المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

نهج هجين جديد يجمع بين طريقتي ديكسون وتشانغ-هاجر للتحسين غير المقيد

A New Hybrid Approach Combining Dixon and Zhang–Hager Methods for Unconstrained Optimization

علاء الطيف جاسم

For the Student:

Alaa Alteif Jasim[1,a)]

أ.د. زياد محمد عبد الله

Supervised by:

Prof. Dr. Zeyad M. Abdullah[2, b)]

Author Affiliations

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

[1,2]Department of Mathematics, College of Computer Science and Mathematics, Tikrit University, Iraq

انتماءات المؤلفين

[1]قسم الرياضيات، كلية علوم الحاسوب والرياضيات، جامعة تكريت

[١,٢] قسم الرياضيات، كلية علوم الحاسوب والرياضيات، جامعة تكريت

Author Emails

[a)] **Corresponding author**: aa230030pcm@st.tu.edu.iq.

[b)] zeyaemoh1978@tu.edu.iq.

ملخص:

تم تهجين خوارزميتين من خلال دمج معاملين قياسيين ببنية محددة. قمنا بدمج خوارزمية HZ مع DIXON، وتميزت الخوارزمية الناتجة بتدرج كافٍ وتقارب شامل باستخدام افتراضات معينة. اختبرنا كفاءتهما عدديًا على مجموعة من دوال الاختبار غير الخطية وغير المقيدة، وأثبتنا فعاليتهما مقارنةً بأداء الخوارزميتين الأساسيتين.

الكلمات المفتاحية:

(الخوارزميات، التدرج المترافق، الشرط المترافق

Print ISSN 2710-0952 Electronic ISSN2790-1254

Abstract: Two algorithms were hybridized by combining two standard parameters with a convex structure. We combined the HZ algorithm with DIXON, and the resulting algorithm possessed sufficient gradient and global convergence using certain assumptions. We tested their efficiency numerically on a set of nonlinear, unconstrained test functions and demonstrated their effectiveness when compared to the performance of the two basic algorithms.

Keyword: Algorithms, conjugate gradient, conjugate condition

Introduction

A basic premise in our world is the pursuit of the optimal condition, namely the selection of the most advantageous test among available options and decisions in a real-world context. Optimisation is a method for attaining optimal outcomes under defined circumstances. Consequently, optimisation approaches have been employed across several industrial sectors, including railways, automotive, aerospace, and electrical industries, among others. Optimisation is ubiquitous in reality, manifesting in natural events that permeate our daily existence. It manifests in several bodily scenarios. The reflection of light on a mirror represents the shortest path that intersects the mirror. Consequently, optimisation may serve as a representation of physical reality.

It includes finding the best possible solutions to the given problem. Mathematically, this means finding the minimum or maximum value of a function consisting of n variables $\boldsymbol{f(x_1, x_2, x_3, \dots\dots.., x_n)}$, where n can be any integer greater than zero [1].

$$min_{x \in R^n} f(x)$$

(1)

The point $x^* \in R^n$ is said to be the point of stability or (critical point) of a differentiable function f if $\nabla f(x^*) = 0$.

Conjugate gradient (CG) methods are recognised techniques for addressing large-scale optimisation challenges. These methodologies have extensive applications across many domains, including control science, several engineering disciplines, management science, economics, operations research, and intelligent technologies [1].

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

The step length $\alpha_k > 0$ is determined by line search and direction formulation as follows [2]:

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

$$d_{k+1} = \begin{cases} -g_k & if\ k = 1 \\ -g_{k+1} + \beta_k d_k & if k > 1 \end{cases} \quad (3)$$

Where $g_k = \nabla f(x_k)$and β_k is a constant, the step length is consistently determined in accordance with Wolff's criterion. In this context, we use the robust Wolff condition, and the step length complies with [3]:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \ ,0 \le \delta \le \frac{1}{2} \quad (4)$$

$$\left| d_k^T\, g(x_k + \alpha_k d_k) \right| \le -\sigma d_k^T g_k \ \ ,\delta \le \sigma \le 1 \quad (5)$$

Several properties differentiate traditional algorithms, primarily based on the selection of the conjugation parameterβ_k , including [4]:

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \quad [1] \quad (6)$$

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k} \quad [5] \quad (7)$$

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k{}^T d_k} \quad [6] \quad (8)$$

$$\beta_k^{LS} = \frac{-g_{k+1}^T y_k}{g_k^T g_k} \quad [7] \quad (9)$$

$$\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{y_k{}^T d_k} \quad [8] \quad (10)$$

$$\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{d_k^T g_k} \quad [9] \quad (11)$$

$$\beta_k^{ZH} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - 2\frac{\|y_k\|^2 g_{k+1}^T d_k}{(d_k^T y_k)^2} \quad [10] \quad (12)$$

$$\beta_k^{DX} = \frac{-g_{k+1}^T g_{k+1}}{g_k^T d_k} \quad [11] \quad (13)$$

Where $y_k = g_{k+1} - g_k$ and $\|\cdot\|$ denotes the Euclidean norm.

Research objective

A unique hybrid approach integrating the Dixon and Zang-Hager algorithms (D.ZH. A) with the Dai-Liao (D-L) condition is devised to compute the conjugate gradient coefficient.

i. Minimizing the iterations necessary for the gradient-related approach and enhancing the speed of solution attainment.
ii. Analyzing the theoretical characteristics of the novel technique, including

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

convergence speed and stability of the search direction (SD). iii. Assessing the numerical efficacy of the novel method in relation to traditional gradient-based techniques utilizing standard functions.

Review of the Literature

A considerable quantity of hybrid conjugate gradient (CG) algorithms has been introduced in recent years [12]. The fundamental concept of these approaches is to amalgamate several conjugate gradient algorithms using convex weighted combinations to improve convergence behavior and prevent stalling problems. Numerous hybrid conjugate gradient (CG) techniques have been proposed, whereby the conjugate gradient parameter β_k is seen as a convex combination of various distinct CG formulae [13]. While assuring the preservation of conjugacy conditions under the stringent Wolfe line search criteria. The hybrid parameters are often delineated by the subsequent equations:

$$\beta_k^{hyb} = (1 - \theta_k)\beta_k^{PRP} + \theta_k \beta_k^{FR} \quad (14)$$

$$\beta_k^{hyb} = (1 - \theta_k)\beta_k^{LS} + \theta_k \beta_k^{CD} \quad (15)$$

$$\beta_k^{hyb} = (1 - \theta_k)\beta_k^{HS} + \theta_k \beta_k^{FR} \quad (16)$$

Moreover, innovative hybrid CG methods have been introduced to address extensive unconstrained optimization challenges [22]. These methods ensure the necessary decreasing condition and incorporate Newton-like directions, which ensures global convergence under standard Wolfe conditions [23]. Some of the hybrid β_k parameters are defined as follows:

$$\beta_k^{hyb} = (1 - \theta_k)\beta_k^{LS} + \theta_k \beta_k^{FR} \quad (17)$$

$$\beta_k^{CLCS} = (1 - \theta_k)\beta_k^{LS} + \theta_k \beta_k^{CD} \quad (18)$$

$$\beta_k^{FG} = (1 - \theta_k)\beta_k^{MMWU} + \theta_k \beta_k^{RMAR} \quad (19)$$

Numerical experiments have demonstrated these hybrid approaches outperform classical CG methods in terms of convergence speed and robustness [18].

Building on the modified BFGS method proposed by introduced a new parameter (t) in the Dai-Liao CG framework [21]. Their method exhibited improved convergence properties, and numerical results confirmed its computational efficiency. In another extension, [20] a proposed has been made to modify the FR method to solve nonlinear equations that are both constrained and ordered, it

Print ISSN 2710-0952 Electronic ISSN2790-1254

shows that the method maintains the property of sufficient decreasing and exhibits strong performance in signal and image processing applications.

For large-scale unconstrained optimization [24]. proposed a Dai-Yuan-based CG method with a spectral CG parameter that guarantees independence form line search conditions. Strong Wolfe conditions were used to establish their method's global convergence, and it we successfully used to solve problems related to the removal of impulse noise. Condition of Dai-Liao conjugacy in hybrid CG methods several recent hybrid CG approaches have utilized the **Dai-Liao (D.L) conjugacy condition** as a fundamental principle for improving search directions. In nonlinear CG metho ds, the classical conjugacy' condition is represented by;

$$d_{k+1}^T y_k = 0 \quad (20)$$

Extended this by incorporating a secant condition from quasi-Newton methods, defining the Hessian approximation B_{k+1} follows [25]:

$$B_{k+1} s_k = y_k \,. \qquad B_{k+1} d_{k+1} = -g_{k+1} \qquad (21)$$

generalized this into the **extended conjugacy condition [29]:**

$$d_{k+1}^T y_k = -t d_{k+1}^T s_k. \quad \text{where} \quad t \geq 0 \,. \qquad (22)$$

The Dai-Liao CG approach has garnered considerable interest owing to its straightforward structure and minimal memory demands. A three-term Dai-Liao CG algorithm was developed that satisfies both the conjugacy and sufficient descent conditions by integrating the Dai-Liao condition with a symmetric modified Perry's matrix. Their approach attained global convergence utilizing Wolfe line search (W.L.S) and exhibited exceptional performance in numerical experiments.

The optimal selection of parameter t remains an active research topic. Several choices have been proposed in prior studies, including those by [25] A new hybrid CG method based on the Dai-Liao (D.L) conjugacy framework and the' Dixon and Zhan- Hager algorithms are presented in this study as a result of these findings. The proposed approach seeks to take advantage of the features of these algorithms, while emphasizing overall convergence and numerical efficiency.

Research Organization

Section 1: Introduction to classical conjugate gradient (C. C.G) methods.

Section 2: Literature review.

Section 3: Presentation of the proposed method with a mathematical formulation and theoretical explanation.

Section 4: Analysis of the theoretical properties of the proposed method, including convergence and stability of the search direction.

Section 5: Numerical experiments and comparisons with traditional approaches on iteration count and processing efficiency.

Section 6: Discussion of the results, conclusions, and future recommendations**.**

Research structure

This section presents a novel hybrid conjugate gradient approach that integrates the update parameters suggested by Zhang-Hager [15,16] and Dixon[11], according to the Dai-Liao (D. L)[19] conjugacy criterion.The subsequent approaches are integrated in a convex manner to formulate the novel hybrid CG parameter:

$$\beta_k^{New2} = (1-\theta_k)\beta_k^{DX} + \theta_k \beta_k^{ZH} \quad (23)$$

From the previous relations we obtain:

$$\beta_k^{New2} = (1-\theta_k)\left(\frac{-g_{k+1}^T g_{k+1}}{g_k^T d_k}\right) + \theta_k\left(\frac{g_{k+1}^T y_k}{d_k^T y_k} - 2\frac{\|y_k\|^2 g_{k+1}^T d_k}{(d_k^T y_k)^2}\right)$$

Where θ_k is the crossbreeding scalar parameter satisfying $\theta_k \in [0.1]$.

If $\theta_k = 0$, then $\beta_k^{New2} = \beta_k^{DX}$.
If $\theta_k = 1$, then $\beta_k^{New2} = \beta_k^{ZH}$.
If $0 < \theta_k < 1$, then β_k^{New2} it is a proper convex combination of β_k^{DX} and β_k^{ZH}.

Therefore; by substituting β_k^{New2} into equation (3), we get:

$$d_{k+1} = -g_{k+1} + \beta_k^{New2} d_k \quad (24)$$

Print ISSN 2710-0952 Electronic ISSN2790-1254

Using relation (24) and by taking the internal strike with the equipment y_k^T, we obtain:

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

By applying the Dai-Liao coupling condition to this relation and following other operations to obtain the following new hybridization parameter:

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k.$$

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

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

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

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

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

$$\theta_k = \frac{\frac{\|g_{k+1}\|^2(d_k^T y_k)+g_{k+1}^T y_k(g_k^T d_k)-t g_{k+1}^T s_k(g_k^T d_k)}{g_k^T d_k}}{\frac{\|g_{k+1}\|^2 (d_k^T y_k)^2+g_{k+1}^T y_k(d_k^T y_k)(g_k^T d_k)-2\|y_k\|^2 g_{k+1}^T d_k(g_k^T d_k)}{(g_k^T d_k)(d_k^T y_k)}}$$

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

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

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

$$\theta_k = \frac{(\|g_{k+1}\|^2-\|g_k\|^2-t g_k^T s_k)(d_k^T y_k)}{\|g_{k+1}\|^2(d_k^T y_k)-\|g_k\|^2(d_k^T y_k) - 2\|y_k\|^2(g_k^T d_k)} \qquad (25)$$

For large-scale optimization problems, methods that avoid direct computation of the Hessian matrix are preferred. Therefore, to maintain efficiency in large -scale problems; this research selects the modulate parameter t based on optimal choices from[16] :

$$t^* = \frac{s_k^T y_k}{s_k^T s_k} \qquad (26)$$

This formulation ensures that the proposed method benefits from the strengths of both Zhang-Hager and Dixon approaches while maintaining global convergence properties under Wolfe line search conditions.

5.1 Dai-Liao (D.L) Hybrid Dixon and Zhang-Hager (New2) Algorithm

Step 1: Initialization; Select an initial point $x_o \in R^n$, a tolerance $\epsilon > 0$,

and parameters $0 < \delta < \sigma < 1$. Compute $f(x_o)$ and g_o.

Step 2: Test for Continuation of Iteration, if $\|g_k\| \leq \epsilon$, then' step.

Step 3: Line Search, Compute $\alpha_k > 0$ satisfied Wolfe conditions (4), (5).

Step 4: Computation of θ_k, If $\|g_{k+1}\|^2(d_k^T y_k) - \|g_k\|^2(d_k^T y_k) -$

Print ISSN 2710-0952 Electronic ISSN2790-1254

$2\|y_k\|^2(g_k^T d_k) = 0$, then set $\theta_k = 0$; otherwise, compute θ_k using the defined formula (25) and (26).

Step 5: Computation of' β_k^{ADM2}; if $0 < \theta_k < 1$, then compute β_k^{ADM2} using the convex combination by (23).

Step 6: Computation of Search Direction (S.D), compute the new search direction by (24) if the Powell restart criterion:

$$|g_{k+1}^T g_k| > c\|g_{k+1}\|^2 \qquad (27)$$

Is satisfied; then set $d_{k+1} = -g_{k+1}$; otherwise; set $d_{k+1} = d$.

Compute α_k; set k= k + 1and go' to step 2.

convergence algorithm Analysis

The convergence outcomes of the proposed hybrid conjugate gradient (C.G) technique are examined under strong Wolff conditions (S.W.C). For convergence, the algorithm must fulfill both the adequate decrease criterion and the global convergence characteristics.

Satisfying Sufficient Descent Condition

Definition; A search in the direction d_k is said to satisfy the decreasing condition only when [17]:

$$d_k^T g_k < 0 \qquad (28)$$

Moreover, the condition of sufficient decrement is satisfied if and only if:

$$d_k^T g_k \le -c\|g_k\|^2, \quad \forall k \ge 0 \qquad (29)$$

Where c is a non-negative constant.

6.1. Theorem.

Let the two chains be $\{g_k\}$ and $\{d_k\}$ they were created by β_k^{New2} the style. Then, the search path d_k satisfies the sufficient descent condition as (27).

With $c = a_1 c_2 + (1 - a_2)c_1$

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

Proof:

We show that the search direction d_k satisfies the required decreasing condition. For the case when $k = 0$, the proof is straightforward, as we can consider $d_0 = -g_0$, which leads to $g_0^T d_0 = -\|g_0\|^2$. From this, we conclude that equation (27) holds for k = 0 . Next, we proceed to the prove that this condition remains valid for $k > 0$.

Distinctly $$d_{k+1} = -g_{k+1} + \beta_k^{New2} d_k$$

$$d_{k+1} = -g_{k+1} + ((1-\theta_k)\beta_k^{DX} + \theta_k \beta_k^{ZH})d_k$$

We can modify the search path in the following way:

$$d_{k+1} = -(\theta_k g_{k+1} + (1-\theta_k)g_{k+1} + ((1-\theta_k)\beta_k^{DX} + \theta_k \beta_k^{ZH})d_k$$

It follows that

$$d_{k+1} = \theta_k(-g_{k+1} + \beta_k^{ZH} d_k) + (1-\theta_k)(-g_{k+1} + \beta_k^{DX} d_k)$$

Where from

$$d_{k+1} = \theta_k d_{k+1}^{ZH} + (1-\theta_k) d_{k+1}^{DX} \quad (28)$$

Multiply equation (28) from the left side by g_{k+1}^T , we get:

$$g_{k+1}^T d_{k+1} = \theta_k g_{k+1}^T d_{k+1}^{ZH} + (1-\theta_k) g_{k+1}^T d_{k+1}^{DX} \quad (29)$$

Firstly; let $\theta_k = 0$; then $d_{k+1} = d_{k+1}^{DX}$. Recollect that

$$d_{k+1}^{DX} = -g_{k+1} + \beta_k^{DX} d_k$$

$$g_{k+1}^T d_{k+1}^{DX} = g_{k+1}^T(-g_{k+1} + \beta_k^{DX} d_k)$$

$$g_{k+1}^T d_{k+1}^{DX} = g_{k+1}^T\left(-g_{k+1} + \left(\frac{-g_{k+1}^T g_{k+1}}{g_k^T d_k}\right) d_k\right)$$

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

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

$$g_{k+1}^T d_{k+1}^{DX} = -\|g_{k+1}\|^2 \left(\frac{g_k^T d_k + g_{k+1}^T d_k}{g_k^T d_k}\right) \quad (30)$$

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

For (30) satisfy sufficient descent condition we have

$$\left|\frac{g_k^T d_k + g_{k+1}^T d_k}{g_k^T d_k}\right| \leq \mu \|g_{k+1}\|^2, \text{ where } 0 < \mu < 1$$

So that

$$g_{k+1}^T d_{k+1}^{DX} \leq -\|g_{k+1}\|^2 + \mu \|g_{k+1}\|^2$$

$$g_{k+1}^T d_{k+1}^{DX} \leq -(1-\mu)\|g_{k+1}\|^2$$

We denote $c_1 = (1-\mu)$,then we can write

$$g_{k+1}^T d_{k+1}^{DX} \leq -c_1 \|g_{k+1}\|^2$$

We are done with $\theta_k = 0$,

Now; let $\theta_k = 1$,,then $d_{k+1} = d_{k+1}^{ZH}$. Recollect that

$$d_{k+1}^{ZH} = -g_{k+1} + \beta_k^{ZH} d_k$$

$$g_{k+1}^T d_{k+1}^{ZH} = g_{k+1}^T(-g_{k+1} + \beta_k^{ZH} d_k)$$

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

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

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

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

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

$$g_{k+1}^T d_{k+1}^{ZH} = -\|g_{k+1}\|^2 \left(\frac{-2\|y_k\|^2 \|d_k\|^2}{(d_k^T y_k)^2}\right) \quad (31)$$

For equation (31) to satisfy the sufficient' descent condition, the following must hold:

$$\left|\frac{-2\|y_k\|^2 \|d_k\|^2}{(d_k^T y_k)^2}\right| \leq k \quad , k > 0$$

So that

$$g_{k+1}^T d_{k+1}^{ZH} \leq -k\|g_{k+1}\|^2$$

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

We denote $c_2 = k > 0\, ,then\ we\ can\ write$

$$g_{k+1}^T d_{k+1}^{ZH} \le -c_2 \|g_{k+1}\|^2$$

Now, we suppose that $0 < \theta_k < 1\, ,such\ that\, , 0 < a_1 < \theta_k < a_2 < 1$. From (29) we conclude that

$$g_{k+1}^T d_{k+1} \le a_1 g_{k+1}^T d_{k+1}^{ZH} + (1 - a_2) g_{k+1}^T d_{k+1}^{DX}$$

$$\Rightarrow g_{k+1}^T d_{k+1} \le a_1(-c_2\|g_{k+1}\|^2) + (1 - a_{2,})(-c_{1,}\|g_{\,k+1}\|^2)$$

$$g_{k+1}^T d_{k+1} \le -(a_1 c_2 + (1 - a_2) c_1)\|g_{k+1}\|^2$$

Denote $c = a_1 c_2 + \left(1 - a_{2,}\right) c_1$,then we eventually get:

$$g_{k+1}^T d_{k+1} \le -c\|g_{k+1}\|^2$$

Convergence Analysis.

Every conjugate gradient (C.G) method employing a deep Wolfe line search retains its consistency. However, for the method to function at a basic level, only the weak form of the **Zoutendijk condition (Z.C)** is needed [14]. For the subsequent analysis, we adopt the following assumption.

7.1- Assumption;

The level' set's $Q = \{x \epsilon R^n : f(x) \le f(x_0)\}\ at\ x_{\,0,}$ Defined because x_0 is the first point, which means there is a value $M > 0$, such that $\|x\| \le M, \forall\ x \epsilon Q$.

7.2- Assumption:

In a neighborhood' $N\ of\ Q$,, the function f is constantly separable, and it's the gradient is Lipschitz constant, implying that a constant exists. $L > 0$, such that:

$$\|\nabla f(x) - \nabla f(y)\| \le L\|\, x - y\, \|\, , \forall\, x, y, \in N$$

It can be determined based on hypotheses (5.1) and (5.2) where there is a positive constant value $\bar{y}$, s.t:

$$\|g_{k+\,1}\| \le\ \bar{y}\, , \forall x \in Q$$

7. 3. Lemma

Let assumptions (5.1) and (5.2) hold. Consider the process (2) and (3), where d_k is a descent direction, and α_k is determined using SWP [26].

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

$$\sum_{k\geq 1}\frac{1}{\|d_k\|^2}=0 \quad (32)$$

Thereafter

$$\lim_{k\to\infty} inf\|g_k\|=0 \quad (33)$$

7 .4. Theorem:

If proposition (5. 1), and (5. 2) hold, and the algorithm is applied using β_k^{New2} where $0\leq\theta_K\leq 1$, withα_k determined according to the strong Wolfe line search, and d_k being the descent direction by (33).

Proof:

Since the decreasing condition is satisfied, we obtain $d_{k+1}\neq 0$. thus, using lemma (5.3), it is sufficient to prove that $\|d_{k+1}\|$ is bounded from above. This is based on equation (23).

$$d_{k+1}=-g_{k+1}+\beta_k^{New2}d_k$$

$$\|d_{k+1}\|\leq\|g_{k+1}\|+|\beta_k^{New2}|\|d_k\|$$

$$\|d_{k+1}\|\leq\|g_{k+1}\|+\left[|1-\theta_k|.|\beta_k^{DX}|+|\theta_k||\beta_k^{ZH}|\right].\|d_k\|$$

$$\leq\|g_{k+1}\|+\left[|1-\theta_k|.\left|\frac{-g_{k+1}^Tg_{k+1}}{g_k^Td_k}\right|+|\theta_k|\left|\frac{g_{k+1}^Ty_k}{d_k^Ty_k}-\frac{2\|y_k\|^2g_{k+1}^Td_k}{(d_k^Ty_k)^2}\right|\right].\|d_k\|$$

By taking $\|g_{k+1}\|$ common, and we replace θ_k from (17), and suppose that

$$a=|1-\theta_k|.\left|\frac{-g_{k+1}^Tg_{k+1}}{g_k^Td_k}\right|\|d_k\|$$

and

$$b=|\theta_k|\left|\frac{g_{k+1}^Ty_k}{d_k^Ty_k}-\frac{2\|y_k\|^2g_{k+1}^Td_k}{(d_k^Ty_k)^2}\right|\|d_k\|$$

Then we get

$$\|d_{k+1}\|\leq(1+a+b)\|g_{k+1}\|$$

We denote $c=(1+a+b)$

$$\|d_{k+1}\|\leq c\,\|g_{k+1}\|$$

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

Hence, $\sum_{k\geq 1}\frac{1}{\|d_k\|^2} \geq \frac{1}{c^2\bar{y}^2}\sum_{k\geq 1} 1 = \infty$

Where $\|g_{k+1}\| \leq \bar{y}$, from Assumption (5.2)

$\therefore \lim_{k\to\infty} inf\|g_{k+1}\| = 0$

Numerical Resulted

This section presents the numerical performance of our newly suggested β_k^{New2} conjugate gradient method, implemented in FORTRAN, on a collection of 68 unconstrained large-scale optimization problems sourced from references [27] and [28]. These issues are examined in their expanded or generalized versions. Each test issue was addressed with differing quantities of variables, ranging from n = 100 to 1000. The performance of the β_k^{New2} algorithm is compared with two well-known methods $(\beta^{HZ})and\ (\beta^{DX})$. All algorithms were implemented under the same conditions, using the standard Wolfe line search, with parameters $\rho = 10^{-4}$ and $\sigma = 0.9$ the initial step size was defined as $\alpha_1 = \frac{1}{\|g_1\|}$, for all subsequent iterations $k > 1$, the step size was updated using $\alpha_k = \alpha_{k-1}\left(\frac{\|d_{k-1}\|}{\|d_k\|}\right)$.

All implementations were written in double precision FORTRAN (2000) and compiled using F77 with default compiler settings. The original code, written by Andrei, was modified by us to suit our proposed method. The stopping criterion used in all cases was the $\|g_k\|_2 \leq 10^{-6}$ and the maximum number of iterations is 1000. The comparison focused on three main performance indicators:

iter-Number of iterations.
fg-Number of function and gradient evaluations.
COP time-Execution time in seconds.
Figures 1,2, and 3 illustrate the performance of the compared methods across all 68 test problems. The results are presented using performance profiles following the methodology of Dolan and More [29], which plot the proportion pp of problems for which a method is within a factor τ\tau of the best result for each metric (iterations, evaluations, and time).

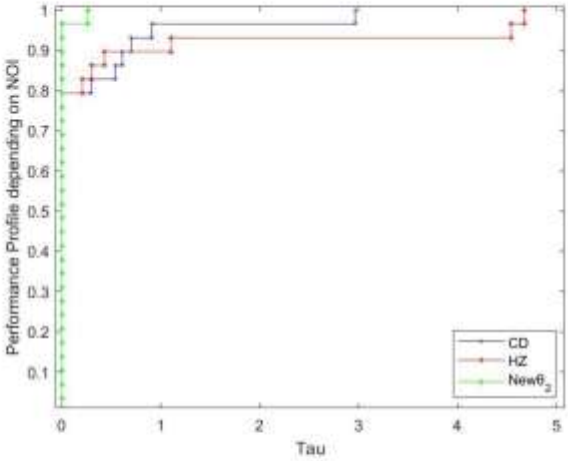

FIGURE 1. Performance profiles derived from iterations (NOI).

The left side of the figure illustrates the percentage of test problems for which each method exhibited the fastest performance, whereas the right side displays the percentage of problems successfully solved that achieved the highest number of solutions within a factor τ of optimal performance, considering iterations, function and gradient evaluations, and computational time.

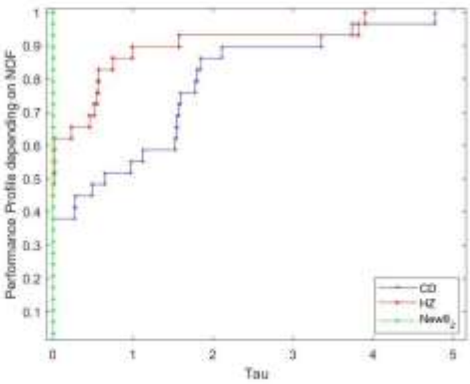

FIGURE 2- Performance profiles based on number of function evaluations (NOF).

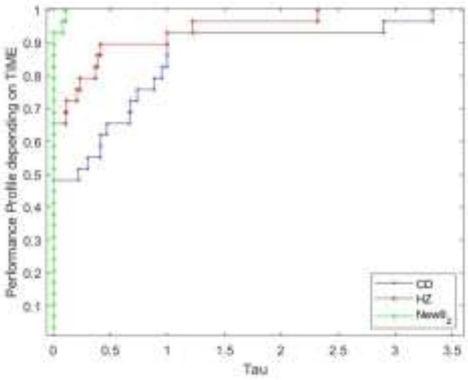

FIGURE 3- Performance profiles based on processing time

(CPU time spent - TIME).

conclusion

This study assessed the efficacy of the Dixon and Zhang-Hager algorithms in addressing unconstrained optimization issues using conjugate gradient techniques. The numerical findings indicated that both methods are efficient, with the Zhang-Hager algorithm excelling in some functions regarding iteration count and computing time. This work underscores the significance of selecting the suitable algorithm according to the function's characteristics and paves the way for future advancements aimed at reducing computing costs and enhancing stability.

References

[1] Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The Computer Journal*, *7*(2), 149-154.

[2] Babaie-Kafaki, S. (2023). A survey on the Dai–Liao family of nonlinear conjugate gradient methods. *RAIRO-Operations Research*, *57*(1), 43-58.

[3] Wolfe, P. Convergence condition for ascent methods. SIAM review, 11(2),226-235(1969).

[4] Sakai, H., & Iiduka, H. (2021). Sufficient descent Riemannian conjugate gradient methods. *Journal of Optimization Theory and Applications*, *190*(1), 130-150.

[5] Polak, E., & Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. Revue Française d'informatique et de Recherche Opérationnelle, 3(16), 35-43.

[6] Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, *49*(6), 409-436.

[7] Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of Optimization Theory and Applications*, *69*(1), 129-137.

[8] Dai, Y. H., & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on Optimization*, *10*(1), 177-182.

[9] Perry, J. M. (1977). A class of conjugate gradient algorithms with a two-step variable-metric memory. *Discussion Paper No. 269, Center for Mathematical Studies in Economics and Management Science, Northwestern University*.

[10] Hager, W. W., & Zhang, H. (2005). A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on optimization*, *16*(1), 170-192.

[11] Dixon, L. C. W. (1972). Conjugate gradient algorithms: quadratic termination without exact line searches. *Journal of the Institute of Mathematics and Its Applications*, *10*(1), 35-42.

[12] Djordjevic, B., Ivanović-Đukić, M., Lepojević, V., & Milanović, S. (2021). Job satisfaction and organizational citizenship behaviour of employees in companies in the Republic of Serbia. *Strategic Management-International Journal of Strategic Management and Decision Support Systems in Strategic Management*, *26*(3).

[13] Hu, Q., Zhu, L., & Chen, Y. (2024). Alternative extension of the Hager–Zhang conjugate gradient method for vector optimization. *Computational Optimization and Applications*, *88*(1), 217-250.

[14] Zoutendijk, G. (1970). Nonlinear programming, computational methods. *Integer and nonlinear programming*, 37-86.

[15] Hager, W. W., & Zhang, H. (2006). Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent. *ACM Transactions on Mathematical Software (TOMS)*, *32*(1), 113-137.

[16] Hager, W. W., & Zhang, H. (2005). A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on optimization*, *16*(1), 170-192.

[17] Al-Baali, M. (1985). Descent property and global convergence of the Fletcher—Reeves method with inexact line search. *IMA Journal of Numerical Analysis*, *5*(1), 121-124.

[18] Andrei, N. (2008). A hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan algorithms. *Studies in Informatics and Control*, *17*(4), 373-392.

[19] Babaie-Kafaki, S., & Ghanbari, R. (2014). The Dai–Liao nonlinear conjugate gradient method with optimal parameter choices. *European Journal of Operational Research*, *234*(3), 625-630.

[20] Hassan, B. A., Abdullah, Z. M., & Jabbar, H. N. (2019). A descent extension of the Dai-Yuan conjugate gradient technique. *Indonesian Journal of Electrical Engineering and Computer Science*, *16*(2), 661.

[21] Abdullah, Z. M., Hameed, M., Hisham, M., & Khaleel, M. A. (2019). Modified new conjugate gradient method for Unconstrained Optimization. *Tik. J. of Pure Sci.*, *24*(5), 86-90.

[22] Abdullah, Z. M., & Dhyab, F. S. (2021). New Hybrid Conjugate Gradient Method as a Convex Combination of Dai–Liao and Wei–Yao–Liu. *Tikrit Journal of Pure Science*, *26*(6), 100-105.

[23] Salihu, N., Odekunle, M. R., Saleh, A. M., & Salihu, S. (2021). A dai-liao hybrid hestenes-stiefel and fletcher-Revees methods for unconstrained optimization. *International Journal of Industrial Optimization*, *2*(1), 33-50.

[24] Baban do, H. A., Barma, M., Nasiru, S., & Suraj, S. (2024). A dai-liao hybrid prep and dy schemes for unconstrained optimization. *International Journal of Development Mathematics (IJDM)*, *1*(1).

[25] Yao, S., Feng, Q., Li, L., & Xu, J. (2019). A class of globally convergent three-term Dai-Liao conjugate gradient methods. *Applied Numerical Mathematics*, 151 354-366.

[26] Sabi'u, J., Shah, A., Waziri, M. Y., & Ahmed, K. (2021). Modified Hager–Zhang conjugate gradient methods via singular value analysis for solving

Print ISSN 2710-0952 Electronic ISSN2790-1254

monotone nonlinear equations with convex constraint. *International Journal of Computational Methods*, *18*(04), 2050043.

[27] Arazm, M. R., Babaie-Kafaki, S., & Ghanbari, R. (2017). An extended Dai-Liao conjugate gradient method with global convergence for nonconvex functions. *Glasnik mathematic*, *52*(2), 361-375.

[28] Bongartz, I., Conn, A. R., Gould, N., & Toint, P. L. (1995). CUTE: Constrained and unconstrained testing environment. *ACM Transactions on Mathematical Software (TOMS)*, *21*(1), 123-160.

[29] Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming*, *91*, 201-213.