(IJRST) 2015, Vol. No. 5, Issue No. III, Jul-Sep **e-ISSN: 2249-0604; p-ISSN:2454-180X**

# ON STABILITY CONJUGATE GRADIENT METHOD WITH "DIXON UPDATE"

**Saba Noori Majeed**

*Department of Mathematic, College of Education for Pure Science*

*Ibn Al-Haitham, University of Baghdad*

## ABSTRACT

*In this paper we discuss and study the stability of conjugate gradient (CG) method with Dixon update, which is suitable and sufficient to large problems since has less storage space.*

***Key Words:*** *Stability, conjugate gradient method, Dixon update.*



## 1. INTRODUCTION

Conjugate Gradient (CG) is the most popular iterative method for solving large systems of linear and nonlinear equations, it is effective and sufficient for such systems of the form $Ag(x) = b$ where $g(x) = \begin{bmatrix} g(x_1) \\ g(x_2) \\ \vdots \\ g(x_n) \end{bmatrix}$ is unknown vector, A is a square n×n matrix, and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ is a known vector.

The method of (CG) was developed independently by Stiefel of the institute of applied mathematics at Zurich and by Hestenes with the cooperation of Rosser, Forsythe, and Paige of the institute of Numerical analysis, National Bureau of standards in 1951. In [1] Magnus and Edward Stiefel studied the method of (CG) for solving linear systems, in [2] Jonathan introduced an introduction to (CG) method without agonizing pain, in [3] Ake, Tommy and Zdenek studied stability of (CG) and Lanczos method for linear least squares problems, in [4] Lumma and Naoum present Dixon update on training of artificial neural networks, in [5] Miro and Julien study numerical stability of iterative methods, in [6] Krystyna study the (CG) method in finite precision computations, in [7] Jabber study (CG) on training feed forward neural networks for approximation problem in his M.Sc. thesis, in [8] Lumma study stability of back propagation training algorithm for neural networks. This research is an improvement of the work in [2] and the results was successfully proved for general large system with Dixon update.

## 2. IMPORTANT CONCEPTS AND DEFINITIONS OF (CG) METHOD

The (CG) method is an iterative method to solve large systems,

$$Ag(x) = b \quad \ldots(1)$$

Which terminates in at most n steps, starting with an initial point $x_0$ of the exact solution x, one determines successive steps $x_0, x_1, \ldots$ of x, the step $x_i$ being closer to the solution x then $x_{i+1}$ and at each step we encountered the following definitions.

**2.1 A quadratic form:** is simply a scalar quadratic function of a vector with the form

$$f(x) = \frac{1}{2} g(x)^T A g(x) - b^T g(x) + c \quad \ldots(2)$$

**2.2 The Gradiant of a quadratic form:** is defined by

$$f'(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix} = \frac{1}{2} A^T g(x) + \frac{1}{2} A g(x) - b \quad \ldots(3)$$

If A is symmetric then

$$f'(x) = Ag(x) - b \quad \ldots(4)$$

If A is not, we denoted that $\frac{1}{2}(A^T + A)$ is symmetric matrix.

We can take a step, we choose a direction in which of decreases most quickly, which is the direction apposite $f'(x_i)$, according to equation (4) this direction is

$$-f'(x) = b - Ax_i \quad \ldots(5)$$

**2.3 The Error**

$$e_i = x_i - x \quad \ldots(6)$$

is a vector that indicates how for we are from the solution.

**2.4 The Residual**

$$r_i = b - Ax_i \quad \ldots(7)$$

Indicates how far we are from the correct value of b.

It is easy to see that

$$r_i = -Ae_i \quad \ldots(8)$$

and

$$r_i = -f'(x_i) \quad \ldots(9)$$

(IJRST) 2015, Vol. No. 5, Issue No. III, Jul-Sep **e-ISSN: 2249-0604; p-ISSN:2454-180X**

putting all the above definitions together we get the method:

$$x_{i+1} = x_i + \alpha_i r_i \qquad \ldots(10)$$

where $\alpha_i = \dfrac{-r_i^T r_i}{r_i^T A r_i}$.

## 3. DIXON UPDATE FORMULA

Dixion update formula is the search direction and defined by

$$\alpha_i = \frac{-r_i^T r_i}{r_{i-1}^T A r_{i-1}} \qquad \ldots(11)$$

Equation (10) is called gradient dessent method or weight updates, $\alpha_i$ is a parameter governing the speed of the $x_i$ convergence and direction to the solution x, controlling the distance between $x_i$ and $x_{i+1}$, $r_i$ is the gradient of the error surface at $x_i$, the convergence condition is satisfied by choosing $\alpha_i = \dfrac{1}{\lambda_{max}}$, where $\lambda_{max}$ is the largest eigenvalue of the matrix A.

## 4. STABILITY OF (CG) METHOD WITH DIXON UPDATE

To understand the convergence of (CG) method its very important skill on linear algebra, if A is a matrix then there exists a set of n-linearly independent eigenvectors of A, denoted by $v_1, v_2, \ldots, v_n$ each eigenvectors correspond eigenvalues denoted $\lambda_1, \lambda_2, \ldots, \lambda_n$, these eigenvalue may or may not be equal together, on repeated application from equation (1) $Ag(x) = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$, if magnitudes of all eigenvalues are smaller than one, i.e. $\lambda_i < 1$, then x converge to zero, if one of the eigenvalues are greater than one, $\lambda_i > 1$, then x will diverge from zero to infinity.

Stability refers to the equilibrium behavior of the activation state of (CG) method where convergence behavior of the gradient during moving from step to other, there are several mathematical definitions of the tearm "stability" the one due to Lyapunou is most useful here.

### 4.1 Lypunove Stability

The equilibrium state x = 0 is stable if for any $\varepsilon > 0$ there exist $\delta(\varepsilon) > 0$, such that $|x(0)| < \delta$ implies $|x(t)| < \varepsilon$ for all $t > 0$.

## 5. CONVERGENCE OF (CG) FOR ONE STEP ONLY

Consider the case where $e_i$ is an eigenvector with eigenvalue $\lambda_e$, see figure (1), the residual

$$r_i = -Ae_i = -\lambda_e e_i \qquad \ldots(12)$$

(IJRST) 2015, Vol. No. 5, Issue No. III, Jul-Sep **e-ISSN: 2249-0604; p-ISSN:2454-180X**

is also an eigenvector, from the equation (10), choosing $\alpha_i = \frac{-1}{\lambda_e}$, $e_{i+1} = e_i + \alpha_i r_i$, take $\alpha_i$ from equation (11)

$$e_{i+1} = e_i + \frac{-r_i^T r_i}{r_{i-1}^T A r_{i-1}} r_i$$

$$e_{i+1} = e_i + \left(\frac{-1}{\lambda_e}\right) \lambda_e e_i$$

$$e_{i+1} = e_i - e_i$$

$$e_{i+1} = 0.$$

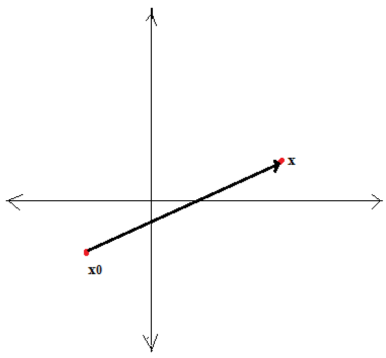


fig.1 one step convergent

## 5.1 General Convergence of (CG) Method with Equal Eigenvalues

More general analysis, we must express $e_i$ as a linear combination of eigenvectors, and we shall furthermore require this eigenvectors to be orthogonal, see figure (2).

Express the error term as a linear combination of eigenvectors $e_i = \sum_i t_i v_i$, let us choos the eigenvectors $v_i$ of length one, i.e.

$$v_i^T v_i = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$$

$$r_i = -Ae_i = -\lambda_i e_i = -\sum_i \lambda_i t_i v_i$$

$$\|e_i\|^2 = e_i^T e_i = \sum_i t_i^2$$

$$r_i^T r_i = \|r_i\|^2 = 1$$

$$r_{i-1}^T r_{i-1} = \|r_{i-1}\|^2 = 1$$

Choose $\alpha_i = \frac{-1}{\lambda_e}$, $\lambda_1 = \lambda_2 = \ldots = \lambda_i = \lambda_e$ "equal eigenvalues", equation (10) gives

(IJRST) 2015, Vol. No. 5, Issue No. III, Jul-Sep **e-ISSN: 2249-0604; p-ISSN:2454-180X**

$$e_{i+1} = e_i + \alpha_i r_i$$

$$= e_i + \frac{r_i^T r_i}{r_{i-1}^T A r_{i-1}} r_i$$

$$= \frac{\|r_i\|^2}{\|r_{i-1}\| \lambda_e} - \lambda_e e_i$$

$$= e_i + \frac{1}{\lambda_e} - \lambda_e e_i$$

$$= 0.$$

This mean the convergence exist and stability is working very well on the convergence region.

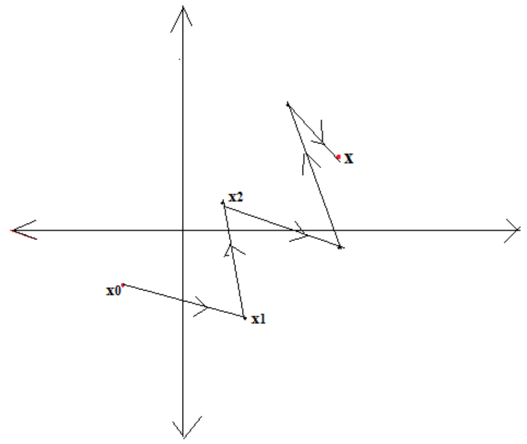


fig.2 general convergent

## 6. GENERAL CONVERGENCE

To bound the convergence of (CG) method with Dixon update; we shall define the "energy norm" $\|e_i\| A = (e_i^T A e_i)^{\frac{1}{2}}$, from equation (2) the minimizing $\|e_{i+1}\| A$ is equivalent to $\|e_{i+1}\| A = (e_{i+1}^T A e_{i+1})^{\frac{1}{2}}$, from equation (10) we get

$$e_{i+1} = e_i + \alpha_i r_i$$

$$\|e_{i+1}\|^2 = \|e_i + \alpha_i r_i\|^2$$

$$\|e_{i+1}\| = \|e_i + \alpha_i r_i\| \le \|e_i\| + \|\alpha_i\| + \|r_i\|$$

$$\leq \|e_i\| - \left(\frac{\lambda_i}{\lambda_{i+1}}\right)^{\frac{1}{2}} \|e_i\|$$

$$\leq \|e_i\| (1-\sqrt{k})$$

K is called the spectral condition number of A, see [3], the ratio of the largest and smallest eigenvalue, $k = \frac{\lambda_{max}}{\lambda_{min}}$ for $\lambda_{max} > \lambda_1 > \lambda_2 > \ldots > \lambda_{min}$.

The convergent result is

$$\|e_{i+1}\| \leq \|e_0\| (1-\sqrt{k})$$

From equation (2)

$$\frac{f(x_i)-f(x)}{f(x_0)-f(x)} = \frac{\frac{1}{2} e_i^T A e_i}{\frac{1}{2} e_0^T A e_0} = \frac{\|e_{i+1}\|}{\|e_0\|}$$

$$\leq 1-\sqrt{k}$$

Then we get the existence of convergence and stability of region.

## 7. CONCLUSION

In this work we successfully proved the stability of (CG) method with Dixon update for large systems of the form Ag(x) = bwhich it's an improvement of (CG) method in [2].

## REFERENCES


1. Magnus R.H. and Eduard S., "Methods of Conjugate Gradients for Solving Linear Systems", Journal of Research of the National Bureau of Standards, Vol.49, No.6, 1952.
2. Jonathan R.S., "An Introduction to the Conjugate Gradient Method without the Agonizing Pain", School of Computer Science, Carnegi Mellon University, Pittusburgh, 1994, PA15213.
3. Ake B., Tommy E. and Zdenek S., "Stability of Conjugate Gradient and Lanezos Methods for Linear Least Squares Probelems", SIAM Journal on Matrix Analysis and Applications, Vol.19:ISSUE.3:Pages.720-736, 1998, July.
4. L.N.M.Tawfiq and R.S.Naom, "On Training of Artifical Neural Networks", Al-Fath Journal, No.23, 2005.
5. Miro R. Christopher C.P. and Julien L., "Numerical Stability of Iterative Methods", Institute of Computer Science Czech Academy of Science, CZ-182, 07 Prague, Czech

(IJRST) 2015, Vol. No. 5, Issue No. III, Jul-Sep **e-ISSN: 2249-0604; p-ISSN:2454-180X**

Republic, GAMM-SIAM Conferena on Applied Linear Algebra, 2006, Dusseldorf, Germany, July.

6. Krystyna Z., "The Conjugate Gradient Method in Finite Precision Computations", Warctaw University of Technology, Institute of Mathematics and Computer Science, 2006, Contribution of Wozniakowski, Strakos.
7. Jabber, A.K., (2009), "On Training Feed Forward Neural Networks for Approximation Problem", M.Sc. Thesis, University of Baghdad, College of Education Ibn-AlHaitham.
8. Luma N.M.T., "Stability of Back Probagation Training Algorithm for Neural Networks", Journal of Baghdad Science, Vol.9 (4), 2012, pp.713-719.