A New Non Quadratic Model For Unconstrained Non Linear Optimization

Basim A. Hassan **Nidhal H. Al-Assady**

basimabas39@gmial.com

College of Computer Sciences and Mathematics
University of Mosul, Iraq

Received: 15/04/2002 **Accepted: 20/07/2002**

ABSTRACT

A new non-quadratic model is proposed for solving unconstrained optimization problems, which modifies and develops the classical conjugate gradient methods. The technique has the same properties as the classical conjugate gradient method that can be applied to a quadratic function. An algorithm is derived and evaluated numerically for some standard test functions .The results indicate that, in general, the new algorithm is an improvement on the previous methods so it remains to be investigated.

Keywords: Non-quadratic model, Conjugate gradient methods, Numerical experimentsn.

حل مسائل الامثلية غير المقيدة باستخدام خوارزميات معتمدة على النموذج العام

باسم عباس حسن **نضال حسين الاسدي**

كلية علوم الحاسبات والرياضيات، جامعة الموصل

تاريخ استلام البحث: 2002/04/15 **تاريخ قبول البحث: 2002/07/20**

الملخص

يتركز هذا البحث على اقتراح نموذج جديد أكثر عمومية من النماذج التربيعية لحل مسائل الامثلية اللاخطية وغير المقيدة والذي يحور ويطور خوارزميات التدرج المترافق الكلاسيكية وله نفس خواص خوارزميات التدرج المترافق المطبقة على دالة تربيعية. إن النموذج الجديد المستخدم في خوارزميات التدرج المترافق الكلاسيكية هو النموذج دالة مثلثية الذي تم اشتقاقه وحله عددياً باستخدام بعض الدوال المختارة لاختبار كفاءة الخوارزمية الجديدة.

وبصــورة عامــة تشــير النتــائج إلــى أن الخوارزميــة الجديــدة هــي تحســين للخوارزميات السابقة.

الكلمات المفتاحية: النموذج غير التربيعي، طرائق التدرج المترافق, التجارب العددية.

1. Introduction

A more general model than the quadratic one is proposed in this paper as a basis for a CG algorithm. If q(x) is a quadratic function, then a function f is defined as a nonlinear scaling of q(x) if the following condition holds :

f = F(q(x)), dF/dq = F′ > o and q(x) > 0………. (1)

where x* is the minimizer of q(x) with respect to x spedicate[12].

The following properties are immediately derived from the above condition:

i) Every contour line to q(x) is a contour line of f .

ii) If x* is a minimizer of q(x), then it is a minimizer of f.

iii) That x* is a global minimum of q(x) does not necessarily mean that it is a global minimum of f.

Various authors have puplished related works in the area:
A conjugate method which minimizers the function
$f(x) = (q(x))^{\rho}$, and $x \in R^n$ in at most step has been described by Fried[9].

Another special case, namely $F(q(x)) = \varepsilon_1 q(x) + \frac{1}{2}\varepsilon_2 q^2(x)$

Where ε_1 and ε_2 are scalars, has been investigated by Boland et al[5].

Another model has been developed by Tassopoulos and Storey [13]as follows: **$F(q(x)) =(\varepsilon_1 q(x) + 1)/\varepsilon_2\ q(x)$: $\varepsilon_2 > 0$**

AL-Assady[3]developed a model as follows (F(q(x)) = In (q(x))

Al-Bayati [1] has developed a new rational model which is defined as follows: **$F(q(x)) = \varepsilon_1 q(x)/1-\varepsilon_2 q(x)$.**

Al-Bayati[4] has also developed an extended CG algorithm which is based on a general logarithmic model

$$F(q(x)) = \log(\varepsilon q(x) - 1),\ \varepsilon > 0$$

And Al-Assady [2] described there ECG algorithm which is based on the natural log function for the rational q(x) function

$$\mathbf{F(q) = log}\ \left[\varepsilon_1 q(x) \big/ \varepsilon_2 q(x) + 1\right]\ ,\ \varepsilon_2 < 0$$

2. The New Non - Quadratic Model:

In this paper, a new sine model is investigated and tested on a set of a standard test function, and, assumed? That condition (1) holds. An extended conjugate gradient algorithm is developed which is based on this new model which scales q(x) by the natural sine function for the rational q(x) functions.

$$\mathbf{F(q(x)) = sin\ (\varepsilon\ q(x))} \qquad \ldots\ldots\ldots\ldots\ (2)$$

We first observe that q(x) and F(q(x)) given by (2) have identical contours, though with different function values, and they have the same unique minimum point denoted by x^*.

2.1 The Algorithm:

Given $x_0 \in R^n$ an initial estimate of the minimizer x^*.

Step (1): set $d_0 = - g_0$.

Step (2) : For i = 1, 2, ….

Compute $x_i = x_{i-1} + \lambda_{i-1} d_{i-1}$

Where λ_{i-1} is the optimal step size obtained by the line search procedure.

Step (3) : compute

$$\rho_i = \left[\frac{ \dfrac{\left[if_{i-1} + \sqrt{1 - f_{i-1}^2} \right]^2 + 1}{ if_{i-1} + \sqrt{1 - f_{i-1}^2} } }{ \dfrac{\left[if_i + \sqrt{1 - f_i^2} \right]^2 + 1}{ if_i + \sqrt{1 - f_i^2} } } \right]$$

Where the derivations of scaling p_i will be presented below.

Step (4): calculate the new direction

$d_i = - g_i + \beta_i d_i$.

Where β_i is defined by different formulae according to variation and it is expressed as follows:

$\beta_i = \rho_i (\| g_i \|^2 / \| g_{i-1} \|^2)$ [modified Fletcher and Reeves[8], F/R]

$\beta_i = g_i^T (\rho_i g_i - g_{i-1}) / d_{i-1}^T (\rho_i g_i - g_{i-1})$ [Modified Hestenes and Stiefle [10], H/S]

$\beta_i = g_i^T (\rho_i g_i - g_{i-1}) / d_{i-1}^T g_i$ [modified Polak and Ribiera [11],P/R]

$\beta_i = \rho_i \|g_{i+1}\|^2 / d_i^T g_i$ [modified Dixon [7]]

conjugate gradient methods are usually implemented by restarts in order to aviod an accumulation of errors affecting the search directions.

It is, therefore, generally agreed that restarting is very helpful in practices, so we have used the following restarting criterion in our practical investigations. If the new direction satisfies:

$$\mathbf{d_i^T g_i} \geq -0.8 \|\mathbf{g_i}\|^2$$

Then a restart is also initiated. This new direction is sufficiently downhill.

2.2 The Derivation of p_i for the New Model:

The implementation of the extended CG method has been performed for general function F(q(x)) of the form of equations(2).

The unknown quantities ρ_i were expressed in terms of available quantities of the algorithm .

The new sinh(q(x)) model can now be written as :

f(x) = F(q(x) = sin (ε q(x)) ……. Solving equation (2) for q

$2if = e^{i\varepsilon q} - e^{-i\varepsilon q} \Longrightarrow e^{2i\varepsilon q} - 2if\, e^{i\varepsilon q} - 1 = 0$

Let $e^{i\varepsilon q} = x$

$x^2 - 2ifx - 1 = 0 \Longrightarrow$

$$x = \frac{2if \pm \sqrt{-4f^2 + 4}}{2}$$

$$x = if \pm \sqrt{1 - f^2} \Longrightarrow e^{i\varepsilon q} = if \pm \sqrt{1 - f^2}$$

$$i\varepsilon q = In\left[if \pm \sqrt{1 - f^2} \right] \Longrightarrow q = \frac{In\left[if \pm \sqrt{1 - f^2} \right]}{i\varepsilon}$$

and using the expression for $\rho_i = f'_{i-1} / f'_i$

$$\rho_i = \frac{\cos(\varepsilon q_{i-1})\varepsilon}{\cos(\varepsilon q_i)\varepsilon}$$

$$\rho_i = \frac{e^{i\in\left[1n[i f_{i-1} \pm \sqrt{1 - f_{i-1}^2}]/i\in \right]} + e^{-i\in\left[1n[i f_{i-1} \pm \sqrt{1 - f_{i-1}^2}]/i\in \right]}}{e^{i\in\left[\ln[i f_i + \sqrt{1 - f_i^2}]/i\in \right]} + e^{-i\in\left[\ln[i f_i + \sqrt{1 - f_i^2}]/i\in \right]}}$$

From the above equation we have

$$\frac{\left[if_{i-1} + \sqrt{1 - f_{i-1}^2} \right] + \dfrac{1}{\left[if_{i-1} + \sqrt{1 - f_{i-1}^2} \right]}}{\left[if_i + \sqrt{1 - f_i^2} \right] + \dfrac{1}{\left[if_i + \sqrt{1 - f_i^2} \right]}} =$$

This is our result

$$\rho_i = \left[\frac{\dfrac{\left[if_{i-1} + \sqrt{1 - f_{i-1}^2} \right]^2 + 1}{if_{i-1} + \sqrt{1 - f_{i-1}^2}}}{\dfrac{\left[if_i + \sqrt{1 - f_i^2} \right]^2 + 1}{if_i + \sqrt{1 - f_i^2}}} \right]$$

3. The Numerical Experiments:

In order to test the effectiveness of the new algorithm that has been used to extend the CG method, a number of functions have been chosen and solved numerically by utilizing the new and established method.

The same line search was employed for all the methods. This was the cubic interpolation procedure described in Bunday[6].

It is found that the NEW method, which modifies CG-algorithm, is better than the previous algorithm shown in Tables (1) and (2).

Table (1) which uses the H/S formula, presents a comparison between the results of the NEW methods and the classical CG-method. So we can show that the NEW method has less (NOI) and (NOF) than the classical CG. Method and that NEW method improves the two measures of performance, vis (NOI) and (NOF) (66)% and the (64) % for the H/S formula.

Table (1): The comparison between the different ECG – methods by using H/S formula .

Test Function	N	New NOI (NOF)	Classical CG NOI (NOF)
CUBIC	2	18 (51)	19 (53)
	100	13 (37)	14 (40)
	200	13 (36)	14 (40)
	400	14 (40)	14 (40)
POWELL	4	59 (148)	65 (170)
	100	126 (271)	105 (275)

	200	126 (261)	202 (462)
	400	57 (129)	401 (860)
WOOD	10	28 (76)	35 (77)
	40	48 (107)	59 (126)
	200	103 (216)	107 (221)
CANTRAL	4	22 (137)	25 (148)
	40	21 (112)	20 (132)
	100	21 (112)	20 (132)
	200	21 (113)	20 (132)
ROSEN	2	28 (74)	34 (87)
	20	18 (48)	17 (52)
	100	18 (51)	17 (52)
Non Diagonal	10	22 (59)	25 (67)
	40	24 (71)	22 (73)
Total	NOI (NOF)	800 (2140)	1235 (3239)

Table (2), which uses the P/R formula, presents a comparison between the results of the NEW methods and the classical CG-method. So we can show that the NEW method has less (NOI) and (NOF) than the classical CG. Method and the NEW method improves the two measures of performance, vis (NOI) and (NOF) by (53)% and the (69) % for the P/R formula.

Table (2): The comparison between the different
ECG – methods by using P/R formula.

Test Function	N	New NOI (NOF)	Classical CG NOI (NOF)
CUBIC	2	18 (52)	19 (53)
	4	14 (42)	16 (42)
	40	13 (33)	15 (40)

ROSEN	2	28 (72)	33 (85)
	100	18 (52)	22 (61)
	200	19 (56)	22 (61)
	400	18 (52)	22 (61)
POWELL	60	65 (157)	84 (186)
	200	105 (253)	205 (427)
	400	90 (218)	405 (826)
WOOD	40	50 (110)	68 (144)
	100	101 (212)	103 (213)
	200	97 (204)	107 (221)
	400	52 (114)	108 (223)
Non Diagonal	10	20 (53)	23 (63)
	20	22 (65)	18 (53)
	200	20 (57)	25 (68)
	400	19 (56)	24 (68)
CANTRAL	200	21 (115)	19 (115)
	400	20 (132)	22 (157)
Total	NOI (NOF)	812 (2005)	1382 (2879)

Appendix

1. Cubic Function :

$$\mathbf{F(x)} = 100\left(\mathbf{x}_2 - \mathbf{x}_1^3\right)^2 + \left(1 - \mathbf{x}_1\right)^2, \qquad \mathbf{x_0} = (-1.2,-1.)^{\mathbf{T}}$$

2. Non – Diagonal Variant of Rosenbrock Function :

$$\mathbf{F(x)} = \sum_{\mathbf{i}=2}^{\mathbf{n}} \left[100\left(\mathbf{x_i} - \mathbf{x_i}^2\right)^2 + \left(1 - \mathbf{x_i}\right)^2\right], \qquad \mathbf{n > 1,}$$

3. Wood Function

$$F(x) = \sum_{i=1}^{n/4} 100\Big[\left(x_{4i-2} + x_{4i-3}^2\right)^2 + \left(1 - x_{4i-3}\right)^2 + 90\left(x_{4i} - x_{4i-1}^2\right)^2$$
$$+ \left(1 - x_{4i-1}\right)^2 + 10.1\left(x_{4i-2} - 1\right)^2 + \left(x_{4i} - 1\right)^2 + 19.8\left(x_{4i-2} - 1\right)\left(x_{4i} - 1\right)\Big]$$

$x_0 = (-3.0;1.0;-3.0;-1.0;\ldots\ldots\ldots)^{T}$

4. Generalized Powell Quartics Functions :

$$\mathbf{F(x)} = \sum_{\mathbf{i}=1}^{\mathbf{n}/4} \left[\left(\mathbf{x}_{4\mathbf{i}-3} + 10\mathbf{x}_{4\mathbf{i}-2}\right)^2 + 5\left(\mathbf{x}_{4\mathbf{i}-1} - \mathbf{x}_{4\mathbf{i}}\right)^2 + \left(\mathbf{x}_{4\mathbf{i}-2} - 2\mathbf{x}_{4\mathbf{i}-1}\right)^4 + 10\left(\mathbf{x}_{4\mathbf{i}-3} - \mathbf{x}_{4\mathbf{i}}\right)^4\right]$$

$\mathbf{x}_0 = \mathbf{(3.0;-1.0;0.0;1.0)}^{\mathbf{T}}$

5. Rosenbrock Function :

$$\mathbf{F(x)} = \sum_{\mathbf{i}=1}^{\mathbf{n}/2} \left[100\left(\mathbf{x}_{2\mathbf{i}} - \mathbf{x}_{2\mathbf{i}-1}^2\right)^2 + \left(1 - \mathbf{x}_{2\mathbf{i}-1}\right)^2\right]$$

$\mathbf{x}_0 = (-1.2;1.0;\ldots\ldots..)^{\mathbf{T}}$

6. Miele Function :

$$F(x) = \sum_{i=1}^{n/4} \left[\exp\left(x_{4i-3}\right) - x_{4i-2}\right]^2 + 100\left(x_{4i-2} - x_{4i-1}\right)^6 +$$

$$\left[\tan\left(x_{4i-1} - x_{4i}\right)\right]^4 + x_{4i-3}^8 + \left(x_{4i} - 1\right)^2,$$

$$x_0 = (1.0;\, 2.0; 2..0; 2.0, \ldots\ldots)^T$$

7. Cantral Function :

$$F(x) = \sum_{i=1}^{n/4} \left[\exp\left(x_{4i-3}\right) - x_{4i-2}\right]^4 + 100\left(x_{4i-2} - x_{4i-1}\right)^6 +$$

$$\left[a\tan\left(x_{4i-1} - x_{4i}\right)\right]^4 + x_{4i-3}^8 .$$

$$x_0 = (1.0;\, 2.0; 2..0; 2.0, \ldots\ldots)^T$$

REFERECES

[1] Al – Bayati, A.Y; (1993), New Non – Quadratic Model Unconstrained Non- linear optimization method", Natural and Applied Series. Mutah Journal for Research and Studies. Ma'th University, Jordan Vol 8, No.1, PP 133-155.

[2] Al-Assady N.H.; Huda, k.M; (1997), "A rational logarithmic model for Unconstrained Nonlinar Optimization. Raf. Journal Sci, Vol 8, No 2, PP 107-117; Mosul University. Iraq.

[3] Al- Assady, N.H. (1991), "new QN and PCG algorithms Based on Non, Quadratic Properties, J. Ed- and Sc., 12.

[4] Al- Bayati, A.Y., (1995), "New Extended CG. Methods for Non Linear Optimization, "Natural and Applied Series, Mu'th Journal for search and studies. Ma'th university, Jordan Vol. 10, No. 6 PP 69 –87.

[5] Boland, W. R and kowalik, J.S . (1979), " Extended Conjugate Gradient method with Restarts". Journal of Opt. Theory and Application, 28, PP. 1-9.

[6] Bunday, b., (1984), "Basic Optimization methods", Edward Arnold, Bedford square, London.

[7] Dixon, L.C.W., (1972), "Nonlinear Optimization", English University press, London.

[8] Flecher, R. And Reeves, C.M., (1964), "Function Minimization by conjugate Gradient" Computer journal, 7, PP. 149-154.

[9] Fried, I., (1971), "N-step Conjugate Gradient Minimization Scheme for Non – Quadratic Function,", AIAA Journal, 19, PP. 2286-2287.

[10] Hestences, M.R. and stiefel, E, (1952), "Mechods of conjugate Gradient for solving Linear systems", j. Res .B.S., 49.

[11] Polak, E. and Ribiere, G., (1969). Note Surla Convergence des methods de directions conjugate,; Rev. Fr in fr, Rech Oper, 16, R1.

[12] Spedicato, E, (1976), "A variable Metric Method for Function Minimization Perived from Invariance and App 1., 20.

[13] Tassopouls, A. and Storey, c., (1984), "Use of Non, Quadrate Model in A conjugate Gradient Model of Optimization with inexact line search", Jota, 43.