On the positive definite solutions of the nonlinear matrix equations $X^p = A \pm M^T(X^{-1}\#B)^{-1}M$

Jie Meng

FFM Industrial Mathematics Center on Big Data
Pusan National University, Korea

Jul. 2, 2018

Outline

- Introduction and preliminary
- Solvability of the nonlinear matrix equations
  - Existence and uniqueness
  - Lower and upper bound
  - Iteration method
- Numerical examples

Introduction

Recently, varieties of nonlinear matrix equations have been studied, for example

- $X - A^* e^X A = I$, [1]
- $X \pm \sum_{i=1}^{m} A_i^* X^{-q} A_i = Q$, [2,3,4]
- $A_0 + \sum_{i=1}^{k} \sigma_i A_i^* X^{p_i} A_i = 0, (\sigma_i = \pm 1)$, [5,6]

---

[1] D.J. Gao, On Hermitian positive definite solutions of the nonlinear matrix equation $X - A^* e^X A = I$, J. Appl. Math. Comput. 50 (2016) 109–116.

[2] X. Yin, R. Wen, L. Fang, On the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-q} A_i = Q$ $(0 < q \leq 1)$, Bull. Korean Math. Soc. 51 (2014) 739–763.

[3] X.Y. Yin, S.Y. Liu, Positive definite solutions of the matrix equations $X \pm A^* X^{-q} A = Q$ $(q \geq 1)$, Comput. Math. Appl. 59 (2010) 3727–3739.

[4] L. Zhang, An improved inversion-free method for solving the matrix equation $X + A^* X^{-\alpha} A = Q$, J. Comput. Appl. Math. 253 (2013) 200–203.

[5] M. Konstantinov, P. Petkov, I.Popchev, V. Angelova, Sensitivity of the matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i X^{p_i} A_i = 0$ $(\sigma_i = \pm 1)$, Appl. Comput. Math. 10 (2011) 409–427.

[6] I. Popchev, M. Konstantinov, P. Petkov, V. Angelova, Norm-wise, mixed and component-wise condition numbers of matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i X^{p_i} A_i = 0$ $(\sigma_i = \pm 1)$, Appl. Comput. Math. 14 (2014) 18–30.

- $X - \sum_{i=1}^{m} A_i^* X^{-1} A_i = Q$, [7,8]
- $X - A^*(X^{-1} + M)^{-1} A = Q$, [9]
- $X^s + \sum_{i=1}^{m} A_i^* X^{-t_i} A_i = Q$, [10]
- $\cdots$

---

[7] J. Li, Y. Zhang, Sensitivity analysis of the matrix equation from interpolation problem, J. Appl. Math. 2013 (2013) ID 518269, 8 pages.

[8] X.Y. Yin, F. Liang, Perturbation analysis for the positive definite solution of the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i X^{-1} A_i = Q$, J. Appl. Math. Comput. 43 (2013) 199–211.

[9] B.C. Levy, M. Zorzi, A contraction analysis of the convergence of risk-sensitive filters. SIAM J. Optimization Control. 54 (2016) 2154–2173.

[10] A.J. Liu, G.L. Chen, On the Hermitian positive definite solutions of nonlinear matrix equation $X^s + \sum_{i=1}^{m} A_i X^{-t_i} A_i = Q$, Appl. Math. Comput. 243 (2014) 950–959.

We consider the nonlinear matrix equations

$$X^p = A + M^T(X^{-1}\#B)^{-1}M, \tag{1}$$

and

$$X^p = A - M^T(X^{-1}\#B)^{-1}M, \tag{2}$$

where $p \geq 1$ is a positive integer, $A, B, M \in \mathbb{R}^{n\times n}$, M is an nonsingular matrix, A is positive semidefinite and B is positive definite.

We call $C\#D$ the geometric mean of positive definite matrices C and D, which is extended by T. Ando [11] from the case of two positive scalars to the case of positive semidefinite operators.

- The geometric mean of positive real numbers a and b is $\sqrt{ab}$.
- The geometric mean of positive definite matrices C and D is defined as[12,13]

$$C\#D := C^{1/2}(C^{-1/2}DC^{-1/2})^{1/2}C^{1/2}.$$

- This definition could be extended to positive semidefinite matrices C and D by a limit from above[14]

$$C\#D := \lim_{\varepsilon\to 0}(C+\varepsilon I)\#(D+\varepsilon I).$$

[11] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl. 26 (1979) 203–241.

[12] R. Bhatia, On the exponential metric increasing property, Linear Algebra Appl. 375 (2003) 211-220.

[13] J.D. Lawson, Y. Lim, The geometric mean, matrices, metrics, and more, Am. Math. Monthly 108 (2001) 797-812.

[14] R. Bhatia, Positive Definite Matrices, Princeton Univrsity Press, Princeton, 2007.

For positive definite matrices A and B, the Riccati equation $XA^{-1}X = B$ has a unique positive definite solution X, which can be expressed as $X = A \# B$.

For $p = 1$, we find that X is a positive definite solution of equation (1) (equation (2)) if and only if $Y = B^{\frac{1}{2}} X B^{\frac{1}{2}}$ is a positive definite solution of the nonlinear matrix equation

$$Y = A_1 + M_1^T Y^{\frac{1}{2}} M_1 \ (Y = A_1 - M_1^T Y^{\frac{1}{2}} M_1),$$

where $A_1 = B^{-\frac{1}{2}} A B^{-\frac{1}{2}}$ and $M_1 = B^{\frac{1}{2}} M B^{-\frac{1}{2}}$. This matrix equation has been studied in [15,16,17].

[15] M. Konstantinov, P. Petkov, I.Popchev, V. Angelova, Sensitivity of the matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i X^{p_i} A_i = 0$ $(\sigma_i = \pm 1)$, Appl. Comput. Math. 10 (2011) 409–427.

[16] I. Popchev, M. Konstantinov, P. Petkov, V. Angelova, Norm-wise, mixed and component-wise condition numbers of matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i X^{p_i} A_i = 0$ $(\sigma_i = \pm 1)$, Appl. Comput. Math. 14 (2014) 18–30.

[17] A.J. Liu, G.L. Chen, On the Hermitian positive definite solutions of nonlinear matrix equation $X^s + \sum_{i=1}^{m} A_i X^{-t_i} A_i = Q$, Appl. Math. Comput. 243 (2014) 950-959.

We also found that if A is a positive definite matrix and the nonlinear matrix equation

$$X = A(X^{-1} \# A^2)^{-1} A \tag{3}$$

has a positive definite solution X, then X also solves equation $AXA = XAX$, a special case of the Yang-Baxter equation which plays an important role in topics of statistical mechanics, braid groups and knot theory, see [18,19,20,21] and the references therein.

Note that equation (3) is a special case of equation (1).

[18] C.N. Yang, Some exact results for the many-body problem in one dimension with repulsive delta-function interaction, Phys. Rev. Lett. 19 (1967) 1312-1315.

[19] R.J. Baxter, Partition function of the eight-vertex lattice model, Ann. Phys. 70 (1972) 193-228.

[20] F. Felix, Nonlinear Equations, Quantum Groups and Duality Theorems: A Primer on the Yang-Baxter Equation, VDM Verlag, 2009.

[21] C. Yang, M. Ge, Braid Group, Knot Theory, and Statistical Mechanics, World Scientific, 1989.

Introduction

It is well-known that for positive definite matrices A and B,

$$(A\#B)^{-1} = A^{-1}\#B^{-1}.$$

Hence, the nonlinear matrix equations (1) and (2) can be written as

$$X^p = A + M^T(X\#B^{-1})M, \tag{4}$$

and

$$X^p = A - M^T(X\#B^{-1})M, \tag{5}$$

Note that B^{-1} in equations (4) and (5) is still positive definite, without loss of generality, we study the nonlinear matrix equations

$$X^p = A + M^T(X\#B)M, \tag{6}$$

and

$$X^p = A - M^T(X\#B)M, \tag{7}$$

where $p \geq 1$ is a positive integer, M is an $n \times n$ nonsingular matrix, A is a positive semidefinite matrix and B is a positive definite matrix.

Notations

- $\mathbb{R}^{n\times n}$: the set of $n \times n$ matrices with elements on field $\mathbb{R}$.
- $P(n)$: the set of $n \times n$ symmetric positive definite matrices.
- $\|\cdot\|$: the spectral norm.
- For a matrix H, $\lambda_{\max}(H)$ ($\lambda_{\min}(H)$) denotes the maximal (minimal) eigenvalue of H.
- For Hermitian matrices X and Y, $X \geq Y (X > Y)$ means that $X - Y$ is positive semidefinite (definite).

The Thompson metric on $P(n)$ is defined by

$$d(A, B) = \max\{\log M(A/B), \log M(B/A)\},$$

where $M(A/B) = \inf\{\lambda > 0 : A \leq \lambda B\} = \lambda_{\max}(B^{-\frac{1}{2}} AB^{-\frac{1}{2}})$.

It can be obtained from [19] that $P(n)$ is a complete metric space with respect to the Thompson metric.

[22] R.D. Nussbaum, Hilbert's projective metric and iterated nonlinear maps, Memoirs of Amer. Math. Soc. 75 (1988), no. 391.

Lemma [Thompson, 1963]

For any $X, Y \in P(n)$ and any $n \times n$ nonsingular matrix M, it holds

$$d(X, Y) = d(X^{-1}, Y^{-1}) = d(M^T XM, M^T YM),$$

and

$$d(X^r, Y^r) \leq |r| d(X, Y), \quad r \in [-1, 1].$$

Lemma [Lim, 2011]

For any $A, B, C, D \in P(n)$,

$$d(A + B, C + D) \leq \max\{d(A, C), d(B, D)\}.$$

Especially,

$$d(A + B, A + C) \leq d(B, C).$$

Preliminaries

lemma [Harmonic-geometric-arithmetic mean inequality]

For two positive definite matrices A and B, the harmonic-geometric-arithmetic mean inequality holds

$$2(A^{-1} + B^{-1})^{-1} \leq A \# B \leq (A + B)/2.$$

Solvability of the nonlinear matrix equation $X^p = A + M^T(X\#B)M$

Theorem

Matrix equation (6) always has a unique positive definite solution X_+. The matrix sequence $\{X_k\}$ generated by the iteration

$$\forall X_0 \in P(n), \quad X_{k+1} = \left(A + M^T(X_k\#B)M\right)^{\frac{1}{p}} \tag{8}$$

converges to X_+.

Proof: Define a map $g : P(n) \to P(n)$ by

$$g(X) = \left(A + M^T(X\#B)M\right)^{\frac{1}{p}}.$$

For any $X, Y \in P(n)$, under Thompson metric, we have

$$
\begin{aligned}
d(g(X), g(Y)) &= d\big((A + M^T(X\#B)M)^{\frac{1}{p}}, (A + M^T(Y\#B)M)^{\frac{1}{p}}\big) \\
&\leq \frac{1}{p} d\big(M^T(X\#B)M, M^T(Y\#B)M\big) \\
&\leq \frac{1}{p} d\big(X\#B, Y\#B\big) \\
&\leq \frac{1}{2p} d(X, Y),
\end{aligned}
$$

which shows that the map g is a strict contraction for Thompson metric with the contraction constant $\frac{1}{2p}$. Based on the Banach fixed point theorem, there is a unique $X_+ \in P(n)$ such that $X_+ = g(X_+)$, that is, X_+ is the unique positive definite solution of equation (6), and for every $X_0 \in P(n)$, the iterative sequence $\{X_k\}$ generated by (8) converges to X_+. $\square$

Numerical example

Let

$$A = \begin{pmatrix} 0.0120 & -0.0030 & 0.0010 \\ -0.0030 & 0.0210 & 0.0020 \\ 0.0010 & 0.0020 & 0.0070 \end{pmatrix},\ B = \begin{pmatrix} 1.1231 & 0.4497 & 0.9024 \\ 0.4497 & 0.8283 & 0.7254 \\ 0.9024 & 0.7254 & 1.0292 \end{pmatrix},$$

$M = \begin{pmatrix} 0.7922 & 0.0357 & 0.6787 \\ 0.9594 & 0.8491 & 0.7577 \\ 0.6557 & 0.9339 & 0.7431 \end{pmatrix}$, and $p = 2, 3, 4, 5, 6, 7$. We apply the fixed-point iteration (8) with starting point $X_0 = I$ to equation (6). The results are shown in Figure 1.

Numerical example

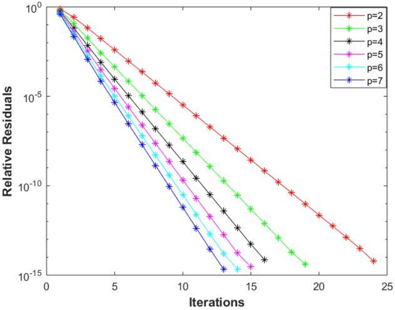

Figure 1: Convergence behaviour of the fixed-point iteration (8).

A lower and an upper bound

We obtain a lower and an upper bound of the unique positive definite solution of equation (6). To this end, we consider the following two auxiliary nonlinear matrix equations:

$$X^p = A + M^T(X^{-1} + B^{-1})^{-1}M \tag{9}$$

and

$$X^p = A + M^T BM + M^T XM, \tag{10}$$

where p, A, M and B are defined the same as in equation (6).

For $p \geq 2$ is a positive integer, it can be deduced directly from Corollary 3.6 in [23] that equations (9)and (10) have a unique positive definite solution X_* and X_{**}, respectively.

[23] C. Jung, H.-M. Kim, Y. Lim, On the solution of the nonlinear matrix equation $X^n = f(X)$, Linear Algebra Appl. 430 (2009) 2042-2052.

A lower and an upper bound

Theorem

Suppose $p \geq 2$ is a positive integer, then the unique positive definite solution X_+ of equation (6) satisfies

$$X_* \leq X_+ \leq X_{**}. \tag{11}$$

A lower and an upper bound

Theorem

For $p = 1$, let X_* and X_{**} be the unique positive definite solution of equation (9) and equation (10), respectively, we have the following results:

(i) *if both A and B are positive definite matrices, then $X_* \leq X_+$;*

(ii) *if $\rho(M^T M) < 1$, then $X_+ \leq X_{**}$.*

Although inequality (11) gives a lower and an upper bound of the unique positive definite solution X_+ of equation (6), the computation of X_* and X_{**} is nontrivial. Hence, finding a more explicit lower and upper bound which can be easily to compute is interesting.

Theorem

Suppose $p \geq 2$ is a positive integer, then the unique positive definite solution X_+ of equation (6) satisfies

$$\alpha_1 I \leq X_+ \leq \alpha_2 I, \tag{12}$$

where α_1 and α_2 are, respectively, the unique positive roots of functions

$$f_1(x) = x^p - \lambda_{\min}(M^T M)(x^{-1} + \lambda_{\min}^{-1}(B))^{-1} - \lambda_{\min}(A)$$

and

$$h_2(x) = x^p - \lambda_{\max}(M^T M)x - \lambda_{\max}(A + M^T BM).$$

Numerical example

Let $B = I_n$, $M = rand(n)$ and $A = \begin{pmatrix} 3 & -1 & & & \\ -1 & 3 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 3 & -1 \\ & & & -1 & 3 \end{pmatrix}_{n\times n}$.

Set $p = 3$ and $n = 3, 4, \ldots, 8$, we compute the lower and upper bound (12) of the unique positive definite solution of equation (6). The results are displayed in Table 1.

Numerical example

n	lower bound $\alpha_1 I$	$\lambda_{\min}(X_+)I$	$\lambda_{\max}(X_+)I$	upper bound $\alpha_2 I$
3	$1.3010I$	$1.7223I$	$2.0196I$	$2.5338I$
4	$1.2088I$	$1.5668I$	$2.1393I$	$2.6423I$
5	$1.1548I$	$1.4617I$	$2.2467I$	$3.0514I$
6	$1.1207I$	$1.4200I$	$2.5952I$	$3.6310I$
7	$1.0928I$	$1.3258I$	$2.8757I$	$4.1136I$
8	$1.0747I$	$1.3536I$	$3.1530I$	$4.5773I$

Table 1: Estimate (12) of the unique positive definite solution of equation (6).

Perturbation analysis

For the case when both A and B are perturbed by ΔA and ΔB, respectively, we have the following theorem.

Theorem

Suppose the matrices A and B are perturbed by ΔA and ΔB, respectively, where $A + \Delta A$ is symmetric positive semidefinite and $B + \Delta B$ is symmetric positive definite. Let X and $\tilde{X}$ be, respectively, the unique positive definite solution of equation (6) and the corresponding perturbed matrix equation. Then,

$$\frac{\|\tilde{X} - X\|}{\|X\|} \leq e^{\sigma} - 1, \tag{13}$$

where $\sigma = \frac{1}{p} d(\hat{A}, A) + \frac{1}{2p-1} d(\hat{B}, B)$.

Example

$$A = \begin{pmatrix} 0.0120 & -0.0030 & 0.0010 \\ -0.0030 & 0.0210 & 0.0020 \\ 0.0010 & 0.0020 & 0.0070 \end{pmatrix}, \; B = \begin{pmatrix} 1.1231 & 0.4497 & 0.9024 \\ 0.4497 & 0.8283 & 0.7254 \\ 0.9024 & 0.7254 & 1.0292 \end{pmatrix},$$

$$M = \begin{pmatrix} 0.7922 & 0.0357 & 0.6787 \\ 0.9594 & 0.8491 & 0.7577 \\ 0.6557 & 0.9339 & 0.7431 \end{pmatrix}.$$

Let

$$\Delta A = \begin{pmatrix} 0.5 & 0.1 & 0.2 \\ 0.1 & 0.7 & 0.3 \\ 0.2 & 0.3 & 1 \end{pmatrix} \times 10^{-j}, \quad \Delta B = \begin{pmatrix} 2 & 0.5 & 0.2 \\ 0.5 & 1 & 0.5 \\ 0.2 & 0.5 & 2 \end{pmatrix} \times 10^{-j}.$$

Example

p	$j=3$		$j=7$		$j=10$	
	$\frac{\|\Delta X\|}{\|X\|}$	$e^{\sigma}-1$	$\frac{\|\Delta X\|}{\|X\|}$	$e^{\sigma}-1$	$\frac{\|\Delta X\|}{\|X\|}$	$e^{\sigma}-1$
2	9.38e-04	7.84e-02	9.59e-08	8.03e-06	9.59e-11	8.03e-09
3	1.35e-03	5.10e-02	1.37e-07	5.29e-06	1.37e-010	5.29e-09
4	1.36e-03	378e-02	1.38e-07	3.95e-06	1.38e-010	3.95e-09
5	1.27e-03	3.01e-02	1.30e-07	3.15e-06	1.30e-010	3.15e-09
6	1.17e-03	2.49e-02	1.19e-07	2.62e-06	1.19e-010	2.62e-09

Table 2: Perturbation bounds (13) for the different p and j.

Solvability of the nonlinear matrix equation $X^p = A - M^T(X\#B)M$

Theorem

If $M^T(A^{\frac{1}{p}}\#B)M < A$, then equation (7) has a positive definite solution. Let $\eta = \lambda_{\max}(A)$, $\zeta = \lambda_{\min}(A) - \eta^{\frac{1}{2p}}\lambda_{\max}(M^TM)\lambda_{\max}^{\frac{1}{2}}(B)$, if $p\zeta^{1-\frac{1}{2p}} > \frac{\|M\|^2\|B\|^{3/2}\|B^{-1}\|}{2}$, then the positive definite solution of (7) is unique and it satisfies that

$$\alpha\zeta^{\frac{1}{p}}I \le X_+ \le \beta\eta^{\frac{1}{p}}I,$$

where X_+ is the unique positive definite solution of equation (7), α and β are positive solutions of the following equations

$$\begin{cases} \alpha = \frac{\lambda_{\min}^{\frac{1}{p}}(A-\beta^{\frac{1}{2}}\eta^{\frac{1}{2p}}M^TB^{\frac{1}{2}}M)}{\lambda_{\min}^{\frac{1}{p}}(A-\zeta^{\frac{1}{2p}}M^TB^{\frac{1}{2}}M)}, \\ \beta = \frac{\lambda_{\max}^{\frac{1}{p}}(A-\alpha\zeta^{\frac{1}{2p}}M^TB^{\frac{1}{2}}M)}{\eta^{\frac{1}{p}}}. \end{cases}$$

Consider the iteration

$$\begin{cases} X_0 = \beta\eta^{\frac{1}{p}} I, \\ X_{k+1} = (A - M^T(X_k \# B)M)^{\frac{1}{p}}. \end{cases} \quad (14)$$

Theorem

If $\zeta = \lambda_{\min}(A) - \lambda_{\max}(M^T M)\lambda_{\max}^{\frac{1}{2p}}(A)\lambda_{\min}^{\frac{1}{2}}(B) > 0$ *and* $p\zeta^{1-\frac{1}{2p}} \geq \frac{\|M\|^2\|B\|^{\frac{3}{2}}\|B^{-1}\|}{2}$*, then the subsequences* $\{X_{2k}\}$ *and* $\{X_{2k+1}\}$ *generated by iteration (14) both converge to the unique positive definite solution* X_+ *of equation (7), and it satisfies that*

$$X_{2k+1} \leq X_+ \leq X_{2k}, \;\; k = 0, 1, \ldots$$

Thank you for your attention!