Applied Mathematics and Computation 209 (2009) 230–237

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Solvability and sensitivity analysis of polynomial matrix equation $X^s + A^T X^t A = Q$ ☆

Zhigang Jia [a,*], Musheng Wei [b]

[a] Department of Mathematics, East China Normal University, No. 5800, Hongmei South Road, Shanghai 200241, PR China
[b] Department of Mathematics, Shanghai Normal University, Scientific Computing Key Laboratory of Shanghai Universities, Shanghai 200234, PR China

ARTICLE INFO

Keywords:
Polynomial matrix equation
Symmetric positive definite solution
Linear operator
Condition number
Algebraic perturbation

ABSTRACT

In this paper, we present a sufficient condition for the existence of the symmetric positive definite solution of polynomial matrix equation $X^s + A^T X^t A = Q$ where s, t are both nonnegative integers, $A, Q \in \mathbb{R}^{n\times n}$ and $Q > 0$. We firstly define the condition number of the unique SPD solution and reduce its representation form. We also give the algebraic perturbation analysis of the unique SPD solutions with respect to perturbations of matrices A and Q.

1. Introduction

The most important class of nonlinear matrix equations is undoubtedly the class of algebraic Riccati equations, either continuous-time or discrete-time. This class of equations is generally quadratic in nature, and analogues of this were studied in papers by Ferrante and Levy [7], and by Engwerda et al. [6] in the nineties. Later on, more general classes of nonlinear matrix equations were studied by Ivanov [11,12], Hasanov [9], El-Sayed [4], and Liu [14]. El-Sayed and Ran firstly brought some general theory to the subject in [5], followed by papers of Ran and Reurings [16], and Reurings [19], and the Ph.D. thesis of Reurings [20]. Recently there has been a lot of activity again in this area, focusing on particular types of equations, such as [3,13,22], etc.

In this paper, we will consider the solvability and sensitivity of the polynomial matrix equation:

$$X^s + A^T X^t A = Q, \tag{1.1}$$

where s, t are both nonnegative integers, $A, Q \in \mathbb{R}^{n\times n}$ and $Q > 0$. First of all, it should be remarked that there is a reduction possible to the case where $Q = I$ and A is invertible. This goes back to an approach outlined in [6], see also [17]. Next, consider three separate cases: $s > t$, $s = t$, and $s < t$.

Case 1, $s > t$. Then $\frac{t}{s} < 1$. Put $X^s = Y$, then the Eq. (1.1) are equivalent to

$$Y + A^T Y^{\frac{t}{s}} A = Q. \tag{1.2}$$

Because $\mathscr{F}(X) = X^{t/s}$ with $0 < t/s < 1$ is order preserving (see Theorem V.1.9 in [1]), it follows that Eq. (1.2) is an example of the equation treated in [5] with the requiring condition $A^T \mathscr{F}(Q) A < Q$. After reducing to $Q = I$ this case has been treated in Chapter 5 of the Ph.D. thesis of Reurings [20], see also her paper [19].

Case 2, $s = t$. This case is trivial, as the equation reduces to a linear matrix equation for X^s, which were treated in several sources. See, e.g., [18].

☆ Foundation item: National Science Foundation of China under Grant 10771073 and Shanghai Leading Academic Discipline Project under Grant S30405.
* Corresponding author.
E-mail address: jiazhigang1980@126.com (Z. Jia).

0096-3003/$ - see front matter
doi:10.1016/j.amc.2008.12.048

Case 3, $s < t$. Put $Y = X^t$. If A is invertible, the equation reduces to

$$Y + A^{T-1}Y^{\frac{s}{t}}A = A^{T-1}QA^{-1}$$

and since $\frac{s}{t} < 1$ this is again a case that reduces to Chapter 5 of the Ph.D. thesis of Reurings. As far as our knowledge, there are no elegant results for (1.1)($s < t$) with singular coefficient matrix A in the literature.

In Section 2, we present a sufficient condition under which (1.1) has an unique symmetric positive definite (SPD) solution. In Section 3, we define the condition number of the unique SPD solution of (1.1) and reduce its representation form. We also discuss the algebraic perturbation analysis of the unique SPD solutions with respect to perturbations of matrices A and Q, for the polynomial matrix Eq. (1.1).

We begin with some notations used throughout this paper. $\mathbb{R}^{m\times n}$ stands for the set of $m \times n$ matrices with elements on field $\mathbb{R}$. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote, respectively, the spectral norm and the Frobenius norm. If H is a SPD matrix on $\mathbb{R}^{n\times n}$, $\lambda_{\min}(H)$ and $\lambda_{\max}(H)$ stand for the minimal eigenvalue and the maximal eigenvalue, respectively. For $A = (a_1, a_2, \ldots, a_n) = (a_{ij}) \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$, $A \otimes B = (a_{ij}B)$ is the Kronecker product of A and B; vec(A) is the vector defined by $\mathrm{vec}(A) = (a_1^T, a_2^T, \ldots, a_n^T)^T$. Denote the singular values of a matrix $A \in \mathbb{R}^{m\times n}$ by $\sigma_1 \geqslant \cdots \geqslant \sigma_l \geqslant 0$, where $l = \min\{m, n\}$. Suppose that X and Y are SPD matrices, we write $X \geqslant Y (X > Y)$ if $X - Y$ is positive semi-definite (definite), and denote the matrices set $\{X | X - \alpha I \geqslant 0 \text{ and } \beta I - X \geqslant 0\}$ by $[\alpha I, \beta I]$.

2. Solvability analysis of (1.1)

In this section, a sufficient condition is given to guarantee the existence and uniqueness of the SPD solution of polynomial matrix Eq. (1.1).

Define

$$g_1(x) = x^s + \lambda_{\min}(A^TA)x^t - \lambda_{\max}(Q)$$

and

$$g_2(x) = x^s + \lambda_{\max}(A^TA)x^t - \lambda_{\min}(Q).$$

Since $\lambda_{\max}(A^TA) \geqslant \lambda_{\min}(A^TA) \geqslant 0$, $g_1(x)$ has a unique positive root $\alpha_1 \in (0, (\lambda_{\max}(Q))^{\frac{1}{s}}]$ (if $\lambda_{\min}(A^TA) = 0$, $\alpha_1 = (\lambda_{\max}(Q))^{\frac{1}{s}}$), and $g_2(x)$ has a unique positive root $\alpha_2 \in (0, (\lambda_{\min}(Q))^{\frac{1}{s}}]$ (if $\lambda_{\max}(A^TA) = 0$, $\alpha_2 = (\lambda_{\min}(Q))^{\frac{1}{s}}$). $g_2(x) - g_1(x) \geqslant 0$ on $[0, +\infty)$ implies $\alpha_2 \leqslant \alpha_1$.

Firstly, the upper and lower bounds for a SPD solution of (1.1) in Proposition 2.1.

Proposition 2.1. *Suppose that $X \in \mathbb{R}^{n\times n}$ is a SPD solution of (1.1), then for any eigenvalue $\lambda(X)$ of X,*

$$\beta_1 \leqslant \lambda(X) \leqslant \alpha_1. \tag{2.1}$$

Proof. From Theorem 3.3.16(d) [10],

$$\sigma_i(A^TX^tA) \quad \leqslant \quad \sigma_i(X^t)(\sigma_1(A))^2,$$

when A is nonsingular,

$$\sigma_i(X^t) = \sigma_i((A^{-1})^TA^TX^tAA^{-1}) \leqslant (\sigma_{\min}(A))^{-2}\sigma_i(A^TX^tA),$$

which implies

$$\sigma_i(A^TX^tA) \geqslant \sigma_i(X^t)\sigma_{\min}(A^TA), \text{ i.e., } \lambda_i(A^TX^tA) \geqslant \lambda_i(X^t)\lambda_{\min}(A^TA). \tag{2.2}$$

When A is singular, (2.2) holds still because $\lambda_{\min}(A^TA) = 0$ From $X^s = Q - A^TX^tA$, we have the following inequalities:

$$g_1(\lambda(X)) \leqslant 0 \quad \text{and} \quad g_2(\lambda(X)) \geqslant 0.$$

Then (2.1) is obtained. □

Proposition 2.2. *Suppose that X is a SPD solution of (1.1), $d_1 \geqslant \cdots \geqslant d_n > 0$ are eigenvalues of X. Then*

$$[\lambda_{\min}(Q) - d_1^s]d_1^{-t} \leqslant |\lambda(A)|^2 \leqslant [\lambda_{\max}(Q) - d_n^s]d_n^{-t}. \tag{2.3}$$

Proof. Let $(\lambda(A), y)$ be an eigenpair of A with $y^Ty = 1$. Multiplying y^T from the left and y from the right of the both sides of (1.1), we have

$$y^TX^sy + |\lambda(A)|^2y^TX^ty = y^TQy.$$

Obviously,

$$d_n^s + |\lambda(A)|^2 d_n^t \leqslant \lambda_{\max}(Q) \quad \text{and} \quad d_1^s + |\lambda(A)|^2 d_1^t \geqslant \lambda_{\min}(Q),$$

from which we obtain (2.3). □

The following two famous inequalities will be used frequently in the remaining of this paper:

Lemma 2.3 (Löwner–Heinz inequality, Theorem 1.1 [23]). *If $A \geqslant B \geqslant 0$ and $0 \leqslant r \leqslant 1$ then $A^r \geqslant B^r$.*

Lemma 2.4. Theorem 2.1 [8]*Let A and B be positive operators on a Hilbert space H, such that $M_1 I \geqslant A \geqslant m_1 I > 0$, $M_2 I \geqslant B \geqslant m_2 I > 0$ and $0 < A \leqslant B$. Then*

$$A^t \leqslant (M_1/m_1)^{t-1} B^t, \; A^t \leqslant (M_2/m_2)^{t-1} B^t,$$

hold for any $t \geqslant 1$.

Now we give an existence and uniqueness theorem of a SPD solution of (1.1).

Theorem 2.5. *Suppose that $\lambda_{\max}(A^T A) \leqslant \lambda_{\min}(Q)(\lambda_{\max}(Q))^{-\frac{t}{s}}$.*

(1) *(1.1) has a SPD solution $X \in [\beta_1 I, \alpha_1 I]$.*
(2) *Moreover, if $\lambda_{\max}(A^T A) < (s\beta_1^{s-1})(t\alpha_1^{t-1})^{-1}$, (1.1) has a unique SPD solution $X \in [\beta_1 I, \alpha_1 I]$.*

Proof

(1) Define a continuous mapping $h(X) = (Q - A^T X^t A)^{\frac{1}{s}}$ on $[0, (\lambda_{\max}(Q))^{\frac{1}{s}} I]$. By Lemma 2.4, the inequalities $\sigma_1(X) I \geqslant X \geqslant \sigma_n(X) I > 0$ and $0 < X \leqslant (\lambda_{\max}(Q))^{\frac{1}{s}} I$ imply $X^t \leqslant (\lambda_{\max}(Q))^{\frac{t}{s}} I$. Applying Lemma 2.3, we obtain

$$h(X) \leqslant Q^{\frac{1}{s}} \leqslant (\lambda_{\max}(Q))^{\frac{1}{s}} I, \tag{2.4}$$

and from $\lambda_{\max}(A^T A) \leqslant \lambda_{\min}(Q)(\lambda_{\max}(Q))^{-\frac{t}{s}}$,

$$h(X) \geqslant (\lambda_{\min}(Q) - \lambda_{\max}(A^T A)\lambda_{\max}(X^t))^{\frac{1}{s}} I, \tag{2.5}$$

$$\geqslant [\lambda_{\min}(Q) - \lambda_{\max}(A^T A)(\lambda_{\max}(Q))^{\frac{t}{s}}]^{\frac{1}{s}} I \geqslant 0. \tag{2.6}$$

Combining with (2.4)–(2.6), Brouwer's fixed-point Theorem and Proposition 2.1 imply (1).

(2) Assume that both $X \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{n\times n}$ are SPD solutions of (1.1) on $[\beta_1 I, \alpha_1 I]$, and $\|X - Y\|_F > 0$. Then

$$X^s - Y^s = A^T (Y^t - X^t) A \quad \text{and} \quad \|X^s - Y^s\|_F \leqslant \|A\|_2^2 \|Y^t - X^t\|_F. \tag{2.7}$$

It is known that

$$X^s - Y^s = \sum_{k=0}^{s-1} X^k (X - Y) Y^{s-1-k}$$

and

$$\|X^s - Y^s\|_F = \left\| \left[\sum_{k=0}^{s-1} (Y^{s-1-k})^T \otimes X^k \right] \mathrm{vec}(X - Y) \right\|_2. \tag{2.8}$$

Suppose $X = U\Lambda_x U^T, Y = V\Lambda_y V^T$, where $U, V \in \mathbb{R}^{n\times n}$ are unitary matrices and $\Lambda_x = \mathrm{diag}(\lambda_1(x), \ldots, \lambda_n(x))$, $\Lambda_y = \mathrm{diag}(\lambda_1(y), \ldots, \lambda_n(y))$, then

$$\sum_{k=0}^{s-1} (Y^{s-1-k})^T \otimes X^k = \sum_{k=0}^{s-1} (\bar{V}\Lambda_y^{s-1-k} V^T) \otimes (U\Lambda_x^k U^T) = (\bar{V} \otimes U) \left[\sum_{k=0}^{s-1} (\Lambda_y^{s-1-k} \otimes \Lambda_x^k) \right] (\bar{V} \otimes U)^T,$$

from which we see that $\sum_{k=0}^{s-1} (Y^{s-1-k})^T \otimes X^k$ is SPD, with eigenvalues

$$\sum_{k=0}^{s-1} \lambda_j(x)^k \lambda_i(y)^{s-1-k}, \quad 1 \leqslant i,j \leqslant n.$$

Therefore, $X, Y \in [\beta_1 I, \alpha_1 I]$ implies

$$\|X^s - Y^s\|_F \geqslant s\beta_1^{s-1} \|X - Y\|_F \quad \text{and} \quad \|X^t - Y^t\|_F \leqslant t\alpha_1^{t-1} \|X - Y\|_F. \tag{2.9}$$

Moreover, if $\lambda_{\max}(A^TA) < (s\beta_1^{s-1})(t\alpha_1^{t-1})^{-1}$, then from (2.7) and (2.9) we have

$$\|X - Y\|_F \leqslant t\alpha_1^{t-1}(s\beta_1^{s-1})^{-1}\lambda_{\max}(A^TA)\|X - Y\|_F < \|X - Y\|_F,$$

which is impossible. Hence, $X = Y$. □

Corollary 2.6. *If $s = 2, t = 1$, $\lambda_{\max}(A^TA) \leqslant 1$ and $Q = I$, then (1.1) has a unique SPD solution.*

Proof. Under the conditions, we have $2\beta_1 = (\lambda_{\max}(A^TA)^2 + 4\lambda_{\max}(A^TA))^{\frac{1}{2}} - \lambda_{\max}(A^TA) > \lambda_{\max}(A^TA)$. □

3. Sensitivity analysis for (1.1)

Let $A(\tau) = A + \tau E$ and $Q(\tau) = Q + \tau G$, where $\tau \in \mathbb{R}$, $E, G \in \mathbb{R}^{n\times n}$, and G is symmetric semi-positive definite (SSD). Frequently we drop the variable τ of $X(\tau)$ if no cause of confusing exists; thus we write X instead of $X(\tau)$. Consider the equation

$$X^s + A^T(\tau)X^tA(\tau) = Q(\tau). \tag{3.1}$$

Let $\mathscr{F}(X,\tau) = X^s + A^T(\tau)X^tA(\tau) - Q(\tau)$, and X_0 be the unique SPD solution of Eq. (1.1) under the assumption that $\lambda(A^TA) < \min\left\{\frac{s\beta_1^{s-1}}{t\alpha_1^{t-1}}, \frac{\lambda_{\min}(Q)}{(\lambda_{\max}(Q))^{\frac{t}{s}}}\right\}$ in Theorem 2.5. Then

(1) $\mathscr{F}(X_0, 0) = 0$;
(2) For X^s and X^t are both polynomials of elements of X, thus $\mathscr{F}(X,\tau)$ is arbitrarily differentiable in neighborhood of $(X_0, 0)$;
(3) $\frac{\partial\mathscr{F}}{\partial X}|_{(X_0,0)} = \sum_{k=0}^{s-1} X_0^k \otimes X_0^{s-1-k} + (A^T \otimes A^T)(\sum_{k=0}^{t-1} X_0^k \otimes X_0^{t-1-k})$. (The introduction of $\frac{\partial\mathscr{F}}{\partial X}$ is similar to [21].)

Applying the implicit function theory [15], $\det(\frac{\partial\mathscr{F}}{\partial X}|_{(X_0,0)}) \neq 0$ implies that there exists $\delta > 0$ s.t. if $\tau \in (-\delta, \delta)$ there is a unique $X(\tau)$ satisfying

(1) $\mathscr{F}(X(\tau),\tau) = 0, X(0) = X_0$;
(2) $X(\tau)$ is arbitrarily differentiable with regard to τ.

So $X(\tau)$ has the expansion

$$X(\tau) = X_0 + \dot{X}(0)\tau + O(\tau^2).$$

We take derivative for both sides of Eq. (3.1) with regard to τ, then let $\tau = 0$,

$$\sum_{k=0}^{s-1} X_0^k\dot{X}(0)X_0^{s-1-k} + \sum_{k=0}^{t-1} A^TX_0^k\dot{X}(0)X_0^{t-1-k}A = G - (E^TX_0^tA + A^TX_0^tE)$$

Similar to [14], we have

Definition 3.1. The condition number of the unique SPD solution X_0 of Eq. (1.1)

$$\mathscr{K}_X = \lim_{\tau\to 0+} \sup_{(E,G)\neq 0, E\in\mathbb{R}^{n\times n}, G\in\mathbb{S}^{n\times n}} \left\{\frac{\|X(\tau) - X_0\|_F}{\|X_0\|_F} \cdot \left(\frac{\tau\|(E,G)\|_F}{\|(A,Q)\|_F}\right)^{-1}\right\}.$$

It is clear that

$$\begin{aligned}\mathscr{K}_X &= \sup_{(E,G)\neq 0, E\in\mathbb{R}^{n\times n}, G\in\mathbb{S}^{n\times n}} \left\{\frac{\|\dot{X}(0)\|_F}{\|(E,G)\|_F} \cdot \frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\}, \\ &= \max_{(E,G)\neq 0, E\in\mathbb{R}^{n\times n}, G\in\mathbb{S}^{n\times n}} \left\{\frac{\|\dot{X}(0)\|_F}{\|(E,G)\|_F} \cdot \frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\},\end{aligned}$$

Define the linear operator $\mathscr{L}$: $\mathbb{S}^{n\times n} \to \mathbb{S}^{n\times n}$ by

$$\mathscr{L}(Z) = \sum_{k=0}^{s-1} X_0^kZX_0^{s-1-k} + \sum_{k=0}^{t-1} A^TX_0^kZX_0^{t-1-k}A, \quad Z \in \mathbb{S}^{n\times n}.$$

Obviously, $\mathscr{L}$ is invertible. Moreover, define the linear operator $\mathscr{P}$: $\mathbb{R}^{n\times n} \times \mathbb{S}^{n\times n} \to \mathbb{S}^{n\times n}$ by

$$\mathscr{P}(E,G) = \mathscr{L}^{-1}(G - (E^TX_0^tA + A^TX_0^tE)), \quad E \in \mathbb{R}^{n\times n}, \quad G \in \mathbb{S}^{n\times n}, \tag{3.2}$$

and define the linear operator $\mathscr{Q}$: $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ by

$$\mathscr{Q}(M,N) = \mathscr{T}^{-1}(N - (M^TX_0^tA + A^TX_0^tM)), \quad M, N \in \mathbb{R}^{n\times n}, \tag{3.3}$$

where $\mathscr{T}$ is the invertible linear operator defined by

$$\mathscr{T}(W)=\sum_{k=0}^{s-1}X_0^kWX_0^{s-1-k}+\sum_{k=0}^{t-1}A^TX_0^kWX_0^{t-1-k}A,\quad W\in\mathbb{R}^{n\times n}.$$

Then

$$\mathscr{K}_X=\max_{(E,G)\neq 0,E\in\mathbb{R}^{n\times n},G\in\mathbb{S}^{n\times n}}\left\{\frac{\|\mathscr{P}(E,G)\|_F}{\|(E,G)\|_F}\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\} \tag{3.4}$$

$$\leqslant \max_{(M,N)\neq 0,M,N\in\mathbb{R}^{n\times n}}\left\{\frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\}. \tag{3.5}$$

Lemma 3.1. [2] *If* $W_1,W_2\in\mathbb{H}^{n\times n}$ *satisfy*

$$W_2\geqslant W_1\geqslant -W_2,$$

then

$$\|W_1\|_F\leqslant\|W_2\|_F.$$

We now prove that the equality of (3.5) holds.

Proposition 3.2

$$\max_{(E,G)\neq 0,E\in\mathbb{R}^{n\times n},G\in\mathbb{S}^{n\times n}}\frac{\|\mathscr{P}(E,G)\|_F}{\|(E,G)\|_F}=\max_{(M,N)\neq 0,M,N\in\mathbb{R}^{n\times n}}\frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}.$$

Proof. Let $\widehat{M},\widehat{N}$ be the matrices such that

$$\max_{(M,N)\neq 0,M,N\in\mathbb{R}^{n\times n}}\frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}=\frac{\|\mathscr{Q}(\widehat{M},\widehat{N})\|_F}{\|(\widehat{M},\widehat{N})\|_F}. \tag{3.6}$$

The relation (3.6) means that $(\widehat{M},\widehat{N})$ is a singular vector of $\mathscr{Q}$ corresponding to its largest singular value. Let

$$\widehat{Z}=\mathscr{Q}(\widehat{M},\widehat{N})\in\mathbb{R}^{n\times n}.$$

By the definition (3.2), we have

$$\sum_{k=0}^{s-1}X_0^k\widehat{Z}X_0^{s-1-k}+\sum_{k=0}^{t-1}A^TX_0^k\widehat{Z}X_0^{t-1-k}A=\widehat{N}-(\widehat{M}^TX_0^tA+A^TX_0^t\widehat{M}).$$

Equivalently,

$$\sum_{k=0}^{s-1}X_0^k\widehat{Z^T}X_0^{s-1-k}+\sum_{k=0}^{t-1}A^TX_0^k\widehat{Z^T}X_0^{t-1-k}A=\widehat{N}^T-(\widehat{M}^TX_0^tA+A^TX_0^t\widehat{M}).$$

Then,

$$\widehat{Z}^T=\mathscr{P}(\widehat{M},\widehat{N}^T).$$

Since $\|\widehat{Z}\|_F=\|\widehat{Z}^T\|_F$, $\|(\widehat{M},\widehat{N}^T)\|_F=\|(\widehat{M},\widehat{N})\|_F$, Eq. (3.6) implies

$$\max_{(M,N)\neq 0,M,N\in\mathbb{R}^{n\times n}}\frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}=\frac{\|\mathscr{Q}(\widehat{M},\widehat{N}^T)\|_F}{\|(\widehat{M},\widehat{N}^T)\|_F},$$

which means that $(\widehat{M},\widehat{N}^T)$ is also a singular vector of $\mathscr{Q}$ corresponding to its largest singular value.

Let

$$\widehat{E}=2\widehat{M}\in\mathbb{R}^{n\times n},\ \widehat{G}=\widehat{N}+\widehat{N}^T\in\mathbb{S}^{n\times n}.$$

If $(\widehat{E},\widehat{G})\neq 0$, then it is also a singular vector of $\mathscr{Q}$ corresponding to its largest singular value. Then

$$\max_{(M,N)\neq 0,M,N\in\mathbb{R}^{n\times n}}\frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}=\frac{\|\mathscr{Q}(\widehat{E},\widehat{G})\|_F}{\|(\widehat{E},\widehat{G})\|_F}=\frac{\|\mathscr{P}(\widehat{E},\widehat{G})\|_F}{\|(\widehat{E},\widehat{G})\|_F}, \tag{3.7}$$

in which $\widehat{E}\in\mathbb{R}^{n\times n}$, $\widehat{G}\in\mathbb{S}^{n\times n}$.

If $(\widehat{E},\widehat{G})=0$, then (3.6) becomes

$$\max_{(M,N)\neq 0, M,N\in\mathbb{R}^{n\times n}} \frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}=\frac{\|\mathscr{Q}(0,\widehat{N})\|_F}{\|(0,\widehat{N})\|_F}, \tag{3.8}$$

in which $\widehat{N}\in\mathbb{R}^{n\times n}$ is skew-symmetric.

We now prove that there exists $\widetilde{G}\in\mathbb{S}^{n\times n}$ such that

$$\frac{\|\mathscr{Q}(0,\widehat{N})\|_F}{\|(0,\widehat{N})\|_F}\leqslant\frac{\|\mathscr{P}(0,\widetilde{G})\|_F}{\|(0,\widetilde{G})\|_F}. \tag{3.9}$$

Since the matrices $\widehat{N}$ is real skew-symmetric, there is a real orthogonal matrix U, and $\lambda_i\geqslant 0$ $(i=1,2,\ldots)$ such that

$$\widehat{N}=U\left[\begin{pmatrix}0 & \lambda_1\\ -\lambda_1 & 0\end{pmatrix}\oplus\begin{pmatrix}0 & \lambda_2\\ -\lambda_2 & 0\end{pmatrix}\oplus\cdots\right]U^T,$$

where the canonical forms of $\widehat{N}$ has odd numbers of 1-by-1 blocks of zero in the diagonal when $n=2k+1$ for a positive integer k. Let

$$\widetilde{G}=U\left[\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_1\end{pmatrix}\oplus\begin{pmatrix}\lambda_2 & 0\\ 0 & \lambda_2\end{pmatrix}\oplus\cdots\right]U^T.$$

Obviously,

$$\widetilde{G}\in\mathbb{S}^{n\times n},\quad \|\widetilde{G}\|_F=\|\widehat{N}\|_F,\quad \widetilde{G}\geqslant 0,\quad \widetilde{G}\geqslant i\widehat{N}\geqslant-\widetilde{G}.$$

Suppose that $\widetilde{Z}$ satisfy

$$\sum_{k=0}^{s-1}X_0^k\widetilde{Z}X_0^{s-1-k}+\sum_{k=0}^{t-1}A^TX_0^k\widetilde{Z}X_0^{t-1-k}A=\widetilde{G}.$$

Since

$$\sum_{k=0}^{s-1}X_0^k i\widehat{Z}X_0^{s-1-k}+\sum_{k=0}^{t-1}A^TX_0^k i\widehat{Z}X_0^{t-1-k}A=i\widehat{N},$$

it is clear that

$$\widetilde{Z}+i\widehat{Z}\geqslant 0,\quad \widetilde{Z}-i\widehat{Z}\geqslant 0,$$

i.e.

$$\widetilde{Z}\geqslant i\widehat{Z}\geqslant-\widetilde{Z}.$$

Then by Lemma 3.1 we have

$$\|i\widehat{Z}\|_F\leqslant\|\widetilde{Z}\|_F.$$

Consequently,

$$\begin{aligned}\frac{\|\mathscr{Q}(0,\widehat{N})\|_F}{\|(0,\widehat{N})\|_F}&=\frac{\|\widehat{Z}\|_F}{\|(0,\widehat{N})\|_F}=\frac{\|i\widehat{Z}\|_F}{\|(0,i\widehat{N})\|_F}\\&\leqslant\frac{\|\widetilde{Z}\|_F}{\|(0,\widetilde{G})\|_F}=\frac{\|\mathscr{P}(0,\widetilde{G})\|_F}{\|(0,\widetilde{G})\|_F},\end{aligned}$$

where $\widetilde{G}\in\mathbb{S}^{n\times n}$. Then Eq. (3.9) is proved. Combining Eq. (3.9) with Eq. (3.8) shows that

$$\max_{(M,N)\neq 0, M,N\in\mathbb{R}^{n\times n}} \frac{\|\mathscr{Q}(M,N)\|_F}{\|(M,N)\|_F}=\frac{\|\mathscr{P}(0,\widetilde{G})\|_F}{\|(0,\widetilde{G})\|_F}. \tag{3.10}$$

From Eqs. (3.7) and (3.10), Proposition 3.2 holds. □

Then we have

Theorem 3.3

$$\mathscr{K}_X=\|(\mathscr{U}+\mathscr{V})^{-1}(\mathscr{W},I)\|_2\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F},$$

in which

$$\mathscr{U}=\sum_{k=0}^{s-1}(X_0^{s-1-k})^T\otimes X_0^k,$$
$$\mathscr{V}=\sum_{k=0}^{t-1}(X_0^{t-1-k}A)^T\otimes(A^TX_0^k),$$
$$\mathscr{W}=-(I_n\otimes(A^TX_0^t)+((X_0^tA)^T\otimes I)\pi),$$

and π *is a permutation matrix and* $\pi\mathrm{vec}(E)=\mathrm{vec}(E^T)$.

Proof. The equality of (3.5), combined with (3.3) and (3.4), implies

$$\mathscr{K}_X=\max_{(M,N)\neq 0, M,N\in\mathbb{R}^{n\times n}}\left\{\frac{\|\mathscr{T}^{-1}(N-(M^TX_0^tA+A^TX_0^tM))\|_F}{\|(M,N)\|_F}\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\}$$

Obviously,

$$\mathscr{K}_X=\max_{(M,N)\neq 0, M,N\in\mathbb{R}^{n\times n}}\left\{\frac{\|(\mathscr{U}+\mathscr{V})^{-1}(\mathscr{W},I)\mathrm{vec}(M,N)\|_2}{\|\mathrm{vec}(M,N)\|_2}\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F}\right\}$$
$$=\|(\mathscr{U}+\mathscr{V})^{-1}(\mathscr{W},I)\|_2\cdot\frac{\|(A,Q)\|_F}{\|X_0\|_F}.\quad\square$$

Now we discuss the algebraic perturbation analysis of the unique SPD solutions with respect to perturbations of matrices A and Q, for the polynomial matrix Eq. (1.1).

The perturbed polynomial matrix equation of (1.1) is

$$\widehat{X}^s+\widehat{A}^T\widehat{X}^t\widehat{A}=\widehat{Q},\tag{3.11}$$

where $\widehat{A}=A+\Delta A\in\mathbb{R}^{n\times n}$, $\widehat{Q}=Q+\Delta Q\in\mathbb{R}^{n\times n}$. Let

$$\hat{g}_1(\hat{x})=\hat{x}^s+\lambda_{\min}(\widehat{A}^T\widehat{A})\hat{x}^t-\lambda_{\max}(\widehat{Q})$$

and

$$\hat{g}_2(\hat{x})=\hat{x}^s+\lambda_{\max}(\widehat{A}^T\widehat{A})\hat{x}^t-\lambda_{\min}(\widehat{Q}).$$

Denote the unique positive roots of $\hat{g}_1(\hat{x})$ and $\hat{g}_2(\hat{x})$ by $\hat{\alpha}_1$ and $\hat{\beta}_1$, respectively. It is obvious that $\hat{\alpha}_1\in[0,(\lambda_{\max}(\widehat{Q}))^{\frac{1}{s}}]$, $\hat{\beta}_1\in[0,(\lambda_{\min}(\widehat{Q}))^{\frac{1}{s}}]$ and $\hat{\beta}_1\leqslant\hat{\alpha}_1$. Recalling that α_1 and β_1 are unique positive roots of $g_1(x)$ and $g_2(x)$, respectively, and that $\alpha_1\in(0,(\lambda_{\max}(Q))^{\frac{1}{s}}]$, $\beta_1\in(0,(\lambda_{\min}(Q))^{\frac{1}{s}}]$ and $\beta_1\leqslant\alpha_1$. Let

$$\tilde{\alpha}_1=\max\{\alpha_1,\hat{\alpha}_1\}\quad\text{and}\quad\tilde{\beta}_1=\min\{\beta_1,\hat{\beta}_1\}.$$

For an arbitrary $\varepsilon>0$, define $\eta(\varepsilon)$ by

$$\eta(\varepsilon)=\|A\|_2+\left(\|A\|_2^2+\frac{2\varepsilon}{3}\tau(\tilde{\beta}_1,\tilde{\alpha}_1)\|\widehat{X}\|_2^{-t}\right)^{\frac{1}{2}},\tag{3.12}$$

in which

$$\tau(\tilde{\beta},\tilde{\alpha})=s\tilde{\beta}^{s-1}-t\tilde{\alpha}^{t-1}\|A\|_2^2.\tag{3.13}$$

Theorem 3.4. *Suppose that both (1.1) and (3.11) have unique solutions denoted by* $X\in[\beta_1I,\alpha_1I]$ *and* $\widehat{X}\in[\hat{\beta}_1I,\hat{\alpha}_1I]$, *respectively; and* $\|A\|_2^2<s\tilde{\beta}_1^{s-1}(t\tilde{\alpha}_1^{t-1})^{-1}$. *Then for an arbitrary* $\varepsilon>0$, *if* $\|\Delta A\|_F<2\varepsilon\tau(\tilde{\beta}_1,\tilde{\alpha}_1)(3\eta(\varepsilon)\|\widehat{X}\|_2^{t})^{-1}$ *with* $\eta(\varepsilon)$ *defined by (3.12) and* $\|\Delta Q\|_F<\frac{1}{3}\tau(\tilde{\beta}_1,\tilde{\alpha}_1)\varepsilon$, *we have*

$$\|\widehat{X}-X\|_F<\varepsilon.\tag{3.14}$$

Proof. (1.1) and (3.11) imply

$$\widehat{X}^s-X^s=-(\widehat{A}^T\widehat{X}^t\widehat{A}-A^TX^tA)+\widehat{Q}-Q$$
$$=-[A^T(\widehat{X}^t-X^t)A+A^T\widehat{X}^t\Delta A+\Delta A^T\widehat{X}^tA+\Delta A^T\widehat{X}^t\Delta A]+\Delta Q.$$

By formula (2.8), we have

$$\|\widehat{X}^s - X^s\|_F \geqslant \left(\sum_{k=0}^{s-1} \hat{\beta}_1^{s-1-k}\beta_1^k\right)\|\widehat{X} - X\|_F \geqslant s\tilde{\beta}_1^{s-1}\|\widehat{X} - X\|_F \tag{3.15}$$

and

$$\|\widehat{X}^t - X^t\|_F \leqslant \left(\sum_{k=0}^{t-1} \hat{\alpha}_1^{t-1-k}\alpha_1^k\right)\|\widehat{X} - X\|_F \leqslant t\tilde{\alpha}_1^{t-1}\|\widehat{X} - X\|_F. \tag{3.16}$$

Then

$$\begin{aligned}\|\widehat{X}^s - X^s\|_F &\leqslant \|A^T(\widehat{X}^t - X^t)A\|_F + \|A^T\widehat{X}^t\Delta A\|_F + \|\Delta A^T\widehat{X}^t A\|_F \\ &\quad + \|\Delta A^T\widehat{X}^t\Delta A\|_F + \|\Delta Q\|_F \\ &\leqslant \|A\|_2^2\|\widehat{X}^t - X^t\|_F + 2\|A\|_2\|\widehat{X}\|_2^t\|\Delta A\|_F \\ &\quad + \|\widehat{X}\|_2^t\|\Delta A\|_F^2 + \|\Delta Q\|_F.\end{aligned}$$

Applying (3.15) and (3.16) with (3.13), we can obtain that

$$\tau(\tilde{\beta}_1, \tilde{\alpha}_1)\|\widehat{X} - X\|_F \leqslant 2\|A\|_2\|\widehat{X}\|_2^t\|\Delta A\|_F + \|\widehat{X}\|_2^t\|\Delta A\|_F^2 + \|\Delta Q\|_F.$$

Because $\|A\|_2^2 < s\tilde{\beta}_1^{s-1}(t\tilde{\alpha}_1^{t-1})^{-1}$, $\tau(\tilde{\beta}_1, \tilde{\alpha}_1) > 0$. Then

$$\|\widehat{X} - X\|_F \leqslant (\tau(\tilde{\beta}_1, \tilde{\alpha}_1))^{-1}(2\|A\|_2\|\widehat{X}\|_2^t\|\Delta A\|_F + \|\widehat{X}\|_2^t\|\Delta A\|_F^2 + \|\Delta Q\|_F).$$

Then for an arbitrary $\varepsilon > 0$, if $-\eta(\varepsilon) < \|\Delta A\|_F < 2\varepsilon\tau(\tilde{\beta}_1, \tilde{\alpha}_1)(3\eta(\varepsilon)\|\widehat{X}\|_2^t)^{-1}$ where $\eta(\varepsilon)$ defined by (3.12) and $\|\Delta Q\|_F < \frac{1}{3}\tau(\tilde{\beta}_1, \tilde{\alpha}_1)\varepsilon$, we have (3.14). $\square$

4. Conclusions

This paper presents a sufficient condition for the unique symmetric positive define solution of an polynomial matrix equation, firstly define the condition number for such SPD solution and reduce its representation form. At last the algebraic perturbation analysis is also given.

Acknowledgement

We thank an anonymous referee for helpful comments that improved the paper.

References

[1] R. Bhatia, Matrix Analysis, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, 1997.
[2] R. Byers, S. Nash, On the singular vectors of the Lyapunov operator, SIAM J. Alg. Disc. Methods 8 (1987) 59–66.
[3] M. Cheng, S. Xu, Perturbation analysis of the Hermitian positive definite solution of the matrix equation $X - A^*X^{-2}A = I$, Linear Algebra Appl. 394 (2005) 39–51.
[4] S.M. El-Sayed, M.G. Petkov, Iterative methods for nonlinear equations $X + A^*X^{-\alpha}A = I$, Linear Algebra Appl. 403 (2005) 45–52.
[5] S.M. El-Sayed, A.C.M. Ran, On an iteration method for solving a class of nonlinear matrix equations, SIAM J. Matrix Anal. Appl. 23 (2001) 632–645.
[6] J.C. Engwerda, A.C.M. Ran, A.L. Rijkeboer, Necessary and sufficient conditions for the existence of a positive solution of the matrix equation $X + A^*X^{-1}A = I$, Linear Algebra Appl. 184 (1993) 255–275.
[7] A. Ferrante, B.C. Levy, Hermitian solutions of the equation $X = Q + NX^{-1}N^*$, Linear Algebra Appl. 247 (1996) 359–373.
[8] T. Furuta, Operator inequalities associated with Holder–McCarthy and Kantorovich inequalities, J. Inequal. Appl. 6 (1998) 137–148.
[9] V.I. Hasanov, Positive definite solutions of the matrix equations $X \pm A^*X^{-q}A = Q$, Linear Algebra Appl. 404 (2005) 166–182.
[10] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
[11] I.G. Ivanov, S.M. Sayed, Properties of positive definite solutions of the equations $X + A^*X^{-2}A = I$, Linear Algebra Appl. 279 (1998) 303–316.
[12] I.G. Ivanov, V.I. Hasavov, B.V. Minchev, On matrix equations $X \pm A^*X^{-2}A = I$, Linear Algebra Appl. 326 (2001) 27–44.
[13] I.G. Ivanov, On positive definite solutions of the family of matrix equations $X + A^*X^{-n}A = Q$, J. Comput. Appl. Math. 193 (2006) 277–301.
[14] X.-G. Liu, H. Gao, On the positive definite solutions of the matrix equations $X^s \pm A^TX^{-t}A = I_n$, Linear Algebra Appl. 368 (2003) 83–97.
[15] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equation in Several Variables, Academic Press, New York, 1970.
[16] A.C.M. Ran, M.C.B. Reurings, A fixed point theorem in partially ordered sets and some applications to matrix equations, Proceedings of the AMS 132 (2004) 1435–1443.
[17] A.C.M. Ran, M.C.B. Reurings, A nonlinear matrix equation connected to interpolation theory, Linear Algebra Appl. 379 (2004) 289–302.
[18] A.C.M. Ran, M.C.B. Reurings, The symmetric linear matrix equation, Electronic Linear Algebra 9 (2002) 93–107.
[19] M.C.B. Reurings, Contractive maps on normed linear spaces and their applications to nonlinear matrix equations, Linear Algebra Appl. 418 (2006) 292–311.
[20] M.C.B. Reurings: Symmetric Matrix Equations, Ph.D. thesis, Vrije Universiteit Amsterdam, 2003. <http://www.math.vu.nl/proefschriften/Reurings>.
[21] J.g. Sun, Eigenvalues and eigenvectors of a matrix dependent on several parameters, J. Comput. Math. 3 (1985) 351–364.
[22] Y. Yang, The iterative method for solving nonlinear matrix equation $X^s + A^*X^{-t}A = Q$, Appl. Math. Comput. 188 (1) (2007) 46–53.
[23] X. Zhan, Matrix Inequalities, Lecture Notes in Mathematics, vol. 1790, Springer-Verlag, Berlin, Heidelberg, 2002.