ANZIAM J. 44(2003), 145–152

A TWO-SIDED ITERATIVE METHOD FOR COMPUTING POSITIVE DEFINITE SOLUTIONS OF A NONLINEAR MATRIX EQUATION

SALAH M. EL-SAYED[1]

(Received 3 May, 2000; revised 18 January, 2003)

Abstract

In several recent papers, a one-sided iterative process for computing positive definite solutions of the nonlinear matrix equation $X + A^*X^{-1}A = Q$, where Q is positive definite, has been studied. In this paper, a two-sided iterative process for the same equation is investigated. The novel idea here is that the two sequences obtained by starting at two different values provide (a) an interval in which the solution is located, that is, $X_k \le X \le Y_k$ for all k and (b) a better stopping criterion. Some properties of solutions are discussed. Sufficient solvability conditions on a matrix A are derived. Moreover, when the matrix A is normal and satisfies an additional condition, the matrix equation has smallest and largest positive definite solutions. Finally, some numerical examples are given to illustrate the effectiveness of the algorithm.

1. Introduction

We consider the nonlinear matrix equation

$$X + A^*X^{-1}A = Q, \tag{1.1}$$

where Q is a positive definite matrix of order n and A is a nonsingular matrix of order n. Equation (1.1) can be reduced to

$$X + A^*X^{-1}A = I \tag{1.2}$$

(see [6, 11]), where I is the identity matrix. We can see that (1.1) is a special case of a discrete-time algebraic Riccati equation

$$0 = Q + F^*XF - X - (F^*XB + A^*)(R + B^*XB)^{-1}(B^*XF + A),$$

[1]Department of Mathematics, Faculty of Science, Benha university, Benha 13518, Egypt; e-mail: mselsayed@frcu.eun.eg.

https://doi.org/10.1017/S1446181100013201 Published online by Cambridge University Press

where Q is a positive definite matrix, see [9]. This equation can be reduced to (1.1), by setting $F = 0$, $B = I$ and $R = 0$.

The existence of positive definite solutions of (1.1) arises in a number of applications such as system theory, control theory, ladder networks, dynamic programming, stochastic filtering and statistics, see [10] and references therein. Finding an efficient numerical solution for (1.1) is a problem which has been extensively studied by several authors (see [1, 5, 6] for example). Zhan and Xie [11] obtained necessary and sufficient conditions for the existence of the positive definite solution of (1.2). Zhan [10] discussed a new algorithm that avoids matrix inversion for solving (1.2). Guo and Lancaster [7] studied several iterative forms to find the maximal positive definite solutions of the two matrix equations (1.2) and $X - A^*X^{-1}A = Q$. In [8], some properties of a positive definite solution of the equation $X + A^*X^{-2}A = I$ were investigated. A set of equations of the form $X + A^*F(X)A = Q$, where F maps positive definite matrices either into positive or negative definite matrices, and satisfies some monotonicity property were studied in [4]. It was proved that the iteration method converges to a positive definite solution under some conditions. The properties of a positive definite solution of the matrix equation $X - A^*X^{-n}A = I$ were investigated in [3].

In this paper we continue to discuss (1.2) with a two-sided iterative process starting with two different values. In Section 2, we obtain a sufficient condition for the existence of a positive definite solution (1.2). In Section 3, we find a sufficient condition for the existence of the smallest and largest positive definite solutions (1.2), when the matrix A is normal. Some numerical examples are given in Section 4 to illustrate the effectiveness of the algorithm. Conclusions drawn from the results obtained in this paper are in Section 5.

The notation $X > 0$ is here taken to mean that X is a positive definite matrix and $A > B$ is used to indicate that $A - B > 0$. Throughout the paper, $\|.\|$ will be the spectral norm for square matrices unless otherwise noted.

2. The existence solution of the general case

In this section, we will obtain a sufficient condition for the existence of solutions of the matrix equation (1.2). Also, we will prove that the two iterative processes converge to the same limit.

THEOREM 2.1. *If the spectral norm $q = \|A\| < 1/2$, then* (1.2) *has a positive definite solution X, which is a limit of a two-sided iterative process.*

PROOF. Let us consider the following two iterative sequences:

$$X_{k+1} = I - A^*X_k^{-1}A, \qquad X_0 = (1/2)I, \tag{2.1}$$

$$Y_{k+1} = I - A^* Y_k^{-1} A, \qquad Y_0 = I, \tag{2.2}$$

where $k = 0, 1, 2, \ldots$.

In order to prove the theorem, we shall show that

$$X_k < X_{k+1} < Y_{k+1} < Y_k, \quad k = 0, 1, 2, \ldots \quad \text{and} \tag{2.3}$$

$$\|Y_k - X_k\| \to 0 \quad \text{if } k \to \infty. \tag{2.4}$$

Let us begin with (2.1). We have $X_1 = I - 2A^*A$. But since $\|A\| = q < 1/2$, we have $X_1 > X_0 = (1/2)I$.

Suppose that for a fixed k the inequality $X_{k-1} < X_k$ is fulfilled. Then, using the inductive argument and the fact that $X_k > 0$ for any k, we have

$$X_{k-1}^{-1} > X_k^{-1}, \quad I - A^*X_{k-1}^{-1}A > I - A^*X_k^{-1}A.$$

Then we get $X_k < X_{k+1}$. We can prove that $Y_{k+1} < Y_k$ [6] in a similar manner.

Next we shall show that

$$X_k < Y_k \quad \text{and} \quad \|Y_k - X_k\| < \frac{1}{2}\left(\frac{q^2}{(1-2q^2)^2}\right)^k, \tag{2.5}$$

when $k = 0, 1, 2, \ldots$. Indeed,

$$Y_0 - X_0 = (1/2)I > 0, \quad \text{that is,} \quad X_0 < Y_0,$$
$$Y_1 - X_1 = A^*A > 0, \quad \text{that is,} \quad X_1 < Y_1.$$

Let us assume that for a fixed $k > 1$, we have $Y_{k-1} - X_{k-1} > 0$. Thus we have $Y_k - X_k = A^*\left(X_{k-1}^{-1} - Y_{k-1}^{-1}\right)A > 0$, that is, $X_k < Y_k$. Now we have

$$\begin{aligned}
\|Y_k - X_k\| &= \|A^*X_{k-1}^{-1}\,(Y_{k-1} - X_{k-1})\,Y_{k-1}^{-1}A\| \\
&\le \|A\|^2\|X_1^{-1}\|^2\|Y_{k-1} - X_{k-1}\| \\
&\le q^2\frac{1}{(1-2q^2)^2}\|Y_{k-1} - X_{k-1}\| \\
&\le \left(\frac{q^2}{(1-2q^2)^2}\right)^k \|Y_0 - X_0\| \\
&= \frac{1}{2}\left(\frac{q^2}{(1-2q^2)^2}\right)^k,
\end{aligned}$$

since $q^2/(1-2q^2)^2 < 1$.

So we have proved (2.3) and (2.4) hold. This implies that

$$\lim_{k\to\infty} X_k = \lim_{k\to\infty} Y_k = X = I - A^*X^{-1}A > 0.$$

The theorem is proved.

https://doi.org/10.1017/S1446181100013201 Published online by Cambridge University Press

3. The existence solution of the special case

In this section, we propose in addition that $AA^* = A^*A$, that is, that the matrix A is normal.

LEMMA 3.1. *If the matrix A is normal, then $AX = XA$, where X is the solution of* (1.2).

PROOF. From [3, Lemma 4], we have $AX_s = X_sA$, $s = 0, 1, \dots$, where the sequence $\{X_s\}$ is generated from (2.1) or (2.2). Furthermore, from Theorem 2.1 the matrix X is the limit of X_s and Y_s. Then the lemma is proved.

In this special case, by some matrix manipulation it can be shown that

$$X_\infty = \frac{I + \sqrt{I - 4A^*A}}{2} \quad \text{and} \quad X_{-\infty} = \frac{I - \sqrt{I - 4A^*A}}{2} \tag{3.1}$$

are always positive definite if $\|A\| \le 1/2$. These expressions clearly generalize the scalar case.

THEOREM 3.2. *If $q = \|A\| < 1/2$ and A is a normal matrix, then the two iterative processes* (2.1) *and* (2.2) *converge to X_∞.*

PROOF. To prove the theorem, it is sufficient to show that $X_k < X_\infty < Y_k$ for every $k \ge 0$ and to apply Theorem 2.1.

Indeed, we have

$$X_0 = \frac{I}{2} < \frac{I + \sqrt{I - 4A^*A}}{2} = X_\infty.$$

Suppose that for a fixed $k \ge 1$, we have $X_{k-1} < X_\infty$. Hence

$$\begin{aligned} X_{k-1}^{-1} &> X_\infty^{-1}, \\ A^*X_{k-1}^{-1}A &> A^*X_\infty^{-1}A, \\ I - A^*X_{k-1}^{-1}A &< I - A^*X_\infty^{-1}A. \end{aligned}$$

Then we get $X_k < X_\infty$. So, by induction, we have proved that $X_k < X_\infty$.

Similarly we can prove that $X_\infty < Y_k$, $k = 0, 1, 2, \dots$. From these properties of the two sequences X_k and Y_k and Theorem 2.1, we get $X_k \to X_\infty$ and $Y_k \to X_\infty$. This concludes the proof of the theorem.

If we know that X_∞ ($X_{-\infty}$), it is easy to find $X_{-\infty}$ (X_∞) from $X_\infty X_{-\infty} = X_{-\infty}(X_\infty) = A^*A$ or from $X_{-\infty} = I - X_\infty$. Nevertheless it is interesting to construct a two-sided iterative process for finding the positive definite solution $X_{-\infty}$ of (1.2). In order to do this, we rewrite (1.2) in the form

$$X = A(I - X)^{-1}A^*. \tag{3.2}$$

Now we shall show that

$$X_{k+1} = A(I - X_k)^{-1}A^*, \quad k = 0, 1, 2, \dots, \quad X_0 = 0, \tag{3.3}$$

$$Y_{k+1} = A(I - Y_k)^{-1}A^*, \quad k = 0, 1, 2, \dots, \quad Y_0 = I/2, \tag{3.4}$$

give a two-sided iterative process, which tends to $X_{-\infty}$.

THEOREM 3.3. *If $q = \|A\| < 1/2$ and A is a normal matrix, then the two iterative processes* (3.3) *and* (3.4) *tend to $X_{-\infty}$.*

PROOF. In order to prove the theorem, it suffices to show that

$$X_k < X_{k+1} < X_{-\infty} < Y_{k+1} < Y_k,$$

for $k = 0, 1, 2, \dots$. Let us start to prove that $X_k < X_{k+1}$. We have from (3.3) that

$$X_1 = AA^* > 0 = X_0,$$

since $\det A \neq 0$. If we now assume that for a fixed k we have $X_{k-1} < X_k$, it is easy to show that $X_k < X_{k+1}$ by Theorem 2.1. So we get that the last inequality will be fulfilled for every k.

Continue with $X_k < X_{-\infty}$. For $k = 0$, the last inequality holds, because $X_0 = 0 < X_{-\infty}$. Let us now suppose that for a fixed k, we have $X_{k-1} < X_{-\infty}$. From the last inequality we get

$$(I - X_{k-1})^{-1} < (I - X_{-\infty})^{-1},$$
$$A(I - X_{k-1})^{-1}A^* < A(I - X_{-\infty})^{-1}A^*.$$

This leads to $X_k < X_{-\infty}$. In such a way we prove that $X_k < X_{k+1} < X_{-\infty}$. The proof for $X_{-\infty} < Y_{k+1} < Y_k$ is similar. The proof of Theorem 3.3 is now complete.

REMARK 1. IF $\|A\| = 1/2$, it is easy to prove that

(i) The two iterative processes (2.1) and (2.2) converge to the largest solution which is $X_\infty = X_0 = (1/2)I$.

(ii) If we use the two iterative processes (3.3) and (3.3), they converge to the smallest solution which is $X_{-\infty} = Y_0 = (1/2)I$.

4. Numerical experiments

In this section, the numerical examples in [10] are given to illustrate the effectiveness of the present algorithms. In the following tables we denote $\varepsilon(X) = \|X + A^T X^{-1} A - I\|_\infty$, $\varepsilon_{X_k} = \|X_\infty - X_k\|_\infty$ and $\varepsilon = \|X_k - Y_k\|_\infty$. First, we will obtain the solution X by the iterative methods (2.1) and (2.2) in the general case (that is, A is nonnormal).

EXAMPLE 1. We define the nonnormal matrix

$$A = \frac{1}{40}\begin{pmatrix} 2 & -1 & 3 & 4 \\ 7 & 6 & -5 & 9 \\ 4 & 8 & 10 & 6 \\ -3 & 5 & 2 & 8 \end{pmatrix}.$$

The maximum solution is

$$X_\infty = \begin{pmatrix} .946873 & -.0448677 & -.00670385 & -.0571869 \\ -.0448677 & .898174 & -.0431112 & -.119047 \\ -.00670385 & -.0431112 & .90855 & -.0354448 \\ -.0571869 & -.119047 & -.0354448 & .827281 \end{pmatrix}.$$

Experimentally, if a scalar multiple of the matrix A is less than 35, then the equation has no positive definite solution, but if it is greater than or equal to 36, then the equation has a positive definite solution. Table 1 gives the error analysis of Example 1.

TABLE 1.

iter	$\varepsilon(X)$	$\varepsilon(Y)$	ε
2	2.70912E-02	1.08620E-02	2.34963E-02
6	3.66368E-04	7.14647E-05	1.64376E-04
10	1.96534E-06	3.71656E-07	8.52334E-07
14	9.99701E-09	1.88586E-09	4.32361E-09
18	5.06288E-11	9.54858E-12	2.18914E-12
22	2.56350E-13	4.84057E-14	1.10800E-13
26	1.33227E-15	2.22045E-16	5.55112E-16
30	1.38778E-17	0.0E00	1.38778E-17
34	0.0E00	0.0E00	1.38778E-17

EXAMPLE 2. We define the normal matrix

$$A = \begin{pmatrix} 0.1 & -0.15 & -0.2598076 \\ 0.15 & 0.2125 & -0.0649519 \\ 0.2598076 & -0.0649519 & 0.1375 \end{pmatrix}.$$

From (3.1), the maximum solution is

$$X_\infty = \begin{pmatrix} 0.88729835 & 0.0 & 0.0 \\ 0.0 & 0.92158407 & -0.01979489 \\ 0.0 & -0.01979489 & 0.89872694 \end{pmatrix}.$$

Here we will use the iterative processes (2.1) and (2.2) to find the largest solution of the matrix equation. We can obtain the same results by using the iterative processes (3.3) and (3.4) to find the smallest solution $X_{-\infty} = I - X_\infty$.

TABLE 2.

iter	ε_{X_k}	ε_{Y_k}	ε
2	1.58405E-03	2.01664E-04	1.78571E-03
4	2.56088E-05	3.25126E-06	2.88600E-05
6	4.14565E-07	5.50941E-08	4.65618E-07
8	3.98173E-08	4.17305E-08	7.51193E-09
10	4.14895E-08	4.15200E-08	1.21192E-10
12	4.15161E-08	4.15166E-08	1.95532E-12
14	4.15165E-08	4.15165E-08	3.15303E-14
16	4.15165E-08	4.15165E-08	4.44089E-16
18	4.15165E-08	4.15165E-08	1.11022E-16
20	4.15165E-08	4.15165E-08	1.11022E-16

5. Conclusion

In this paper we consider a nonlinear matrix equation (1.2). The equation can be viewed as a natural extension of the scalar equation $x + a^2/x = 1$. This scalar problem is equivalent to equation $\varphi(x) = a^2$, where $\varphi(x) = x(1-x)$. This equation has a positive solution x so that $0 < x < 1$ if

$$a^2 \leq \max \varphi(x) = \varphi(1/2).$$

In the case when A is a normal matrix, (1.2) can be reduced to $X^2 - X + A^*A = 0$ using Lemma 3.1. We introduce a two-sided recursion algorithm from which a positive definite solution can be calculated. We calculate the extremal positive definite solutions of the matrix equation. The numerical experiments demonstrate that the described iterative methods are efficient. We observed that the iteration (2.1) is equivalent to the iteration (3.3) when $X = I - X$.

The two-sided iteration method described above possesses some advantages. We can compute X_{k+1} and Y_{k+1} in parallel. We can obtain and prove that X_∞ and $X_{-\infty}$ are the solutions of the equation when the matrix A is normal, while this cannot be proved for one-sided iteration methods. It is also easy to propose a stopping criteria, using

$$\max(\|Y_k - X\|, \|X - X_k\|) < \|Y_k - X_k\|,$$

which is not applicable for one-sided iteration methods.

Here we consider the case when A is a nonsingular. If A is singular, the problem of obtaining the extremal solutions is not solved. This problem is still a topic for future research. The problem of how to compute the iterative positive definite solution of (1.1) without calculating the matrix inversion is currently under consideration.

References

[1] W. N. Anderson, Jr., T. D. Morley and G. E. Trapp, "Positive solution to $X = A - BX^{-1}B^*$", *Linear Algebra Appl.* **134** (1990) 53–62.

[2] B. L. Buzbee, G. H. Golub and C. W. Nielson, "On direct methods for solving Poisson's equations", *SIAM J. Numer. Anal.* **7** (1970) 627–656.

[3] S. M. El-Sayed, "On the two iteration processes for computing positive definite solutions of equation $X - A^*X^{-n}A = I$", *Comput. Math. Appl.* (submitted).

[4] S. M. El-Sayed and A. C. M. Ran, "On the iteration methods for solving a class of nonlinear matrix equations", *SIAM J. Matrix Anal. Appl.* (to appear).

[5] J. C. Engwerda, "On the existence of a positive definite solution of the matrix equation $X + A^TX^{-1}A = I$", *Linear Algebra Appl.* **194** (1993) 91–108.

[6] J. C. Engwerda, A. C. M. Ran and A. L. Rijkeboer, "Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^*X^{-1}A = Q$", *Linear Algebra Appl.* **186** (1993) 255–275.

[7] C. H. Guo and P. Lancaster, "Iterative solution of two matrix equations", *The Mathematics of Computation* **68** (1999) 1589–1603.

[8] I. G. Ivanov and S. M. El-Sayed, "Properties of positive definite solutions of the equation $X + A^*X^{-2}A = I$", *Linear Algebra Appl.* **279** (1998) 303–316.

[9] P. Lancaster and L. Rodman, *The algebraic Riccati equation* (Oxford University Press, Oxford, 1995).

[10] X. Zhan, "Computing the extremal positive definite solutions of a matrix equation", *SIAM J. Sci. Comput.* **17** (1996) 1167–1174.

[11] X. Zhan and J. Xie, "On the matrix equation $X + A^*X^{-1}A = I$", *Linear Algebra Appl.* **247** (1996) 337–345.