A STUDY ON CONJUGATE GRADIENT METHODS AND THEIR MODIFICATIONS

A Thesis submitted to Assam University, Silchar in partial fulfilment of the requirement for the Degree of Doctor of Philosophy in Mathematics

By

BIPRAJIT NATH

ABSTRACT

Ph.D Registration Number

Ph.D/1881/2012 Date-09/04/2012

DEPARTMENT OF MATHEMATICS

ALBERT EINSTEIN SCHOOL OF PHYSICAL SCIENCES

ASSAM UNIVERSITY,SILCHAR

Silchar-788011, India

Year of Submission-2016

Objectives of the work:

In this thesis we have considered some open problems associated to the nonlinear conjugate gradient algorithms for unconstrained optimization and we have solved some of the problems associated to the nonlinear conjugate gradient algorithms for unconstrained optimization. These problems mainly refer to the initial direction, the conjugacy condition and the step length computation etc. The general nonlinear method always considers a critical search direction to find a convergent solution. In the present thesis our aim is to establish the fact that the convergent solution can also be obtained for nonlinear function,even if the search direction is different from the initial search direction taken in most of the algorithms of the nonlinear conjugate gradient method to get a convergent solution and with the help of this new initial search direction we have got a new modified form of conjugacy condition and using this the step length can also be determined.

Introduction:

Because of the advances in Science, Engineering, Economics etc studies on global and local optimization for unconstrained problems have become a topic of great concern. In recent years there has been the great deal of interest in the development of optimization algorithms that deal with the problems of finding a global or local minimum of a given problem.. Unconstrained optimization problem arise in virtually in areas in Science and Engineering, and in many areas of the Social Sciences. A significant percentage of real world optimization problems are data fitting problem. The size of real world unconstrained optimization problem is widely distributed, varying from small problems to large problems. One method is mentioned below to solve for the unconstrained optimization problems

Conjugate gradient method

The Conjugate gradient method represents major contribution to panoply of methods for solving large scale optimization problems. They are characterised by

- Low memory requirements .
- Strong local and global convergence propertyIn this survey, we focus on conjugate gradient methods applied to the linear unconstrained optimization problem

$$\min\{f(x): x \in R^n\} \tag{1.1}$$

Where $f: R^n \to R$ is a continuously differentiable function especially if the dimension n is large.

They are of the form $x_{k+1} = x_k + \alpha_k d_k$ (1.2)

Where α_k is a step size obtained by a line search and d_k is the search direction botanised by

$$d_k = \begin{cases} -g_k, k = 1 \\ -g_k + \beta_k d_{k-1}, k \geq 2 \end{cases} \tag{1.3}$$

Where β_k is a parameter and g_k denotes $\nabla f(x_k)$ where the gradient $\nabla f(x_k)$ of f at x_k is a row vector and g_k is a column vector .Different C.G methods correspond to different choices for the scalar β_k .

It is known from (1.2) and (1.3) that only the step size α_k and the parameter β_k remain to be determined in the definition of Conjugate Gradient method. In this case that if f is a convex quadratic, the choice of β_k should be such that the method (1.2)-(1.3) reduces to the linear Conjugate Gradient method if the line search is exact namely

Survey of Literature:

The following chronological list gives us an idea about some choices for the conjugate gradient update parameter.

In 1952, Hestenes and stiefel[45] proposed the formula as $\beta_k^{HS} = \dfrac{g_k^T + y_k}{d_k^T y_k}$ in the original (linear) Conjugate Gradient paper.

First nonlinear Conjugate Gradient method ,proposed by Fletcher and Reeves in 1964 and proposed the formula as $\beta_k^{DY} = \dfrac{\|g_{k+1}\|^2}{d_k^T y_k}$.

In 1967 Denial proposed a new method which requires evaluation of the Hessian $\nabla^2 f(x)$,and the formula used by him is $\beta_k^D = \frac{g_{k+1}^T \nabla^2 f(x) d_k}{d_k^T \nabla^2 f(x) d_k}$

In 1969, Polak and Riebiere proposed the formula for β as $\beta_k^{PRP} = \max\{0, \frac{g_{k+1}^T y_k}{g_k^T g_k}\}$

Powell proposed a formula and analysed by Gilbert and Nocedal for β as

$$\beta_k^{PRP} = \max\{0, \frac{g_{k+1}^T y_k}{g_k^T g_k}\}$$

Fletcher proposed a formula for β as $\beta_K^{CD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}$

Liu and Store proposed a formula for $\beta_K^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}$

Hybrid Liu and Storey-conjugate descent

$$\beta_k^{Hu-Storey} = \max\{0, \min\{\beta_k^{PRP}, \beta_K^{FR}\}\}$$

Hu and Storey proposed a formula for $\begin{aligned} \beta_k^{TA-S} &= \beta_k^{PRP} if\, 0 \le \beta_k^{PRP} \le \beta_k^{FR} \\ &= \beta_k^{FR} otherwise \end{aligned}$

Dai and Liao proposed a formula for $\beta_k^{LD} = \frac{g_{k+1}^T (y_k - tS_k)}{d_k^T y_k}$,t>o

In 1987,Fletcher ,CD stands for “Conjugate Descent” $\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k}$

In 1991,Liu and Storey proposed a formula for $\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}$

In 1999,Dai and Yuan proposed a formula for $\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}$

Or as $d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_k S_K$ where the parameter θ_{k+1} is a scalar approximation of the inverse Hessian of the function f and β_k is selected as suggested by Bringin and Martinez

$$\beta_k^{BM} = \frac{g_{k+1}^T(\theta_k y_k - s_k)}{y_k^T s_k}$$ scaled form of Perry

$$\beta_k^{BM+} = \max\left\{0, \frac{\theta_k g_{k+1}^T y_k}{y_k^T y_k}\right\} - \frac{g_{k+1}^T s_k}{y_k^T s_k}$$ scaled form of Perry

$$\beta_k^{SPRP} = \frac{\theta_k g_{k+1}^T y_k}{\alpha_k \theta_{k-1} g_k^T g_k}$$ scaled form of Polak-Ribiere- Polyak

Present Work:

The objective of the first problem(In the thesis,chapter-4) is to search an initial search direction other than $d_0 = -g_0$ for general unconstrained nonlinear Conjugate Gradient method and to establish the fact that the general unconstrained nonlinear conjugate gradient algorithm can be used with this new search direction. The other objective of this problem is to test the convergence of the Conjugate Gradient algorithm for this new search direction.

In this problem, we will assume that the initial search direction for the nonlinear conjugate gradient algorithm is slightly deflect from the direction - g_0.

In second problem(In the thesis,chapter-5) we have derived a new expression of the Conjugate Gradient parameter β for the non linear Conjugate Gradient method. The convergence of the CG method with the proposed β has been discussed .Also, the proposed β has been tested for a good number of nonlinear test function

In the third problem(In the thesis,chapter-6) we have found a new procedure to find the sequence of step lengths which tend to vary in an unpredictable manner. The convergence of the C.G method with this proposed step length has been discussed.

Results and discussion

The followings are the results and conclusion for the work:

1. From the investigation and discussion for the first problem(in chapter 4) it is clear that the convergent solution can be obtained for unconstrained non-linear programming problems using non-linear Conjugate Gradient methods by taking an initial search direction other than the direction $d_0 = -g_0$. In this problem we have taken the initial search direction slightly deflect from $d_0 = -g_0$ and with this deflected direction also we have succeeded to achieve a convergent solution. The future prospect of the present work is open as the nature of γ needs more investigation.

2. For the second problem(In chapter 5) a new β has been proposed and the efficiency of this new proposed β is analysed in light of the experimental values.

3.From the above investigation and discussion for the third problem(in chapter-6), it is clear that if we take the value of the difference between the consecutive pair of step lengths from the sequence $\{\alpha_k\}$, we observe that the value of each pair is different from the value of the next consecutive pair for different values of $x_0, x_1, x_2, ...$ and $g_0, g_1, g_2, ...$ i.e. the sequence of step length tends to vary in a totally unpredictable manner.

References:

[1] Sumit Saha, Biprajit Nath, A new initial search direction for nonlinear Conjugate Gradient method, Internal Journal of Mathematics Research, ISSN 0976-5840, Volume 6,Number 2 (2014), pp. 183-190.

[2] WEI, Z. X.,HUANG, H. D., TAO,Y.R., A Modified Hestenes-Stiefel Conjugate Gradient Method and Its Convergence, Journal of Mathematical Research & Exposition Mar., 2010, Vol. 30, No. 2, pp. 297–308

[3] Yuan, G. ,Lu, S. and Wei, Z., A line search algorithm for un constrained optimization, J. Software Engg. And applications,2010,3:503 -509

[4] Bayati ,A.Y. AL and Muhammad ,R. S., New Scaled Sufficient Descent Conjugate Gradient Algorithm for Solving Unconstraint Optimization Problems, Journal of Computer Science 6 (5): 511-518, 2010 ISSN 1549-3636

[5]Kiwiel, K., Improved convergence result for the discrete gradient and Secant methods for nonsmooth optimization, J. Optim Theory Appl. (2010) 144: 69-75

[6]Goh, B.S., Convergence of algorithms in optimization and solutions of non linear equations, J. Optim Theory Appl. (2010) 144: 43-55

[7] Bayati ,A. Y. AL, Salim ,A.J. and Abbo ,K. K.,Two-Versions of Conjugate Gradient-Algorithms Based on Conjugacy Conditions for Unconstrained Optimization, American Journal of Economics and Business Administration 1 (2) 97-104, 2009,ISSN 1945-5488

[8] N. Andrei, Acceleration of conjugate gradient algorithms for unconstrained optimization, Appl.Math. Comput. 213 (2009), no. 2, 361-369.