

Conjugate Gradient Algorithm for the Symmetric Arrowhead Solution of Matrix Equation $AXB=C$

Minghui Wang, Luping Xu, Juntao Zhang

Abstract—Based on the conjugate gradient (CG) algorithm, the constrained matrix equation $AXB=C$ and the associate optimal approximation problem are considered for the symmetric arrowhead matrix solutions in the premise of consistency. The convergence results of the method are presented. At last, a numerical example is given to illustrate the efficiency of this method.

Keywords—Iterative method, symmetric arrowhead matrix, conjugate gradient algorithm.

I. INTRODUCTION

LET $R^{m \times n}$ be the set of $m \times n$ real matrices, $SAR^{n \times n}$ be the set of $n \times n$ real symmetric arrowhead matrices and I_n be the identity matrix of order n . For any $A \in R^{m \times n}$, A^T , A^\dagger , $\|A\|_F$ and $\|A\|_2$ denote the transpose, Moore-Penrose generalized inverse, Frobenius norm and Euclid norm, respectively.

For any $A, B \in R^{m \times n}$, $\langle A, B \rangle = \text{trace}(B^T A) = 0$ denotes the inner product of A and B . Therefore, $R^{m \times n}$ is a complete inner product space endowed with $\|A\|^2 = \langle A, A \rangle$. For any non-zero matrices $A_1, A_2, \dots, A_k \in R^{m \times n}$, if $\langle A_j, A_i \rangle = \text{trace}(A_i^T A_j) = 0 (i \neq j)$, then it is easy to verify that A_1, A_2, \dots, A_k are linearly independent and orthogonal.

Proposition 1. Let $A, B \in R^{n \times n}$, then
 $\text{trace}(A) = \text{trace}(A^T)$; $\text{trace}(AB) = \text{trace}(BA)$
 $\text{trace}(A+B) = \text{trace}(A) + \text{trace}(B)$

Definition 1. If a matrix $A = (a_{ij}) \in R^{n \times n}$ satisfies the following form:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

then we denote that A is arrowhead matrix, this type of matrix

This work is supported by the Natural Science Foundation of Shandong Province in China (ZR2013FL014, ZR2016AB07) and the Science and Technology Program of Shandong Universities of China (J13LN73)

M. Wang, L. Xu and J. Zhang are with Department of Mathematics, Qingdao University of Science & Technology, Shandong 266061, China (e-mail: mhwang@yeah.net; qustxuluping@163.com; qustzhangjuntao@163.com).

set is denoted as $AR^{n \times n}$. If $a_{ii} = a_i (i = 1, 2, \dots, n)$, then we denote that A is symmetric arrow-head matrix, this type of matrix set is denoted as $SAR^{n \times n}$. $\text{vec}_i(A)$ denotes a vector $(a_{11}, a_{21}, \dots, a_{n1}, a_{22}, a_{33}, \dots, a_{nn})^T$.

Symmetric arrowhead matrices have many applications in the modern control theory which can represent the parameter matrix of nonlinear control systems or the large sparse matrix in the linear systems [1], [5]-[7]. With the development of electromagnetic compatibility, the mathematical representation of the influence factors of electromagnetic interference also has potential application value. In this paper, we consider the following constrained matrix equation

$$AXB = C \quad (1)$$

in which $A \in R^{m \times n}$, $B \in R^{n \times s}$, $C \in R^{m \times s}$. The above matrix equation and other constrained matrix equations have been studied in [2], [3], [8], [9], etc. Peng et al. [4] analyzed CG algorithm to obtain corresponding symmetric solutions, skew-symmetric solutions, centro-symmetric solutions and so on. Based on the classical method, we will utilize the operable iterative method to find the symmetric arrowhead matrix solution of the matrix equation (1).

II. THE CONJUGATE GRADIENT ALGORITHM

In this section, by means of the study of the classical CG algorithm for solving the linear matrix equation in [4], we propose the following algorithm to solve the matrix equation (1) for the symmetric arrowhead solution and give some main results in detail.

Firstly, we define the following linear operator:

$$\Gamma : \begin{cases} R^{n \times n} & \rightarrow AR^{n \times n} \\ X & \rightarrow \Gamma(X) \end{cases} \quad (2)$$

in which $\Gamma(X) = E_{11}X + XE_{11} + \text{diag}(X) - 2E_{11}XE_{11}$. According to the properties of the inner product matrix, it is easy to verify that

$$\langle X, Y \rangle = \langle X, \Gamma(Y) \rangle = \langle \Gamma(X), Y \rangle$$

in which $X \in R^{n \times n}$, $Y \in AR^{n \times n}$. Here we discuss the iterative algorithm of the matrix equation (1) as:

Algorithm CG-W.

Step1. Initialization. For initial matrix $X_1 \in \mathbf{SAR}^{n \times n}$, compute

$$\begin{aligned} R_1 &= C - AX_1B, \\ P_1 &= \frac{A^T R_1 B^T + (A^T R_1 B^T)^T}{2}, \\ Q_1 &= \Gamma(P_1) = E_{11}P_1 + P_1E_{11} + \text{diag}(P_1) - 2E_{11}P_1E_{11}. \end{aligned}$$

Step2. Iteration. For $k = 1, 2, \dots$, compute

$$X_{k+1} = X_k + \frac{\|R_k\|^2}{\|Q_k\|^2} Q_k,$$

Step3. Compute

$$\begin{aligned} R_{k+1} &= C - AX_{k+1}B, \\ P_{k+1} &= \frac{A^T R_{k+1} B^T + (A^T R_{k+1} B^T)^T}{2}, \\ Q_{k+1} &= \Gamma(P_{k+1}) - \frac{\text{trace}(P_{k+1}^T Q_k)}{\|Q_k\|^2} Q_k. \end{aligned}$$

if $R_{k+1} = 0$ or $R_{k+1} \neq 0, Q_{k+1} = 0$, stop; otherwise continue to step (2).

By algorithm CG-W it is clear that:

$$P_i \in \mathbf{SR}^{n \times n}, Q_i \in \mathbf{SAR}^{n \times n}, X_i \in \mathbf{SAR}^{n \times n}, i = 1, 2, \dots$$

The following will demonstrate that the algorithm CG-W is terminated by a finite iterative step.

Lemma 1. The sequences $\{R_i\}$ and $\{Q_i\}$ generalized by Algorithm CG-W satisfy

$$\text{trace}(R_{i+1}^T R_j) = \text{trace}(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} \text{trace}(Q_i P_j), \quad i, j = 1, 2, \dots \quad (3)$$

Proof: Since $P_i^T = P_i, Q_i^T = Q_i$, by Algorithm CG-W, we have

$$\begin{aligned} \text{trace}(R_{i+1}^T R_j) &= \text{trace}\left[(C - AX_{i+1}B)^T R_j\right] = \text{trace}\left[\left(C - A\left(X_i + \frac{\|R_i\|^2}{\|Q_i\|^2} Q_i\right)B\right)^T R_j\right] \\ &= \text{trace}\left[(C - AX_i B)^T R_j + \frac{\|R_i\|^2}{\|Q_i\|^2} B^T Q_i A^T R_j\right] \\ &= \text{trace}(R_i^T R_j) + \frac{\|R_i\|^2}{\|Q_i\|^2} \text{trace}\left(\frac{Q_i A^T R_j B^T + (Q_i A^T R_j B^T)^T}{2}\right) \\ &= \text{trace}(R_i^T R_j) + \frac{\|R_i\|^2}{\|Q_i\|^2} \text{trace}\left(\frac{Q_i [A^T R_j B^T + (A^T R_j B^T)^T]}{2}\right) \\ &= \text{trace}(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} \text{trace}(Q_i P_j). \quad \square \end{aligned}$$

Lemma 2. For $k \geq 2$, the sequences $\{R_i\}, \{Q_i\}$ generalized by Algorithm CG-W satisfy

$$\text{trace}(R_i^T R_j) = 0, \text{trace}(Q_i^T Q_j) = 0, \quad i, j = 1, 2, \dots, k, \quad i \neq j \quad (4)$$

Proof: We shall prove this lemma by induction.

First, notice that $P_i \in \mathbf{SR}^{n \times n}, Q_i \in \mathbf{SAR}^{n \times n}$, by Lemma 1 and Algorithm CG-W, we obtain

$$\begin{aligned} \text{trace}(R_2^T R_1) &= \text{trace}(R_1^T R_1) - \frac{\|R_1\|^2}{\|Q_1\|^2} \text{trace}(Q_1 P_1) \\ &= \|R_1\|^2 - \frac{\|R_1\|^2}{\|Q_1\|^2} \text{trace}(Q_1 Q_1) = \|R_1\|^2 - \frac{\|R_1\|^2}{\|Q_1\|^2} \text{trace}(Q_1^T Q_1) = 0, \end{aligned}$$

as well as

$$\begin{aligned} \text{trace}(Q_2^T Q_1) &= \text{trace}\left[\left(\Gamma(P_2) - \frac{\text{trace}(P_2^T Q_1)}{\|Q_1\|^2} Q_1\right)^T Q_1\right] \\ &= \text{trace}\left((\Gamma(P_2))^T Q_1\right) - \text{trace}(P_2 Q_1) = 0. \end{aligned}$$

Suppose that (4) holds for $k = s \geq 2$ and notice that $\text{trace}(Q_s^T Q_{s-1}) = 0$. According to Lemma 1, we have

$$\begin{aligned} \text{trace}(R_{s+1}^T R_s) &= \text{trace}(R_s^T R_s) - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s P_s) = \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s \Gamma(P_s)) \\ &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}\left[Q_s \left(Q_s + \frac{\text{trace}(P_s^T Q_{s-1})}{\|Q_{s-1}\|^2} Q_{s-1}\right)\right] \\ &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \left[\text{trace}(Q_s^T Q_s) + \frac{\text{trace}(P_s^T Q_{s-1})}{\|Q_{s-1}\|^2} \text{trace}(Q_s^T Q_{s-1})\right] \\ &= \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s^T Q_s) = 0, \end{aligned}$$

and

$$\begin{aligned} \text{trace}(Q_{s+1}^T Q_s) &= \text{trace}\left[\left(\Gamma(P_{s+1}) - \frac{\text{trace}(P_{s+1}^T Q_s)}{\|Q_s\|^2} Q_s\right)^T Q_s\right] \\ &= \text{trace}\left((\Gamma(P_{s+1}))^T Q_s\right) - \text{trace}(P_{s+1}^T Q_s) = 0. \end{aligned}$$

Thus, by Lemma 1, it is clear that $\text{trace}(R_{s+1}^T R_j) = 0$ when $j = 1$. And we notice that $\text{trace}(R_s^T R_j) = 0, \text{trace}(Q_s^T Q_j) = 0, \text{trace}(Q_s^T Q_{j-1}) = 0$ for $j = 2, 3, \dots, s-1$. By Lemma 1 and Algorithm CG-W, we have

$$\begin{aligned} \text{trace}(R_{s+1}^T R_j) &= \text{trace}(R_s^T R_j) - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s P_j) = -\frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s \Gamma(P_j)) \\ &= -\frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace} \left[Q_s \left(Q_j + \frac{\text{trace}(P_j^T Q_{j-1})}{\|Q_{j-1}\|^2} Q_{j-1} \right) \right] \\ &= -\frac{\|R_s\|^2}{\|Q_s\|^2} \left[\text{trace}(Q_s^T Q_j) + \frac{\text{trace}(P_j^T Q_{j-1})}{\|Q_{j-1}\|^2} \text{trace}(Q_s^T Q_{j-1}) \right] = 0, \end{aligned}$$

and

$$\begin{aligned} \text{trace}(Q_{s+1}^T Q_j) &= \text{trace} \left[\left(\Gamma(P_{s+1}) - \frac{\text{trace}(P_{s+1}^T Q_s)}{\|Q_s\|^2} Q_s \right)^T Q_j \right] \\ &= \text{trace} \left[(\Gamma(P_{s+1}))^T Q_j \right] - \frac{\text{trace}(P_{s+1}^T Q_s)}{\|Q_s\|^2} \text{trace}(Q_s^T Q_j) \\ &= \text{trace}(P_{s+1} Q_j) = \text{trace}(Q_j P_{s+1}). \end{aligned}$$

which gives

$$\begin{aligned} \text{trace}(Q_{s+1}^T Q_j) &= \text{trace}(Q_j P_{s+1}) = \frac{\|R_j\|^2}{\|Q_j\|^2} \left[\text{trace}(R_j^T R_{s+1}) - \text{trace}(R_{j+1}^T R_{s+1}) \right] \\ &= \frac{\|R_j\|^2}{\|Q_j\|^2} \left[\text{trace}(R_{s+1}^T R_j) - \text{trace}(R_{s+1}^T R_{j+1}) \right] = 0, \end{aligned}$$

Thus, (4) holds when $k = s + 1$. By the induction, we know that (4) holds for when $i, j = 1, 2, \dots, k, i \neq j$.

Lemma 3. Suppose that the equation is consistent and X^* is one solution of (1), then the sequences $\{R_i\}$ and $\{Q_i\}$ generalized by Algorithm CG-W satisfy

$$\text{trace} \left[(X^* - X_k) Q_k \right] = \|R_k\|^2, \quad k = 1, 2, \dots \quad (5)$$

Proof: We shall also prove this lemma by induction. First of all, when $k = 1$ we have

$$\begin{aligned} \text{trace} \left[(X^* - X_1) Q_1 \right] &= \text{trace} \left[(X^* - X_1) \Gamma(P_1) \right] = \text{trace} \left[(X^* - X_1) P_1 \right] \\ &= \text{trace} \left[(X^* - X_1) \left(\frac{A^T R_1 B^T + (A^T R_1 B^T)^T}{2} \right) \right] \\ &= \text{trace} \left[\frac{(B R_1^T A (X^* - X_1))^T + ((X^* - X_1) B R_1^T A)}{2} \right] \\ &= \text{trace} \left[\frac{(A (X^* - X_1) B R_1^T)^T + (A (X^* - X_1) B R_1^T)}{2} \right] \\ &= \text{trace} \left[A (X^* - X_1) B R_1^T \right] = \text{trace} \left[(C - A X_1 B) R_1^T \right] = \text{trace} (R_1 R_1^T) = \|R_1\|^2, \end{aligned}$$

Suppose that (5) holds when $k = s$. Owing to

$$\begin{aligned} \text{trace} \left[(X^* - X_{s+1}) Q_s \right] &= \text{trace} \left[\left(X^* - X_s - \frac{\|R_s\|^2}{\|Q_s\|^2} Q_s \right) Q_s \right] \\ &= \text{trace} \left[(X^* - X_s) Q_s \right] - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s Q_s) = \|R_s\|^2 - \frac{\|R_s\|^2}{\|Q_s\|^2} \text{trace}(Q_s^T Q_s) = 0, \end{aligned}$$

we obtain

$$\begin{aligned} \text{trace} \left[(X^* - X_{s+1}) Q_{s+1} \right] &= \text{trace} \left[(X^* - X_{s+1}) \left(\Gamma(P_{s+1}) - \frac{\text{trace}(P_{s+1}^T Q_s)}{\|Q_s\|^2} Q_s \right) \right] \\ &= \text{trace} \left[(X^* - X_{s+1}) \Gamma(P_{s+1}) \right] - \frac{\text{trace}(P_{s+1}^T Q_s)}{\|Q_s\|^2} \text{trace} \left[(X^* - X_{s+1}) Q_s \right] \\ &= \text{trace} \left[(X^* - X_{s+1}) P_{s+1} \right] = \text{trace} \left[(X^* - X_{s+1}) \left(\frac{A^T R_{s+1} B^T + (A^T R_{s+1} B^T)^T}{2} \right) \right] \\ &= \text{trace} \left[\frac{(B R_{s+1}^T A (X^* - X_{s+1}))^T + ((X^* - X_{s+1}) B R_{s+1}^T A)}{2} \right] \\ &= \text{trace} \left[(C - A X_{s+1} B) R_{s+1}^T \right] = \text{trace} (R_{s+1} R_{s+1}^T) = \|R_{s+1}\|^2. \end{aligned}$$

By the induction, we know that (5) holds for $k = 1, 2, \dots$.

Theorem 1. Suppose that the matrix equation (1) is consistent and for any initial matrix $X_1 \in \mathbf{SAR}^{n \times n}$, the sequence $\{X_k\}$ generated by Algorithm CG-W converges to a solution of (1) after finite-steps.

Proof: The proof is by contradiction. Assume that $R_i \neq 0$, $i = 1, 2, \dots, mp$, then by Lemma 3, we have $Q_i \neq 0$, $i = 1, 2, \dots, mp$ and can further obtain X_{mp+1} and R_{mp+1} . If $R_{mp+1} \neq 0$, then according to Lemma 2 we get the orthogonal basis matrix set $\{R_1, R_2, \dots, R_{mp}, R_{mp+1}\}$ of $R_{m \times p}$, which contradicts the assumption. Thus, $R_{mp+1} = 0$, and X_{mp+1} is the exact solution of (1).

Theorem 2. Suppose that the matrix equation (1) is consistent, then we take the initial matrix

$$X_1 = \Gamma(P_1), \quad P_1 = (A^T H B^T + B H^T A) / 2,$$

with any $H \in \mathbf{R}^{m \times s}$ (or specially, for $X_1 = 0 \in \mathbf{R}^{n \times n}$), the Algorithm CG-W converges to the minimum norm solution of (1) after finite-steps.

Proof: If we take $X_1 = \Gamma(P_1)$, $P_1 = (A^T H B^T + B H^T A) / 2$ with any $H \in \mathbf{R}^{m \times s}$, by Algorithm CG-W, we can get a solution \hat{X} of the matrix $AXB = C$ after finite-steps, and there exists the matrix $\hat{H} \in \mathbf{R}^{m \times s}$, such that $\hat{X} = \Gamma(\hat{P})$ with $\hat{P} = (A^T \hat{H} B^T + B \hat{H}^T A) / 2$. From $A(\hat{X} + \tilde{X})B = A\hat{X}B + A\tilde{X}B$ we know that all the symmetric arrowhead solution of matrix equation $AXB = C$ can be expressed as $\hat{X} + \tilde{X}$ with $\tilde{X} \in \mathbf{SAR}^{n \times n}$, satisfying $A\tilde{X}B = 0$. For $\tilde{X} = \Gamma(\tilde{X})$, we

get $A\Gamma(\tilde{X})B = 0$, thus we have

$$\begin{aligned} \langle \hat{X}, \tilde{X} \rangle &= \langle \Gamma(\hat{P}), \tilde{X} \rangle = \left\langle \Gamma\left(\frac{A^T \hat{H}^T B^T + B \hat{H} A}{2}\right), \tilde{X} \right\rangle \\ &= \langle \hat{H}, A\Gamma(\tilde{X})B \rangle = 0. \end{aligned}$$

thus we get

$$\|\hat{X} + \tilde{X}\|^2 = \|\hat{X}\|^2 + \|\tilde{X}\|^2 \geq \|\hat{X}\|^2$$

\hat{X} is the symmetric arrowhead minimum norm solution of (1). It is not difficult to verify the solution set of (1) is a closed convex set, therefore, the symmetric arrowhead minimum norm solution of (1) is unique.

III. NUMERICAL EXPERIMENTS

In this section, under the compatibility condition of the constrained matrix equation $AXB = C$, we give an example to illustrate the efficiency and investigate the performance of Algorithm CG-W which has been shown to be numerically reliable in various circumstances. All functions are defined by Matlab 7.0 and all codes are calculated with machine precision around 10^{-9} .

Example 1. Given $A = [\text{toeplitz}(1:30*i), \text{zeros}(30*i, 11*i)]$ of row full rank, $B = [\text{eye}(40*i); \text{ones}(i, 40*i)]$ of column full rank for $i = 1, 2, \dots, 5$ and $C = AXB$. Given $\bar{Y} = 0.5 \text{ones}(n, n)$ and $\Gamma(X) = E_{11}\bar{Y} + P_1\bar{Y} + \text{diag}(\bar{Y}) - 2E_{11}\bar{Y}E_{11}$. Notice that in this case, the matrix equation $C = AXB$ is consistent and has a unique minimum norm solution.

TABLE I
 THE ITERATIVE STEPS, ITERATIVE TIME AND RESIDUAL NORM OF THE ALGORITHM CG-W

		CG-W
$i = 1$	Iter	94
	CPU	0.282
	$\ R_k\ $	3.8626e-008
$i = 2$	Iter	249
	CPU	1.690
	$\ R_k\ $	4.1348e-008
$i = 3$	Iter	420
	CPU	7.995
	$\ R_k\ $	8.8017e-008
$i = 4$	Iter	609
	CPU	23.669
	$\ R_k\ $	9.5693e-008
$i = 5$	Iter	820
	CPU	62.574
	$\ R_k\ $	9.7150e-008

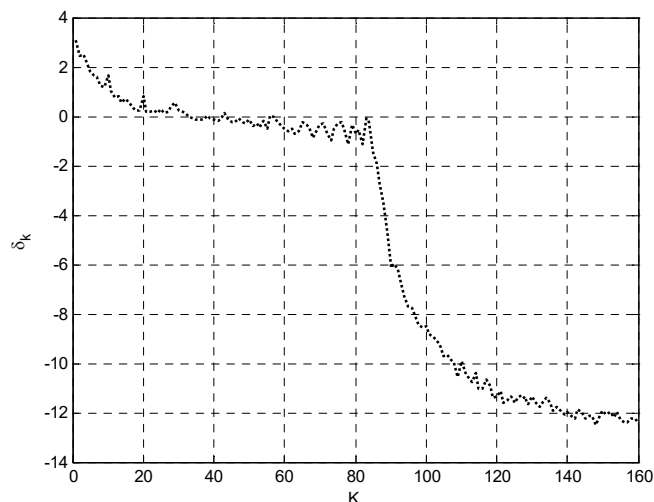


Fig. 1 Relation between error γ_k and iterative number K when $i = 1$

In Table I, we obtain iterative steps (Iter), Iterative time (CPU) and residual norm $\|R_k\| = \|AX_k B - C\|_F$ of the algorithm respectively. We set a stop criterion for $\|R_k\| = \|AX_k B - C\|_F \leq 10^{-7}$. Then, Fig. 1 plots the relation between error $\gamma_k = \log_{10}(\|AXB - C\|)$ and the iterative number K when $i = 1$.

We choose the initial matrix $X_0 = \text{zeros}(41*i, 41*i)$, the unique minimal norm solution of the matrix equation (1) is obtained by the algorithm CG-W. It can be seen from the Table I, when the order of the matrices A and B is growing exponentially, the iterative steps of the algorithm CG-W is growth multiples.

REFERENCES

- [1] Y.F. Xu, An inverse eigenvalue problem for a special kind of matrices. *Math. Appl.*, 1(1996)68-75.
- [2] C.J. Meng, X.Y. Hu, L. Zhang, The skew symmetric orthogonal solution of the matrix equation $AXB=C$, *Linear Algebra Appl.* 402(2005)303-318.
- [3] Li Jiaofen, Zhang Xiaoning, Peng Zhenyun, Alternative projection algorithm for single variable linear constraints matrix equation problems, *Mathematica Numerical Since*, 36(2014)143-162.
- [4] Y.X. Peng, X.Y. Hu, L. Zhang, An iteration method for the symmetric solutions and the optimal approximation solution of the matrix equation $AXB=C$, *Applied Mathematics and Computation*, 160(2005)763-777.
- [5] Z.Y. Peng, A matrix LSQR iterative method to solve matrix equation $AXB=C$, *International Journal of Computer Mathematics*, 87(2010)1820-1830.
- [6] J.F. Li, X.F. Duan, L. Zhang, Numerical solutions of $AXB=C$ for mirror symmetric matrix X under a specified submatrix constraint. *Computing*, 90(2010) 39-56.
- [7] Von Neumann J., *Functional Operators. II. The Geometry of Spaces*, *Annals of Mathematics Studies*, vol.22, Princeton University Press, Princeton, 1950.
- [8] Cheney W., Goldstein A. Proximity maps for convex sets, *Proceedings of the American Mathematical Society*, 10(1959)448-450.
- [9] Hongyi Li, Zongsheng Gao, Di Zhao, Least squares solutions of the matrix equation with the least norm for symmetric arrowhead matrices. *Appl. Math. Comput.*, 226(2014)719-724.