New Search Direction in Nonlinear Programming problem

Eman T Hamed
Elaf Sulaiman Khaleel
Hamsa Th. Saeed chilmeran
Huda I. Ahmed
Karam Adel Abed,
1dr.emantarik@uomosul.edu.iq 2 elaf.sulaiman@uomosul.edu.iq
3Hamsathrot@uomosul.edu.iq 4dr.hudaea@uomosul.edu.iq
5karamadel@uomosul.edu.iq

[a] *Department of Mathematics college of Computers Sciences and Mathematics University of Mosul, Mosul, Iraq*
[b]*Department of Operation Research and Intelligent Techniques college of Computers Sciences and Mathematics University of Mosul, Mosul, Iraq*

Abstract
The present research introduces an innovative conjugate direction for the resolution of non-constrained optimization problem. The proposed direction possesses the descent property and under suitable assumptions we proven to be globally convergent. When compared to classical conjugate gradient techniques the Computational findings show the efficiency and superiority of the suggested method. We utilize the performance profiles developed by Dolan and Moré. it three key metrics is based on execution time (CPU time), numeral of iterations (NOI) and the numeral of function evaluations. Abstract. Quasi-Newton method is one of the most popular methods for solving unconstrained single and multiobjective optimization problems. In a quasi-Newton method, the search direction is computed based on a quadratic model of the objective function, where some approximations replace the true Hes- sian at each iteration. Abstract. Quasi-Newton method is one of the most popular methods for solving unconstrained single and multiobjective optimization problems. In a quasi-Newton method, the search direction is computed based on a quadratic model of the objective function, where some approximations replace the true Hes- sian at each iteration.
Keywords : Nonlinear Programming

تجاه بحث جديد في مسائل البرمجة غير الخطية

إيمان ت. حامد
إيلاف سليمان خليل
حمزة ث. سعيد تشيلميران
هدى إ. أحمد
كرم عادل عابد

أ- قسم الرياضيات، كلية علوم الحاسب والرياضيات، جامعة الموصل، الموصل، العراق

ب- قسم بحوث العمليات والتقنيات الذكية، كلية علوم الحاسب والرياضيات
جامعة الموصل، الموصل، العراق

الملخص: يقدم هذا البحث اتجاهًا مبتكرًا مترافقًا لحل مسائل التحسين غير المقيدة. يتميز الاتجاه المقترح بخاصية الانحدار، وقد أثبتنا، في ظل افتراضات مناسبة، تقاربه العالمي. عند مقارنته بتقنيات التدرج المترافق الكلاسيكية، تُظهر النتائج الحسابية كفاءة وتفوق الطريقة المقترحة. نستخدم ملفات تعريف الأداء التي طورها دولان وموريه. تعتمد مقاييسه الرئيسية الثلاثة على وقت التنفيذ (وقت وحدة المعالجة المركزية)، وعدد التكرارات، وعدد تقييمات الدالة. تُعد طريقة شبه نيوتن إحدى أكثر الطرق شيوعًا لحل مسائل التحسين غير المقيدة أحادية الهدف ومتعددة الأهداف. في طريقة شبه نيوتن، يُحسب اتجاه البحث بناءً على نموذج تربيعي لدالة الهدف، حيث تحل بعض التقريبات محل مصفوفة هيسيان الحقيقية في كل تكرار. تُعد طريقة شبه نيوتن إحدى أكثر الطرق شيوعًا لحل مسائل التحسين غير المقيدة أحادية الهدف ومتعددة الأهداف. في طريقة شبه نيوتن، يتم حساب اتجاه البحث بناءً على نموذج تربيعي لدالة الهدف، حيث تحل بعض التقريبات محل مصفوفة هيسيان الحقيقية في كل تكرار.

كلمات مفتاحية : البرمجة غير الخطية

1. Introduction

Conjugate Gradient (CG) techniques are vital and crucial choice for resolving High-dimensional optimization. They are among the most High-performing technique for minimizing a Target functions

Abstract. Quasi-Newton method is one of the most popular methods for solving unconstrained single
and multiobjective optimization problems. In a quasi-Newton method, the search direction is computed
based on a quadratic model of the objective function, where some approximations replace the true Hes-
sian at each iteration.

Abstract. Quasi-Newton method is one of the most popular methods for solving unconstrained single
and multiobjective optimization problems. In a quasi-Newton method, the search direction is computed
based on a quadratic model of the objective function, where some approximations replace the true Hes-
sian at each iteration.

$$\min f(x) \tag{1}$$

The CG method produces $\{x_k\}$ as follows:

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

Somewhere $x_k$is Current Iterate, α_k >0 is length step.

$$d_{k+1} = \begin{cases} -g_{k+1} & \text{for } k = 0 \\ -g_{k+1} + \beta_k d_k & \text{for } k \geq 1 \end{cases} \tag{3}$$

Somewhere g_k is the gradient of function then the coefficient of algorithm is β_k . [1 ,2 ,3 ,4, 5 ,6,7,8].

Print ISSN 2710-0952 Electronic ISSN 2790-1254

We utilized the weak wolf condition for determined the search line of CG algorithm:

$$f(x_k+\alpha_k d_k)-f(x_k)\le \delta\alpha_k \nabla f(x_k)^T d_k \qquad (4)$$

$$\nabla f(x_k+\alpha_k d_k)^T d_k \ge \sigma \nabla f(x_k)^T d_k \qquad (5)$$

Moreover, robust Wolfe conditions necessitate the Implementation of condition (4) and

$$\left|g(x_k+\alpha_k d_k)^T d_k\right| \le -\sigma\ g_k^T d_k \qquad (6)$$

2-The new formula:

Despite the well-known efficiency of the CG algorithm, its practical performance strongly depends on the choice of the search direction at each step. The search direction directly impacts the convergence speed, the numeral of iterations and the algorithmic stability. Motivated by these critical factors, this work focuses on developing a new search direction within the conjugate gradient framework. In reference [10] a direction is proposed, define as:

$$d_{k+1} = -(1+\beta_k \frac{s_k^T g_{k+1}}{g_{k+1}^T g_{k+1}})g_{k+1} + \beta_k s_k \qquad (7)$$

Where the direction in (3) is replace by eq. (7) and β_k is one of the values [11] or [12].

The present research introduces a modified search direction utilizing the same underlying framework, characterized by the following construction:

$$d_{k+1} = -\left[1+\beta_k (\frac{d_k^T g_{k+1}}{\|d_k\|^2})\right] g_{k+1} + \frac{g_{k+1}^T y_k}{\|d_k\|^2} s_k \qquad (8)$$

Where the parameter

$\beta_k^{RMLI} = \frac{g_{k+1}^T y_k}{\|d_k\|^2}$ is defined according to references [13]

2. 1 Theorem:

The direction given in equation (8) is a sufficient descent direction

Proof:

we shall prove that through pummel both sides of eq. (8)

المجلة العراقية للبحوث الإنسانية والإجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

by $\left(\frac{g_{k+1}}{\|g_{k+1}\|^2}\right)$then we get:

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} \le -\left[1+\frac{g_{k+1}^T y_k}{\|d_k\|^2}\frac{d_k^T g_{k+1}}{\|d_k\|^2}\right]\frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}+\frac{g_{k+1}^T y_k}{\|d_k\|^2}\frac{s_k^T g_{k+1}}{\|g_{k+1}\|^2} \quad (9)$$

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2}+1 \le -\left[\frac{g_{k+1}^T y_k}{\|d_k\|^2}\frac{d_k^T g_{k+1}}{\|d_k\|^2}\right]+\frac{g_{k+1}^T y_k}{\|d_k\|^2}\frac{s_k^T g_{k+1}}{\|g_{k+1}\|^2} \quad (10)$$

since $s_k^T g_{k+1} \le s_k^T y_k$, $y_k^T g_{k+1} \le \|y_k\|\|g_{k+1}\|$

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2}+1 \le -\left[\frac{\|y_k\|\|g_{k+1}\|}{\|d_k\|^2}\frac{s_k^T y_k}{\|d_k\|^2}\right]+\frac{\|y_k\|\|g_{k+1}\|}{\|d_k\|^2}\frac{s_k^T y_k}{\|g_{k+1}\|^2} \quad (11)$$

Since $s_k^T y_k$, $\|y_k\|$, $\|g_{k+1}\|$and $\|d_k\|^2$ are positive constant, then the first them of eq. (11) is negative

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2}+1 \le -\left[\frac{\|y_k\|\|g_{k+1}\|}{\|d_k\|^2}\frac{s_k^T y_k}{\|d_k\|^2}\right]+\frac{\|y_k\|\|g_{k+1}\|}{\|d_k\|^2}\frac{s_k^T y_k}{\|g_{k+1}\|^2} \le \frac{\|y_k\|\|g_{k+1}\|}{\|d_k\|^2}\frac{s_k^T y_k}{\|g_{k+1}\|^2} = \varsigma$$

$$d_{k+1}^T g_{k+1} \le -(1-\varsigma)\|g_{k+1}\|^2 \quad (12)$$

The proof is complete.

ASSUMPTION (A) [14]:

1) The boundedness of set S implies that b>0 is a positive parameter as $\|x\| \le b, \forall x \in s$. For all k. This confirms the set is contained within a finite range

2) The function *f* demonstrates continuous differentiability inside a neighborhood N of S, and its satisfy The condition of Lipschitz, as given by the equation:

$$\|g(x) - g(y)\|L\|x - y\|, \forall x, y \in N\, N \quad (13)$$

On the basis of these assumptions concerning f, we are able to arrive at the conclusion when the positive constant represented by $\gamma>0$ occurs in such a way that:

2.2 Theorem

Assume that the previous assumptions are true. If function f is convex on R^n then $\lim_{k\to\infty}(\inf \| g_k \| = 0)$

Assuming the preceding hypotheses are satisfied, the convexity of the objective function f implies the following convergence property: $\lim_{k\to\infty}(\inf \| g_k \| = 0)$

Proof:

we shall take the absolute value of eq. (7), we obtain:

$$\|d_{k+1}\| \le \left\| \left[1 + \frac{g_{k+1}^T y_k}{\|d_k\|^2} \frac{d_k^T g_{k+1}}{\|d_k\|^2} \right] d_k \right| + \left| \frac{g_{k+1}^T y_k}{\|d_k\|^2} \right|$$

$$\|d_{k+1}\| \le \left\| \left[1 + \frac{g_{k+1}^T y_k}{\|d_k\|^2} \frac{d_k^T g_{k+1}}{\|d_k\|^2} \right] d_k \right| + \left| \frac{g_{k+1}^T y_k}{\|d_k\|^2} \right| \text{``}$$

since $y_k^T g_{k+1} \le \|y_k\| \|g_{k+1}\|, s_k^T g_{k+1} \le s_k^T y_k$,we get:

$$\|d_{k+1}\| \le \left| 1 + \frac{\|y_k\| \|g_{k+1}\|}{\|d_k\|^2} \frac{s_k^T y_k}{\|d_k\|^2} \right| \|d_k\|^2 + \left| \frac{\|y_k\| \|g_{k+1}\|}{\|d_k\|^2} \right| = D$$

$$\|d_{k+1}\| \le \|g_{k+1}\| + \varpi \|s_k\| + \rho \|y_k\| = D$$

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

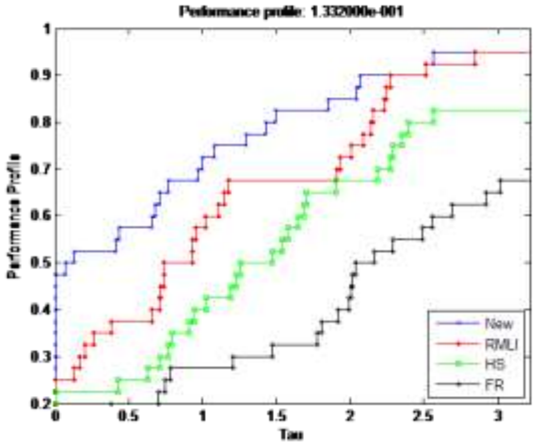
i.e. $\lim_{k\to\infty} \|g_k\| = 0.$

3. Numerical results

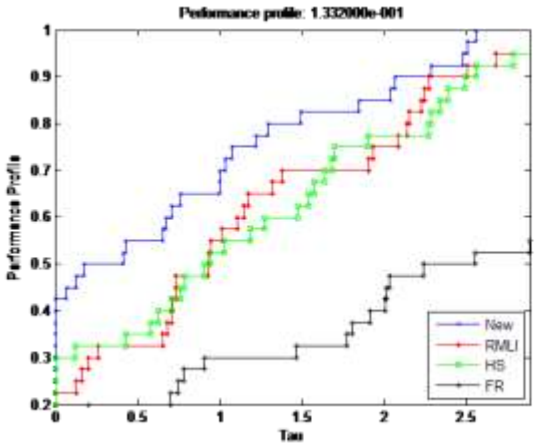
This part of the paper details the computational performance evaluated against a suite of standard unconstrained optimization benchmarks derived from [15]. All algorithms were implemented using the Fortran programming language. In the numerical experiments, the parameter values were set to = 0.001 and σ = 0.99, also the stopping criterion defined as $\|g_{k+1}\| \le 10^{-5}$.

To evaluate the efficiency of the proposed method defined in eq. (8), its performance was compared with several classical CG algorithm, namely the Hestenes-Stiefel (HS), the Fletcher-Reeves (FR), and the RMLI methods. As illustrated in the subsequent figures evident the numerical results indicated that the proposed technique significantly outperforms.

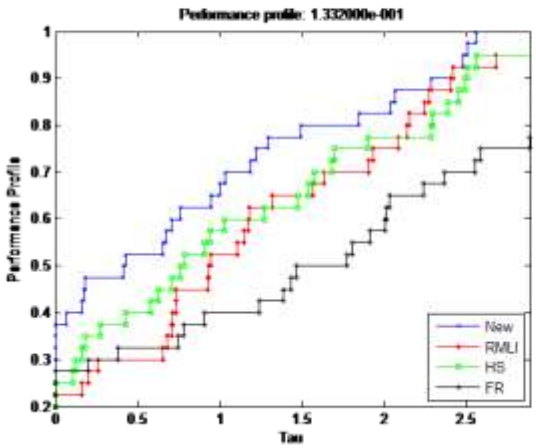
Furthermore, to evaluate the comparative efficiency of the algorithms, we utilized the Dolan-Moré performance profiles [16]. The assessment was conducted based on three primary metrics: CPU time, the total count of iterations (NOI), and the numeral of objective function calculations (NOF). Detailed performance trends for various problem sizes (n=500, 1000, and 10000) are presented in Figures (1) through (3)



Figures (1): the numeral of objective function calculation



Figures (2): the numeral of iterations (NOI),



Figures (3): the CPU time Moreover, robust Wolfe conditions necessitate the (4) Implementation of condition

Reference:

[1] M. R. Hestenes, and E. Stiefel, "Methods of conjugate gradients for solving linear systems" J. Res. Natl. Bur. Stand. (1934)., vol. 49, no. 6, p. 409, 1952.

[2] Dai, Y. H., & Liao, L. Z. (2002). R-linear convergence of the Barzilai and Borwein gradient method. IMA Journal of Numerical Analysis, 22(1), 1-10.

[3] Dai, Yu-Hong, (2003). A Family of Hybrid Conjugate Gradient Method for Unconstrained Optimization. Journal of American Mathematical Society, 72 (243), 1317-1328.

[4] Fletcher, R. & Reeves, C. M., (1964). Function Minimization by Conjugate Gradients. Computer Journal, 7, 149-154.

[5] Polak, E. & Ribiére, G., (1969). Note Sur la Convergence de Directions Conjuguée. Revue Francaise Information, Recherche. Operationnelle, (16), pp.35-43.

[6] Altaee Satia A. & Chilmeran Hamsa Th. (2013). "Improve A Three Terms Conjugate", Gradient Algorithms for Solving The Nonlinear optimization.

[7] Ban A. Mitras, & Nada F. Hassan (2013). New Three–Term CG- Method for Unconstrained Optimization. AL-Rafidain Journal of Computer Sciences and Mathematics, 10(3), 29-35.

[8] Basim A. Hassan[1] and Ranen M. Sulaiman, (2020). "A new Type of self-scaling Quasi – Newton for Unconstrained optimization", International Journal of Advanced Science and Technology, Vol. 29. No. 4, PP. 10822-10827.

[9] Chilmeran, H. S., Ahmed, H. I., Hamed, E. T., & Al-Bayati, A. Y. (2021). A new modification of the quasi-newton method for unconstrained optimization. Indonesian Journal of Electrical Engineering and Computer Science, 21(3), 1683-1691.

[10] L.Zhang, W. Zhou (2008),"Tow descent hybrid conjugate gradient method for optimization", J. of Computational and Applied Mathematics 216, pp.215-264.

[11] E.G. Birgin, J.M. Martinez (2001),"A spectral conjugate gradient method for unconstrained optimization", Applied Mathematics and Optimization, 43, pp.117-128.

[12] A. G. Buckley(1978),"Extending the relationship between the conjugate gradient and BFGS algorithms", Mathematical Programming, 15, pp. 343-348.

[13] Rivaie, M., Mamat, M., June, L.W., Mohd, I.: A new class of nonlinear conjugate gradient coefficients with global convergence properties. Appl. Math. Comput. 218(22), 11323–11332 (2012).

[14] Yabe, H., & Takano, M. (2004). Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. Computational optimization and Applications, 28(2), 203-

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

[16] E. D. Dolan, and J. J. Moré," Benchmarking optimization software with performance profiles," Mathematical programming, vol. 91, no. 2, pp. 201-213, 2002.