A New Non Quadratic Algorithm for Solving Non-Linear Optimization Problems

Adham A. Ali
College of Sciences –University of Kirkuk
Accepted: 2011/6/19, Received: 2010/9/1

Abstract

This paper proposes a new algorithm for non-linear optimization to modify and develop the conjugate gradient (CG) methods and to obtain a strong global convergence. This algorithm is derived and evaluated numerically against the standard (P/R and H/S)-CG algorithms and T/S algorithm using more than (20) standard well-known test functions. The numerical results show that, Non –quadratic models are very beneficial in most of the problems especially when the dimensionality of the problem increases.

Introduction

Conjugate gradient methods (CG) were proposed by Hestenses and Stiefel (Hestences & Stiefle, 1952) for solving systems of linear equations. The use of this method for unconstrained optimization was prompted by the fact that the minimization of a positive-define quadratic function is equivalent to solving the linear equation system that results when its gradient is set at zero. Conjugate gradient methods as applied to quadratic functions are described first. Actually, the extension of conjugate gradient methods for solving non-linear equation systems and its use in solving general unconstrained minimization problems was first done by Fletcher and Reeves (Fletcher & Revees, 1964). We will show these methods can be extended to minimize general non-linear functions.

The conjugate gradient method have in general the following basic properties (Dragica Vasileska, 2006):

1) The conjugacy condition.
2) The orthogonally condition
3) The descent direction
4) The quadratic termination condition with exact line search (ELS).

Concept of the extended CG-methods (ECG)

A function f is defined as a non-linear scaling of the quadratic function q(x) if the following condition holds:

$$f = F(q(x)), \frac{dF}{dq} = \bar{F} > 0 \text{ and } q(x) > 0 \qquad \ldots (1)$$

This property is called invariancy to non-linear scaling (Spedicato, 1976).

The following properties are immediately derived from the above condition:

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

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

Various authors have published related work in this area:

i) A CG methods which minimizes the general function

$$f(x) = ((q(x))^{\rho}, \rho > 0 \;\; x \in R^n \qquad \text{. . . (2)}$$

In at most n steps has been proposed in (Fried, 1971).

ii) Two CG methods which minimize the following polynomial model

$$f(x) = \varepsilon_1 q(x) + \frac{1}{2}\varepsilon_2 q^2(x) \qquad \text{. . . (3)}$$

where ε_1 and ε_2 are scalars, have been investigated with two different restarting critieriain in (Boland & Kowalik , 1977a) and (Boland & Kowalik , 1979b) .

iii) Also two different rational models have been developed in (Tassopoulos and Storey, 1984a) and (Tassopoulos & Storey, 1984b) namely:

$$f(x) = \frac{\varepsilon_1 q(x) + 1}{\varepsilon_2 q(x)}, \varepsilon_2 > 0 \qquad \text{. . . (4)}$$

and

$$f(x) = \frac{\varepsilon q(x)}{1 + q(x)}, \varepsilon > 0 \qquad \text{. . . (5)}$$

iv) Another new CG method which based on general logarithmic model

$$f(x) = \varepsilon(\log q(x) - 1), \varepsilon > 0 \qquad \text{. . . (6)}$$

have been implemented by (Al-Bayati , 1995).

v) And (AI-Assady and Al-Ta'ai, 2002) described their algorithm which based on the trigonometric function:

$$F(q(x)) = \sin(\varepsilon q(x)) \qquad \text{… (7)}$$

where ε is scalar.

vi) Also (Al-Mashhadany, 2002) has been developed a new rational model which is defined as following:

$$F(q(x)) = \exp(\frac{\varepsilon_1 q(x)}{\varepsilon_2 q(x) - 1}); \; \varepsilon_2 < 0 \qquad \text{. . . (8)}$$

vii) Finally, another specific rational model was considered by (Taqi & Adham, 2009) which is defined as following:

$$f(x) = \cos^{-1}(\frac{1 - \varepsilon_1 q(x)}{\varepsilon_2 q(x)}) \qquad \text{. . . (9)}$$

where ε_1 and ε_2 are scalars.

We have suggested in this paper a new algorithm based on the non-quadratic model to use a successful convenient technique for solving unconstrained optimization problems which develops the classical CG method. This new model is as follow:

$$F(q(x)) = \coth^{-1}(\varepsilon q(x)) \qquad \text{. . . (10)}$$

Where $q(x) = \frac{1}{2}x^T Gx + x^T b + c$ is quadratic function and ε is a scalar. It is assumed that:

$$\frac{dF}{dq} = f' > 0 \text{ for } q > 0 \qquad \text{. . . (11)}$$

holds

However, we first observe that q(x) and f(x) given in the above new model have identical contours through with different function values, and they have the same unique minimum point x^*.

The derivation of new Algorithm

The implementation of the extended conjugate gradient method has been performed for general function $F(q(x))$ of the form of equation (10)

The new model is

$$F(q(x)) = \coth^{-1}(\varepsilon q(x))$$

Take the inverse function relationship we get

$$q = \frac{1}{\varepsilon}\coth f \quad = \frac{1}{\varepsilon}\left(\frac{e^f + e^{-f}}{e^f - e^{-f}}\right) \qquad \text{… (12)}$$

The unknown quantities ρ_k was expressed in term of available quantities of the algorithm (i.e. function and the gradient values of the objective function).

Since ρ_k is a parameter, which is defined as

$$\rho_k = \frac{f'_{k-1}}{f'_k} \qquad \text{… (13)}$$

Where, $f'_k = \frac{-\varepsilon(e^{f_k} - e^{-f_k})^2}{4}$ and $f'_{k-1} = \frac{-\varepsilon(e^{f_{k-1}} - e^{-f_{k-1}})^2}{4}$

Hence

$$\rho_k^{New} = \left(\frac{e^{f_{k-1}} - e^{-f_{k-1}}}{e^{f_k} - e^{-f_k}}\right)^2 \qquad \text{… (14)}$$

The outline of the new algorithm

Given $x_0 \in R^n$ the initial point, and scalar ε.

Step (0): Set $d_\circ = -g_\circ$

Step (1): For k = 1,2,…

Compute $x_k = x_{k-1} + \lambda_{k-1} d_{k-1}$

where λ_{k-1} is the minimizer of f on d_{k-1}.

Step (2): Check for convergence

If $\|g_k\| \le \varepsilon$, then stop, Otherwise continue.

Step (3): Calculate

$$Exp(f) = f + f' + \frac{f^2}{2!} + \frac{f^3}{3!} + \ldots$$

Step (4): Compute

$$\rho_k^{New} = \left[\frac{e^{f_{k-1}} - e^{-f_{k-1}}}{e^{f_k} - e^{-f_k}} \right]^2$$

Step (5): Check if $0 \le \rho_k \le 1$, then go to step (6). Otherwise set $\rho_k = 1$ and go to step(6).

Step (6): Calculate the new direction

$$d_k = -g_k + \beta_k d_{k-1}$$

where β_k is defined as follows:

$$\beta_k = \frac{g_k^T(\rho_k g_k - g_{k-1})}{[d_{k-1}^T(\rho_k g_k - g_{k-1})]}$$ modified H/S in (Hestences and Stiefle ,1952).

$$\beta_k = \frac{g_k^T(\rho_k g_k - g_{k-1})}{(g_{k-1}^T g_{k-1})}$$ modified P/R in (Polak and Ribier, 1969).

Step (7): Check for restarting criterion

If k = n, then stop, Otherwise set k = k +1 and go to step (1).

Numerical Computation

Twenty four non linear test functions with dimensions $2 \le n \le 200$ (see Appendix), were chosen to test the effectiveness of the new algorithm. All computations are performed by using (Pentium 4 computer) by using "FORTRAN PROGRAMS" and for all cases the stopping criterion requires $\|g_k\| \le 5 \times 10^{-5}$ to be satisfied. In order to compare the new algorithm with some established algorithms the identical linear search was used, namely, a cubic fitting procedure described in (Bunday, 1984).

All the results given in the tables specifically count the number of function calls (NOF) and the number of the iterations calls (NOI).

Results in table I and II give the comparison of new algorithm with standard CG methods and Tassopoulos and Storey (T/S) algorithm.

Table (I)

Test Function	n	H/S		
		NEW algorithm NOI(NOF)	T/S algorithm NOI(NOF)	Standard CG NOI(NOF)
Dixon	١٠	35(82)	35(60)	22(46)
	١٠٠	107(618)	130(860)	120(860)
Cantreal	4	26(155)	33(230)	33(230)
	10	20(105)	21(115)	18(122)
	20	17(109)	22(132)	18(123)
	200	21(112)	26(189)	19(137)
Rosen	2	29(97)	23(73)	31(73)
	4	18(42)	24(58)	24(58)
	80	25(62)	23(56)	26(56)
Cubic	٢	12(36)	17(48)	17(48)
	٤	13(36)	16(41)	16(42)
	١٠٠	13(36)	14(36)	14(37)
OSP	٢	5(31)	3(17)	3(17)
	٤	8(36)	5(24)	5(24)
	10	12(60)	10(24)	10(48)
Bigg	40	29(92)	25(73)	43(141)
	100	29(92)	25(73)	45(142)
	200	29(92)	25(73)	45(142)
Non Diagonal	2	17(41)	18(43)	19(46)
	4	22(50)	22(52)	21(50)
Powell	٤٠	65(138)	72(158)	72(158)
	٦٠	60(183)	92(198)	62(198)
Wolfe	4	22(68))	12(27)	12(27)
	20	48(200)	39(79)	39(79)
Total NOI(NOF)		682(2573)	732(2739)	734(2904)

Table (II)

Test Function	n	P/R		
		NEW algorithm NOI(NOF)	T/S algorithm NOI(NOF)	Standard CG NOI(NOF)
Dixon	١٠ ١٠٠	22(46) 227(460)	37(73) 250(490)	35(72) 249(490)
Cantreal	4 20 100 200	19(102) 18(109) 29(102) 20(97)	21(107) 17(100) 17(100) 17(100)	21(107) 19(119) 19(119) 20(131)
Rosen	2 4	26(66) 31(85)	31(73) 33(62)	31(73) 33(72)
Cubic	٢ 10 80	19(49) 25(60) 24(57)	17(48) 24(35) 24(35)	17(48) 24(37) 24(37)
OSP	2 4 10	5(31) 8(38) 12(57)	3(17) 5(23) 11(52)	3(17) 5(23) 11(52)
Bigg	4 20 100	20(64) 25(78) 39(49)	20(65) 16(49) 49(70)	20(65) 23(75) 49(73)
Non Diagonal	2 4	15(37) 21(48)	25((49) 26(47)	19(46) 20(47)
Powell	4 20 60	69(162) 48(115) 64(143)	77(183) 48(115) 78(169)	77(183) 48(115) 85(182)
Wolfe	10 100	48(71) 64(68)	32(63) 49(99)	31(63) 49(99)
Total NOI(NOF)		898(2194)	926(2224)	932(2345)

From comparing new algorithm with standard CG methods using (H/S) formula, see table (I) we obtained the following results:

Table (III)

Tools	Standard H/S-CG	NEW Algorithm
NOI	100%	92.9%
NOF	100%	88.6%

It is clear from the above table, that the new algorithm improve the standard H/S-CG algorithm in about (7.1%) NOI and (11.4%) NOF.

And from comparing new algorithm with standard CG-methods using (P/R) formula, see table (II) we obtained the following results:

Table (IV)

Tools	Standard P/R-CG	NEW Algorithm
NOI	100%	96.3%
NOF	100%	93.5%

It is clear from the above table, that the new algorithm improve the standard P/R-CG algorithm in about (3.7%) NOI and (6.5%) NOF.

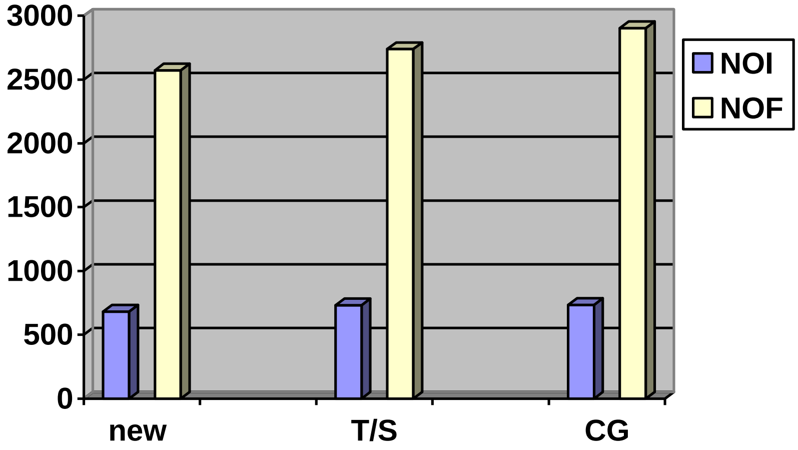
Moreover, we can see from table (I) and table (II) that the new algorithm is better than the T/S algorithm since it has less (NOI) and (NOF) than the T/S algorithm.

Graphics

In this section we are going to illustrate our numerical results by figures. "Microsoft Graph 2003" has been utilized to draw the figures. The comparison between the standard CG algorithms, T/S algorithm, and the new algorithm proposed. with regard to NOI and NOF has been demonstrated by drawing figures.

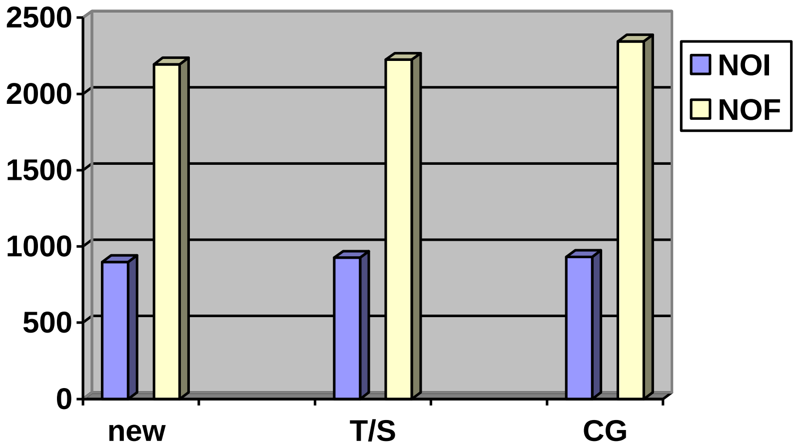
Figures (V) and (VI) show that the new algorithm is better than the established algorithms which are T/S algorithm and standard CG algorithms with respect to the NOI and NOF when ELS are used. These figures draw the total of the NOI and NOF for the numerical results in tables (I) and (II).

Finally, all the above tables and figures show that the new algorithm requires less NOI and NOF than other established algorithms to achieve convergence for solving all those test functions with the use of ELS.



The TOTAL OF (NOI) AND (NOF)

Figure (V) : The Comparison among the different CG algorithms with ELS by using H/S formula.



The TOTAL OF (NOI) AND (NOF)

Figure (VI) : The Comparison among the different CG algorithms with ELS by using P/R formula.

Appendix (Test functions)

1- Bigg Function:

$$f = \sum_{i=1}^{10}\left(\exp(-x_1 z_i) - 5\exp(-x_2 z_i) - \exp(z_i) + 5\exp(-10 z_i)\right)^2, z_i = \frac{i}{10}$$

$x_\circ = (1,2)^T$

2-Cubic Function:

$$f = 100(x_2 - x_1^3)^2$$

$x_\circ = (-1.2,1)^T$

3- Dixon Function:

$$f = (1 - x_1)^2 + (1 - x_{10})^2 + \sum_{i=1}^{9}(x_i^2 - x_{i+1})^2$$

$x_\circ(-1,-1)^T$

4-Generalized Cantreal Function:

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

$x_\circ = (1,2,2,2)^T$

5-Non –Diagonal variant of Rosenbrock Function:

$$f = \sum_{i=1}^{n} 100\left(x_i - x_\circ^2\right)^2 + \left(1 - x_\circ\right)^2$$

$x_\circ = (-1,.....)^T$

6-Oren and Spedicato Power Function (OSP):

$$f = \sum_{i=1}^{n}\left(i x_i^2\right)^r, r = 2$$

$x_\circ = (1,....)^T$

7-Rosenbrock Function:

$$f = \sum_{i=1}^{\frac{n}{2}}\left(100(x_{2i} - x_{2i}^2) + (1 - x_{2i-1})^2\right)$$

$x_\circ = (-1.2,1)^T$

8-Wolfe Function:

$$f = \left[-x_1(3 - \frac{x_1}{2}) + 2x_2 - 1\right]^2 + \sum_{i=1}^{n-1}\left[x_{i-1} - x_i(3 - \frac{x_i}{2}) + 2x_{i-1} - 1\right]^2 + \left[x_{n-1} - x_n(3 - \frac{x_n}{2}) - 1\right]^2$$

$x_\circ = (-1,............)^T$

9-Generalized Powell Function:

$$f = \sum_{i=1}^{\frac{n}{4}}\left[(x_{4i-3} - 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4\right]$$

$x_\circ = (3,-1,0,1,.....)^T$

References

- AI-Assady, N. H. and Al-Ta'ai, B. A (2002): A New non-quadratic model for unconstraind non-linear Optimization, MS thesis University of Mosul.
- Al-Bayati, A. Y. (1995): New Extended CG. Methods for Non Linear Optimization, Natural and Applied Series, Mu'tah Journal for Research and Studies. Mu'tah University, Jordan Vol. 10. No. 6, PP. 69-87.
- Al-Mashhadany, H.(2002): A Rational Exponential Model for Unconstrained Non-Linear Optimization, Qatar University .Sc.J,pp.5-16
- Boland, W. R. Kamagnia, E. R. and Kowalik, J. S. (1977a): A Conjugate Gradient Optimization Method in Variant to Non-Linear scalling, Journal of Optimization Theory and Application, 27, PP. 221-230.
- Boland, W. R. and kowalik, J. S. (1979b): Extended Conjugate Gradient Method with Restarts, Journal of Opt. Theory and Application, 28, PP. 1-9.
- Bunday, B. (1984): Basic Optimization Methods, Edward Arnold, Bedford Square, London, (1984).
- Dragica Vasileska. (2006): Conjugate gradient method

  http://mathworld.wolfram.com/ConjugateGradientMethod.html
- Fletcher, R. and Revees. (1964): Function Minimization by Conjugate Gradient, Computer J. 7, 147-154.
- Fried, I. (1971): N-Step Conjugate Gradient Minimization Scheme for Non - Quadratic Function, AIAA J., a.
- Hestences, M. R. and Stiefle, E. (1952): Methods of Conjugate Gradient for Solving Linear Systems, J. Res. N. B. S., 49.
- Polak, E. and Ribier, G. (1969): Note Surla Convergence Des Methods De Directions Conjugate, Rev. Fr. infr, Rech Oper, 16, R.
- Spedicato, E. (1976): A Variable Metric Method for Function Minimization Derived from Invarianey to Non-linear scaling, journal of optimization theory and applications 20, pp.315-329.

- Tassopoulos, A. and Storey, C. (1984a): A Conjugate Direction Method Based on Non-Quadratic Model, Journal of Optimization theory and Applications, 43, pp. 371-381.
- Tassopoulos, A. and Storey, C. (1984b): Use of a New-Quadratic Model in a CG searches, JOTA, 43.
- Taqi, A. H. and Adham, A. Ali. (2009): New Extended CG Algorithm for Non-Linear Optimization, Journal of kirkuk university/scientific studies, Vo 4. No 2, PP. 83-93.

خوارزمية غير تربيعية جديدة لحل مسائل الامثلية اللاخطية

ادهم عبد الوهاب علي

قسم الرياضيات

كلية العلوم _ جامعة كركوك

تاريخ الاستلام:2010/9/1 , تاريخ القبول: 2011/6/19

الخلاصة

في هذا البحث تم اقتراح خوارزمية جديدة في مجال الامثلية اللاخطية لتطوير وتحوير طرق التدرج المترافق للحصول على القيمة المطلقة بشكل أسرع. هذه الخوارزمية قد اشتقت وحسبت عدديا مقارنة بخوارزمية (CG)القياسية وخوارزمية T/S المقترحة سابقا في هذا المجال باستخدام أكثر من (٢٠) دالة قياسية معروفة عالميا ومن ذوات الأبعاد المختلفة. وأظهرت النتائج العددية إن استخدام النماذج غير التربيعية ذات فائدة كبيرة في معظم المسائل وعلى وجه الخصوص عندما تتزايد أبعاد المسألة.