Iraqi Journal of Statistical Science (8) 2005
P.P. [1-14]

COMBINED NON-QUADRATIC MODELS AS A BASIS FOR CG-A LGORITHMS

Abbas Y. AL-Bayati* **Abbas H. Taqi****

ABSTRACT

In this paper,a new extended (CG) algorithms is proposed. It is in fact, a particular type of the Conjugate Gradient (CG) method which employs non-quadratic rational model, and based on inexact line searches. The Flecther and Reeves restarting criterion was employed to the standard and New versions and gave dramatic savings in computational time. The new algorithms is were promising in general, seven non linear tests function with different versions were used.

ربط نماذج غير تربيعية كأساس الخوارزمية في التدرج المترافق

الملخص

في هذا البحث تم استحداث خوارزمية جديدة موسعة في مجال التدرج المترافق لربط نماذج غير تربيعية .والنموذج الجديد نموذج نسبي غير تربيعي معتمد على خطوط بحث غير تامة . ولقد تم استعمال وسيلة استرجاع Fletcher and Reeves في الصيغتين القديمة والجديدة وكان لذلك اثر كبير في توفير الزمن المطلوب للحل. اثبتت الحسابات العددية ا كفاءة الخوارزمية الجديدة مقارنة بالأخرى وباستعمال دوال غير خطية وبأبعاد مختلفة .

*Department of Mathematics/ College of Science/ Mosul University
**Department of Mathematics/College of Science/Kirkuk University

Received: 3/ 9 /2004 Accepted: 27/ 11 / 2004

1.Introduction:

Conjugate gradient (CG) algorithms form a class of CG-algorithms for minimizing a general differentiable function f (x) $x \in R^n$. whose gradient g (x) can be calculated , are based on the following concept of conjugacy:

If Q is a positive definite symmetric n×n matrix , then the directions $d_1, d_2, \ldots\ldots\ldots d_n$ where $d_k \neq 0$ for k=1,2,……n, are mutually Q-conjugate if

$$d_i^T Q d_k = 0 \qquad for \;\; i \neq k$$

The classical algorithm in this category proposed by Flectcher and Reeves (9) and is based on the iterative scheme:

$$x_{k+1} = x_k + \lambda_k d_k, \;\; k = 1,2,\ldots\ldots\ldots\ldots, n$$

Where the scalar λ_k is the smallest positive local minimizer of the One dimensional problem

$$\min_\lambda f(x_k + \lambda d_k)$$

is a search direction generated by the equations : d_k and

$$\begin{aligned} & for \;\; k = 1,2,\ldots\ldots\ldots\ldots\ldots\ldots, n \\ & d_1 = -g_1, \\ & d_{k+1} = -g_{k+1} + \beta_k d_k \;, \\ & \beta_k = g_{k+1}^T y_k / d_k^T y_k \end{aligned}$$

If the method is applied to the quadratic function

$$q(x) = \frac{1}{2} x^T Q x$$

2. Extended Conjugate Gradient (ECG) method:

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( ).

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 minimzer of q(x), then it is a minimizer of f.

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 minimizes the function **$f(x) = (q(x))^{\rho}$** , and $x \in R^n$ in at most step has been described by Fried [10].

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 &Kowalik [7].

Another model has developed by Tassopoulos and Storey ,[14] as follows:

$$\mathbf{F(q(x) = \varepsilon_1 q(x) + 1/\varepsilon_2 q(x):}\ \varepsilon_2 > \mathbf{0}$$

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

Al-Bayati ,[1] has been developed a new rational models which is defined as follows: **$F(q(x)) = \varepsilon_1\ q(x)/1-\varepsilon_2\ q(x),\ \varepsilon_2 < 0$.**

Also Al-Bayati ,[2] developed an extended CG algorithm which is based on a general logarithmic model

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

Al-Assady &Huda,[5] described there ECG algorithm which is based on the natural log function for the rational q(x) function

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

3.CONJUGATE GRADIENT METHOD WITW INEXACT LINE SEARCH:

In order to improve the local rate of convergence and the efficiency of the traditional CG-method several well-Know methods are discussed in this area . A mong these methods Sloboda (13) defines a new generalized CG.a lgorithm for minimizing a strictly convex function of the general form
f(x) = F(q (x))…………(2)
We now list out-lines of sloboda extended CG. method :
Algorithm (sloboda 1980)

Step (1): set k=1; $\overline{g_k} = g_k$ and $d_k = -\overline{g_k}$

Step (2): compute λ_k by exact line search and $x_{k+1} = x_k + \lambda_k d_k$

Step (3): compute $g^{+}_{k+1/2} = g(x_k + \lambda_k d_k / 2)$

Step (4): Test for convergence if achieved stop. If not continue

Step (5): If k=0 mod (n) go to step (1) else continue

Step (6): compute $\overline{g_{k+1}} = w_k g^{+}_{k+1} - \overline{g_{k+1}}$ where

$w_k = d_k^T \overline{g_k} / d_k^T g^{+}_{k+1/2}$

If $\overline{g_{k+1}} = 0$ set i=i+1 go to step (1)

Step (7): $d_{k+1} = -\overline{g_{k+1}} + \beta_k d_k$; $\beta = (\overline{g_{k+1}} - \overline{g_k})\overline{g_{k+1}} / d^T (\overline{g_{k+1}} - g_k)$

Step (8): set k=k+1 and go to step (2)

Step (9): If k. EQ .n go to Step (1). If not continue.

4.The Derivation p_i :

The implementation of the extended CG method has been performed for general function F(q(x) of the form of equation(2). The unknown quantities P_i were expressed in terms of available quantities of the algorithm .The authors , introduced in [4] a new model , which can be written as

The ne $\sin\left(\frac{\varepsilon_1 q(x)+1}{\varepsilon_2 q(x)}\right)$ model can now be written as

$$\mathbf{f(x) = F(q(x)} \sin\left(\frac{\varepsilon_1 q(x)+1}{\varepsilon_2 q(x)}\right)$$

Solving equation (2) for q

$$\mathrm{Sin}^{-1} f(x \left(\frac{\varepsilon_1 q(x)+1}{\varepsilon_2 q(x)}\right)$$

$$\ln\left[if(x)+\sqrt{1-f(x)^2}\right] = \frac{\varepsilon_1 q(x)+1}{\varepsilon_2 q(x)} \qquad {}^{1}\!/_{\varepsilon_2} \ln\left[if(x)+\sqrt{1-f(x)^2}\right]-\varepsilon_1$$

And using the expression for $\mathbf{p_i = f'_{i-1}/f'_i}$

$$\rho_i = -\frac{\cos(\varepsilon_1 q_{i-1} + 1/\varepsilon_2 q_{i-1})\left(-1\!/_{\varepsilon_2 q^2_{i-1}}\right)}{\cos(\varepsilon_1 q_i + 1/\varepsilon_2 q_i)\left(-1\!/_{\varepsilon_2 q^2_i}\right)}.$$

From the above equation we have

$$\rho_i = \left[\frac{\dfrac{\left[\left[if_{i-1} + \sqrt{1 - f_{i-1}^2}\right]^2 + 1\right]\left[\ln(if_{i-1} + \sqrt{1 - f_{i-1}^2}) - \dfrac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_{i-1} + \sqrt{1 - f_{i-1}^2}}}{\dfrac{\left[\left[if_i + \sqrt{1 - f_i^2}\right]^2 + 1\right]\left[\ln(if_{i-1} + \sqrt{1 - f_i^2}) - \dfrac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_i + \sqrt{1 - f_i^2}}} \right]$$

In terms of the known quantities such a function and gradient values, from

$g_i = F_i^{'} Q(x_i - x^*)$

$g_{i-1} = F_{i-1}^{'} Q(x_{i-1} - x^*)$

Where Q is the Hessian Matrix and x^* is the minimum point , we have :

$$\rho_i = \left[\frac{\dfrac{\left[\left[if_{i-1} + \sqrt{1 - f_{i-1}^2}\right]^2 + 1\right]\left[\ln\left(if_{i-1} + \sqrt{1 - f_{i-1}^2}\right) - \dfrac{\varepsilon_1}{\varepsilon_2}\right.}{if_{i-1} + \sqrt{1 - f_{i-1}^2}}}{\dfrac{\left[\left[if_i + \sqrt{1 - f_i^2}\right]^2 + 1\right]\left[\ln\left(if_i + \sqrt{1 - f_i^2}\right) - \dfrac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_i + \sqrt{1 - f_i^2}}} \right.$$

Furthermore

$$g_{i-1}^T (x_i - x^*) = g_{i-1}^T (x_{i-1} + \lambda_{i-1} d_{i-1} - x^*)$$
$$= g_{i-1}^T (x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1} \text{.........................(3)}$$

$$g_i^T (x_i - x^*) = g_i^T (x_i + \lambda_i d_i - x^*)$$
$$= g_i^T (x_i - x^*)$$

Since $g_i^T d_{i-1} = 0$ Therefore, we can express ρ_i us follows: -

$$\rho_i = \frac{g_{i-1}^T (x_{i-1} + \lambda_{i-1} d_{i-1} - x^*)}{g_i^T (x - x*)}$$

From (3) and (4) , it follows that : ……..(4)

$$\rho_i = \rho_i\left[\frac{q_{i-1}}{q_i}\right] + \lambda_{i-1} g_{i-1}^T d_{i-1} / 2F_i^{'} q_i$$

Where q $and\ f^{'} = \dfrac{\left[\left[if + \sqrt{1-f^2}\right]^2 + 1\right] - \varepsilon_2\left[\ln\left(if + \sqrt{1-f^2}\right) - \frac{\varepsilon_1}{\varepsilon_2}\right]^2}{2\left[if + \sqrt{1-f^2}\right]}$

The quantities q_{i-1}/q_i And $f_i^{'} q_i$ can be rewritten as :

$$\frac{q_{i-1}}{q_i} = \frac{\ln\left[if_i + \sqrt{1-f_i^2}\right] - \frac{\varepsilon_1}{\varepsilon_2}}{\ln\left[if_{i-1} + \sqrt{1-f_{i-1}^2}\right] - \frac{\varepsilon_1}{\varepsilon_2}}$$

$$f_i^{'} q_i = \frac{\left[\left[if_i + \sqrt{1-f_i^2}\right]^2 + 1\right]\left[\ln\left(if_i + \sqrt{1-f_i^2}\right) - \frac{\varepsilon_1}{\varepsilon_2}\right]}{2\left[if_i + \sqrt{1-f_i^2}\right]}$$

From the definition of ρ_i we have :

$$\left[\frac{\dfrac{\left[\left[if_{i-1} + \sqrt{1-f_{i-1}^2}\right]^2 + 1\right]\left[\ln\left(if_{i-1} + \sqrt{1-f_{i-1}^2}\right) - \frac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_{i-1} + \sqrt{1-f_{i-1}^2}}}{\dfrac{\left[\left[if_i + \sqrt{1-f_i^2}\right]^2 + 1\right]\left[\ln\left(if_i + \sqrt{1-f_i^2}\right) - \frac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_i + \sqrt{1-f_i^2}}}\right] =$$

$$\left[\frac{\dfrac{\left[\left[i f_{i-1}+\sqrt{1-f_{i-1}^{2}}\right]^{2}+1\right]\left[\ln\left(i f_{i-1}+\sqrt{1-f_{i-1}^{2}}\right)-\dfrac{\varepsilon_1}{\varepsilon_2}\right]}{if_{i-1}+\sqrt{1-f_{i-1}^{2}}}}{\dfrac{\left[\left[i f_{i}+\sqrt{1-f_{i}^{2}}\right]^{2}+1\right]\left[\ln\left(i f_{i}+\sqrt{1-f_{i}^{2}}\right)-\dfrac{\varepsilon_1}{\varepsilon_2}\right]^{2}}{if_{i}+\sqrt{1-f_{i}^{2}}}}\right]-\frac{(\lambda_{i-1}g_{i-1}^{T}d_{i-1}}{\dfrac{\left[\left[i f_{i-1}+\sqrt{1-f_{i-1}^{2}}\right]^{2}+1\right]\left[\ln\left(i f_{i-1}+\sqrt{1-f_{i-1}^{2}}\right)-\dfrac{\varepsilon_1}{\varepsilon_2}\right]}{if_{i-1}+\sqrt{1-f_{i-1}^{2}}}}$$

using the following transformation:

$$\frac{\left[if_i+\sqrt{1-f_i^2}\right]^2+1}{if_i+\sqrt{1-f_i^2}}=x, \qquad \ln\left[\mathrm{if}_{\mathrm{i-1}}+\sqrt{1-f_{i-1}^2}\right]-\frac{\varepsilon_1}{\varepsilon_2}=y$$

$$\ln\left[\mathrm{if}_{\mathrm{i}}+\sqrt{1-f^2}\right]-\frac{\varepsilon_1}{\varepsilon_2}=y+w \quad \text{and} \quad \ln\left[\mathrm{if}_{\mathrm{i}}+\sqrt{1-f_i^2}\right]-\ln\left[if_{i-1}+\sqrt{1-f_{i-1}^2}\right]=w$$

$$c=\lambda_{i-1}g_{i-1}^{T}d_{i-1}$$

Then y=cw/xw+c

Therefore

$$\overline{-f_{i-1}^2}\Bigg]-\frac{\left[\ln\left(if_i+\sqrt{1-f_i^2}\right)\right]-\ln\left[if_i+\sqrt{1-f_i^2}\right].\left[-\lambda_{i-1}g_{i-1}d_{i-1}\right]}{\dfrac{\left[if_{i-1}+\sqrt{1-f_{i-1}^2}\right]^2+1}{\left[if_{i-1}+\sqrt{1-f_{i-1}^2}\right]}.\left[\ln\left(if_{i-1}+\sqrt{1-f_{i-1}^2}\right)-\ln\left(if_{i-1}+\sqrt{1-f_{i-1}^2}\right)\right]+\lambda_{i-1}g_{i-1}d_{i-1}}$$

5.New modification for the Sloboda method :

In order to improve the global rate of convergence of minimization algorithms when applied to more general functions than the quadratic variable metric (**VM**) matrices which may be used to accelerate the CG-algorithm (see for example Al-Byati and Al-assady (6)). In this section, a new expression for the new search direction d_{k+1} of the **Sloboda method** is suggested which is invariant to non-linear scaling of non-quadratic function. Al-assady and Hassan (4) (to appear) are used to extend the **sloboda method.**

In use the new suggested algorithms require less vector storage's than the sloboda algorithms.

We now give the outline of the new proposed modifications:

New Algorithm:

Step (1): set k=1; $\overline{g}_k = g_k$ and $d_k = -\overline{g_k}$

Step (2): compute λ_k by ELS and $x_{k+1} = x_k + \lambda_k d_k$

Step (3): compute $g^{+}_{k+1} = g(x_k + \lambda_k d_k / 2)$

Step (4): Test for convergence if achieved stop.

If not continue

Step (5): If k=0 mod (n) go to step (1) else continue

Step (6): compute $\overline{g_{k+1}} = w_k g^{+}_{k+1} - \overline{g_{k+1}}$ where

$$w_k = d_k^T \overline{g_k} / d_k^T g^{+}_{k+1/2}$$

If $\overline{g_{k+1}} = 0$ set I=I+1 go to step (1)

Step (7)compute

$$\rho_i = \left[\frac{\dfrac{\left[\left[if_{i-1} + \sqrt{1-f_{i-1}^2}\right]^2 + 1\right]\left[\ln\left(if_{i-1} + \sqrt{1-f_{i-1}^2}\right) - \dfrac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_{i-1} + \sqrt{1-f_{i-1}^2}}}{\dfrac{\left[\left[if_i + \sqrt{1-f_i^2}\right]^2 + 1\right]\left[\ln\left(if_i + \sqrt{1-f_i^2}\right) - \dfrac{\varepsilon_1}{\varepsilon_2}\right]^2}{if_i + \sqrt{1-f_i^2}}} \right]$$

Step (8): $d_{k+1} = -\overline{g_{k+1}} + \beta_k d_k$; $\beta = (\overline{\rho_k g_{k+1}} - \overline{g_k})\overline{g_{k+1}} / d^T (\rho_k \overline{g_{k+1}} - g_k)$

Step (9): set k=k+1 and go to step (2)

Step (10): If k. EQ .N go to Step (1)

6. Numerical Results and conclusion :

In order to test the effectiveness of the new algorithms that have been used to extent the Sloboda method, a number of function 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(8).

It is found that the NEW method which modifies Sloboda algorithm is better than the previous algorithm shown in table.

Table (1): Which uses the H/S formula, presents a comparison between the results of the NEW methods and the Sloboda method. So we can show that the NEW method has less (NOI) and (NOF) than the classical Sloboda method. The NEW method improves the two measures of performances, vis (NOI) and (NOF) (67.17)% and the (80.25) % for the H/S formula.

Table: Comparison between the different ECG – methods by using H/S formula.

Test Function	N	New NOI (NOF)	Sloboda NOI (NOF)
SHALLO	2	7 (20)	9 (23)
	4	7 (21)	8 (21)
	10	10 (30)	8 (20)
DIXON	2	6 (13)	6 (17)
	4	10 (25)	12 (26)
	10	18 (40)	19 (41)
POWELL	20	49 (126)	60 (158)
	100	120 (270)	112 (384)
	200	178 (410)	119 (251)
WOOD	100	102 (210)	205 (423)
	200	150 (310)	402 (817)
	400	105 (214)	103 (213)
NON-DIAGONL	4	20 (60)	23 (63)
	40	21 (61)	21 (61)
	400	16 (45)	23 (67)
CANTRAL	100	35 (230)	43 (347)
	200	47 (270)	46 (393)
	400	42 (180)	47 (410)
WOLFE	4	11(26)	12(27)
	20	30(71)	35(71)
Total	NOI (NOF)	837 (2560)	1246 (3190)

Appendix

1.Generalized Powell Functions:

$$F(x)=\sum_{i=1}^{n/4}(x_{4i-9}+10x_{4i-2})^2+5(x_{4i-1}-x_{4i})^2+(x_{4i-2}-2x_{4i-1})^4+10(x_{4i-9}-x_{4i})^4$$

$x_0=(3,1,0,1)^T$

2. Generalized Cantreal Functions:

$$F(x)=\sum_{i=1}^{n/4}\left[\exp(x_{4i-3})-x_{4i-2}\right]^2+100(x_{4i-2}-x_{4i-1})^6+\left[(a\tan(x_{4i-1}-x_{4i}))\right]^4+x_{4i-3}^8$$

$x_0=(1,2,2,2)^T$ 3.Wo

od Functions:

$$F(x)=\sum_{i=1}^{n/4}100(x_{4i-2}+x_{4i-3}^2)^2+(1-x_{4i-3})^2+90(x_{4i}-x_{4i-1}^2)^2+(1-x_{4i-1})^2+1.0$$

$x_0=(-3,-1,-3,-1)^T$ 4

.Non-Diagonal Functions:

$$F(x)=\sum_{i=2}^{n}100(x_i-x_0^2)^2+(1-x_i)^2$$

$x_0=(-1,..........)^T$

5.Dixon Functions:

$$F(x)=(1-x_1)^2+(1-x_0)^2+\sum_{i=2}^{9}(x_i-x_{i-1})$$

$x_0=(-1,............)^T$

6.Wolfe Functions:

$$F(x)=(-x_1(3-x_1/2)+2x_2-1)^2+\sum_{i=1}^{n-1}(x_{i-1}-x_i(3-x_i/2)+2x_{i+1}-1)^2+(x_{n-1}-x_n(3x_n/2)-1)^2$$

$x_0=(-1,.............)^T$

7.Shallo Functions:

$$F(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2-x_{2i})^2+(1-x_{2i-1})^2$$

$x_0=(-2,-2,.........)^T$

REFRENCES

[1] **AL-Bayati: A.Y. (1993),** "new non. Quadratic model unconstrained linear optimization method ", natural and applied Series .Mutah journal for research and studies .ma"th university .jordan vo1 8,no1., pp. 133-155 .

[2] **AL Bayati, A.Y. (1995),"**new extendedcg methods for non linear optiization ", nataral and applied series, mu"th journal for search and studies .ma"th university, Jordan vol10, no.6, pp69 –87.

[3] **AL -Assady, N. H. (1991),**" new qn &pcg al gorithms based on non, quadratic properties. j. ed –and sc . , 12.

[4] A**L-Bayati, A. Y. &Basim A. H. to be apper**

[5] **AL – Assady, N. H. & Huda K. M . (1997),"** arational logarithmic model for unconstrained no linear Optimization .raf . Journal sci , vol 8 ,no 2 , pp 107 –117 mosul university . iraq .

[6] **AL - Bayati, A. Y. &AL -Assady, N. H. (1997), "** pcg methods for non linear optimization ", mu"th Journal for research and studies, natural and applied sc. Iences seines jourdan, vol .12, pp: 13-29.

[7] **Boland .W. R. & Kowalik. J.S. (1979).** "Extended conjugate gradient method with restarts ". Journal of op.t.. theory and application , 28. PP: 1-9

[8] **Bunday .B. (1984), "**BASIC OPTIMIZATION METHOD S ", Edward Arnold, Bedford square, London.

[9] **Flecher. R. & Reeves, C.M. (1964) ,**" function minimization by conjugate gradient " , computer Journal ,7,pp:149 -154 .

[10] **Fried .i. (1971),** " n-step conjugate gradient mini mization scheme for non – quadratic function " , aiaa journal ,19 ,PP: 2286 –2287 .

[11] **Hestences, M. R. & Stiefel, E. (1952),** " methods of conjugate gradient for solving linear systems ", j. Res. b. s. , 49

[12] **Spedicato ,E. (1976) ,** **”** avariable metric method for function minimization perived from invariance & app 1 . , 20

[13] **Sloboda (1980), “**Ageneralized Conjugate gradint algorithm for minimization” Numer Math.,35.

[14] **Tassopouls, A. & Storey, C. (1984),** ” vse of non quadratic model in a conjugate gradient model of opt. with in exact line search “ ,jota ,43