International Journal of Multidisciplinary Research and Growth Evaluation.

Convex Combination of Parameter CG Algorithms for Dai and Liao Methods in Unconstrained Optimization

Khalil K Abbo [1*], Aynur J Namik [2]
[1] Professor, Department of Mathematics, College of Computers Sciences and Mathematics, Mosul University, Iraq
[2] Department of Mathematics, College Of Computers Sciences and Mathematics, Mosul University, Iraq

* Corresponding Author: **Khalil K Abbo**

Article Info

ISSN (online): 2582-7138
Volume: 06
Issue: 04
July - August 2025
Received: 02-05-2025
Accepted: 05-06-2025
Published: 10-07-2025
Page No: 1009-1013

Abstract
This paper studies the two – term conjugate gradient method for unconstrained optimization problem. Our idea is based on work scaled as convex combination of the Dai and Liao CG method. The conjugacy condition and global convergence for the proposed algorithms are satisfied under some assumptions. For our numerical results we have applied this algorithm this algorithm on the set of well-known test nonlinear functions in the unconstrained optimization.

Keywords: Unconstrained Optimization, Conjugate Gradient Methods, Global Convergence

1. Introduction

We consider the minimization problem of a smooth nonlinear function $f: R^n \to R$,

$$\min_{x \in R^n} \; f(x), \qquad (1)$$

In the case where the number of variables *n* is large, and where analytic expressions for the function *f* and the gradient g are available

The iterative formula of a CG method is given by

$$x_{k+1} = x_k + s_k, \; s_k = \alpha_k d_k, \; k = 1, 2, \ldots, \qquad (2)$$

in which α_k is a step-length to be computed by a line search procedure and d_k is the search direction defined by

$$d_1 = -g_1, \; d_{k+1} = -g_{k+1} + \beta_k d_k, \; k = 1, 2, \ldots, \qquad (3)$$

where $g_k = \nabla f(x_k)$ and β_k is a parameter called the conjugacy condition. The step-length α_k is usually chosen to satisfy certain line search conditions [6].

For general nonlinear functions, different choices of β_k lead to different conjugate gradient methods. Well-known formulas for β_k are called the Fletcher-Reeves (FR) [2], Hestenes -Stiefel (HS) [7], and Polak-Ribiere (PR) [8]. are given by:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2},$$

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k},$$

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}$$

Where $y_k = g_{k+1} - g_k$

The line search in conjugate gradient algorithms is often based on the standard Wolfe Conditions (SDWC) [12, 13]:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \rho \alpha_k g_k^T d_k \quad (4)$$

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k, \quad (5)$$

Where $d_k$is a descent direction and $0 < \rho \leq \sigma < 1$. However, for some conjugate gradient algorithms, a Stronger version of the Wolfe line 3

$$|g_{k+1}^T d_k| \leq -\sigma g_k^T d_k, \quad (6)$$

Is needed to ensure the convergence and to enhance the stability.

2. The derivation of the new AK2 conjugate gradient (CG) method

Based on an extended conjugacy condition, one of the essential CG methods has been proposed by [9], with the following CG parameter. The aim of this section is to derive a new two-term conjugate gradient method Aynur and Khalil (AK2 say) by using Dai and Liao CG method. consider the search direction given by Dai and Liao

$$d_{k+1}^{DL} = -g_{k+1} + \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k - t \frac{s_k^T g_{k+1}}{s_k^T y_k} s_k, \quad (7)$$

Where t is a nonnegative parameter and $y_k = g_{k+1} - g_k$, It is remarkable that numerical performance of the DL method is very dependent on the parameter t for which there is no any optimal choice [1]. It has been attempts to find an ideal value for t.

Now in this paragraph we will give a new model for the parameter t, Based on the principle of the convex combination to take advantage of existing properties in parameter the vehicle without neglecting and as well as the fact that this new parameter called AK2 will be used to expand the conjugate gradient method CG used to solve convex general functions.

The proposed parameter as a convex combination of (β_k^{DL})

The parameter proposed AK2 convex combination of parameter β_k^{DL} Based on the convex combination parameter, So the linear combination of the search direction $\boldsymbol{d_{k+1}}$ becomes as follows:

$$\therefore d_{k+1}^{AK2} = -g_{k+1} + \left(\lambda \frac{g_{k+1}^T y_k}{s_k^T y_k} - (1-\lambda) \frac{g_{k+1}^T s_k}{s_k^T y_k}\right) s_k \quad (8)$$

By multiplying both sides of the equation (8) in y_k^T and use conjugacy condition $y_k^T d_{k+1} = -g_{k+1}^T s_k$,

$$\therefore y_k^T d_{k+1}^{AK2} = -g_{k+1}^T y_k + \left(\lambda \frac{g_{k+1}^T y_k}{s_k^T y_k} - (1-\lambda) \frac{g_{k+1}^T s_k}{s_k^T y_k}\right) s_k^T y_k = -g_{k+1}^T s_k$$

Now we simplify the equation and delete similar terms to get the following

$$-g_{k+1}^T y_k + \lambda g_{k+1}^T y_k - g_{k+1}^T s_k + \lambda g_{k+1}^T s_k = -g_{k+1}^T s_k$$

$$\lambda (g_{k+1}^T y_k + g_{k+1}^T s_k) = g_{k+1}^T y_k$$

$$\lambda = \frac{g_{k+1}^T y_k}{g_{k+1}^T y_k + g_{k+1}^T s_k}, \quad (9)$$

$$1 - \lambda = \frac{g_{k+1}^T s_k}{g_{k+1}^T y_k + g_{k+1}^T s_k}; \quad (10)$$

Substitute (9) and (10) in the equation (8) we get:

$$d_{k+1}^{AK2} = -g_{k+1} + \left[\frac{g_{k+1}^T y_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T s_k}{s_k^T y_k}\right] s_k \quad (11)$$

It is the new search direction.

Algorithm (AK2)

Step (1): Select a starting point $x_1 \in domf$ and $\varepsilon > 0$, compute $f_1 = f(x_1)$ and $g_1 = \nabla f(x_1)$. Select some positive values for ρ and σ. Set $d_1 = -g_1$ and $k = 1$.
Step (2): Test for convergence. If $\|g_k\|_\infty \le \varepsilon$, then stop $x_k$is optimal ; otherwise go to step (3).
Step (3): Determine the step length α_k, by using the Wolfe line search conditions (4)-(5).
Step (4): Update the variables as : $x_{k+1} = x_k + \alpha_k d_k$. Compute f_{k+1} and g_{k+1}. Compute $y_k = g_{k+1} - g_k$and $s_k = x_{k+1} - x_k$.
Step (5): Compute the search direction as: d_{k+1}^{AK2} in (11).
Step (6): Set $k = k + 1$ and go to step 2.

In the following theorem we will show that our method generates a descent direction for all k.

3. Convergence analysis

Assume the following.

(1) The level set $S = \{x \in R^n : f(x) \le f(x_0)\}$ is bounded, i. e. there exists positive constant $B > 0$ such that, for all $x \in S, \|x\| \le B$.

(2) **(2)** In a neighborhood N of S the function f is continuously differentiable and its gradient is Lipschitz continuous, i. e. there exists a constant $L > 0$such that $\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$, for all $x, y \in N$.

Under these assumptions on f, there exists a constant $\Gamma \ge 0$ such that $\|\nabla f(x)\| \le \Gamma$, for all $x \in S$. Observe that the assumption that the function f is bounded below is weaker than the usual assumption that the level set is bounded. Although the search directions generated by (2. 4) are always descent directions, to ensure convergence of the algorithm we need to constrain the choice of the step length α_k. The following proposition shows that the Wolfe line search always gives a lower bound for the step length α_k.

Proposition 1 [10]

Suppose that $d_k$is a descent direction and that the gradient ∇f satisfies the Lipschitz condition $\|\nabla f(x) - \nabla f(x_k)\| \le L\|x - x_k\|$ for all x on the line segment connecting x_k and x_{k+1}, where L is a positive constant. If the line search satisfies the Wolfe conditions (4) and (5), then

$$\alpha_k \ge \frac{(1-\sigma)|g_k^T d_k|}{L\|d_k\|^2}. \tag{12}$$

Proposition 2 [3]

Suppose that assumptions **(1)** and **(2)** hold. Consider the algorithm (2) and (11), where d_k is a descent direction andα_k is computed by the general Wolfe line search (4) and (5). Then

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{13}$$

Proposition 3 [11, 12]

Suppose that assumptions **(1)** and **(2)** hold, and consider any conjugate gradient algorithm (2), where d_k is a descent direction andα_k is obtained by the strong Wolfe line search (4) and (1. 6). If

$$\sum_{k\ge1} \frac{1}{\|d_k\|^2} = \infty, \tag{14}$$

then

$$\lim_{k\to\infty} \inf \|g_k\|. \tag{15}$$

For uniformly convex functions, we can prove that the norm of the direction d_{k+1} generated by (11) is bounded above. Therefore, by Proposition 3, we can prove the following result.

Theorem (1): Suppose that assumptions **(1)** and **(2)** hold, and consider the algorithm (2) and (11), where d_k is a descent direction and α_k is computed by the strong Wolfe line search (4) and (6). Suppose that f is a uniformly convex function on S, i. e., there exists a constant $\mu > 0$ such that [4]

$$\left(\nabla f(x) - \nabla f(y)\right)^T (x - y) \ge \mu\|x - y\|^2 \text{ for all } x, y \in N; \tag{16}$$

then

$$\lim_{k\to\infty} \|g_k\| = 0. \tag{17}$$

Proof: The prove is by Contradiction.

$$\|d_{k+1}\| \;\; = \left\| -g_{k+1} + \left(\frac{g_{k+1}^T y_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T s_k}{s_k^T y_k} \right) s_k \right\|$$

$$\|d_{k+1}\| \leq \|g_{k+1}\| + \frac{\|g_{k+1}\|^2 \quad \|y_k\|^2 \quad \|s_k\|}{(\|g_{k+1}\| \quad \|y_k\| + \|g_{k+1}\| \quad \|s_k\|) \quad \mu\|s_k\|^2} + \frac{\|g_{k+1}\|^2 \quad \|s_k\|^2 \quad \|s_k\|}{(\|g_{k+1}\| \quad \|y_k\| + \|g_{k+1}\| \quad \|s_k\|) \quad \mu\|s_k\|^2}$$

$$\|d_{k+1}\| \leq \|g_{k+1}\| + \frac{\|g_{k+1}\|^2 \quad \|y_k\|^2 \quad \|s_k\|}{(\|g_{k+1}\| \quad \|y_k\| + \|g_{k+1}\| \quad \|s_k\|) \quad \mu\|s_k\|^2} \frac{\|g_{k+1}\|^2 \quad \|s_k\|^2 \quad \|s_k\|}{(\|g_{k+1}\| \quad \|y_k\| + \|g_{k+1}\| \quad \|s_k\|) \quad \mu\|s_k\|^2}$$

$$\|d_{k+1}\| \leq \|g_{k+1}\| + \frac{\|g_{k+1}\|^2 \quad \|y_k\|^2}{\|g_{k+1}\|(\|y_k\| + \|s_k\|) \quad \mu\|s_k\|} + \frac{\|g_{k+1}\|^2 \quad \|s_k\|}{\|g_{k+1}\|(\|y_k\| + \|s_k\|) \quad \mu}$$

By assumption **2** and Lipschitz continuity, we have $\|y_k\| \leq L\|s_k\|$. we get

$$\|d_{k+1}\| \leq \Gamma + \frac{\Gamma \quad L^2\|s_k\|^2}{(L\|s_k\| + \|s_k\|) \quad \mu\|s_k\|} + \frac{\Gamma \quad \|s_k\|}{(L\|s_k\| + \|s_k\|) \quad \mu}$$

$$\|d_{k+1}\| \leq \Gamma + \frac{\Gamma L^2}{(L+1)\mu} + \frac{\Gamma}{(L+1)\mu}$$

$$\|d_{k+1}\| \leq \Gamma + \frac{\Gamma L^2 + \Gamma}{(L+1)\mu}$$

$$\|d_{k+1}\| \leq \Gamma + \frac{\Gamma(L^2+1)}{\mu(L+1)}$$

$$\|d_{k+1}\| \leq \frac{\Gamma\mu(L+1) + \Gamma(L^2+1)}{\mu(L+1)}$$

It is a bove relation we get

$$\sum_{k\geq 1} \frac{1}{\|d_{k+1}\|^2} \geq \left(\frac{\mu(L+1)}{\Gamma\mu(L+1) + \Gamma(L^2+1)} \right)^2 \quad \sum_{k\geq 1} 1 = \infty$$

And using Proposition (3) then satisfy $\lim_{k\to\infty} \|g_k\| = 0$

4. Numerical results and comparisons

In this section, we report some numerical results on 75 nonlinear unconstrained test problems. For each test problem, the dimension *n*=100, …, 1000. The Fortran expression of its function and gradient can be downloaded from N. Andrei's website: The following CG methods in the form of (2) and (3), only different in the choice of the CG parameter, are test:

1. The DL method [3]:

$$\beta^{DL} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - t\frac{s_k^T g_{k+1}}{d_k^T y_k}, t = 1$$

2. The AK2 method (Algorithm AK2):

$$\beta^{AK2} = \frac{g_{k+1}^T y_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{g_{k+1}^T y_k + g_{k+1}^T s_k} \cdot \frac{g_{k+1}^T s_k}{s_k^T y_k}$$

Here we utilize the source code Fortran 77 on N. Andrei's website. All the parameters, including the parameters $\rho = 0.\ 0001$, σ = 0. 9, are set as default. The implementations are run on PC with 1. 3 GHz CPU
processor and 760 MB RAM memory. We stop the iterations if the inequality $\|g_k\|_\infty \leq 10^{-6}$ is satisfied.
We adopt the performance profiles by Dolan and More´ [5] to compare the performance among the tested methods. For n_s and $n_p$problems, the performance profile $P: \Re \to [0, 1]$ is defined as follows:
Let P and S be the set of problems and the set of solvers, respectively. For each problem $p \in P$ and for each solver $s \in S$, we define $t_{p,s} :=$ (computing time or (number of iterations, etc.) required to solve problem $p$by solver s). The performance ratio is given by $r_{p,s} := t_{p,s} / \min_{s\in S} t_{p,s}$. Then the performance profile is defined by:

$P(\tau) = \frac{1}{n_p} size\ \{p \in P : r_{p,s} \leq \tau\}, \forall\tau \in \Re$ where size $\{p \in P : r_{p,s} \leq \tau\}$ stands
For the number of elements of the set $\{p \in P : r_{p,s} \leq \tau\}$. Note that if the performance profile of a method is over the performance profiles of the other methods, then this method performed better than the other methods.

Figures 1-3 are the performance profiles measured by the number of iterations, the number of function and gradient evaluations, and CPU time respectively. From Figures1-3, we can observe that our proposed method (AK2) numerically outperforms with slight superiority to the other methods, since the figures graphically illustrate that the curves of AK2 are always the top performer for almost all values of τ. The possible reason is that our method suggests optimal value for the parameter t which is an open question.

5. Conclusion

In this paper, a new two –term conjugate gradient algorithm, as a modification of the DL methods which generates conjugate directions. Under suitable assumptions our method have been shown to converge globally. In numerical experiments, we have confirmed the effectiveness of the proposed method by using performance profile.

6. References

1. Andrei N. An unconstrained optimization test function collection. Adv Model Optim. 2008;10:147-61.
2. Andrei N. Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization. Bull Malays Math Sci Soc. 2011;34:319-30.
3. Dai YH, Liao LZ. New conjugacy conditions and related nonlinear conjugate gradient methods. Appl Math Optim. 2001;43:87-101.
4. Dolan ED, Moré JJ. Benchmarking optimization software with performance profiles. Math Program. 2002;91:201-13.
5. Hassan BA, Jabbar HN, Laylani YA, Moghrabi IAR, Alissa AJ. An enhanced Fletcher-Reeves-like conjugate gradient methods for image restoration. Int J Electr Comput Eng. 2023;13:6268-76.
6. Hestenes MR, Stiefel E. Methods of conjugate gradients for solving linear systems. J Res Natl Bur Stand. 1952;49(5):409-36.
7. Polak E, Ribiére G. Note sur la convergence de directions conjuguée. Rev Fr Inform Rech Oper. 1969;16:35-43.
8. Powell MJD. Nonconvex minimization calculations and the conjugate gradient method. In: Numerical Analysis (Dundee, 1983). Lecture Notes in Mathematics, vol. 1066. Berlin: Springer-Verlag; 1984. p. 122-41.
9. Sun W, Yuan Y. Optimization theory and methods: Nonlinear programming. New York: Springer Science, Business Media, LLC; 2006.
10. Wolfe M. Numerical methods for unconstrained optimization: An introduction. New York: Van Nostrand Reinhold; 1978.
11. Laylani Y, Khudhur HM, Nori EM, Abbo KK. A hybridization of the Hestenes-Stiefel and Dai-Yuan conjugate gradient methods. Eur J Pure Appl Math. 2023;16(2):1059-67.
12. Watkins DS. Fundamentals of matrix computations. 2nd ed. New York: John Wiley and Sons; 2002.
13. Zoutendijk G. Nonlinear programming, computational methods. In: Abadie J, editor. Integer and nonlinear programming. Amsterdam: North-Holland; 1970. p. 37-86.