A New Inversion-free Method for Hermitian Positive Definite Solution of Matrix Equation

Minghui Wang, Juntao Zhang

Abstract—An inversion-free iterative algorithm is presented for solving nonlinear matrix equation with a stepsize parameter t. The existence of the maximal solution is discussed in detail, and the method for finding it is proposed. Finally, two numerical examples are reported that show the efficiency of the method.

Keywords—Inversion-free method, Hermitian positive definite solution, Maximal solution, Convergence.

I. INTRODUCTION

I N this paper, we will consider the nonlinear matrix equation

$$X + A^*X^{-\alpha}A = I, \qquad (1)$$

where $A \in \mathbf{C}^{n \times n}, \alpha \in (0,1)$ and X is an Hermitian positive definite (HPD) matrix which will be found. Equation (1) has many applications in dynamic programming, statistics, stochastic filtering, control theory, Kalman filtering and so on, see [1]–[3], [9] and references therein.

In [1], Engwerda discussed the equation $X + A^*X^{-1}A = I$ and gave the algorithm

$$X_0 = I,$$
$$X_{n+1} = I - A^*X_n^{-1}A, n = 0, 1, \cdots.$$

which has a solution if and only if $X_n > AA^*$ for all n.

For (1), there also have been many results, see [4]–[8], [10], [11]. Xinguo Liu and Hua Gao [10] analyzed the equation

$$X^s \pm A^T X^{-t}A = I_n$$

and proved the existence of the symmetric positive definite solutions based on the fixed-point theory, then proposed iterative method for computing the positive solutions. In [6], Marlliny Monsalve and Marcos Raydan developed a inversion-free method for obtaining the minimal HPD solution of the matrix rational equation

$$X + A^*X^{-1}A = I$$

where I is the identity matrix and A is a given nonsingular matrix, then discussed stability properties when the method starts from the available matrix AA^*.

Minghui Wang and Musheng Wei [7] investigated the matrix equation

$$X^s + A^*X^{-q}A = I$$

M. Wang and J. Zhang are with the Department of Mathematics, Qingdao University of Science & Technology, Shandong 266061, China, (e-mail: mhwang@yeah.net).

where $A \in C^{m \times n}$, s, $q \in R^+$ and discussed the existence of the maximal solution X_L and minimal solution X_S that $X_S \le X \le X_L$ for any HPD solutions X. Here $X \ge Y$ means that $X - Y$ is Hermitian positive semidefinite and if a Hermitian matrix Z satisfies $Y \le Z \le X$, then we say $Z \in [Y, X]$.

Inspired by the methods proposed by Peng, El-Sayed [5] and L. Zhang [12], we propose an inversion-free iterative method for (1), discuss the existence of the maximal solution and the method to find it.

In the remainder of this section, we review some important results for later discussion:

Lemma 1. *If $A > B > 0$ (or $A \ge B > 0$), then $A^\alpha > B^\alpha$ (or $A^\alpha \ge B^\alpha > 0$) for all $\alpha \in (0, 1]$, and $A^\alpha < B^\alpha$ (or $0 < A^\alpha \le B^\alpha$) for all $\alpha \in [-1, 0)$.*

Lemma 2. *If C and P are Hermitian matrices of the same order with $P > 0$, then $CPC + P^{-1} \ge 2C$.*

Lemma 3. *If $0 < \alpha \le 1$, and P and Q are positive definite matrices of the same order with $P, Q \ge bI > 0$, then $\|P^\alpha - Q^\alpha\| \le \alpha b^{\alpha-1}\|P - Q\|$.*

For $A \in \mathbf{C}^{n \times n}$, $\lambda_{\max}(A)$, $\lambda_{\min}(A)$ and $\|A\|$ denote the maximal eigenvalue, the minimal eigenvalue and the spectral norm of A, respectively.

II. A NEW ITERATIVE METHOD

In this section, we propose an inversion-free iterative algorithm to find the maximal solution for the nonlinear matrix equation $X + A^*X^{-\alpha}A = I$. Based on the relevant results in [7], it is easy to obtain the following results.

Theorem 1. *Suppose that $\|A\|_2 \le \left(\frac{\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{1}{1+\alpha}\right)^{\frac{1}{2}}$ and X is an HPD solution of (1). Then we have*
1. when A is nonsingular matrix,

$$X \in [\alpha_1 I, \beta_1 I] \cup [\beta_2 I, \alpha_2 I] \cup \{X : \alpha_1 \le \lambda_{min}(X) \le \beta_1, \beta_2 \le \lambda_{max}(X) \le \alpha_2\};$$

2. when A is singular matrix,

$$X \in \{X : 0 \le \lambda_{min}(X) \le \beta_1, \lambda_{max}(X) = 1\} \cup \{X : \beta_2 \le \lambda_{min}(X) \le 1, \lambda_{max}(X) = 1\};$$

where α_1, α_2 and β_1, β_2 are the solutions of

$$x^\alpha(1 - x) = \lambda_{min}(A^*A)$$

and

$$x^\alpha(1 - x) = \lambda_{max}(A^*A)$$

respectively, which satisfy

$$0 < \alpha_1 \le \beta_1 \le \frac{\alpha}{1 + \alpha} \le \beta_2 \le \alpha_2 < 1.$$

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:8, No:6, 2014

Theorem 2. *1. When $\|A\|_2 \leq \left(\frac{\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{1}{1+\alpha}\right)^{\frac{1}{2}}$, (1) has an HPD solution $X_l \in [\beta_2 I, \alpha_2 I]$;*

2. when $\|A\|_2 < \left(\frac{\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{1}{1+\alpha}\right)^{\frac{1}{2}}$, X_l is unique and can be obtained by the following algorithm:

$$\begin{cases} X_0 \in \left[\frac{\alpha}{1+\alpha} I, I\right] \\ X_{n+1} = I - A^* X_n^{-\alpha} A, n = 0, 1, \cdots . \end{cases} \quad (2)$$

Moreover, these doesn't exist other solution X such that $X \geq X_l$.

Now we propose the following algorithm with a stepsize paramenter $t \in (0, 1]$ and give the proof of convergence in detail.

Algorithm 1.
Step 1: Select $Y_0 = I, t \in (0, 1]$ and let $k = 1$.
Step 2: Let

$$\begin{cases} X_k = I - A^* Y_k^\alpha A \\ Y_{k+1} = (1+t)Y_k - t Y_k X_k Y_k. \end{cases} \quad (3)$$

Set $k := k + 1$ and go to Step 2.

When $t = 1$, the above algorithm reduces to Algorithm 3.1 of [7]. So the former can be viewed as the generalization of the latter.

For Algorithm 1, we have the following results.

Theorem 3. *If given A is a nonsingular matrix, $\|A\|_2 \leq \left(\frac{\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{1}{1+\alpha}\right)^{\frac{1}{2}}$ and $\alpha \in (0, 1)$, then (1) has the maximal solution X_l, which can be generated by Algorithm 1.*

Proof. We prove this theorem by induction. From Theorem 2, we know that (1) has the HPD solution.

Let X is an Hermitian positive definite solution of (1). Then $X \leq I$ and hence

$$A^* X^{-\alpha} A \geq A^* A,$$

$$I - A^* X^{-\alpha} A \leq I - A^* A,$$

that is

$$X \leq I - A^* A.$$

It follows from (3) we have

$$X_0 = I - A^* A \geq X$$

and

$$X_0^{-1} \leq X^{-1}, X_0 \leq I, Y_0 \leq X_0^{-1} \leq X^{-1}.$$

By Lemma 1 and Lemma 2, we have

$$\begin{aligned} Y_1 &= (1+t)Y_0 - t Y_0 X_0 Y_0 \\ &= (1-t)Y_0 + t(2Y_0 - Y_0 X_0 Y_0) \\ &\leq (1-t)X_0^{-1} + t X_0^{-1} \\ &= X_0^{-1} \leq X^{-1} \end{aligned}$$

and

$$\begin{aligned} Y_1 - Y_0 &= t Y_0 - t Y_0 X_0 Y_0 \\ &= t Y_0 (Y_0^{-1} - X_0) Y_0 \geq 0, \end{aligned}$$

that is $Y_1 \leq X^{-1}, Y_1 \geq Y_0$ and we know that

$$Y_0 \leq Y_1 \leq X_0^{-1} \leq X^{-1}.$$

It follows from Lemma 1 that

$$\begin{aligned} X_1 &= I - A^* Y_1^\alpha A \\ &\geq I - A^* X^{-\alpha} A = X \end{aligned}$$

and

$$\begin{aligned} X_1 - X_0 &= I - A^* Y_1^\alpha A - (I - A^* Y_0^\alpha A) \\ &= A^* (Y_0^\alpha - Y_1^\alpha) A \leq 0, \end{aligned}$$

which mean $X_0 \geq X_1 \geq X$.

From the above, we get

$$Y_0 \leq Y_1 \leq X_0^{-1} \leq X_1^{-1} \leq X^{-1}.$$

Assume that

$$Y_{k-1} \leq Y_k \leq X_{k-1}^{-1} \leq X_k^{-1} \leq X^{-1},$$

we can find that

$$\begin{aligned} Y_{k+1} &= (1-t)Y_k + t(2Y_k - Y_k X_k Y_k) \\ &\leq (1-t)Y_k + t X_k^{-1} \\ &\leq (1-t)X_k^{-1} + t X_k^{-1} \\ &\leq X_k^{-1} \end{aligned}$$

and

$$\begin{aligned} X_{k+1} &= I - A^* Y_k^\alpha A \\ &\geq I - A^* X^{-1} A = X. \end{aligned}$$

Further, based on the assumption, we obtain

$$Y_{k+1} - Y_k = t Y_k (Y_k^{-1} - X_k) Y_k \geq 0$$

and

$$\begin{aligned} X_{k+1} - X_k &= (I - A^* Y_{k+1}^\alpha A) - (I - A^* Y_k^\alpha A) \\ &= A^* (Y_k^\alpha - Y_{k+1}^\alpha) A \leq 0. \end{aligned}$$

Therefore, we obtain that

$$Y_k \leq Y_{k+1} \leq X_k^{-1} \leq X_{k+1}^{-1} \leq X^{-1}$$

and

$$Y_0 \leq Y_1 \leq \cdots Y_k \leq Y_{k+1} \leq X^{-1},$$

$$X_0 \geq X_1 \geq \cdots X_k \geq X_{k+1} \geq X$$

hold for all $k = 0, 1, 2 \ldots$ and so the limits of $\{X_k\}$ and $\{Y_k\}$ exist.

Taking the limit in the Algorithm 1, we can obtain that $\lim_{k \to \infty} X_k$ is an HPD solution of (1). Moreover, since

$$X_0 \geq \cdots X_k \geq X$$

holds for any HPD solution X of (1), then we have

$$\lim_{k \to \infty} X_k = X_l$$

and since

$$Y_0 \leq \cdots Y_k \leq X^{-1}$$

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:8, No:6, 2014

holds for any X^{-1}, we also have

$$\lim_{k \to \infty} Y_k = X_l^{-1}. \square$$

Theorem 4. *Suppose* $\|A\|_2 \le \left(\frac{\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{1}{1+\alpha}\right)^{\frac{1}{2}}$ *for the case* $\alpha \in (0,1)$ *and after k steps of Algorithm 1, we have* $\|I - X_k Y_k\| < \varepsilon$, *then*

$$\|X_k + A^* X_k^{-\alpha} A - I\| < \varepsilon \alpha \beta_2^{-1} \|A\|^2. \tag{4}$$

Proof. From Theorem 1 and Theorem 2, we know that $X_l \in [\beta_2 I, \alpha_2 I]$. Thus, according to the proof of Theorem 3, we can obtain

$$I \le Y_k \le Y_{k+1} \le X_k^{-1} \le X_{k+1}^{-1} \le X^{-1} \le \beta_2^{-1} I.$$

Since $X_{k+1} + A^* Y_{k+1}^\alpha A = I$ and

$$
\begin{aligned}
X_k &+ A^* X_k^{-\alpha} A - I \\
&= X_k - X_{k+1} + A^* (X_k^{-\alpha} - Y_{k+1}^\alpha) A \\
&= A^* (Y_{k+1}^\alpha - Y_k^\alpha) A + A^* (X_k^{-\alpha} - Y_{k+1}^\alpha) A \\
&= A^* (X_k^{-\alpha} - Y_k^\alpha) A,
\end{aligned}
$$

by Lemma 3 we have

$$
\begin{aligned}
\|X_k + A^* X_k^{-\alpha} A - I\| &\le \|A\|^2 \|X_k^{-\alpha} - Y_k^\alpha\| \\
&\le \alpha \|A\|^2 \|X_k - Y_k\| \\
&\le \alpha \|A\|^2 \|X_k^{-1}\| \|I - X_k Y_k\| \\
&< \varepsilon \alpha \beta_2^{-1} \|A\|^2. \square
\end{aligned}
$$

III. NUMERICAL EXAMPLE

In this section, we give two examples to illustrate the efficiency and investigate the performance of Algorithm 1 with different t and α values for (1). Example 1 and Example 2 will give the maximal solution and the performance with different t and α. All codes are calculated by MATLAB with machine precision around 10^{-16} and let the residual

$$\eta_k = \|X + A^* X^{-\alpha} A - T\|_F < 10^{-16} \tag{5}$$

as the practical stopping criterion, where $\|\cdot\|_F$ stands for the Frobenius norm.

Example 1. Given nonsingular matrix

$$
A = \begin{pmatrix}
0.071 & 0.028 & 0.08 & 0.034 \\
-0.065 & 0.031 & 0.047 & 0.056 \\
0.023 & -0.05 & 0.019 & 0.025 \\
-0.012 & 0.035 & -0.06 & 0.041
\end{pmatrix}.
$$

When $\alpha = 0.5$ and $t = 0.8$, using Algorithm 1 and iterating 20 steps, we have the matrix solution of (1) which is $X \approx X_{20} =$

$$
\begin{pmatrix}
0.9900 & 0.0016 & -0.0038 & 0.0011 \\
0.0016 & 0.9945 & -0.0006 & -0.0029 \\
-0.0038 & -0.0006 & 0.9874 & -0.0034 \\
0.0011 & -0.0029 & -0.0034 & 0.9934
\end{pmatrix},
$$

with the residual $\eta_{20} = \|X_{20} + A^* X_{20}^{-0.5} A - I\|_F < 2.0032 \times 10^{-17}$.

From Algorithm 1, when $\alpha = 0.5$, we choose different t values with $t \in (0, 2)$ and list the numerical results in the following Table I. Then Fig. 1 plots the relation between η_k with different t values and iteration number K.
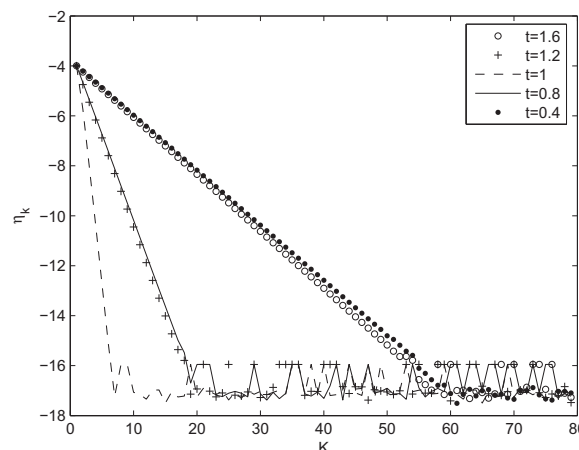

Fig. 1 THE CONVERGENCE OF SOLUTIONS WITH DIFFERENT t VALUES WHEN $\alpha = 0.5$

TABLE I
INTERATIVE STEP AND PRECISION FOR $\alpha = 0.5$

T	ITERATION	RESIDUAL
1.6	56	2.7575×10^{-17}
1.4	32	3.6981×10^{-17}
1.2	20	7.2342×10^{-17}
1.1	14	2.3183×10^{-17}
1	8	0.5887×10^{-17}
0.9	15	1.5036×10^{-17}
0.7	25	8.6194×10^{-17}
0.5	42	6.5984×10^{-17}
0.4	56	7.8284×10^{-17}

When $\alpha = 0.25$, we list the numerical results in the following Table II.

TABLE II
INTERATIVE STEP AND PRECISION FOR $\alpha = 0.25$

T	ITERATION	RESIDUAL
1.6	54	5.3183×10^{-17}
1.4	32	2.4819×10^{-17}
1.2	19	2.6263×10^{-17}
1.1	14	1.4509×10^{-17}
1	7	1.7200×10^{-17}
0.9	14	5.3063×10^{-17}
0.7	25	3.3580×10^{-17}
0.5	41	6.1947×10^{-17}
0.4	55	5.9060×10^{-17}

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:8, No:6, 2014

Example 2. Given nonsingular matrix

$$B = \begin{pmatrix} 0.08 & 0.02 & -0.03 & 0.04 & 0 & 0.07 \\ -0.07 & 0.03 & 0.04 & -0.06 & 0.02 & 0.08 \\ 0.02 & 0.03 & 0.04 & 0.05 & 0.01 & 0.03 \\ -0.01 & 0.02 & 0.03 & -0.03 & 0.04 & 0.05 \\ 0.02 & 0.01 & -0.02 & 0.07 & 0.06 & 0.03 \\ 0.03 & -0.05 & 0.06 & 0.04 & 0.2 & 0.06 \end{pmatrix}.$$

When $\alpha = 0.5$ and $t = 0.8$, using Algorithm 1 and iterating 22 steps, we have the matrix solution of (1) which is $X \approx X_{22} =$

$$\begin{pmatrix} 98.68 & 0.14 & 0.33 & -1.15 & -0.58 & -0.26 \\ 0.14 & 99.48 & 0.08 & 0.14 & 0.77 & -0.30 \\ 0.33 & 0.08 & 99.09 & 0.15 & -1.33 & -0.69 \\ -1.15 & 0.14 & 0.15 & 98.46 & -1.07 & -0.27 \\ -0.58 & 0.77 & -1.33 & -1.07 & 95.35 & -1.81 \\ -0.26 & -0.30 & -0.69 & -0.27 & -1.81 & 98.05 \end{pmatrix} \times 0.01,$$

with the residual $\eta_{22} = \|X_{22} + A^* X_{22}^{-0.5} A - I\|_F < 9.1243 \times 10^{-17}$.

From Algorithm 1, when $\alpha = 0.5$, we choose different t values and list the numerical results in the following Table III. Then Fig. 2 plots the relation between η_k with different t values and iteration number K.
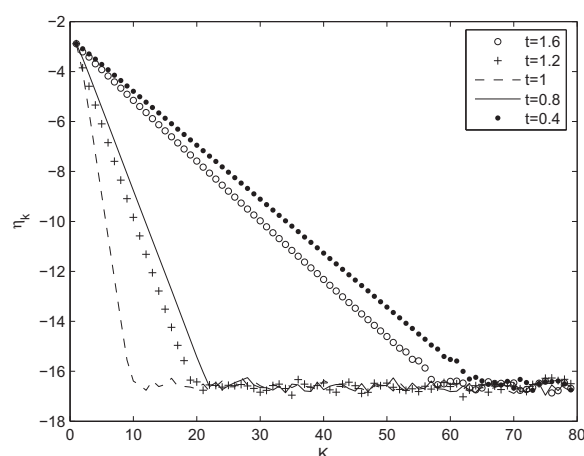

Fig. 2 THE CONVERGENCE OF SOLUTIONS WITH DIFFERENT t VALUES WHEN $\alpha = 0.5$

TABLE III
INTERATIVE STEP AND PRECISION FOR $\alpha = 0.5$

T	ITERATION	RESIDUAL
1.6	58	4.7411×10^{-17}
1.4	33	5.4275×10^{-17}
1.2	20	4.4258×10^{-17}
1.1	15	3.3239×10^{-17}
1	11	4.0280×10^{-17}
0.9	17	4.8098×10^{-17}
0.7	28	9.1601×10^{-17}
0.5	47	4.8974×10^{-17}
0.4	63	9.8793×10^{-17}

When $\alpha = 0.75$, we list the numerical results in the following Table IV.

TABLE IV
INTERATIVE STEP AND PRECISION FOR $\alpha = 0.75$

T	ITERATION	RESIDUAL
1.6	60	3.0437×10^{-17}
1.4	34	4.9502×10^{-17}
1.2	20	2.1350×10^{-17}
1.1	16	5.0963×10^{-17}
1	12	4.0159×10^{-17}
0.9	18	5.1166×10^{-17}
0.7	30	7.8826×10^{-17}
0.5	49	6.1751×10^{-17}
0.4	65	7.8643×10^{-17}

From Table I and Table II, we know that different stepsize parameter t affects the performance of Algorithm 1 significantly. With given fixed constant α for all $\alpha \in (0, 1)$, we can find that the algorithm get better performance when t tends to 1. In this paper we only prove that Algorithm 1 converges when $t \in (0, 1]$, but we can run Algorithm 1 when t decreasing towards 1 in practical computation, which provides more choices for practical problems.

REFERENCES

[1] J.C. Engwerda, A.CM. Ran, A.L. Rijkeboer, *Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^* X^{-1} A = I$*, Linear Algebra Appl., vol. 186, 1993, pp.255-275.
[2] X. Zhan, J. Xie, *On the matrix equation $X + A^T X^{-1} A = I$*, Linear Algebra Appl., vol. 247, 1996, pp.337-345.
[3] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, Spring-Berlin 1997
[4] X. Liu, H. Gao, *On the positive definite solutions of the equation $X^s \pm A^T X^{-t} A = I$*, Linear Algebra Appl., vol. 368, 2003, pp.83-97.
[5] Z. Peng, S. M. EL-Sayed and X. Zhang, *Iterative methods for the extremal positive definite solution of the matrix equation $X + A^* X^{-\alpha} A = Q$*, Comput. Appl. Math., vol. 200, 2007, pp.520-527.
[6] M. Monsalve and M. Raydan, *A new inversion-free method for a rational matrix equation*, Linear Algebra Appl., vol. 433, 2010, pp.64-71.
[7] Minghui Wang, Musheng Wei, Shanrui Hu, *The extremal solution of the matrix equation $X^s + A^* X^{-q} A = I$*, Appl. Math. Comput., vol. 200, 2013, pp.193-199.
[8] J.F. Wang, Y.H. Zhang, B.R. Zhu, *The Hermitian positive definite solutions of matrix equation $X + A^* X^{-q} A = I(q > 0)$*, Mathematica Numerica Sinica, vol. 26, 2004, pp.31-73.
[9] C.H. Guo, P. Lancaster, *Iterative solution of two matrix equations*, Math. Comput., vol. 228, 1999, pp.1589-1603.
[10] X. Liu, H. Gao, *On the positive definite solutions of the equation $X^s \pm A^T X^{-t} A = I$*, Linear Algebra Appl., vol. 368, 2003, pp.83-97.
[11] Y. Yang, *The iterative method for solving nonlinear matrix equation $X^s + A^* X^{-t} A = Q$*, Appl. Math. Comput.,vol. 188, 2007, pp.46-53.
[12] L. Zhang, *An improved inversion-free method for the matrix equation $X + A^* X^{-\alpha} A = Q$*, J. Comput. Appl. Math., vol. 253, 2013, pp.200-203.