International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 4 Issue 5, May-2015, pp: (37-40), Impact Factor: 1.252, Available online at: www.erpublications.com

Modification of Flecher-Reeves Method with Wolfe Condition

Salah Gazi Shareef[1], Bayda Ghanim Fathi[2]
[1,2]Dept. of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region-Iraq

Abstract: In this paper, we proposed a new conjugate gradient method for unconstrained optimization problems by using Logistics Equation, One of the remarkable properties of the conjugate gradient method is its ability to generate so they are widely used for large scale unconstrained optimization problems.

Keywords: Unconstrained optimization, conjugate gradient method, Armijo condition, logistic equation.

Introduction

Consider the form of the method for nonlinear conjugate gradient to find the minimum for unconstrained optimization problems

$$min\{f(x): x\epsilon\mathbb{R}^n\}(1)$$

where$f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable, the conjugate gradient method (CG) is one for solving large scale problem, because it does not need any matrices and it is an iterative method of the form

$$x_{k+1} = x_k + \alpha_k d_k, \text{ k=0, 1,..., n,} \tag{2}$$

where x_k is the current iterate, $\alpha_k > 0$ is a step size, and d_k is the search direction defined by

$$d_{k+1} = \begin{cases} d_0 = -g_0 \\ -g_{k+1} + \beta_k d_k , \quad k \geq 0, \end{cases} \tag{3}$$

where g_k is the gradient $\nabla f(x_k)$ of$f(x)$ at the point x, and $\beta_k \in R$ which determines the different conjugate gradient methods is a scalar. There are many standard parameters such as:

$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}$proposedby Fletcher and Reeves [1](4)

$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}$proposed by Polak and Ribière [2] and by Polyak [3] (5)

$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k{}^T d_k}$proposedby Hestenes and Stiefel [4](6)

$\beta_k^{DY} = \frac{\|g_k\|^2}{y_k{}^T d_k}$proposed by Dai and Yuan [5] (7)

$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k^T}$proposed by Fletcher [6], CD indicates 'Conjugate Descent' (8)

whereg_k and $g_{k+1}$are gradients $\nabla f(x_k)$ and $\nabla f(x_{k+1})$ of $f(x)$at the point x_k andx_{k+1}, respectively,$\|.\|$ denotes the Euclidian norm of vectors. The CG method is a powerful line search method for solving optimization problems, and it remains very popular for engineers and mathematicians who are interested in solving large–scale problems. This method can avoid, like steepest descent method, the computation and storage of some matrices associated with the Hessain of objective function. Then there are many new formulas that have been studied by many authors.

New Conjugate Gradient

In this section, we propose our new β_k known as β_k^{New} . To find a new conjugate gradient algorithm, we will use the equation of Logistic [7] and conjugate gradient method of Flecher-Reeves (FR).

The Logistic Mapping method is used extensively. Its equation is as follows:

$$\delta_{k+1} = \mu\ \delta_k(1-\delta_k), \tag{9}$$

where $\mu\epsilon(0,4)$ is a control parameter.

Now, we can rewrite (9) as follows

$$\beta_k^{New} = \mu\,\beta_k^{FR}(1-\beta_k^{FR}). \tag{10}$$

Since $\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}$, we have

$$\beta_k^{New} = \mu\frac{\|g_{k+1}\|^2}{\|g_k\|^2}(1-\frac{\|g_{k+1}\|^2}{\|g_k\|^2}). \tag{11}$$

We can add the term $\frac{d_k^T g_{k+1}}{\|g_k\|^2}$ for achieving, (11) becomes

$$\beta_k^{New} = \mu\ \frac{\|g_{k+1}\|^2}{\|g_k\|^2}(1\text{-}\frac{d_k^T g_{k+1}}{\|g_k\|^2}(\frac{\|g_{k+1}\|^2}{\|g_k\|^2})). \tag{12}$$

Algorithm 1: New Conjugate Gradient

Step 1: Set $k=0$, select the initial point x_0 and compute $g_0 = \Delta f(x_0)$
Step 2: If $g_k = 0$, then stop
else $d_k = -g_k$
Step 3: Compute the step size (α_k) to minimize $f(x_{k+1})$
Step 4: $x_{k+1} = x_k + \alpha_k d_k$,
Step 5: If $g_{k+1} = 0$, then stop
Step 6: Compute $\beta_k^{New} = \mu\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\left(1-\frac{d_k^T g_{k+1}}{\|g_k\|^2}\left(\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\right)\right)$
Step 7: $d_{k+1} = -g_{k+1} + \beta_k^{New} d_k$
Step 8: If $|g_k^T g_{k+1}| > 0.2\|g_{k+1}\|^2$, then go to step 2
else $k = k+1$ and go to step 3

Theorem 1: Assume that the sequence $\{x_k\}$ is generated by the Algorithm 1, then the modified of CG-method in (12) is satisfied the descent condition, i.e. $d_{k+1}^T g_{k+1} \le 0$ in exact and inexact line searches.

Proof: The proof is by induction over k. The result clearly holds for $k=0$

$$g_0^T d_0 = \|g_0\|^2 \le 0,$$

Now, we prove the current search direction is 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 using (12), we get

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

Impliesthat

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \mu\frac{\|g_{k+1}\|^2(d_k^T g_{k+1})}{\|g_k\|^2} - \mu\frac{(d_k^T g_{k+1})^2\|g_{k+1}\|^4}{\|g_k\|^2\|g_k\|^4}. \tag{13}$$

If the step length α_k is chosen by an exact line search which requires $d_k^T g_{k+1} = 0$, then the proof is completed.

Now, if the step length α_k is chosen by an inexact line search which requires $d_k^T g_{k+1} \neq 0$, then it is clear that the first and second terms of (13) are less than or equal to zero because the parameter of the FR method is satisfies the descent condition, i.e.

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

and we know that μ and $\frac{(g_{k+1}^T y_k)^2\|g_{k+1}\|^2}{\|g_k\|^2\|g_k\|^2}$ are positive, therefore

$$\mu\frac{(d_k^T g_{k+1})^2\|g_{k+1}\|^4}{\|g_k\|^2\|g_k\|^4} \le 0,$$

Finally, we get

$$d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 + \mu \frac{\|g_{k+1}\|^2 (d_k^T g_{k+1})}{\|g_k\|^2} - \mu \frac{(d_k^T g_{k+1})^2 \|g_{k+1}\|^4}{||g_k||^2 \|g_k\|^4} \leq 0$$

The proof is completed. ■

Theorem 2: Assume that the sequence $\{x_k\}$ is generated by the Algorithm1, then the modified of CG-method as in (12) is satisfied the sufficient descent condition, i.e.

$d_{k+1}^T g_{k+1} \leq -c\|g_{k+1}\|^2,$

where$c$ is a small positive real number.

Proof. From (3) and (12), we have

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

Multiplying both sides of above equation byg_{k+1}, we obtain

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

which gives

$$d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 + \mu \frac{\|g_{k+1}\|^2 (d_k^T g_{k+1})}{\|g_k\|^2} - \mu \frac{(d_k^T g_{k+1})^2 \|g_{k+1}\|^4}{||g_k||^6},$$

or

$$d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 \left[1 + \mu \frac{(d_k^T g_{k+1})^2 \|g_{k+1}\|^2}{||g_k||^6} - \mu \frac{(d_k^T g_{k+1})}{\|g_k\|^2}\right]. \quad (14)$$

By curvature condition

$$g_{k+1}^T d_k \geq c_1 g_k^T d_k, \qquad c_1 \in (0,1). \quad (15)$$

We can rewrite the curvature condition as follows

$$-g_{k+1}^T d_k \leq c_1 \|g_k\|^2. \quad (16)$$

From (14) and (16), we get

$$d_{k+1}^T g_{k+1} \leq -||g_{k+1}||^2 \left[1 + \mu \frac{(d_k^T g_{k+1})^2 \|g_{k+1}\|^2}{||g_k||^6} + \mu c_1\right]. \quad (17)$$

Let $c = 1 + \mu \frac{(d_k^T g_{k+1})^2 \|g_{k+1}\|^2}{||g_k||^6} + \mu c_1$, then (17) gives:

$$d_{k+1}^T g_{k+1} \leq - c||g_{k+1}||^2.$$

Hence the proof is completed. ■

Numerical Results

This section is devoted to test the implementation of the new method. We compare our modification of the CG method with standard Flecher-Reeves (FR) method, the comparative tests involve well-known nonlinear problems (standard test function) with different dimension 4 ≤n≤3000, all programs are written in FORTRAN95 language and for all cases the stopping condition is$||g_{k+1}||_\infty \leq 10^{-5}$. The results are given in Table 1 is specifically quote the number of functions NOF and the number of iteration NOI. Experimental results in Table 1 confirm that the new CG method is superior to standard CG method with respect to the NOI and NOF.

Table 1: Comparative Performance of the new conjugate gradient method and the FR method

No. of test	Test function	N	Standard Formula (FR)		New Formula (New 1)	
			NOI	NOF	NOI	NOF
1	Rosen	4	30	85	29	83
			30	85	29	83
		100	30	85	29	83
		500	30	85	29	83
		1000	30	85	29	83
		5000				
2	Cubic	4	13	38	13	38
			14	40	13	38
		100	15	44	13	38
		500	15	44	13	38
		1000	15	44	13	38
		5000				
3	G Powell	4	40	109	38	108
			42	123	39	110
		100	43	125	39	110
		500	43	125	39	110
		1000	43	125	34	129
		5000				
4	G Wolfe	4	11	23	11	23
			45	91	44	89
		100	46	93	46	93
		500	52	105	48	97
		1000	141	293	104	220
		5000				
5	Wood	4	26	60	26	61
			27	62	26	61
		100	27	62	26	61
		500	27	62	26	61
		1000	27	62	27	62
		5000				
6	G-central	4	18	123	19	129
			24	194	19	129
		100	28	251	20	146
		500	28	251	20	146
		1000	28	251	26	233
		5000				

Conclusion

In this paper, we have considered the conjugate gradient method with the formula (12).We have also shown that the search direction satisfied the descent condition $d_{k+1}^T g_{k+1} \leq 0$.The new algorithm is superior to standard CG method with respect to the NOI and NOF.

References

[1]. R. Flecher, and C.M. Reeves, ''Function Minimization by Conjugate Gradients'', Computer Journal, 7, 1964, Pp. 149-154.
[2]. E. Polak and G . Ribière, ''Note sur la convergence de directions conjugées'', Rev. Francaise Informat Recherche Opertionelle, 3e Année, 16, 1969, pp. 35-43.
[3]. B. T. Polyak, The conjugate gradient method in extreme problems, USSR Comp. Math. Phys., 9, 1969, pp. 94-112.
[4]. M.R. Hestenes and E. Stiefel, ''Methods of conjugate gradients for solving linear systems'', Journal of research of the National Bureau of Standards, 49, 1952, pp. 409-436.
[5]. Y.H. Dai and Y. Yuan, ''A nonlinear conjugate gradient method with a strong global convergence property'', SIAM J.Optim., 10, 1999, pp. 177-182.
[6]. R. Fletcher, ''Practical Methods of Optimization'', John Wiley and Sons, New York, 1987.
[7]. Li B. and W. S. Jiang, ''Chaos optimization method and its application'', Journal of Control Theory and Applications, (4), 1997, pp. 613-615.