INSTITUTE OF APPLIED MATHEMATICS AND INFORMATICS
Technical University of Sofia
XXIV Summer School
Applications of Mathematics in Engineering
Sozopol June 14-20, 1998

Positive Definite Solutions of the Equation $X - A^*\sqrt{X^{-1}}A = I$ [1]

I. Ivanov, B. Minchev, V. Hasanov

Abstract: The matrix equation $X - A^*\sqrt{X^{-1}}A = I$ in this paper is studied. There is an iterative method for obtaining of a positive definite solution of this equation. Sufficient conditions for existence of positive definite solutions are proved. Results of numerical expiriments are given.

Keywords: matrix equation, iterative method, positive definite solution

1. Introduction

We consider the matrix equation

$$X - A^*\sqrt{X^{-1}}A = I, \tag{1}$$

where I is $n \times n$ a unit matrix and A is $n \times n$ a invertible matrix. We shall study the equation (1) for the existence of a Hermitian positive definite solution X, $(X > 0)$.

In many physical applications we must solve a system of linear equations [1]

$$Mx = f \tag{2}$$

where the positive definite matrix M arises from a finite difference approximation to an elliptic partial differential equation. As an example, let

$$M = \begin{pmatrix} I & A \\ A^* & I \end{pmatrix}.$$

We consider the matrix $M = \tilde{M} + diag[I - X, 2I]$ where

$$\tilde{M} = \begin{pmatrix} X & A \\ A^* & -I \end{pmatrix}.$$

We can decompose the matrix $\tilde{M}$ via the following way

$$\begin{pmatrix} X & A \\ A^* & -I \end{pmatrix} = \begin{pmatrix} I & 0 \\ A^*X^{-1} & I \end{pmatrix} \begin{pmatrix} X & A \\ 0 & -X^2 \end{pmatrix}. \tag{3}$$

In order to exists the decompositon (3) the matrix X must be a solution of the matrix equation $Y - A^*\sqrt{Y^{-1}}A = I,\ X = \sqrt{Y}$.

---

[1]This work is partially supported by Contract MM 521/95 with the Bulgarian Ministry of Education, Sciences and Technologies

We find a LU-decomposition to the matrix M. The solving of the system $\tilde{M}y = f$ is transformed to the solving of two linear systems that have a left block coefficient matrix and a right block coefficient matrix, respectively. For computing the solution of (2) the Woodbury formula [3] can be applied.

In this paper we propose an iterative method which is converged to a positive definite solution of (1). The rate of convergence of these methods depend of the parameter α. Numerical examples are discussed and results of experiments are given. We study the equation (1) of a positive definite solution because the solving of linear systems having a positive definite matrix is numerically stable [7].

2. Solution of the matrix equation

We will describe an iterative method which is suitable for obtaining to a positive definite solution of the equation (1). We start with some properties which will be used throughout this paper.

(i) If $P \geq Q > 0$ then $P^{-1} \leq Q^{-1}$.

(ii) If $P \geq Q > 0$ then $\sqrt{P} \geq \sqrt{Q}$.

Consider the sequence of the following matrices

$$X_0 = \alpha I, \; X_{k+1} = I + A^* \sqrt{X_k^{-1}} A, \quad k = 0, 1, 2, \ldots. \tag{4}$$

We will prove the following theorems

Theorem 1. *If there is a real α so that $\alpha > 1$ and*

(i) $\sqrt{\alpha}(\alpha - 1)I < A^*A$,

(ii) $\frac{\sqrt{\alpha}}{(\alpha-1)^2}(AA^*)^2 - A^*A > \sqrt{\alpha} I$,

(iii) $\|A\|^2 < 2\alpha\sqrt{\alpha}$.

Then the equation (1) has a positive definite solution.

Proof. We consider the sequence (4). For X_1 we have

$$X_1 = I + \frac{1}{\sqrt{\alpha}} A^*A.$$

From the condition (i) we obtain

$$X_0 = \alpha I < I + \frac{1}{\sqrt{\alpha}} A^*A = X_1$$

Hence $X_0 < X_1$.

For X_2 we have

$$\begin{aligned} X_2 &= I + A^* \sqrt{X_1^{-1}} A \\ &= I + A^* \sqrt{(I + \frac{1}{\sqrt{\alpha}} A^*A)^{-1}} A \end{aligned}$$

Applying the condition (ii) yeilds

$$\begin{aligned} \frac{1}{\sqrt{\alpha}} A^*A + I &< \frac{1}{(\alpha-1)^2}(AA^*)^2 \\ \sqrt{(\frac{1}{\sqrt{\alpha}} A^*A + I)^{-1}} &> (\alpha - 1) A^{-*}A^{-1} \\ X_2 = I + A^* \sqrt{(\frac{1}{\sqrt{\alpha}} A^*A + I)^{-1}} A &> \alpha I = X_0. \end{aligned}$$

Consequently $X_0 < X_2$.

Using $X_0 < X_1$ we obtain

$$\begin{array}{rcl} X_0^{-1} & > & X_1^{-1} \\ A^*\sqrt{X_0^{-1}}A & > & A^*\sqrt{X_1^{-1}}A \\ X_1 & > & X_2. \end{array}$$

Hence $X_0 < X_2 < X_1$.

We receive by analogy

$$X_1 > X_3$$

and

$$X_3 > X_2.$$

Consequently $X_0 < X_2 < X_3 < X_1$.

We receive by analogy that for each two integer numbers s, k is satisfied

$$X_0 \leq X_{2k} < X_{2k+2} < X_{2s+3} < X_{2s+1} \leq X_1.$$

Consequently the subsequences $\{X_{2k}\},\ \{X_{2s+1}\}$ are convergent ones to positive definite matrices. These sequences have a common boundary. We have

$$\begin{array}{rcl} \|X_{2k+1} - X_{2k}\| & = & \|A^*(\sqrt{X_{2k}^{-1}} - \sqrt{X_{2k-1}^{-1}})A\| \\ & = & \|A^*\sqrt{X_{2k}^{-1}}(\sqrt{X_{2k-1}} - \sqrt{X_{2k}})\sqrt{X_{2k-1}^{-1}}A\| \\ & \leq & \|A\|^2\|\sqrt{X_{2k}^{-1}}\|\|\sqrt{X_{2k-1}^{-1}}\|\|\sqrt{X_{2k-1}} - \sqrt{X_{2k}}\|. \end{array}$$

We consider the equation

$$\sqrt{X_{2k-1}}(\sqrt{X_{2k-1}} - \sqrt{X_{2k}}) + (\sqrt{X_{2k-1}} - \sqrt{X_{2k}})\sqrt{X_{2k}} = X_{2k-1} - X_{2k}.$$

Since $X_{2k+1} > X_{2s}$ for each k, s then $Y = \sqrt{X_{2k-1}} - \sqrt{X_{2k}}$ is a positive definite soluiton of the matrix equation

$$\sqrt{X_{2k-1}}\ Y + Y\ \sqrt{X_{2k}} = X_{2k-1} - X_{2k}.$$

According to theorem 8.5.2 [4] we have

$$Y = \int_0^\infty e^{-\sqrt{X_{2k-1}}t}(X_{2k-1} - X_{2k})e^{-\sqrt{X_{2k}}t}\ dt. \tag{5}$$

Since $X_0 < X_s < X_1$ are positive definite matrices then

$$\sqrt{X_0^{-1}} > \sqrt{X_s^{-1}},\ s = 0, 1, 2, \ldots$$

and

$$\|\sqrt{X_s^{-1}}\| \leq \frac{1}{\sqrt{\alpha}}.$$

Then

$$\begin{aligned}
\|X_{2k+1}-X_{2k}\| &\leq \frac{1}{\alpha}\|A\|^2\|\int_0^\infty e^{-\sqrt{X_{2k-1}}t}(X_{2k-1}-X_{2k})e^{-\sqrt{X_{2k}}t}\,dt\| \\
&\leq \frac{1}{\alpha}\|A\|^2\frac{1}{2\sqrt{\alpha}}\|X_{2k-1}-X_{2k}\| \\
&\leq (\frac{1}{2\alpha\sqrt{\alpha}}\|A\|^2)^{2k}\|X_1-X_0\| \\
&\leq (\frac{1}{2\alpha\sqrt{\alpha}}\|A\|^2)^{2k}\|\frac{1}{\sqrt{\alpha}}A^*A+(1-\alpha)I\|.
\end{aligned}$$

Consequently

$$\|X_{2k+1}-X_{2k}\| \leq (\frac{1}{2\alpha\sqrt{\alpha}}\|A\|^2)^{2k}\|\frac{1}{\sqrt{\alpha}}A^*A+(1-\alpha)I\|.$$

and

$$\|X_{2k+1}-X_{2k}\| \to 0, \quad k\to\infty.$$

Hence

$$\max(\|X_{2k+1}-X\|,\ \|X-X_{2k}\|) \leq (\frac{1}{2\alpha\sqrt{\alpha}}\|A\|^2)^{2k}\|\frac{1}{\sqrt{\alpha}}A^*A+(1-\alpha)I\|.$$

Remark. In the case $\alpha=1$ the conditions (i) and (ii) of Theorem 1 are satisfied and then the equation (1) has a positive definite solution.

Theorem 2. If there is a real β so that $\beta>1$ and

(i) $A^*A<\sqrt{\beta}(\beta-1)I$,
(ii) $\frac{\sqrt{\beta}}{(\beta-1)^2}(AA^*)^2-\sqrt{\beta}I<A^*A$,
(iii) $\|A\|^2<2\rho\sqrt{\rho}$,

where ρ is the minimal eigenvalue of the matrix $I+\frac{1}{\sqrt{\beta}}A^*A$. Then the equation (1) has a positive definite solution.

Proof. The theorem is proved analogous of the theorem 1 as we consider the iterative method (4) with $X_0=\beta I$.

3. Numerical experiments

We made numerical experiments for computing of a positive definite solution of the equation (1). The solution is computed for different matrices A and different values of n. Denote X the solution which is obtained by the iterative method (4), i.e.

$$X_{k+1}=I+A^*\sqrt{X_k^{-1}}A, \quad X_0=\alpha I, \quad k=0,1,2,\ldots$$

and m_X be the smallest number k, for which

$$\|X_k-X\| \leq \left(\frac{\|A\|^2}{2\alpha\sqrt{\alpha}}\right)^k \|\frac{1}{\sqrt{\alpha}}A^*A+(1-\alpha)I\| \leq 10^{-5}.$$

Denote Y the solution which is obtained by the iterative method (4), in case . $X_0=\beta I$ and m_Y be the smallest number r, for which

$$\|X_r-Y\| \leq \left(\frac{\|A\|^2}{2\rho\sqrt{\rho}}\right)^r \|(\beta-1)I-\frac{1}{\sqrt{\beta}}A^*A\| \leq 10^{-5},$$

where ρ is the minimal eigenvalue of the matrix $I + \frac{1}{\sqrt{\beta}}A^*A$.

Denote ε the norm

$$\varepsilon = \|X_{m_X} - X_{m_Y}\|_\infty.$$

We can consider decomposition of the matrix M,

$$\begin{pmatrix} X & A \\ A^* & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ A^*X^{-1} & I \end{pmatrix} \begin{pmatrix} X & A \\ 0 & X \end{pmatrix}.$$

This decomposition leads to solving the matrix equation $X + A^TX^{-1}A = I$. Furthermore we solve this matrix equation with same matrices A. We compute the positive definite solution [2] by

$$X_0 = I, \quad X_{p+1} = I - A^TX_p^{-1}A, \; p = 0, 1, \ldots. \tag{6}$$

We denote k_X the smallest number p so that

$$\|X_p - X\| \leq \frac{1}{2}\left(4\|A\|^2\right)^p \leq 10^{-5}.$$

This inequality follows immediately from the recursion problem (2b) of [2] and Lemma 4 [2].

Example 1. Let A has the form

$$A = (a_{ij}) = \begin{cases} a_{ij} = \frac{2(2n+i)}{n^3} & i = j \\ a_{ij} = \frac{2(i+j+n)}{n^3} & i \neq j \end{cases}$$

Example 2. Let A has the form

$$A = diag[\frac{1}{2+1}, \frac{2}{2.2+1}, \ldots, \frac{n}{2n+1}]$$

Example 3. Let A has the form

$$A = (a_{ij}) = \begin{cases} a_{ij} = 25(1 - \frac{i(n-i)}{20^2n}) & i = j \\ a_{ij} = \frac{(n-i)}{30^3n} & i \neq j \end{cases}$$

Example 4. Let A has the form

$$A = (a_{ij}) = \begin{cases} a_{ij} = 3(1 - \frac{i}{10n^2}) & i = j \\ a_{ij} = \frac{(i-j)}{10n^3} & i \neq j \end{cases}$$

The results from experiments are given in the following tables.

Table 1.

	Example 1 ($\alpha = 1$)					Example 2 ($\alpha = 1.1$)				
n	k_X	β	m_X	m_Y	ϵ	k_X	β	m_X	m_Y	ϵ
5	–	1.8	14	14	$1.730e-10$	57	1.20	4	4	$4.219e-6$
10	40	1.19	5	5	$5.613e-7$	111	1.21	5	5	$5.362e-7$
15	10	1.09	3	3	$5.262e-6$	165	1.22	5	5	$6.615e-7$
20	7	1.05	3	3	$5.579e-7$	220	1.22	5	5	$7.058e-7$
25	5	1.03	2	2	$6.674e-6$	274	1.22	5	5	$7.339e-7$

Table 2.

	Example 3					Example 4				
n	α	β	m_X	m_Y	ϵ	α	β	m_X	m_Y	ϵ
5	16.05	16.27	15	15	$2.608e-6$	4.89	5	11	11	$4.651e-6$
10	16.05	16.48	16	16	$2.444e-6$	4.95	5.02	11	11	$2.972e-6$
15	4.86	5.32	15	16	$6.444e-7$	4.97	5.02	10	10	$5.305e-6$
20	1	1.17	5	4	$3.798e-7$	4.98	5.02	10	10	$4.243e-6$
25	2.03	2.61	14	16	$4.776e-8$	4.99	5.02	10	10	$3.180e-6$

4. Conclusion

In this paper we consider a nonlinear matrix equation. LU-decompositon (3) leads to the computing of a positive definite solution of the equation (1). We introduced a recursion algorithm from which a positive definite solution can be calculated. When the matrix A satisfy theorem 1 or theorem 2 then we receive the solution of the equation (1) faster than the solution of the equation $X + A^*X^{-1}A = I$. There are matrices A (examples 3 and 4) for wich the iterative method (4) is convergence but the iterative method (6) is not convergence.

In theorems 2 and 3 bounds are given in term of a parameter α. The rate of convergence of the described iterative method (4) depends of the parameter α.

References

[1] Buzbee B.L., Golub G.H. and Nielson C.W., On Direct Methods for Solving Poisson's Equations, *SIAM J. Numer. Anal.* 7:627-656 (1970).

[2] Engwerda J.C., On the Existence of a Positive Definite Solution of the Matrix Equation $X + A^TX^{-1}A = I$, *Linear Algebra Appl.* 194:91-108 (1993).

[3] Housholder A.S., *The theory of matrices in numerical analysis.* Blaisdell, NewYork (1964).

[4] Lankaster P., *Theory of Matrices.* Academic Press. New York (1969).

[5] Parodi M., *La localisation des valeurs caracteristiques des matrices et ses applications.* Paris, Gauthier-Villars, 1959.

[6] Petkov M. On the matrix equations $A_0X^m + A_1X^{m-1} + \ldots + A_m = 0$, $(A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m)x = 0$. *Annuaire of Sofia University "Kl. Okhridski",* 72:159-164 (1978) (in Bulgarian).

[7] Wilkinson J.H., *The algebric eigenvalue problem.* Oxford University Press, London (1965).

Ivan Gantchev Ivanov
Borislav Ventcislavov Minchev
Vezhdi Ismailov Hasanov
Faculty of Mathematics and Informatics
Shoumen University, Shoumen 9712, Bulgaria
e-mail: i.gantchev@fmi.uni-shoumen.acad.bg