IASJ المجلات الاكاديمية العراقية

Tikrit Journal of Pure Science

ISSN: 1813 – 1662 (Print) --- E-ISSN: 2415 – 1726 (Online)

Journal Homepage: http://tjps.tu.edu.iq/index.php/j

A New Formula for Conjugate Gradient in Unconstrained Optimization

Isam H. Hali[1], Khalil K. Abbo[2], Hassan H. Ebrahim[3]

[1] *Department of Mathematics, College of Science, Kirkuk University, Kirkuk, Iraq*

[2] *Department of Mathematics, College of Basic Education, Telafer University, Telafer, Iraq*

[3] *Department of Mathematics, College of Computer Science and Mathematics, Tikrit University, Tikrit, Iraq*

https://doi.org/10.25130/tjps.v27i1.87

ARTICLE INFO.

Article history:

-Received: 4 / 10 / 2021

-Accepted: 22 / 11 / 2021

-Available online: / / 2022

Keywords: Unconstrained Optimization; (CG) Method; Wolfe Line Search; Global Convergence

Corresponding Author:

Name: Isam H. Hali

E-mail:

isam.h.halil@uokirkuk.edu.iq

dr.khalil-khuder@uotelafer.edu.iq

Hassan1962pl@tu.edu.iq

Tel:

ABSTRACT

The method of Conjugate Gradient (CG) is a key component of optimization methods that aren't bound by local convergence characteristics. In this study, we created KHI3, a novel search direction in the (CG) Algorithm. The novel approach satisfies the regression criterion. The overall convergence of the proposed technique has also been proved utilizing Wolff search line words. A new algorithm for solving the large-scale unconstrained optimization issue is particularly successful.

1. Introduction

We look at the problem of unbounded nonlinear optimization.

$\min(f(x)) : x \in R^n$, ...(1)

where $f: R^n \to R$ is a bounded continuously differentiable function from below. There are numerous approaches to solving the problem (1).

We're interested in (CG) algorithms since they need less memory and have good local and global convergence qualities [1].

Starting from the beginning point $x_0 \in R^n$, solve the issue (1), The (CG) approach generates the sequent $\{x_n\} \subset R^n$such that:

$x_{k+1} = x_k + \alpha_k d_k$; .....(2)

where $\alpha_k > 0$ is a step size founded via the line search, and $d_k$are directions given by [2-4].

$d_0 = -g_0,\ d_{k+1} = -g_{k+1} + \beta_k s_k$. .....(3)

β_k is the (CG) parameter in the last relation,

$s_k = x_{k+1} - x_k,\ g_k = \nabla f(x_k)$.

Assume that $\|.\|$ is the Euclidean norm.

We now use the term

$y_k = g_{k+1} - g_k$, ...(4)

In (2) and (3) define the classical (CG) algorithms, where the parameter β_k is founded in one of the following ways:

$\beta^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^t g_k}$ ......(5)

$\beta^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k}$ .....(6)

$\beta^{CD} = \frac{-g_{k+1}^T g_{k+1}}{g_k^T d_k}$ .....(7)

$\beta^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}$ .....(8)

$\beta^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$ ......(9)

$\beta^{LS} = \frac{g_{k+1}^T y_k}{g_k^T d_k}$ ....(10)

The first CG methods for nonlinear functions [(5) FR,] were introduced by[5]. [(6) DY] technique introduced by [6], and (7) introduced the CD conjugate in [7].With g_{k+1}^T, g_{k+1} in the number of β_k having strong convergence theory. The [(HS) CG] proposed by [8] described in (8), the [(PR) CG] established by [9] defined in (9) and the [(LS)CG] obtained by [10] defined in (10) all use $g_{k+1}^T y_k$ in the number of parameters β_k.

2. The New (KHI3) Algorithm for (CG)

In this search, we suggest a new beta as follows:

Consider $d_{k+1} = -g_{k+1} + \beta_k s_k$ ....(11)

Now a new

$$\beta_{k+1}^{KHI3}=\frac{y_k^Tg_{k+1}}{y_k^Ts_k}-\frac{s_k^Ts_ks_k^Tg_{k+1}}{(g_k^Tg_k+y_k^Ts_k)^2} \quad \text{..... (12)}$$

Then by substitute equation (12) in equation (11) we get :

$$d_{k+1}=-g_{k+1}+\left\{\frac{y_k^Tg_{k+1}}{y_k^Ts_k}-\frac{\|s_k\|s_k^Tg_{k+1}}{(g_k^Tg_k+y_k^Ts_k)^2}\right\}s_k \quad \text{....(13)}$$

2.1 Algorithm

Step 1. The initial point $x_0\in R^n$, and the accurate solution $\in>0$, and we find $d_0=-g_0$, and , $k=0$

Step 2: Checking for iteration continuation. If $\|g_{k+1}\|\le 10^{-6}$ stop. If not, proceed to the next step.

Step 3: Using Wolfe line search conditions, calculate the step size α_k

$f_{k+1}\le f_k+c_1\alpha_1d_k^Tg_k$ .....(14)

$d_k^Tg_{k+1}\ge c_2d_k^Tg_k$ ......(15)

and go to step (4).

Step 4. Calculate $x_{k+1}=x_k+\alpha_kd_k$.

Step 5. we compute the search direction by the equation (13).

Repeat step 2 with the set $k=k+1$.

3. Descent Property

We shall prove the descent property $(d_k^Tg_k<0)$ for the new formula which was introduce in equation (13) and that the sufficient descent of the (CG) algorithm is written as follows$(d_k^Tg_k<-c\|g_k\|^2)$.

3.1 Theorem

Assume that Lipchitz condition with $0<L\le 1$ hold and $y_k^TS_k>0$ for all k ,then the search direction defined in (12) and (13) is descent direction for all k. i.e. $d_k^Tg_k<0 \quad \forall k$.

Proof:-

The Prove is by induction i.e. for $k=1$

$d_1^Tg_1=-\|g_1\|<0$

Suppose that $d_k^Tg_k<0 \quad \forall k$,

And we will prove that the relation is true when $k=k+1$, i.e.

$$d_{k+1}=-g_{k+1}+\left\{\frac{y_k^Tg_{k+1}}{y_k^Ts_k}-\frac{\|s_k\|s_k^Tg_{k+1}}{(\|g_{k+1}\|+y_k^Ts_k)^2}\right\}s_k$$

Multiply both sides by g_{k+1}^T then

$$g_{k+1}^Td_{k+1}=-\|g_{k+1}\|^2+\frac{y_k^Tg_{k+1}s_k^Tg_{k+1}}{y_k^Ts}-\frac{\|s\|^2(s^Tg_{k+1})^2}{(\|g_k\|+y_k^Ts)^2} \quad \text{....(16)}$$

$$g_{k+1}^Td_{k+1}=-\|g_{k+1}\|^2+\frac{y_k^Tg_{k+1}s_k^Tg_{k+1}}{y^Ts}-\frac{\|s\|^2(s^Tg_{k+1})^2}{4(y^Ts)^2} \quad \text{..... (17)}$$

For $\frac{y_k^Tg_{k+1}s_k^Tg_{k+1}}{y^Ts}=\frac{s_k^Ty_ky_k^Tg_{k+1}s_k^Tg_{k+1}}{(y^Ts)^2}$

$\because u^Tv\le\frac{1}{2}\{\|u\|^2+\|v\|^2\}$

Let $u=\frac{1}{\sqrt{2}}y_k^Ts_kg_k$ and $v=\sqrt{2}s^Tg_{k+1}y_k$

Then

$$\frac{y_k^Ts_kg_{k+1}^Ty_ks_k^Tg_{k+1}}{\left(y_k^Ts_k\right)^2}\le\frac{\frac{1}{3}\left\{\frac{1}{2}\left(y_k^Ts_k\right)^2\|g_{k+1}\|^2+2\left(s_k^Tg_{k+1}\right)^2\|y\|^2\right\}}{\left(y^Ts\right)^2} \quad \text{..... (18)}$$

$$=\frac{1}{4}\|g_{k+1}\|^2+\frac{(s^Tg_{k+1})^2\|y\|^2}{(y^Ts)^2}$$

$$\therefore \quad d_{k+1}^Tg_{k+1}\le\frac{-3}{4}\|g_{k+1}\|^2+\frac{\left(s^Tg_{k+1}\right)^2\|y\|^2}{\left(y^Ts\right)^2}-\frac{\|s\|^2\left(s^Tg_{k+1}\right)^2}{4\left(y^Ts\right)^2} \quad \text{..... (19)}$$

By Lipchitz condition with $0\le L\le 1$

$$\therefore \quad d_{k+1}^Tg_{k+1}\le\frac{-3}{4}\|g_{k+1}\|^2+\frac{L\left(s^Tg_{k+1}\right)^2\|s\|^2}{\left(y^Ts\right)^2}-\frac{\left(s^Tg_{k+1}\right)^2\|s\|^2}{4\left(y^Ts\right)^2} \quad \text{..... (20)}$$

$\because y^Ts\le\|s\|\|y\|$

$$\therefore \quad d_{k+1}^Tg_{k+1}\le\frac{-3}{4}\|g_{k+1}\|^2+\left(L-\frac{1}{4}\right)\frac{\|s\|^2\|g_{k+1}\|^2\|s\|^2}{\|y\|^2\|s\|^2} \quad \text{.......(21)}$$

$$=\left\{\frac{-3}{4}+\left(L-\frac{1}{4}\right)\right\}\frac{\|s\|^2}{\|y\|^2}\|g_{k+1}\|$$

$$=\left(\frac{-3}{4}+\frac{(L-1)\|s\|^2}{4\|y_k\|^2}\right)\|g_{k+1}\|^2$$

Hence the search direction are sufficient descent for $0\le L<1$ and descent holds for $L=1$.

4. The Global Convergent

In this section, we'll look at Algorithm (2.1)'s global convergence. To prove the key facts in this study, The following moderate hypotheses are presented first.

4.1 Assumption

i- The level set $S=\{x\in R^n\backslash f(x)\le f(x_0)\}$ is closed and bounded at the start point.

ii- In some N neighborhood of S, $f(x)$ is an constantly applicable and included is a continuous lipchitz, which, there, there is a fixed L> 0 s.t:

$\|g(x)-g(y)\|\le L\|x-y\|, \ \forall x,y \in N$ ....(22)

iii- The goal function is uniformly convex, and there are constants $\lambda>0$ that satisfy the following conditions:

$\left(g(x)-g(y)\right)^T(x-y)\le\lambda\|x-y\|^2,$

for any $x,y\in S$.

Or $y_k^ks_k\ge\lambda\|s_k\|^2$

based on these f assumptions, $\exists$ a constant $l\ge 0$, s.t: $\|g(x)\|\le l$, and by assumption (i) then there exists $\|x\|\le M, \quad \forall x\in S$

$\underline{\eta}\le\|g(x)\|\le\overline{\eta}$, $\forall x\in S$ [11-13]

4.1. 1 Lemma: [14, 15]

Suppose $d_k$is a descent direction and g_k satisfies the Lipchitz condition. If the line search satisfies the Wolfe condition then

$$\alpha_k\ge\frac{1-\sigma}{L}\frac{|d_k^Tg_k|}{\|d_k\|^2}$$

4.1. 2 Lemma: [16, 17]

If assumptions (4.1.i) and (4.1.iii) are true, consider any (CG) method where d_{k+1} is the descent direction and the step size α_k is determined by the strong Wolfe line search. If

$\sum_{k\ge 1}\frac{1}{\|d_{k+1}\|^2}=\infty$ .....(23)

$\Rightarrow lim_{k\to\infty}(inf\|g_k\|)=0$ ....(24)

4.1.3 Theorem:

We assume that the descending property holds for assumptions (4.1.i) and (4.1.iii), the method of CG and the equation (13) and

$x_{k+1}=x_k+\alpha_kd_k$

Where the size of the α_k step was found using a strong search for Wolfe line and a convex target function uniformly, then

$lim_{k\to\infty}(inf\|g_k\|) = 0$

Proof:

The equation (13) is

$$d_{k+1} = -g_{k+1} + \left\{ \frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{\frac{s_k^T s_k\, s_k^T g_{k+1}}{g_k^T g_k + y_k^T s_k}}{g_k^T g_k + y_k^T s_k} \right\} d_k$$

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

And squaring both side of it, We get

$\Rightarrow \|d_{k+1}\|^2 = \|-g_{k+1} + \beta d_k\|^2$ .... (25)

$\Rightarrow \|d_{k+1}\|^2 \le \|g_{k+1}\|^2 + \beta\|d_k\|^2$ ....(26)

$\Rightarrow \|d_{k+1}\|^2 \le \left(1 + \frac{\beta\|d_k\|^2}{\|g_{k+1}\|^2}\right)\|g_{k+1}\|^2$ ....(27)

Let $w = 1 + \frac{\beta\|d_k\|^2}{\|g_{k+1}\|^2}$

$\Rightarrow \|d_{k+1}\|^2 \le w\|g_{k+1}\|^2$ ....(28)

$\Rightarrow \sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} \ge \frac{1}{w}\frac{1}{\|g_{k+1}\|^2}\sum_{k=0}^{\infty} 1 = \infty$ ....(29)

By the lemma (4.2) then

$lim_{k\to\infty}(inf\|g_k\|) = 0$

5. Numerical Experiment

This section displays the numerical results of the new method (KHI3). Different values of d_{k+1} in the proposed algorithm KHI3 were used to test its performance (13). We picked (75) large-scale unregulated optimization problems taken from [18-21] for any of the test problems. We also considered numerical tests with the number of variables n=100, 1000 for every test element. These new versions are compared to the HS, FR algorithms, known (CG) algorithms. With search terms, wolf solid line search terms, all these algorithms are implemented. The stopping conditions in all these situations are $\|g_k\| = 10^{-6}$. All codes are written with F77 default compiler settings in double-precision FORTRAN Language. The test functions typically initially start point normal. Figures 1, 2, and 3 show the numerical summary results, correspondingly. The performance profile is used to view the performance of the KHI3 algorithm of the built A new CG- algorithm [22]. Define P=75 as the entire set of problems with $n_p$test and S = 3 as the set of solutions involved. Let $l_{p,s}$ have a number of evaluations of objective functions needed by solver s for problem p. Determine average results as:

$r_{p,s} = \frac{l_{p,s}}{l_p^*}$ ....(30)

Where $l_p^* = \min\{l_{p,s} : s \in S\}$. $r_{p,s} \ge 1$ for any P,S is self-evident. If solver is unable to resolve the problem, the $r_{p,s}$ ratio is set to a large integer M. The cumulative distribution function is used for the $r_{p,s}$, to create the performance profile for each solver S and after.

$\rho_s(\tau) = \frac{size\{p\in P: r_{p,s} \le \tau\}}{n_p}$ ......(31)

Obviously , $P_s(1)$ denotes the percentage of problems that are Solver S are the most effective.

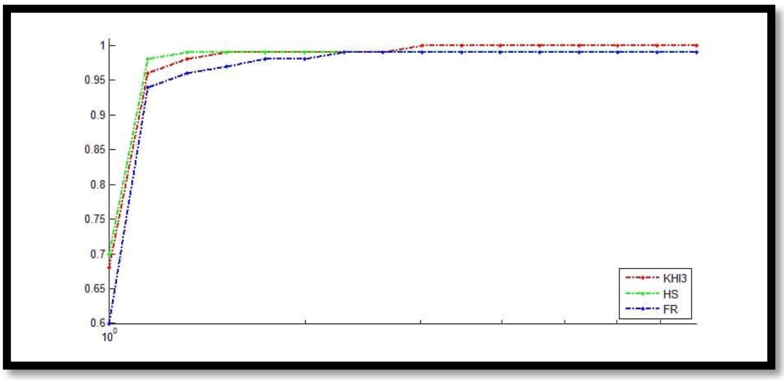

Figure 1: Iteration-Based Performance

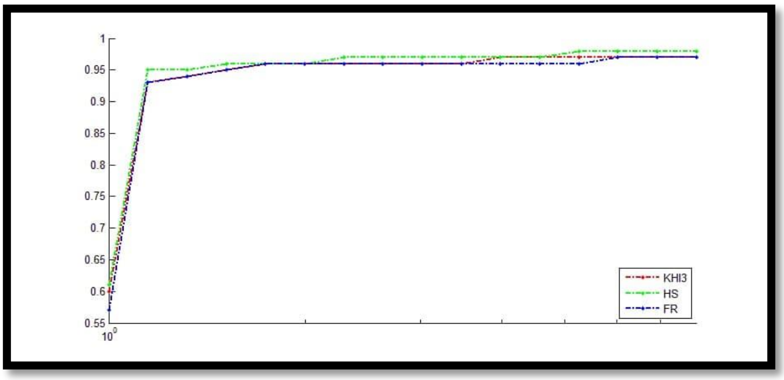

Figure 2: Function-Based Performance

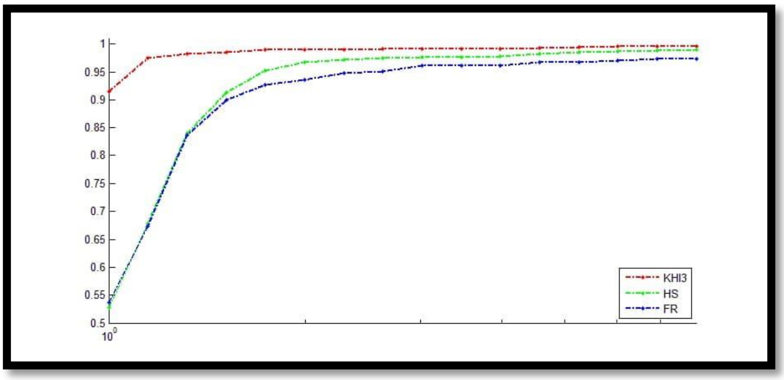

Figure 3: Time-based performance

6. Conclusions

The new algorithm (KHI 3) presented a new form of conjugate gradient (CG) with descent property and global convergence with Wolfe line search. The new algorithm gives a good result compares with the other algorithms.

References

[1] D.F.J.M.o.o.r. Shanno, Conjugate gradient methods with inexact searches, Publisher, City, 1978.

[2] J. Barzilai, J.M. Borwein, Two-point step size gradient methods, Publisher, City, 1988.

[3] J.J.E. Dennis, J.J. Moré, Quasi-Newton Methods, Motivation and Theory, Publisher, City, 1977.

[4] Y.A. Laylani, K.K. Abbo, H.M. Khudhur, Training feed forward neural network with modified Fletcher-Reeves method, Publisher, City, 2018.

[5] R. Fletcher, C.M. Reeves, Function minimization by conjugate gradients, Publisher, City, 1964.

[6] Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, Publisher, City, 1999.

[7] R.J.J. Fletcher, C. Sons, Practical methods of optimization. 1987, Publisher, City, 1987.

[8] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, Publisher, City, 1952.

[9] E. Polak, G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, Publisher, City, 1969.

[10] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, part 1: Theory, Publisher, City, 1991.

[11] G.J.I. Zoutendijk, n. programming, Nonlinear programming, computational methods, Publisher, City, 1970.

[12] H.M. Khudhur, K.K. Abbo, A New Type of Conjugate Gradient Technique for Solving Fuzzy Nonlinear Algebraic Equations, in: Journal of Physics: Conference Series, {IOP} Publishing, 2021.

[13] H.M. Khudhur, K.K. Abbo, New hybrid of Conjugate Gradient Technique for Solving Fuzzy Nonlinear Equations, Publisher, City, 2021.

[14] W.W. Hager, H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, Publisher, City, 2006.

[15] H.M. Khudhur, Numerical and analytical study of some descent algorithms to solve unconstrained Optimization problems, in: University of Mosul college Computer Sciences and Mathematics Department of Mathematics Iraq, 2015, pp. 83-83.

[16] K.K. Abbo, Y.A. Laylani, H.M. Khudhur, A NEW SPECTRAL CONJUGATE GRADIENT ALGORITHM FOR UNCONSTRAINED OPTIMIZATION, Publisher, City, 2018.

[17] K.K. Abbo, Y.A. Laylani, H.M. Khudhur, Proposed new Scaled conjugate gradient algorithm for Unconstrained Optimization, Publisher, City, 2016.

[18] N. Andrei, An Unconstrained Optimization Test Functions Collection, Publisher, City, 2008.

[19] K.K. Abbo, H.M. Khudhur, New A hybrid Hestenes-Stiefel and Dai-Yuan conjugate gradient algorithms for unconstrained optimization, Publisher, City, 2016.

[20] K.K. Abbo, H.M. Khudhur, New A hybrid conjugate gradient Fletcher-Reeves and Polak-Ribiere algorithm for unconstrained optimization, Publisher, City, 2016.

[21] H.N. Jabbar, K.K. Abbo, H.M. Khudhur, Four--Term Conjugate Gradient (CG) Method Based on Pure Conjugacy Condition for Unconstrained Optimization, Publisher, City, 2018.

[22] E.D. Dolan, J.J. Moré, Benchmarking optimization software with performance profiles, Publisher, City, 2002.

صيغة جديدة للتدرج المترافق في الامثلية غير المقيدة

عصام حيدر خليل[1] , خليل خضر عبو[2] , حسن حسين ابراهيم[3]

[1]قسم الرياضيات , كلية العلوم , جامعة كركوك , كركوك , العراق

[2]قسم الرياضيات , كلية التربية الاساسية , جامعة تلعفر , تلعفر , العراق

[3]قسم الرياضيات, كلية علوم الحاسوب والرياضيات , جامعة تكريت , تكريت ,العراق

الملخص

طريقة التدرج المترافق من الطرق الرئيسية في الامثلية .في هذه الدراسة انشانا اتجاه بحث جديد في خوارزمية التدرج المترافق مع تحقيق خاصية الانحدار كما تم اثبات التقارب الشامل للتقنية المقترحة باستخدام خط البحث ولف.

وتعد الخوارزمية الجديدة ناجحة بشكل خاص لحل المسائل الامثلية غير المقيدة وعلى نطاق واسع.