A Constrained Conjugate Gradient Method and its Parallelization

M.H.B.M. SHARIFF
Etisalat College of Engineering
PO Box 980, Sharjah
United Arab Emirates
shariff@ece.ac.ae

Abstract:- A minimization problem with constraints generally leads to solving an indefinite system of equations. Generally algorithms design to solve an indefinite system have convergence properties that are not as good as algorithms design to solve a positive definite system. In this paper we develop a constrained conjugate gradient method which take advantage of the positive definite part of the minimization problem. Proofs of several theorems related to the method are given. Parallelization of the proposed method is given.

Key-Words:- Optimization, Iterative, Large Scale, Constrained.

1 Introduction

Many problems with constraints generally yield indefinite system of equations. Solving such system of equations, especially sparse systems, via a direct method, based on, e.g. Gaussian elimination, on parallel MIMD machine, is generally inefficient and cumbersome. However, direct methods based on matrix-matrix multiplications, matrix-vector multiplications and scalar products are, suitable for MIMD machines and with appropriate housekeeping these methods can be made efficient. An example of an efficient direct method for a sparse system of equations can be found in references [1] and [2]. In this paper we deal with an indirect (iterative) method, although in exact arithmetic the method can be considered as a direct method. Iterative methods to solve an indefinite system of equations (associated with a constrained minimization problem) have been developed, however, the convergence properties of such methods are not as good as iterative methods developed for a positive definite system of equations. For the class of problems discussed here a constrained conjugate gradient method [3] is developed to solve the problems. This method employs, matrix-vector multiplications, outer products and scalar products, hence it is ideal for parallelization.

2 Constrained Conjugate Gradient Method

Consider the problem:
Minimize

$$f = \frac{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}{2} - \boldsymbol{b}^T \boldsymbol{x} \tag{1}$$

subject to constraint

$$\boldsymbol{C}^T \boldsymbol{x} = \boldsymbol{h}\,, \tag{2}$$

where $\boldsymbol{A}$ is a symmetric positive (semi)definite $n \times n$ matrix, $\boldsymbol{b}$ is a given n-vector, $\boldsymbol{C}$ is an $n \times m$ matrix, $\boldsymbol{h}$ is a given m-vector and $m < n$. It is assumed that the columns of $\boldsymbol{C}$ are linearly independent. These constraints restrict the solution to m hyperplanes. The intersection of m linearly independent hyperplanes is, in general, an affine subspace (or flat) of R^n and will be denoted by F_m.

2.1 The proposed Algorithm

The heart of the algorithm is to use a projection matrix $\boldsymbol{H}$ so that the directions of search will always be in the feasible region. The imposition of m linearly independent equality constraints on an n-dimensional problem reduces the dimensionality of the optimization to $n - m$. We note any n-vector $\boldsymbol{x}$ has a unique expansion as a linear combination of the columns of $\boldsymbol{W}$ and $\boldsymbol{Z}$, i.e.,

$$\boldsymbol{x} = \boldsymbol{Z}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{q}$$

for some $n-m$ vector $\boldsymbol{y}$ and m vector $\boldsymbol{q}$. $\boldsymbol{W}$ denotes any matrix whose columns form a basis for the range space of $\boldsymbol{C}$, the columns of $\boldsymbol{Z}$ form a basis for the null space of $\boldsymbol{C}^T$ and, $\boldsymbol{Z}$ and $\boldsymbol{W}$ define complimentary subspaces. For feasible $\boldsymbol{x}$, we have

$$\boldsymbol{C}^T\boldsymbol{x} = \boldsymbol{C}^T(\boldsymbol{Z}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{q}) = \boldsymbol{h}$$

Since $\boldsymbol{C}^T\boldsymbol{Z} = \boldsymbol{0}$, we have

$$\boldsymbol{C}^T\boldsymbol{W}\boldsymbol{q} = \boldsymbol{h} \tag{3}$$

By definition of $\boldsymbol{W}$, the matrix $\boldsymbol{C}^T\boldsymbol{W}$ is non-singular and thus from equation (3), $\boldsymbol{q} = \boldsymbol{q}^*$ is uniquely determined. Hence, the transformation for feasible $\boldsymbol{x}$, i.e.,

$$\boldsymbol{x} = \boldsymbol{Z}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{q}^* \tag{4}$$

can be considered as elimimnation of constraints by transformation of variables. Substituting equation (4) in (1), we have

$$f(\boldsymbol{x}) = f(\boldsymbol{Z}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{q}^*)\,.$$

Any feasible direction $\boldsymbol{p}$ can be written as

$$\boldsymbol{p} = \boldsymbol{Z}\boldsymbol{v}\,,$$

where $\boldsymbol{v}$ is any vector of changes in $\boldsymbol{y}$. The reduced gradient

$$\frac{\partial f}{\partial \boldsymbol{y}} = \boldsymbol{Z}^T\boldsymbol{g}\,,$$

where $\boldsymbol{g} = \dfrac{\partial f}{\partial \boldsymbol{x}}$. $\boldsymbol{y}$ should be changed so that

$$\boldsymbol{v} = -\boldsymbol{Z}^T\boldsymbol{g}\,.$$

Hence,

$$\boldsymbol{p} = -\boldsymbol{Z}\boldsymbol{Z}^T\boldsymbol{g}$$

We note that the projection matrix $\boldsymbol{H} = \boldsymbol{Z}\boldsymbol{Z}^T$ is not unique and in general does not project into itself, i.e., $\boldsymbol{H}\boldsymbol{H} \neq \boldsymbol{H}$. The orthogonal projection matrix $\boldsymbol{H}_m$ is related to $\boldsymbol{Z}$ by the relation

$$\boldsymbol{H}_m = \boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T\,.$$

If $\boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{I}$(the identity matrix), we then have

$$\boldsymbol{p} = -\boldsymbol{H}_m\boldsymbol{g}$$

We now in a position to formulate our algorithm. The proposed algorithm is:
Initialize:
$\boldsymbol{x} = \boldsymbol{x}_0$ (which satisfies equation (2))
$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}$, $\boldsymbol{z}_0 = \boldsymbol{H}\boldsymbol{g}_0$, $\boldsymbol{d}_0 = -\boldsymbol{z}_0$
For $i = 1, 2, \ldots$

$$\alpha_i = \frac{-\boldsymbol{g}_i^T\boldsymbol{d}_i}{\boldsymbol{d}_i^T\boldsymbol{A}\boldsymbol{d}_i} \tag{5}$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha_i\boldsymbol{d}_i \tag{6}$$

$$\boldsymbol{g}_{i+1} = \boldsymbol{g}_i + \alpha_i\boldsymbol{A}\boldsymbol{d}_i \tag{7}$$

$$\boldsymbol{z}_{i+1} = \boldsymbol{H}\boldsymbol{g}_{i+1} \tag{8}$$

$$\beta_i = \frac{\boldsymbol{g}_{i+1}^T\boldsymbol{z}_{i+1}}{\boldsymbol{g}_i^T\boldsymbol{z}_i} \tag{9}$$

$$\boldsymbol{d}_{i+1} = -\boldsymbol{z}_{i+1} + \beta_i\boldsymbol{d}_i \tag{10}$$

Convergence check: if $\boldsymbol{d}_{i+1}^T\boldsymbol{d}_{i+1} < \textit{tolerance} \times \boldsymbol{d}_0^T\boldsymbol{d}_0$ stop
end for

2.2 Special Case

If $\boldsymbol{H} = \boldsymbol{H}_m$, then in exact arithmetic (taking note that $\boldsymbol{H}_m\boldsymbol{H}_m = \boldsymbol{H}_m$)

$$\beta_i = \frac{\boldsymbol{g}_{i+1}^T\boldsymbol{z}_{i+1}}{\boldsymbol{g}_i^T\boldsymbol{z}_i} = \frac{\boldsymbol{z}_{i+1}^T\boldsymbol{z}_{i+1}}{\boldsymbol{z}_i^T\boldsymbol{z}_i} = \frac{\boldsymbol{z}_{i+1}^T(\boldsymbol{z}_{i+1} - \boldsymbol{z}_i)}{\boldsymbol{z}_i^T\boldsymbol{z}_i} \tag{11}$$

Since our formula for β_i is not found in the literature, it is prudent to prove that our algorithm produces mutually conjugate directions of search. We also that the proposed method converges in at most $n-m$ iterations in exact arithmetic.

Proposition 1: *The directions of search determined by the proposed iteration are mutually conjugate w.r.t. the matrix* $\boldsymbol{A}$
Proof: The theorem can be stated in the form

$$\boldsymbol{d}_i^T\boldsymbol{A}\boldsymbol{d}_j = 0\,, \quad i \neq j\,, \quad i, j \geq 0$$

The proof is by mathematical induction. We first assume that the equation (10) produces mutually conjugate directions $\boldsymbol{d}_0, \ldots \boldsymbol{d}_k$. we shall prove

that $\boldsymbol{d}_0, \ldots \boldsymbol{d}_{k+1}$ are mutually conjugate. Using equations (7) and (10), and the relation

$$\boldsymbol{d}_0 = -\boldsymbol{H}\,\boldsymbol{g}_0$$

we get

$$\boldsymbol{d}_j = \sum_{i=0}^{j} \alpha_i (\boldsymbol{H}\,\boldsymbol{A})^i \boldsymbol{H}\,\boldsymbol{g}_0 \quad 0 \le j \le k\,, \tag{12}$$

where $(\boldsymbol{H}\,\boldsymbol{A})^0$ is defined to be $\boldsymbol{I}$. First we show that

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_j = 0\,. \tag{13}$$

For $j = k$

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k = 0$$

since $\boldsymbol{x}_{k+1}$ minimizes f in the direction $\boldsymbol{d}_k$. For $0 \le j \le k-1$

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_j = \left(\boldsymbol{g}_{j+1}^T + \sum_{i=j+1}^{k} \alpha_i \boldsymbol{d}_i^T \boldsymbol{A} \right) \boldsymbol{d}_j = 0$$

This completes the verification of equation (13). In view of equations (12) and (13), we have

$$\boldsymbol{g}_{k+1}^T (\boldsymbol{H}\,\boldsymbol{A})^j \boldsymbol{H}\,\boldsymbol{g}_0 = 0 \quad 0 \le j \le k\,. \tag{14}$$

For $1 \le j \le k$ we can write equation (14) in the form

$$(\boldsymbol{H}\,\boldsymbol{g}_{k+1})^T \boldsymbol{A} (\boldsymbol{H}\,\boldsymbol{A})^{j-1} \boldsymbol{H}\,\boldsymbol{g}_0 = \boldsymbol{0}\,.$$

Hence, $\boldsymbol{H}\,\boldsymbol{g}_{k+1}$ is mutually conjugate to $(\boldsymbol{H}\,\boldsymbol{A})^j \boldsymbol{H}\,\boldsymbol{g}_0\,,\ \ 0 \le j \le k-1$. In view of equation (12), $\boldsymbol{H}\,\boldsymbol{g}_{k+1}$ is also mutually conjugate to $\boldsymbol{d}_0, \ldots \boldsymbol{d}_{k-1}$. Setting $i = k$ in equation (10) and multiplying by $\boldsymbol{A}\,\boldsymbol{d}_j$, $j = 1, \ldots, k-1$, we get

$$\boldsymbol{d}_{k+1} \boldsymbol{A}\,\boldsymbol{d}_j = 0\,,$$

and in the case $j = k$,

$$\boldsymbol{d}_{k+1}^T \boldsymbol{A}\,\boldsymbol{d}_k = 0 \tag{15}$$

since

$$\beta_k = \frac{(\boldsymbol{H}\,\boldsymbol{g}_{k+1})^T \boldsymbol{A}\,\boldsymbol{d}_k}{\boldsymbol{d}_k^T \boldsymbol{A}\,\boldsymbol{d}_k}\,.$$

We have shown that our algorithm produces mutually conjugate directions $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_k$ then it produces mutually conjugate directions $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k+1}$. It only remains to prove that $\boldsymbol{d}_0$ and $\boldsymbol{d}_1$ are conjugate directions and this is clear by putting $k = 0$ in equation (15). The proof of the next theorem will use properties of the method given below; the proofs of these properties are trivial, and hence, are not given.

Properties

(1) $\boldsymbol{C}^T \boldsymbol{d}_k = \boldsymbol{0}$, this will ensure that every iteration point is feasible.
(2) $\boldsymbol{C}^T \boldsymbol{H} = \boldsymbol{0}$.
(3) $\boldsymbol{d}_i$ are linearly independent.

Theorem 1: *The proposed method minimizes f in equation (1) subject to the constraint given in equation (2) in at most $n - m$ iterations.*

Proof: Let $s = n - m$. Applying equation (6) repeatedly, we have

$$\boldsymbol{x}_s = \boldsymbol{x}_{r+1} + \sum_{q=r+1}^{s-1} \alpha_q \boldsymbol{d}_q\,, \quad 0 \le r < s-1$$

Hence

$$\boldsymbol{g}_s^T \boldsymbol{d}_r = \boldsymbol{g}_{r+1}^T \boldsymbol{d}_r + \sum_{q=r+1}^{s-1} \alpha_q \boldsymbol{d}_q^T \boldsymbol{A}\,\boldsymbol{d}_r$$

Using Proposition 1 and the fact that $\boldsymbol{g}_{r+1}^T \boldsymbol{d}_r = 0$, we get

$$\boldsymbol{g}_s^T \boldsymbol{d}_r = 0 \tag{16}$$

$$\boldsymbol{g}_s^T \boldsymbol{d}_{s-1} = 0\,,$$

since α_{s-1} is determined to minimize f along the direction $\boldsymbol{d}_{s-1}$. Let the linear operator $\boldsymbol{T} : R^n \to R^m$ be multiplication by $\boldsymbol{C}^T$, where R^n and R^m are n and m dimensional vector spaces. Properties 1 and 2 show that

$$\boldsymbol{d}_k\,, \boldsymbol{H}\,\boldsymbol{d}_k \in ker\boldsymbol{T}$$

where $ker\boldsymbol{T}$ is the kernal of $\boldsymbol{T}$. Hence we can write

$$\boldsymbol{H}\,\boldsymbol{d}_k = \sum_{i=0}^{s-1} \lambda_i \boldsymbol{d}_i\,, \quad 0 \le k \le s-1 \tag{17}$$

since $\boldsymbol{d}_i$ are linearly independent. By virtue of equations (16) and (17), and using the symmetry of $\boldsymbol{H}$, we have

$$(\boldsymbol{H}\,\boldsymbol{g}_s)^T \boldsymbol{d}_k = 0 \tag{18}$$

Equation (18) implies that

$$\boldsymbol{H}\,\boldsymbol{g}_s = \boldsymbol{0}$$

since $\boldsymbol{H}\boldsymbol{g}_s \in ker\boldsymbol{T}$. The minimum may be obtained earlier if λ_i all turn out to be zero at the end part of the iterations.
We shall now derive a rate of convergence for the constrained conjugate gradient method.
Theorem 2. *For the constrained conjugate gradient method, we have the error estimate*

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_{\boldsymbol{A}} \leq 2\left(\frac{\sqrt{k_2}-1}{\sqrt{k_2}+1}\right)^k \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_{\boldsymbol{A}}\,, \quad (19)$$

where

$$\|\boldsymbol{x}\| = (\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x})^{\frac{1}{2}}$$

$\boldsymbol{x}^*$ is the optimum solution and the condition number $k_2 = \dfrac{\lambda_{max}(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})}{\lambda_{min}(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})}$. Further, if $\rho(\epsilon)$ is defined for any $\epsilon > 0$ to be the smallest integer k such that

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_{\boldsymbol{A}} \leq \epsilon\|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_{\boldsymbol{A}} \quad (20)$$

then

$$\rho(\epsilon) \leq \frac{1}{2}\sqrt{C(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})}ln(\frac{2}{\epsilon}) + 1\,, \quad (21)$$

where $C(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})$ is the condition number of $\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}$.
Proof: Consider the transformation of variable for feasible $\boldsymbol{x}$ by the relation

$$\boldsymbol{x} = \boldsymbol{Z}\boldsymbol{y} + \boldsymbol{W}\boldsymbol{q}^* \quad (22)$$

Substituting (**) in (**) we get

$$f = \frac{\boldsymbol{y}^T\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}\boldsymbol{y}}{2} + \boldsymbol{y}^T\boldsymbol{Z}^T(\boldsymbol{A}\boldsymbol{W}\boldsymbol{q}^* - \boldsymbol{b}) +$$

$$constant \quad (23)$$

The constrained minimization problem given by equations (1) and (2) is equivalent to the unconstrained minimization of f given by equation (23) for $n-m$ variables. If we apply the unconstrained algorithm we have:
Initialise:

$$\boldsymbol{y} = \boldsymbol{y}_0$$

$$\hat{\boldsymbol{g}}_0 = \boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}\boldsymbol{y}_0 + \boldsymbol{Z}^T(\boldsymbol{A}\boldsymbol{W}\boldsymbol{q}^* - \boldsymbol{b})$$

$$\hat{\boldsymbol{d}}_0 = -\hat{\boldsymbol{g}}_0$$

where $\hat{\boldsymbol{g}}_i = \boldsymbol{Z}^T\boldsymbol{g}_i$, $\hat{\boldsymbol{d}}_i = \boldsymbol{Z}\boldsymbol{d}_i$, $i = 0, 1, \ldots$.
For $i = 0, 1, \ldots$

$$\alpha_i = -\frac{\hat{\boldsymbol{g}}_i^T\boldsymbol{d}_i}{\hat{\boldsymbol{d}}_i^T\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}\hat{\boldsymbol{d}}_i}$$

$$\boldsymbol{y}_{i+1} = \boldsymbol{y}_i + \alpha_i\hat{\boldsymbol{d}}_i$$

$$\hat{\boldsymbol{g}}_{i+1} = \boldsymbol{Z}^T\boldsymbol{g}_{i+1}$$

$$\beta_i = \frac{\hat{\boldsymbol{g}}_{i+1}^T\hat{\boldsymbol{g}}_{i+1}}{\hat{\boldsymbol{g}}_i^T\hat{\boldsymbol{g}}_i}$$

$$\hat{\boldsymbol{d}}_{i+1} = -\hat{\boldsymbol{g}}_{i+1} + \beta_i\hat{\boldsymbol{d}}_i$$

if $\hat{\boldsymbol{d}}_{i+1}^T\hat{\boldsymbol{d}}_{i+1} < tolerance \times \hat{\boldsymbol{d}}_0^T\hat{\boldsymbol{d}}_0$ stop
end for.
The above algorithm can be immediately transformed to the proposed algorithm by multiplying the equations of the above algorithm by $\boldsymbol{Z}$ and taking

$$\boldsymbol{x}_0 = \boldsymbol{Z}\boldsymbol{y}_0 + \boldsymbol{W}\boldsymbol{q}^*$$

We note that the $\boldsymbol{y}_k$ values are related to the $\boldsymbol{x}_k$ values of the proposed algorithm via the equation (22).An error extimate of the above algorithm is well known [4], i.e.,

$$\|\boldsymbol{y}_k - \boldsymbol{y}^*\|_{\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}} \leq$$

$$2\left(\frac{\sqrt{k_2}-1}{\sqrt{k_2}+1}\right)^k \|\boldsymbol{y}_0 - \boldsymbol{y}^*\|_{\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}}\,, \quad (24)$$

where $\boldsymbol{y}^*$ is the optimum solution. If $\rho(\epsilon)$ is defined for any $\epsilon > 0$ to be the smallest integer k such that

$$\|\boldsymbol{y}_k - \boldsymbol{y}^*\|_{\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}} \leq \epsilon\|\boldsymbol{y}_0 - \boldsymbol{y}^*\|_{\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}} \quad (25)$$

then

$$\rho(\epsilon) \leq \frac{1}{2}\sqrt{C(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})}ln(\frac{2}{\epsilon}) + 1\,, \quad (26)$$

It can be easily shown that

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_{\boldsymbol{A}} = \|\boldsymbol{y}_k - \boldsymbol{x}^*\|_{\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}}\,. \quad (27)$$

Substitution of equation (27) in equations (25) and (26) completes the proof.

3 Parallelization

In this paper we only concentrate on evaluating $\boldsymbol{H} = \boldsymbol{H}_m$. Since

$$\boldsymbol{H}_m = \boldsymbol{I} - \boldsymbol{C}(\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T$$

We can evaluate $\boldsymbol{H}_m$ iteratively using the algorithm [5]

$$\boldsymbol{H}_i = \boldsymbol{H}_{i-1} - \frac{\boldsymbol{H}_{i-1}\boldsymbol{c}_i\boldsymbol{c}_i^T\boldsymbol{H}_{i-1}}{\boldsymbol{c}_i^T\boldsymbol{H}_{i-1}\boldsymbol{c}_i},$$

$$i = 1, 2 \ldots, m, \tag{28}$$

where $\boldsymbol{H}_0 = \boldsymbol{I}$ and $\boldsymbol{c}_i$ are columns of $\boldsymbol{C}$. If the constrained matrix $\boldsymbol{C}$ is large we can parallelized the iteration (28) easily since it involves matrix-vector multiplications and outer products.
For the constrained conjugate gradient algorithm given in Section 2.1, all the matrix-vector multiplications and the inner products are parallelized. It obvious that these parallelizations can be easily done.

4 A Numerical Example

To give confidence to our readers we give a simple numerical example to indicate that our theory is valid. For simplicity we let

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$\boldsymbol{C}^T = [1, 1, 1]$ and $\boldsymbol{b} = 1$. The proposed algorithm converges in two iterations using tolerance $= 10^{-06}$.

REFERENCES

[1] D.J. Grey, M.H.B.M. Shariff, *"Parallel Inversion of Finite Element Matrices via a Direct Method"*, Advances in Computational Strutures Technology, pp 359-364, Civil-Comp Press, Edinburgh, 1996.

[2] M.H.B.M. Shariff, D.J.Grey, *"Parallel Direct Method for Inversion of Matrices"*, Advances in Engineering Software, **30**, 1999, 839-845.

[3] M.H.B.M. Shariff, *" A Constrained Conjugate Gradient Method and the Solution of Linear Equations"*, Comp. Math. Applic., **30**, 1995, 25-37.

[4] O. Axelsson, V.A. Barker *Finite Element Solution of Boundary Value Problems*, Academic Press, New York, 1984

[5] D. Goldfarb, *"Extension of Davidon's Variable Metric Method to Maximization under Linear Inequality and Equality Constraints"*, SIAM Journal on Applied athematics, **17**, 1969, 739-764.