A new Suggested Conjugate Gradient Algorithm with Logistic Mapping

Dlovan **H**aji **O**mar
Department of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region – Iraq.

(Accepted for publication: December 25, 2016)

Abstract:

In this paper, we will use logistic mapping to find new conjugate gradient coefficients for unconstrained optimization.

Keywords: *Unconstrained Optimization, Conjugate Gradient Method, Logistic Mapping.*

1. Introduction

The conjugate gradient (CG) method is one of the most popular and well known iterative techniques for solving sparse symmetric positive definite (SPD) system of linear equations. It was originally developed as a direct method, but became popular for its properties as an iterative method especially following the development of sophisticated precondition techniques.

There are Two Types of Conjugate Gradient (CG) Methods

Type one: linear Conjugate Gradient Method: Is called quadratic Conjugate Gradient and used to Minimizer Quadratic Method. Type two: nonlinear Conjugate Gradient Method: it is called Non quadratic Conjugate Gradient used it to minimizer general function or nonlinear function

2. Classical Conjugate Gradient Method

Method of linear conjugate gradient is iterative method to solve minimization problem,

$$\min f(x) = \frac{1}{2}x^T G x + b^T + c \quad (2.1)$$

Where b is $nx1$vector, c is constant and G is an nxnpositive symmetric definite matrix, we can show that (2.1) is equivalent to a system of linear equations,

$$Gx = b, \quad (2.2)$$

Then the unique solution of (2.1) is the same as the solution of (2.2).

In this study we consider the unconstrained minimization problem

$$\min f(x) \quad (2.3)$$

And the conjugate gradient method of the form:

$$x_{k+1} = x_k + \alpha_k d_k \quad (2.4)$$

$$d_{k+1} = \begin{Bmatrix} -g_k & for k = 0 \\ -g_{k+1} + \beta_k d_k & for k \geq 0 \end{Bmatrix} \quad (2.5)$$

Where$x_k \in R^n$is the current iterative, d_k is a descent direction of $f(x) at x_k$, $g_k = \nabla f(x_k)$ is step size obtained by a line search andβ_k

is a scalar. The scalar chosen so that the methods (2.4) and (2.5) reduce to the linear conjugate gradient method when f is a strictly convex quadratic and when α_k is the exact one-dimensional minimizer. Various conjugate gradient method have been proposed, and they mainly differ in the choice of the parameter β_k.some well-known formulas for β_k, called the Fletcher-Reeves (FR)[FLETCHER,R.&REEVES, C. (1964).], Polak-Ribiere-Polyak

(PRP)[POLAK,B.&RIBIERE, G. (1969).], Hestenes-Stiefel (HS)[HESTENES,M.R.&STIEFEL, E. L. (1952)] andConjugate Descent(CD)[FLETCHER, R. (1987).] respectively, are given below:

$$\beta_k^{FR} = \|g_{k+1}\|^2 / \|g_k\|^2 \quad (2.6)$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_{k-1}\|^2} \quad (2.7)$$

$$\beta_k^{HS} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)} \quad (2.8)$$

$$\beta_k^{CD} = \frac{-\|g_{k+1}\|^2}{d_k^T g_k} \quad (2.9)$$

Where $\|.\|$ denotes the Euclidean norm. The conjugate gradient method is a very efficient line search method for solving large unconstrained problems, due to its lower storage and simple computation. The conjugate gradient method is still the best choice for solving (2. 3).

In practical computations, it is generally believed that the conjugate gradient method is preferred to the relatively exact line searches. As a result, in the conjugate gradient method, the strong Wolfe conditions, namely,

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \quad (2.9)$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k \quad (2.10)$$

Where $0 < \delta < \sigma < 1$are often on the line search. However, recent studies show that one can analyze the conjugate gradient method under several practical line searches other than the strong Wolfe line search, and good numerical results can be obtained. For example, the nonlinear conjugate gradient method converges globally provided that the step size satisfies the standard Wolfe conditions [Wolfe P. (1969).], namely (2.9s) and

$$g(x_k+\alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \quad (2.11)$$

2.1 Classical Conjugate Gradient Algorithm

Step (1):The initial point x_k ,

Step (2): $g_k = \nabla f(x_k), if g_k = 0$, Then stop

Set $k=0$, select

Else

Set $d_k = -g_k$

Step (3): Compute α_k to minimize $f(x_{k+1})$

Step (4): $x_{k+1} = x_k + \alpha_k d_k$,

Step (5): $g_{k+1} = \nabla f(x_{k+1}), if g_{k+1} = 0$,Then stop .

Step (6): compute β_k ,

Step (7): $d_{k+1} = -g_{k+1} + \beta_k d_k$.

Step (8): If k = n then go to step 2.

Else

$k = k+1$ and go to step 3.

3. New Conjugate Gradient method (β_k^{new})

In this section, to fınd a new conjugate gradient method we will use conjugate Gradient coefficient of (Conjugate Descent)

$$\beta_k^{CD} = \frac{-\|g_{k+1}\|^2}{d_k^T g_k} \quad (3.1)$$

and logistic mapping method which is used extensively[LU Hui-juan, ZHANG Huo-ming, MA Long-hua,(2005). Its equation is as follows:

$$\gamma_{k+1} = \mu\gamma k(1-\gamma k) \quad (3.2)$$

Whereμ is a control parameter (μϵ(0,4)).

Now, from the equation (3.1) and the formula (3.2), we have

$$\beta_k^{new} = \mu\beta_K(1-\beta_K) \quad (3.3)$$

$$Where Y_K = \beta_K$$

Or

$$\beta_k^{new} = \mu\frac{-\|g_{k+1}\|^2}{d_k^T g_k}\left(1+\frac{\|g_{k+1}\|^2}{d_k^T g_k}\right) (3.4)$$

To achieve balance, we will multiply second term of (3.4) by $(k\frac{d_k^T g_{k+1}}{\|y_k\|^2})$

$$\beta_k^{new} = \mu\frac{-\|g_{k+1}\|^2}{d_k^T g_k}\left(1+(k\frac{d_k^T g_{k+1}}{\|y_k\|^2})(\frac{\|g_{k+1}\|^2}{d_k^T g_k})\right) (3.5)$$

Where $k>0$.

3.1 Algorithm of New Conjugate Gradient Method

Step (1): Set k =0, select initial point x_k.

Step (2): $g_k = \nabla f(x_k)$, $If g_k = 0$, $then stop$.

Else

Set $d_k = -g_k$.

Step (3): Compute α_k , to minimize $f(x_{k+1})$.

Step (4): $x_{k+1} = x_k + \alpha_k d_k$.,

Step (5) : $g_{k+1} = \nabla f(x_k+1), If g_{k+1} = 0, then stop$.

Step (6) : Compute β_k^{new}

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

Step (7) : $d_{k+1} = -g_{k+1} + \beta_k^{new} d_k$

Step (8) : If k=n then go to step 2

else

k=k+1 and go to step 3.

Theorem 3.1: Assume that the sequence $\{x_k\}$ is generated by the from $(x_{k+1} = x_k + \alpha_k d_k)$, then the modified CG-method in from (3.6) is satisfied the descent condition, i.e. $d_{k+1}^T g_{k+1} \leq 0$ in both cases: exact and inexact line search.

Proof

The proof is done induction, the result clearly holds for k=0

$$g_0^T d_0 = -\| g_0 \|^2 \leq 0,$$

Now, we prove the current search direction in descent direction at the iteration (k+1) , we have

$$d_{k+1}^T g_{k+1} = \| g_{k+1} \|^2 + \beta_k^{new} d_k^T g_{k+1}$$

By (3.5) , we get

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

Implies that

$$d_{k+1}^T g_{k+1} = -\| g_{k+1} \|^2 - \mu\frac{\|g_{k+1}\|^2}{g_k^T d_k} d_k^T g_{k+1} - \mu k\frac{(\|g_{k+1}\|^2)^2(d_k^T g_{k+1})2}{(g_k^T d_k)2\|y_k\|^2} \quad (3.6)$$

We know that, the first two terms of equation (2.6) are less than or equal to zero because, the formula of Conjugate Descent is satisfies the descent condition, i.e.

$$-\| g_{k+1} \|^2 - \mu\frac{\|g_{k+1}\|^2}{g_k^T d_k} d_k^T g_{k+1} \leq 0$$

The proof is complete if the step length d_k is chosen by an exact line search which requires $d_k^T g_{k+1} = 0$.

Now, if the step length d_k is chosen by an exact line search which requires $d_k^T g_{k+1} \neq 0$.

We know that

$(\|g_{k+1}\|^2)^2$,

$\| y_k \|^2, (g_k^T d_k)^2$ and $(d_k^T g_{k+1})^2$ are positive

So we have

$$-\mu k \frac{(\|g_{k+1}\|^2)2(d_k^T g_{k+1})2}{(g_k^T d_k)2\|y_k\|^2} \leq 0$$

Finally, we have

$$d_{k+1}^T g_{k+1} = -\| g_{k+1} \|^2 - \mu \frac{\|g_{k+1}\|^2}{g_k^T d_k} d_k^T g_{k+1} - \mu k \frac{(\|g_{k+1}\|^2)2(d_k^T g_{k+1})2}{(g_k^T d_k)2 \| y_k \|^2} \leq 0$$

Then the proof is complete

4. Numerical Results

This section is devoted to test the implementation of new method. We compare the modified method with Conjugate Descent, the comparative tests involve Well-known nonlinear problems (standard test function) with different dimensions 4<n<5000, all programs are written in FORTRAN95 language and for all cases the stopping condition is $\|g_{k+1}\|_\infty \leq 10^{-5}$ the results given in table (1) specifically quote the number of function NOF and the number of iteration NOI Experimentalresults in table (1) confirm that the new CG is superior to standard CG method with respect to the NOI and NOF, especially for the test function Mile and G central.

Table (1): Comparative Performance of TheTwo Algorithm (New Conjugate Gradient Method and Conjugate Descent)

Test fun	N	Conjugate descent		New algorithm	
		NOI	NOF	NOI	NOF
Cubic	4	13	38	13	37
	50	14	40	14	39
	100	14	40	14	39
	500	15	44	15	43
	1000	15	44	15	43
	5000	15	44	15	43
Powell	4	32	81	30	74
	50	35	97	30	74
	100	35	97	30	74
	500	35	97	30	74
	1000	35	97	30	74
	5000	35	97	30	74
Powell	4	13	57	13	31
	50	13	57	13	31
	100	13	57	13	31
	500	13	57	14	34
	1000	13	57	14	34
	5000	13	57	14	34
Wood	4	28	65	27	62
	50	28	65	27	62
	100	28	65	27	62
	500	29	68	27	62
	1000	29	68	27	62
	5000	29	68	30	68
G-central	4	12	67	17	114
	50	17	131	17	114
	100	18	142	18	127
	500	23	210	21	169
	1000	23	210	21	169
	5000	28	278	24	214

Rosen	4	30	85	29	80
	50	30	85	30	82
	100	30	85	30	82
	500	30	85	30	82
	1000	30	85	30	82
	5000	30	85	30	82
Mile	4	51	172	48	144
	50	68	246	54	174
	100	68	246	55	176
	500	68	246	61	208
	1000	68	246	62	210
	5000	74	284	68	244
Total		1240	4545	1157	3486

Table (2): Comparing the Rate of Improvement between the New Algorithm and the Standard Algorithm(CD)

TOOLS	Standard Algorithm(CD)	New Algorithm
NOI	100%	93.3%
NOF	100%	76.69%

5. Conclusion

Our method has been analyzed, implemented and tested to some extent, while numerical tests were carried out, on low and high dimensionality problems, and comparisons were made amongst different non-quadratic and quadratic models with exact and inexact line search. The general conclusion that can be drawn from the tests on the gradient model is that.

6. References

Fletcher,R.&Reeves, C. (1964). Function minimization by conjugate gradients. J. Compute., 7, 149–154.

Fletcher, R. (1987). Practical Methods of Optimization, Vol I: Unconstrained Optimization. New York: Wiley.

Hestenes,M.R.&Stiefel, E. L. (1952) Method of conjugate gradient for solving linear systems. J. Res. Natl.Bur. Stand., 49, 409–432.

Hui-juan LU, Zhang Huo-ming, MA Long-hua,(2005). A new optimization algorithm based on chaos,

Polak,B.&Ribiere, G. (1969). Note surla convergence des methods' de directions conjugu´ees. Rev. Fr.Imform. Rech. Oper., 16, 35–43.

Wolfe P. (1969). Convergence conditions for ascent method, SIAM Rev. 11, pp. 226-235Zhejiang University, Hangzhou 310027, China.

كورتيا ليَكولينىَ:

دڤىَ ليَكولينىَ دا، سەپاندنا رەنگيَن نەخشيَن لوجيستى دىَ هيَتە بكارئينان ژ بو ب دەستڤە ئينانا (conjugate gradient coefficients) **نوى بو** (unconstrained optimization).

ملخص البحث:

في هذا البحث، سوف نستخدم رسم الخرائط اللوجستية للعثور على معاملات التدرج المتقارن الجديدة للتحسين غير المقيد.