$$f(x) = \sum_{i=1}^{n} \ln(\exp(x_i) + \exp(-x_i)).$$

The initial point is $x_1 = (1.1,1.1,...,1.1)^T$.

The obtained results are summarized in the following tables:
For $\lambda = 0.25$, we have

n	k	*Time*	$\| g_k \|$
10	96	0.109602	9.5359e-07
100	105	0.400481	9.3323e-07
200	107	0.774303	9.4973e-07
500	114	1.892069	9.9115e-07

For $\lambda = 0.5$, we have

n	k	*Time*	$\| g_k \|$
10	96	0.115135	9.5359e-07
100	105	0.400212	9.3323e-07
200	107	0.786534	9.4973e-07
500	114	1.853797	9.9115e-07

For $\lambda = 0.75$, we have

n	k	*Time*	$\| g_k \|$
10	96	0.146463	9.5359e-07
100	105	0.481576	9.3323e-07
200	107	0.684385	9.4973e-07
500	114	1.831285	9.9115e-07

Example 3: We consider the Diagonal 7 function

$$f(x) = \sum_{i=1}^{n} (\exp(x_i) - 2x_i - x_i^2).$$

The initial point is $x_1 = (1,1,...,1)^T$.
The obtained results are summarized in the following tables:
For $\lambda = 0.25$, we have

n	k	*Time*	$\| g_k \|$
10	35	0.056827	7.0295e-07
100	37	0.092795	2.4251e-07
200	39	0.154000	7.7542e-07
500	36	0.339118	7.9941e-07

For $\lambda = 0.5$, we have

n	k	*Time*	$\| g_k \|$
10	35	0.036467	7.0295e-07
100	37	0.077891	2.4251e-07
200	39	0.144259	7.7542e-07
500	36	0.481349	7.9941e-07

For $\lambda = 0.75$, we have

n	k	*Time*	$\| g_k \|$
10	35	0.023698	7.0295e-07
100	37	0.086676	2.4251e-07
200	39	0.149239	7.7542e-07
500	36	0.434274	7.9941e-07

Example 4: We consider the Diagonal 8 function

$$f(x) = \sum_{i=1}^{n} (x_i \exp(x_i) - 2x_i - x_i^2).$$

The initial point is $x_1 = (1,1,...,1)^T$.
The obtained results are summarized in the following tables:
For $\lambda = 0.25$, we have

n	k	*Time*	$\| g_k \|$
10	37	0.083246	9.3544e-07
100	40	0.127539	9.8513e-07
200	39	0.225010	4.4489e-07
500	39	0.485449	9.8800e-07

For $\lambda = 0.5$, we have

n	k	*Time*	$\| g_k \|$
10	37	0.035746	9.3544e-07
100	40	0.127327	9.8513e-07
200	39	0.225039	4.4489e-07
500	39	0.453066	9.8800e-07

For $\lambda = 0.75$, we have

n	k	*Time*	$\| g_k \|$
10	37	0.042650	9.3544e-07
100	40	0.124961	9.8513e-07
200	39	0.216284	4.4489e-07
500	39	0.564935	9.8800e-07

Commentaries: The numerical tests show clearly the efficiency of the new proposed method for the different values of λ, in terms of number of iterations and computation time.

Indeed, the algorithm gives us the results after a reduced number of iterations for any λ and in a minimal time whatever the dimension of the problem.

6. Conclusion

In this paper, we have proposed a new family of conjugate gradient methods for solving unconstrained optimization problems, where the parameter β_k^{KB} is a combination of β_k^{PRP} and β_k^{RMIL}.

Firstly, we have shown that the proposed methods via the Wolfe line search are globally convergent for nonlinear functions.

Secondly, the numerical test confirm the effectiveness of the new methods in terms of number of iterations and computation time.

Acknowledgements

The authors would like to thank the anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this paper.

References

[1] Andrei, N.: An unconstrained optimization test functions collection. *Adv. Model. Optim.* 10, 147–161 (2008).

[2] Dai, Y.H., Liao, L.Z.: New conjugacy conditions and related nonlinear conjugate gradient methods. *Appl. Math. Optim.* 43(1), 87–101 (2001).

[3] Dai, Y.H., Yuan, Y.: A class of globally convergent conjugate gradient methods. *Sci. China Ser. A-Math.* 46, 251–261 (2003).

[4] Dai, Y.H., Yuan, Y.: An efficient hybrid conjugate gradient method for unconstrained optimization. *Ann. Oper. Res.* 103, 33–47 (2001).

[5] Dai, Y.H., Yuan, Y.: A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.* 10(1), 177–182 (1999).

[6] Delladji, S., Belloufi, M., Sellami, B.: New hybrid conjugate gradient method as a convex combination of FR and BA methods. *Journal of Information and Optimization Sciences,* 42(3), 591-602 (2021).

[7] Fletcher, R.: Practical Methods of Optimization. Unconstrained Optimization, vol. 1. Wiley, New York (1987).

[8] Fletcher, R., Reeves, C.M.: Function minimization by conjugate gradients. *Comput. J.* 7(2), 149–154 (1964).

[9] Gilbert, J.C., Nocedal, J.: Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.* 2(1), 21–42 (1992).

[10] Hassan, B.A., Kahya, M.A.: A new class of quasi-Newton updating formulas for unconstrained optimization. *Journal of Interdisciplinary Mathematics,* 24(8), 2355-2366 (2021).

[11] Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.* 49(6), 409–436 (1952).

[12] Liu, Y., Storey, C.: Efficient generalized conjugate gradient algorithms, part 1: theory. *J. Optim. Theory Appl.* 69(1), 129–137 (1991).

[13] Mtagulwa, P., Kaelo, P.: A convergent modified HS-DY hybrid conjugate gradient method for unconstrained optimization problems. *Journal of Information and Optimization Sciences,* 40(1), 97-113 (2019).

[14] Polak, E., Ribiere, G.: Note sur la convergence des méthodes de directions conjuguées. Rev. *Française Imformat Recherche Opertionelle* 16, 35–43 (1969).

[15] Polyak, B.T.: The conjugate gradient method in extreme problems. *U.S.S.R. Comput. Math. Phys.* 9, 94–112 (1969).

[16] Pu, D., Yu, W.: On the convergence property of the DFP algorithm. *Ann Oper Res* 24, 175–184 (1990).

[17] Rivaie, M., Mustafa, M., Leong W. J., Ismail, M.: A new class of nonlinear conjugate gradient coefficients with global convergence properties. *Applied Mathematics and Computation*, 218(22), 11323–11332 (2012).

[18] Sellami, B., Chaib, Y.: A new family of globally convergent conjugate gradient methods. *Ann. Oper. Res. Springer,* 241, 497–513 (2016).

[19] Sellami, B., Chaib, Y.: New conjugate gradient method for unconstrained optimization. *RAIRO Operations Research*, 50, 1013–1026 (2016).

[20] Yang, X., Luo, Z., Dai, X.: A global convergence of LS-CD hybrid conjugate gradient method. *Advances in Numerical Analysis* (2013).

[21] Zheng, Y., Zheng, B.: Two new Dai-Liao-type conjugate gradient methods for unconstrained optimization problems. *J. Optim. Theory Appl.* 175, 502–509 (2017).

[22] Zoutendijk, G.: Nonlinear programming, computational methods. In: Abadie, *J. (ed.) Integer and Nonlinear Programming*, 37–86 (1970).

Received October, 2021