African Journal of Mathematics and Computer Science Research Vol. 1(2), pp. 028-031, October, 2008
Available online at http://www.academicjournals.org/AJMCSR


*Full Length Research Paper*

# Conjugate gradient method for non-positive definite matrix operator

**Omolehin J.O.[1*], Abdulrahman M.K.A.[1], Rauf K.[1] and Udoh P.J.[2]**

[1]Mathematics Department, University of Ilorin, Ilorin, Nigeria.
[2]Mathematics Department, University of Uyo, Uyo, Nigeria.

Accepted 18 September, 2008

**The conjugate Gradient Method relies on symmetric positive definite property of a matrix operator. In this work, we investigate and establish the case of non-positive symmetric matrix operator using various forms of conversions and transformations on the operator.**

**Key words:** Hilbert Space, Positive Symmetric Matrix, Quadratic Functional, Conjugate, Gradient Method.

## INTRODUCTION

Consider the quadratic functional of the form:

$$f(X) = f_0 + \langle a, X \rangle_H + \frac{1}{2}\langle X, AX \rangle_H \quad (1)$$

where $A$ is an nxn symmetric positive definite constant matrix operator on Hilbert space H, $a$ is a vector in H and $f_0$ is a constant term. The minimization of the above quadratic form by conjugate gradient Method (CGM) relies on the symmetric positive definite of matrix A. See [6, and 9]. Our attention in this work is on a non-positive symmetric $A$ which can be converted to positive definite through series of approaches. Ibiejugba et al. (2003), Hasdorff (1976) and several others have shown the effect of positive definite matrix on the convergence rate of CGM. It was shown that the CGM converges to optimal value in at most n iteration, where n is the dimension of the matrix operator. This in essence shows that a non-positive definite matrix may not terminate in at most n iteration. For further work see (Abdulrahman, (2007), Andrew and Chelsea, 1977, Enid, 1993, Lipschutz, 1987).

### Effect of matrix on the optimal value

The operator arising from the quadratic functional (1) is always symmetric and became a powerful tool to locate the optimal value. This symmetric matrix may be positive definite, negative definite, Singular point (positive-indefinite) and Saddle point (indefinite) (Jonathan, 1994). While the solution of positive and negative definite are unique, the singular point is as a result of indefinite matrix which does not have a solution. The solution to saddle point can represent the maximum and minimum at the same time. The definiteness of the operator plays a vital role in the optimality conditions.

A symmetric positive or negative definite matrix operator of a quadratic function guarantee a strictly convex function which in the first place show that it is an increasing or decreasing function and in turn give rise to a unique minimum or maximum.

A symmetric semi positive (semi negative) definite matrix operator in turn gives a convex function which may possibly be increasing (decreasing). In this case the open ball center at $x^*$ with radius r is also an optimum. This suggests that there may be infinitely many points within the domain of optimum.

Indefinite matrix is the case where the function may be increasing, decreasing or neither. This is a serious case as we can not initially predict the nature of result that will come out of the computation.

The positive definite matrix guarantees an inverse of the matrix which is also unique, but this does not necessarily hold for indefinite matrix (Lawrence, 1976). We then conclude that indefinite matrix can only affect the uniqueness of the solution and does not necessarily mean it has no solution. The generated CGM by Ibiejugba et al.

*Corresponding author. E-mail: omolehin_joseph@yahoo.com

(2003) and Hasdorff (1976) have shown the effect of positive definite operator on the convergence properties of the algorithm.

**Effect of matrix conversion on the objective functional**

A non-positive definite matrix can be converted to positive definite in a number of ways. These include positive definite completion problem due to Wayne et al. (1993), forcing of indefinite matrix due to Thomas (2001) and matrix transformation.

**Positive definite completion problem**

In this case we try to find in a symmetric matrix which pair of $a_{ij} = a_{ji}$ can be replaced in order to get a positive definite matrix. Wyne Baret [12] shows that this replacement must be with respect to the dimension of the matrix. In a $nxn$ matrix the pair of $a_{ij} = a_{ji}$ that will be replaced must be (n - 2), i.e. in a 4 x 4 matrix (4 - 2) pair can be replace. This replacement has a price on both the objective function and the corresponding system of linear equation. With this (n - 2) it implies that the completion problem can not take care of 2 x 2 operators. Consider the problem

$$\min imize\ F_1(X) = 1 - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

This is written in matrix form as

$$Minimize F_1(X) = 1 - \begin{pmatrix}1\\1\\1\end{pmatrix}^T \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} + \frac{1}{2}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}^T \begin{pmatrix}2&2&2\\2&4&0\\2&0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

With Barret's result our (n - 2) here is 1 then it make sense to remove only a pair from the original matrix.

$\begin{pmatrix}2&2&2\\2&4&0\\2&0&2\end{pmatrix}$ Can be replaced by $\begin{pmatrix}2&2&x\\2&4&0\\x&0&2\end{pmatrix}$ the principal minor becomes $H_1 = 2, H_2 = 4$ and $H_3 = 4 - 2x^2$ and all will be greater than zero if $x < \sqrt{2}$. Choosing x = 1, $\begin{pmatrix}2&2&1\\2&4&0\\1&0&2\end{pmatrix}$ is a positive definite completion of $\begin{pmatrix}2&2&2\\2&4&0\\2&0&2\end{pmatrix}$.

The resulting functional is

$$\min imize\ F_2(X) = 1 - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + x_1x_3$$

and Choosing x = 1.4,

$\begin{pmatrix}2&2&1.4\\2&4&0\\1.4&0&2\end{pmatrix}$ is a positive definite completion of $\begin{pmatrix}2&2&2\\2&4&0\\2&0&2\end{pmatrix}$.

The resulting functional is

$$\min imize\ F_3(X) = 1 - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 1.4x_1x_3$$

This conversion to positive definitely affects the last term of the functional and the effect is be shown in the result.

**Forcing of an indefinite matrix to be positive definite**

This case quite unique and has a similar problem on the objective function and its corresponding gradient. All the same, an indefinite matrix can be made positive definite by adding a small positive number to the diagonal element. This is done by $A - I\beta$ where $I$ is the identity matrix and $\beta$ is the smallest positive number. This number will be choosing appropriately so that the resulting matrix will be positive definite (Thomas et al., 2001). In the above example, the smallest possible value is $\beta = 0.5$ and the resulting positive definite matrix is

$$\begin{pmatrix}2.5&2&2\\2&4.5&0\\2&0&2.5\end{pmatrix},$$

the resulting functional becomes

$$\min imize\ F_4(X) = 1 - x_1 - x_2 - x_3 + 1.25x_1^2 + 2.25x_2^2 + 1.25x_3^2 + 2x_1x_2 + 2x_1x_3$$

We can see now that there is effect of 0.25 increases on $x_1^2, x_2^2\ and\ x_3^2$ which is half of the value added to obtain the positive definite matrix.

**Matrix transformation**

Transformation of matrix has played a major role in finding solution to some problem where the original problem may not. In this case a non-positive definite matrix $A$ can be made positive definite by

(2) $B = AA^T$

$B$ is symmetric and positive definite for non-symmetric and/or non-positive definite matrix $A$.

The gradient of functional (1) is

(3) $\nabla f(X) = a^T + AX = 0$

The replacement $A$ by $B$ does not affect the solution to (1) i.e. equation (3) is equivalent to

(4) $\nabla f(X) = a^T + BX = 0$

With equation (4), $B$ is a full matrix which may affect the solution as the dimension of the matrix keeping increasing (John and Karin, 2001). The equivalent of equations (3) and (4) then show that they both minimize the functional (1).

**Table 1.** Numerical result.

| | Optima value | Exact | CGM1 | CGM2 |
|---|---|---|---|---|
| Problem 1 | $x_1^*$ | 0.5 | 0.1875 | 0.5 |
| | $x_2^*$ | 0 | 0.1875 | 0 |
| | $x_3^*$ | 0 | 0.1875 | 0 |
| | $f(X^*)$ | -0.25 | -0.2813 | -0.25 |
| | NITR | - | 1 | 3 |
| Problem 2 | $x_1^*$ | 0 | 0.2143 | 0 |
| | $x_2^*$ | 0.25 | 0.2143 | 0.25 |
| | $x_3^*$ | 0.5 | 0.2143 | 0.5 |
| | $f(X^*)$ | -0.375 | -0.3214 | -0.375 |
| | NITR | - | 1 | 3 |
| Problem 3 | $x_1^*$ | -10 | 0.2027 | -10 |
| | $x_2^*$ | 5.25 | 0.2027 | 5.25 |
| | $x_3^*$ | 7.5 | 0.2027 | 7.5 |
| | $f(X^*)$ | -1.375 | -0.3041 | -1.375 |
| | NITR | - | 1 | 3 |
| Problem 4 | $x_1^*$ | -22 | 0.1714 | -22 |
| | $x_2^*$ | 10 | 0.1714 | 10 |
| | $x_3^*$ | 18 | 0.1714 | 18 |
| | $f(X^*)$ | -3 | -0.2571 | -3 |
| | NITR | - | 1 | 3 |
| Problem 5 | $x_1^*$ | 0.5 | 0.2681 | 0.5 |
| | $x_2^*$ | 4.44E-16 | 0.1814 | 6.66E-15 |
| | $x_3^*$ | 8.88E-16 | 0.2143 | 2.36E-15 |
| | $f(X^*)$ | -0.25 | -0.2680 | -0.25 |
| | NITR | - | 19 | 3 |

## Main result

In this section, we provide various forms of conversions and transformation of a functional with non positive definite matrix operator. Consider.

Problem 1:

$$\min imize\ F_1(X) = f_{0(1)} - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

Problem 2:

$$\min imize\ F_1(X) = f_{0(2)} - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

Problem 3:

$$\min imize\ F_3(X) = f_{0(3)} - x_1 - x_2 - x_3 + x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 1.4x_1x_3$$

Problem 4:

$$\min imize \quad F_4(X) = f_{0(4)} - x_1 - x_2 - x_3 + 1.25x_1^2 + 2.25x_2^2 + 1.25x_3^2 + 2x_1x_2 + 2x_1x_3$$

Problem 5: problem1 with gradient (4)

## Description of CGM algorithm

The conventional conjugate gradient method is used to minimize the quadratic functional (1), the CGM algorithm is describe as follows:

### CGM1

$$i \quad choose \quad X_0$$
$$ii \quad compute \ \nabla F(X_0) = g_0$$
$$p_0 = -g_0$$
$$iii \quad compute \quad X_{i+1} = X_i + \alpha_i p_i \ ; \ \alpha_i = \frac{g_i^T g_i}{p_i^T A p_i}$$
$$g_{i+1} = g_i + \alpha_i A p_i$$
$$if \quad g_{i+1} = 0 \quad or \quad g_{i+1} - g_i \le \varepsilon \quad then$$
$$iv \quad compute \ \beta_{i+1} = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}$$
$$p_{i+1} = -g_i + \beta_{i+1} p_i$$

### CGM2

This is as a result of the modification of the descent $p_{i+1} = -g_{i+1} + \beta_{i+1} p_i$ due to Omolehin et al. (2006), due to these results, the convergence of some functional is expected in at most two iterations.

Using the two CGM algorithms to solve problems 1 - 5 above and comparing the result with the exact value obtained by matrix inverse method. We have the following table of results: (Table 1)

The results in this table represent the optimal value without $f_0$. This is done in order to reconcile the optimal value, since $f_0$ does not take part in the minimization process but in the optimal value. This is the reason why different value of $f_0$ are used for different problem presented. Now, for the exact, we have,

$$F_1(X^*) = -0.25 + f_{0(1)}$$
$$F_2(X^*) = -0.375 + f_{0(2)}$$
$$F_3(X^*) = -1.375 + f_{0(3)}$$
$$F_4(X^*) = -3 + f_{0(4)}$$
$$F_5(X^*) = -0.25 + f_{0(1)}$$

For all the functional to be equal i.e. $F_1(X^*) = F_2(X^*) = F_3(X^*) = F_4(X^*) = F_5(X^*)$, Problem 1 being the original functional will assume $f_{0(1)} = 0$ then, $f_{0(2)} = 0.125$, $f_{0(3)} = 1.125$ and $f_{0(4)} = 2.75$.

## DISCUSSION AND RESULT

The operator of problem1 is non-positive definite, while the operators for the other problems are positive definite. The results of CGM2 obtained at third iteration are exact while that of CGM1 diverges as the first iteration gives the closest value to optimum. The result of CGM1 that is different came from Problem 1, Problem 2, Problem 3, Problem 4, Problem 5 which occurred at 19th iteration without satisfying the termination criteria. In our next work we extend this to higher dimensional operator while at the same time reducing the difference noticed between the functional.

## REFERENCES


Abdulrahman MKA (2007), Conjugate Gradient Method for solving non-symmetric matrix operators (M. Sc. Dissertation).

Andrew PS and Chelsea CW (1977). Optimum System Control (Second Edition), Prentice-Hall, New Jersey.

Enid RP (1993). Optimal Control and the Calculus of Variation, Oxford University Press, New York.

Ibiejugba MA, Omolehin JO, Rubio JE (2003). Foundation Postgraduate Course on Computational Methods and Application in Optimization, National Mathematical Centre, Abuja, Nigeria.

John MH, Karin M (2001). PDMATRIX-Programs to make matrix positive definite.

Jonathan RS (1994). An Introduction to the Conjugate Gradient Method without Agonizing Pain .

Lawrence H (1976). Gradient Optimization and Nonlinear control, John Wiley & sons, London.

Lipschutz S (1987). Theory and Problems of Linear Algebra (Metric Edition), McGraw-Hill, inc. Singapore.

Omolehin JO (2006). Further Result on CGM Algorithm, J. Math. Assoc. Niger, 33(2b): 374-388.

Omolehin JO, Rauf K (2006). Computational Capability of CGM Algorithm, Creative Math. & Inf., 15: 93-106.

Thomas FE (2001), Optimization of Chemical Processes (second edition), McGraw-Hill, New York.

Wayne B, Charles RJ, Pablo T (1993). The Real Positive Definite Completion Problem for a Simple Cycle. J. Linear. Algebra. Appl. 192: 3-31.