Hindawi
Journal of Function Spaces
Volume 2018, Article ID 3291867, 8 pages
https://doi.org/10.1155/2018/3291867

Research Article

Condition Numbers of the Nonlinear Matrix Equation $X^p - A^* e^X A = I$

Chacha Stephen Chacha and Syed Muhammad Raza Shah Naqvi

Department of Mathematics, Pusan National University, Busan 46241, Republic of Korea

Correspondence should be addressed to Chacha Stephen Chacha; chchstephen@yahoo.com

Received 25 April 2018; Accepted 19 July 2018; Published 1 August 2018

Academic Editor: Henryk Hudzik

We explore the condition numbers of the nonlinear matrix equation $X^p - A^* e^X A = I$. Explicit expressions for the normwise, mixed, and componentwise condition numbers are derived. The upper bounds for the mixed and componentwise condition numbers are obtained. The numerical result favors the fact that our estimations are fairly sharp. Also, the relative upper perturbation bounds give satisfactory results for small perturbations in the input data.

1. Introduction

We consider the nonlinear matrix equation

$$X^p - A^* e^X A = I \quad (p > 1), \tag{1}$$

where A is a real or complex square matrix, I is an identity matrix, e^X is the matrix exponential function, and p is a positive integer. The basic general form of (1) is $X + A^* \mathcal{F}(X) A = Q (Q > 0)$, and it occurs in the analysis of ladder networks, the dynamic programming, control theory, stochastic filtering, and statistics [1]. Ran and Reurings in [2] studied the solutions and perturbation theory for a general matrix equation $X + A^* \mathcal{F}(X) A = Q$, where $\mathcal{F}$ represent a map from the set of all positive semidefinite matrices into a space of complex matrices and satisfy some monotonicity properties. Recently, Gao in [3] studied the Hermitian positive definite solution (HPDS) of the nonlinear matrix equation $X - A^* e^X A = I$ which is (1) for $p = 1$ and derived some necessary and sufficient conditions for the existence of the HPDS. In [4], authors derived the explicit expressions for the normwise, mixed, and componentwise condition numbers and their upper bounds for the nonlinear matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i^* X^{p_i} A_i = 0, \sigma_i = \pm 1, k \geq 2$. Authors in [5] presented a perturbation analysis of the matrix equation $C + \sum_{i=1}^{r} \sigma_i A_i X B_i + D X^s E = 0$, for positive integers $r, s \geq 2$, and employed Lyapunov majorant and fixed point-point principle to derive both local and nonlocal bounds. For more details about condition numbers, see ([6, 7]) and the references therein.

To the best of our knowledge, no one has studied the condition numbers of (1). Thus, the objective of this study is to derive the explicit expressions for the normwise, mixed, and componentwise condition numbers as well as the local upper bounds for mixed and componentwise condition numbers of (1). Finally, we give a comparative analysis for the computed condition numbers.

The following notations will be used throughout this paper: K_{rel}^{j} stands for normwise condition number; "$:=$" means equal by definition; $B_{\epsilon}^{0}(a)$ stands for a ball with center a and radius ϵ; $\mathrm{dom}(F)$ stands for domain of F; $\mathbb{C}^{n\times n}$ denotes the set of all complex $n\times n$ Hermitian matrices; if $A, B \in \mathbb{C}^{n\times n}$, then $A \geq 0 (B > 0)$ means that A is positive semidefinite (positive definite) and $A \geq B (A > B)$ means that $A - B \geq 0 (A - B > 0)$; the notation $\rho(\blacksquare)$ stands for spectral radius; A^*, A^T signify the transpose conjugate and transpose of matrix A, respectively; $\|\blacksquare\|_F$ and $\|\blacksquare\|_2$ denote the Frobenius norm and usual spectral norm, respectively; given $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$, the Kronecker product is $A \bigotimes B = [a_{ij} B] \in \mathbb{C}^{mp\times nq}$; the operator $\mathrm{vec} : \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{mn}$ is defined by $\mathrm{vec}(A) = [a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{1n}, \ldots, a_{mn}]^T$;

cond(■) means the ratio of the largest singular value to the smallest; $|A| = [a_{ij}] \in \mathbb{R}_{+}^{n\times n}$ stands for the absolute value of A.

2. Preliminaries

In this section, we provide useful definitions and lemmas that will be applied in our proofs in the next sections.

Definition 1. The condition number of a matrix function $f : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ at a point $X \in \mathbb{C}^{n\times n}$ is defined as

$$\operatorname{cond}(f, X) := \lim_{\epsilon\longrightarrow 0} \sup_{\|E\|\le\epsilon\|X\|} \frac{\|f(X+E) - f(X)\|}{\epsilon \|f(X)\|}, \quad (2)$$

for any matrix norm. The computed condition number measures the stability or sensitivity of a problem. In this case, the problem is said to be well-structured or well-posed or well-conditioned if the condition number is small and ill-conditioned if the condition number is large, where the definition of large or small condition number is problem dependent.

Definition 2. The Fréchet derivative of a matrix function $f : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ at a point $X \in \mathbb{C}^{n\times n}$ is a linear operator

$$\mathbb{C}^{n\times n} \overset{L_f(X,E)}{\longrightarrow} \mathbb{C}^{n\times n} \quad E \longmapsto L_f(X, E) \quad (3)$$

such that $f(X+E) - f(X) - L_f(X, E) = o\|E\|$ for all $E \in \mathbb{C}^{n\times n}$. The operator $L_f(X, E)$ denotes the Fréchet derivative of f at X in the direction E. If such an operator exists, f is said to be Fréchet differentiable and

$$\|L_f(X, E)\| := \max_{E\neq 0} \frac{\|L_f(X, E)\|}{\|E\|}. \quad (4)$$

Lemma 3 (see [8], Lemma 4.3.1). *Suppose that $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times q}$, $C \in \mathbb{C}^{p\times q}$, and $X \in \mathbb{C}^{n\times p}$. Then,*

$$\operatorname{vec}(AXB) = \left(B^T \otimes A\right) \operatorname{vec}(X). \quad (5)$$

Lemma 4 (see [9], pp. 178, Theorem 3.3.16(a, b)). *Let $A, B \in M_{m,n}$ be given and let $q = \min(m, n)$. Then, following inequalities hold for the decreasingly ordered singular values of $A, B, A + B$ and AB^*:*

(I) $\sigma_{i+j-1}(A + B) \le \sigma_i(A) + \sigma_j(B)$,

(II) $\sigma_{i+j-1}(AB^*) \le \sigma_i(A) + \sigma_j(B)$.

Lemma 5 (see [10], Theorem 15). *Let $A \in M_n$ and $B \in M_m$. If λ is an eigenvalue of A and μ is an eigenvalue of B, then $\lambda\mu$ is an eigenvalue of $A \otimes B$.*

For easy expansion and simplification of matrix polynomials, we need Lemma 6.

Lemma 6 (see [11]). *Let $\mathbb{Z}_+$ denote a set of positive integers including zero and $\mathbb{Z}_{++} = \mathbb{Z}_+ \setminus \{0\}$ such that $\Phi : \mathbb{Z}_+ \times \mathbb{Z}_+ \times \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$. Then,*

(i) $\{\Phi(i, 0)(X, Y) = X^i, \Phi(0, j)(X, Y) = X^j, i \in \mathbb{Z}_+, \Phi(i, j)(X, Y) = X\Phi(i-1, j)(X, Y) + Y\Phi(i, j-1)(X, Y), i, j \in \mathbb{Z}_{++}\}$.

(ii) $\Phi(0, 0)(X, Y) = I$, *and for* $\Phi(k, 1)(X, Y) = \sum_{i=1}^{k} X^{k-i} Y X^i$.

(iii) $(X + Y)^n = \sum_{i=0}^{n} \Phi(n - i, i)(X, Y), n \in \mathbb{Z}_+$.

3. Normwise, Mixed, and Componentwise Condition Numbers

In this section, we concentrate on the derivation of the explicit expressions for the normwise, mixed, and componentwise condition numbers. In order to derive the explicit expressions for the normwise, mixed, and componentwise condition numbers of (1), we consider the perturbed equation (6).

$$\widetilde{X}^p - \widetilde{A^*} e^{\widetilde{X}} \widetilde{A} = \widetilde{I}. \quad (6)$$

Replacing e^X by $\sum_{k=0}^{\infty}(1/k!)X^k$ in (6) yields

$$\widetilde{X}^p - \widetilde{A^*} \sum_{k=0}^{\infty} \frac{1}{k!} \widetilde{X}^k \widetilde{A} = \widetilde{I}. \quad (7)$$

Now, let us make small perturbations in the matrices A, X, and I as shown in

$$(X + \Delta X)^p - (A + \Delta A)^* \sum_{k=0}^{\infty} \frac{1}{k!} (X + \Delta X)^k (A + \Delta A) = I + \Delta I. \quad (8)$$

Subtracting (8) by (1) yields

$$(X + \Delta X)^p - X^p = \Delta I + (A + \Delta A)^* \sum_{k=0}^{\infty} \frac{1}{k!} (X + \Delta X)^k (A + \Delta A) - A \sum_{k=0}^{\infty} \frac{1}{k!} (X)^k. \quad (9)$$

Using Lemma 6, we have

$$\begin{aligned}(X + \Delta X)^p - X^p &= \sum_{j=0}^{p} \Phi(p - 1, j)(X, \Delta X) - X^p \\ &= \Phi(p, 0)(X, X) + \Phi(p - 1, 1)(X, \Delta X) + H(\Delta X) \\ &= \sum_{j=0}^{p-1} X^{p-1-j} \Delta X X^j + H(\Delta X),\end{aligned} \quad (10)$$

where $H(\Delta X) = \sum_{j=2}^{p} \Phi(p - j, j)(X, \Delta X)$. Because $H(\Delta X)$ have higher orders of ΔX, we omit it and consider only the first term of (10). Then, replacing $(X+\Delta X)^p - X^p$ in (9) by $\sum_{j=0}^{p-1} X^{p-1-j}\Delta X X^j$ gives

$$\begin{aligned}\sum_{j=0}^{p-1} X^{p-1-j}\Delta X X^j = \Delta I + A^* \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{i=0}^{k-1} X^{k-1-i}\Delta X X^i A \\ + A^* e^X \Delta A + \Delta A^* e^X A \\ + O\left(\|H\|^2\right).\end{aligned} \tag{11}$$

Using the fact that X is symmetric and A is real and applying the vec operator in (11), we get

$$\begin{aligned}&\left(\sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j}\right)\operatorname{vec}(\Delta X) \\ &= \operatorname{vec}(\Delta I) \\ &\quad + \left(\sum_{k=0}^{\infty}\sum_{i=0}^{k-1}\left(A^T X^i\right)\bigotimes\left(A^* X^{p-1-i}\right)\right)\operatorname{vec}(\Delta X) \\ &\quad + \left(I \bigotimes A^* e^X\right)\operatorname{vec}(\Delta A) \\ &\quad + \left(\left(A^T e^X\right)\bigotimes I\right)\operatorname{vec}(\Delta A^*) + O\left(\|H\|^2\right).\end{aligned} \tag{12}$$

Combining the terms with $\operatorname{vec}(\Delta X)$ in (12) yields

$$\begin{aligned}&\left(\sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j}\right)\operatorname{vec}(\Delta X) \\ &\quad - \left(\sum_{k=0}^{\infty}\sum_{i=0}^{k-1}\left(A^T X^i\right)\bigotimes\left(A^* X^{p-1-i}\right)\right)\operatorname{vec}(\Delta X) \\ &= \operatorname{vec}(\Delta I)\left(I \bigotimes \left(A^* e^X\right)\right)\operatorname{vec}(\Delta A) \\ &\quad + \left(\left(A^T e^X\right)\bigotimes I\right)\operatorname{vec}(\Delta A^*) + O\left(\|H\|^2\right).\end{aligned} \tag{13}$$

Then, we have

$$\begin{aligned}&\left(\left(\sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j}\right)\right. \\ &\quad \left. - \left(\sum_{k=0}^{\infty}\sum_{i=0}^{k-1}\left(A^T X^i\right)\bigotimes\left(A^* X^{p-1-i}\right)\right)\right)\operatorname{vec}(\Delta X) \\ &= \operatorname{vec}(\Delta I) + \left(\left(I \bigotimes \left(A^* e^X\right)\right)\right. \\ &\quad \left. + \left(\left(A^T e^X\right)\bigotimes I\right)\Pi\right)\operatorname{vec}(\Delta A) + O\left(\|H\|^2\right).\end{aligned} \tag{14}$$

In (14), the term $O(\|H\|^2) = \|[\Delta I, \Delta A]\|_F^2$ refers to higher order approximation of ΔX with respect to $[\Delta I, \Delta A]$ and $\Pi \in \mathbb{R}^{n^2\times n^2}$ is the vec permutation operator satisfying $\Pi\operatorname{vec}(\mathrm{A}) = \operatorname{vec}(A^T)$ and it is defined as

$$\Pi = \sum_{i=1}^{n}\sum_{j=1}^{n} E_{ij}(n\times n)\bigotimes E_{ji}(n\times n), \tag{15}$$

where $E_{ij} = e_i^n (e_j^n)^T \in \mathbb{R}^{n\times n}$ and e_i^n is the i^{th} column of the identity matrix I.

Let us define a map $\Psi : \mathbb{L}^{2n^2} \longmapsto \mathbb{L}^{n^2}$, where, $\mathbb{L}$ represents a real or complex space. We have $\Psi : \Gamma = [\operatorname{vec}(I)^T, \operatorname{vec}(A)^T]^T \longrightarrow \operatorname{vec}(X)$, where the matrices $I, A \in \mathbb{L}^{n\times n}$. According to the implicit function theorem, it is apparent that $\Delta X \longrightarrow 0$ as $[\Delta I, \Delta A] \longrightarrow 0$, because ΔX is a function of $[\Delta I, \Delta A]$.

3.1. Normwise Condition Numbers. In this subsection, we define the two kinds of normwise condition numbers. According to Rice [12], the two kinds of normwise condition numbers of map Ψ are defined by

$$\begin{aligned}\mathrm{K}_{\text{rel}}^{\mathrm{j}} &= \lim_{\delta\longrightarrow 0}\sup_{\Delta_j\le\delta}\frac{\|\Delta X\|_F}{\delta\,\|X\|_F}, \quad \mathrm{j}=1,2, \\ \text{where } \Delta_1 &= \frac{\|[\Delta I, \Delta A]\|_F}{\|[I, A]\|_F} \\ \text{and } \Delta_2 &= \left\|\left[\frac{\Delta I}{\|I\|_F}, \frac{\Delta A}{\|A\|_F}\right]\right\|_F.\end{aligned} \tag{16}$$

Now, suppose that Ψ is differentiable at $\Gamma = [\operatorname{vec}(I)^T, \operatorname{vec}(A)^T]^T$, then using Theorem 4 in [12], we have

$$\mathrm{K}_{\text{rel}}^{1} = \frac{\left\|\Psi'(\Gamma)\right\|_2 \|[I, A]\|_F}{\|X\|_F}, \tag{17}$$

where $\Psi'(\Gamma)$ is a Fréchet derivative of Ψ at Γ. It follows that

$$\begin{aligned}&B\operatorname{vec}(\Delta X) = \operatorname{vec}(\Delta I) \\ &\quad + \left(\left(I\bigotimes\left(A^* e^X\right)\right) + \left(\left(A^T e^X\right)\bigotimes I\right)\Pi\right)\operatorname{vec}(\Delta A),\end{aligned} \tag{18}$$

where

$$\begin{aligned}B &= \sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j} \\ &\quad - \sum_{k=0}^{\infty}\sum_{i=0}^{k-1}\left(A^T X^i\right)\bigotimes\left(A^* X^{p-1-i}\right).\end{aligned} \tag{19}$$

Denoting

$$\Lambda = \left[I_{n^2}, \left(I\bigotimes\left(A^* e^X\right)\right) + \left(\left(A^T e^X\right)\bigotimes I\right)\Pi\right], \tag{20}$$

we have

$$B\operatorname{vec}(\Delta X) = \Lambda\Gamma. \tag{21}$$

Now, we prove that matrix B is nonsingular in Theorem 7.

Theorem 7. *Suppose that X is the Hermitian positive definite solution of (1) and A is a real matrix with $\rho(A) < 1/e$ such that $(p-1)\lambda^p(X) > (K-1))\lambda^2(A)e^{\lambda(X)}$. Then,*

$$\begin{aligned}B &= \sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j} \\ &\quad - \sum_{k=0}^{\infty}\sum_{i=0}^{k-1}\left(A^T X^i\right)\bigotimes\left(A^* X^{k-1-i}\right)\end{aligned} \tag{22}$$

is nonsingular.

Proof. Denote

$$D := \sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j},$$
$$\text{and } C := \sum_{k=0}^{\infty}\sum_{i=0}^{k-1} \left(A^T X^i\right) \bigotimes \left(A^* X^{k-1-i}\right). \tag{23}$$

Then, we have

$$\lambda (B+C) = \lambda (D). \tag{24}$$

Using Lemmas 4 and 5, it follows that

$$\sum_{j=0}^{p-1} \lambda^j (X) \lambda^{p-1-j} (X)$$
$$\le \lambda (B) + \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^{k} \lambda^i (X) \lambda^2 (A) \lambda^{k-1-i} (X), \tag{25}$$

and

$$\sum_{j=0}^{p-1} \lambda^j (X) \lambda^{p-1-j} (X)$$
$$- \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^{k} \lambda^i (X) \lambda^2 (A) \lambda^{k-1-i} (X) \le \lambda (B). \tag{26}$$

From $X = (I + A^* e^X A)^{1/p}$, we know that $A^* e^X A > 0$; this implies that $I < X$, which means that $1 < \lambda(X)$. Therefore, $\lambda^{-1}(X)((p-1)\lambda^p(X) - (k-1)\lambda^2(A)e^{\lambda(X)}) \le \lambda(B)$.

Since $(p-1)\lambda^p(X) > (k-1)\lambda^2(A)e^{\lambda(X)}$, then we have $0 < \lambda(B)$. Therefore, we conclude that B is invertible.

From Theorem 7 above, since B is invertible, then we have $\mathrm{vec}(\Delta X) = B^{-1}\Lambda\Gamma$, and the Kronecker Fréchet derivative

$$\Psi' (I, A) = B^{-1}\Lambda. \tag{27}$$

□

Explicit expressions for the two kinds of normwise condition numbers are derived in Theorem 8.

Theorem 8. *Suppose that X is the HPD solution of (1) and B is nonsingular. Then,*

① $\mathrm{K}_{\mathrm{rel}}^1 = (\|B^{-1}\Lambda\|_2 \|[I, A]\|_F)/\|X\|_F$,

② $\mathrm{K}_{\mathrm{rel}}^2 = \|[\|I\|_F B^{-1}, \|B\|_F B^{-1}\Xi]\|_2/\|X\|_F$,

where

$$B = \sum_{j=0}^{p-1} X^j \bigotimes X^{p-1-j}$$
$$- \sum_{k=0}^{\infty}\sum_{i=0}^{k-1} \left(A^T X^i\right) \bigotimes \left(A^* X^{p-1-i}\right), \tag{28}$$
$$\Lambda = \left[I_{n^2}, \left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right) \Pi\right],$$

and

$$\Xi = \left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right) \Pi. \tag{29}$$

Proof.

① Using the fact that $\Psi'(I, A) = B^{-1}\Lambda$ and (17), we can easily get

$$\mathrm{K}_{\mathrm{rel}}^1 = \frac{\left\|B^{-1}\Lambda\right\|_2 \|[I, A]\|_F}{\|X\|_F}. \tag{30}$$

② In this case, we rewrite $\mathrm{vec}(\Delta X) = B^{-1}\Lambda\Gamma$ as $\mathrm{vec}(\Delta X) = B^{-1}\Lambda_1\kappa$, where $\Lambda_1 = \Lambda \mathrm{diag}([\|I\|_F, \|A\|_F])$ and $\kappa = \|[\Delta I/\|I\|_F, \Delta A/\|A\|_F]\|_F$, then we have $\|\mathrm{vec}(\Delta X)\|_2 = \|\Delta X\|_F$. It follows that

$$\|\Delta X\|_F = \left\|B^{-1}\Lambda\kappa\right\|_2 \le \left\|B^{-1}\Lambda\right\|_2 \|\kappa\|_2 = \Big\|\Big[\|I\|_F B^{-1},$$
$$B^{-1}\left(\left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right)\Pi\right)\Big]\Big\|_2 \|\kappa\|_2. \tag{31}$$

Finally, using (16) and (31) and $\Delta_2 = \|\kappa\|_2$, we see that

$$\mathrm{K}_{\mathrm{rel}}^2 = \frac{\left\|\left[\|I\|_F B^{-1}, \|A\|_F B^{-1}\Xi\right]\right\|_2}{\|X\|_F}. \tag{32}$$

□

3.2. Mixed and Componentwise Condition Numbers. In this subsection, we derive the explicit expressions for mixed and componentwise condition numbers of (1). The following distance function is introduced before defining mixed and componentwise condition numbers. For any vectors

$$u = [u_1, u_2, \cdots, u_m]^T,$$
$$v = [v_1, v_2, \cdots, v_m]^T,$$
$$\text{we define } \frac{u.}{v} = [w_1, w_2, \cdots, w_m]^T.$$
$$w_i = \begin{cases} \dfrac{u_i}{v_i} & \text{if } v_i \ne 0, \\ 0 & \text{if } u_i = v_i = 0, \\ \infty & \text{otherwise, for all } i = 1, 2, \cdots, m. \end{cases} \tag{33}$$

Now, let us define a distance function

$$d (u, v) = \left\|\frac{(u-v).}{v}\right\|_\infty = \max_i \left\{\frac{|u_i - v_i|.}{|v_i|}\right\}. \tag{34}$$

In the rest of this paper, we assume $d(u, v)$ is finite for any pair (u, v), and we extend the function d to matrices. That is for any matrices X, Y, we have, $d(X, Y) = d(\mathrm{vec}(X), \mathrm{vec}(Y))$. For $\epsilon > 0$, we denote $B_\epsilon^0(a) = \{x \mid d(a, x) \le \epsilon\}$.

Based on the work by Gohberg and Koltracht [13] on mixed, componentwise, and structured condition numbers, we have Definition 9.

Definition 9. Let $F : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ be a continuous mapping defined on an open interval set $\mathrm{dom}(F) \subset \mathbb{R}^p$ such that $0 \notin \mathrm{dom}(F)$ and $F(a) \neq 0$ for a given $a \in \mathbb{R}^p$.

(I) The mixed condition number of F at a is defined by

$$m(F,a) = \lim_{\epsilon \longrightarrow 0} \sup_{\substack{x \in B_\epsilon^0(a) \\ x \neq a}} \frac{\|F(x) - F(a)\|_\infty}{\|F(a)\|_\infty} \frac{1}{d(x,a)}. \tag{35}$$

(II) The componentwise condition number of F at a is defined by

$$c(F,a) = \lim_{\epsilon \longrightarrow 0} \sup_{\substack{x \in B_\epsilon^0(a) \\ x \neq a}} \frac{d(F(x), F(a))}{d(x,a)}. \tag{36}$$

If F is Fréchet differentiable at point a. Then, the explicit expressions of the mixed and componentwise condition numbers of F at a are given by Lemma 10.

Lemma 10 (see [13, 14]). *Assume F is Fréchet differentiable at point a, we have*

(1) *if* $F(a) \neq 0$, *then*

$$m(F,a) = \frac{\left\|F'(a)\,\mathrm{diag}(a)\right\|_\infty}{\|F(a)\|_\infty} = \frac{\left\|\left|F'(a)\right||a|\right\|_\infty}{\|F(a)\|_\infty}, \tag{37}$$

(2) *if* $F(a) = [f_1(a), f_2(a) \cdots, f_q(a)]^T$ *such that* $f_j(a) \neq 0$, *for* $j = 1, 2, \cdots, q$, *then*

$$\begin{aligned} c(F,a) &= \left\|\mathrm{diag}(F(a))^{-1} F'(a)\,\mathrm{diag}(a)\right\|_\infty \\ &= \left\| \frac{\left(\left|F'(a)\right||a|\right).}{|F(a)|} \right\|_\infty . \end{aligned} \tag{38}$$

Now, we derive the explicit expressions for the mixed and componentwise condition numbers and their upper bounds in Theorem 11.

Theorem 11. *Let X be the Hermitian positive definite solution of (1). Define the mapping*

$$\Psi : (A, I) \longmapsto \mathrm{vec}(X) = \Psi(A, I). \tag{39}$$

(I) *Let* $m(\Psi, X)$ *denote the mixed condition number defined by (35). Then* $m(\Psi, X)$ *has the explicit expression*

$$m(\Psi, X) = \frac{\|\Theta\|_\infty}{\|X\|_{\max}}, \tag{40}$$

where

$$\begin{aligned} \Theta &= \left|B^{-1}\right| \mathrm{vec}(|I|) \\ &+ \left|B^{-1}\left(\left(\left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right) \Pi\right)\right| \\ &\cdot \mathrm{vec}(|A|), \end{aligned} \tag{41}$$

(II) *Let* $c(\Psi, X)$ *denote the componentwise condition number defined by (36). Then,* $c(\Psi, X)$ *has the explicit expression*

$$c(\Psi, X) = \left\| \frac{\Theta.}{|\mathrm{vec}(X)|} \right\|_\infty, \tag{42}$$

where Θ *is the same as in item (I).*

Moreover, we define two simple upper bounds for $m(\Psi, X)$ *and* $c(\Psi, X)$ *given by*

$$\begin{aligned} & m_{up}(\Psi, X) \\ &= \|X\|_{\max}^{-1} \left(\left\|B^{-1}\right\|_\infty \left\||I| + 2\left|A^*\right|\left|e^X\right||A|\right\|_{\max} \right) \\ &\geq m(\Psi, X) \end{aligned} \tag{43}$$

and

$$\begin{aligned} c_{up}(\Psi, X) &= \left\|\mathrm{diag}^{-1}(\mathrm{vec}(|X|))\left|B^{-1}\right|\right\|_\infty \\ &\cdot \left\||I| + 2\left|A^*\right|\left|e^X\right||A|\right\|_{\max} \geq c(\Psi, X), \end{aligned} \tag{44}$$

respectively.

Proof. We first prove (I) using (1) of Lemma 10. In this case, for $a = \mathrm{vec}([I, A])$, we obtain that the mixed condition number is $m(\Psi, X) = \||\Psi'(a)||a|\|_\infty / \|\Psi(a)\|_\infty$. It follows from (27) that

$$m(\Psi, X) = \frac{\left\|\left|B^{-1}\Lambda\right||\Gamma|\right\|_\infty}{\|\mathrm{vec}(X)\|_\infty} = \frac{\|\Theta\|_\infty}{\|X\|_{\max}}, \tag{45}$$

where

$$\begin{aligned} \Theta &= \left|B^{-1}\Lambda\right||\Gamma| = \left|B^{-1}\right|\left[I_{n^2},\right. \\ &\left.\left|B^{-1}\left(\left(\left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right)\Pi\right)\right)\right|\right] \\ &\cdot \begin{bmatrix} \mathrm{vec}(I) \\ \mathrm{vec}(A) \end{bmatrix} = \left|B^{-1}\right| \mathrm{vec}(|I|) \\ &+ \left|B^{-1}\left(\left(\left(I \bigotimes \left(A^* e^X\right)\right) + \left(\left(A^T e^X\right) \bigotimes I\right)\Pi\right)\right)\right| \\ &\cdot \mathrm{vec}(|A|). \end{aligned} \tag{46}$$

It also holds that

$$\begin{aligned} \|\Theta\|_\infty &= \left\|\left|B^{-1}\Lambda\right||\Gamma|\right\|_\infty \leq \left\|\left|B^{-1}\right||\Lambda||\Gamma|\right\|_\infty \\ &= \left\|\left|B^{-1}\right|\right\|_\infty \||\Lambda||\Gamma|\|_\infty \\ &= \left\|B^{-1}\right\|_\infty \left\||I| + 2\left|A^*\right|\left|e^X\right||A|\right\|_{\max}. \end{aligned} \tag{47}$$

So it follows from (45) that the upper bound of $m(\Psi, X)$ is given by

$$\begin{aligned} & m(\Psi, X) \\ &\leq \|X\|_{\max}^{-1} \left(\left\|B^{-1}\right\|_\infty \left\||I| + 2\left|A^*\right|\left|e^X\right||A|\right\|_{\max} \right) \\ &= m_{\mathrm{up}}(\Psi, X). \end{aligned} \tag{48}$$

TABLE 1: Condition numbers for Example 1 evaluated using different tridiagonal matrices with size n.

n	K^1_{rel}	K^2_{rel}	$m(\Psi,X)$	$c(\Psi,X)$	cond(X)
4	0.1991	0.5114	0.5066	11.4838	1.0071
6	0.2096	0.5124	0.5066	23.1583	1.0079
8	0.2140	0.5128	0.5066	52.7581	1.0082
10	0.2162	0.5130	0.5066	134.5177	1.0084

Now we prove item (II) for the componentwise condition number $c(\Psi, X)$ of (1). Using (27) together with item (2) of Lemma 10 yields

$$c(\Psi, X) = \left\| \frac{\left(\left|\Psi'(a)\right| |a|\right).}{|\Psi(a)|} \right\|_\infty = \left\| \frac{\left(\left|B^{-1}\Lambda\right| |\Gamma|\right).}{|\operatorname{vec}(X)|} \right\|_\infty = \left\| \frac{\Theta.}{|\operatorname{vec}(X)|} \right\|_\infty, \tag{49}$$

where Θ is the same as in item (I).

Likewise, to estimate the upper bound of $c(\Psi, X)$, we have

$$c_{up}(\Psi, X) = \left\| \operatorname{diag}^{-1}(\operatorname{vec}(|X|)) \left|B^{-1}\right| \right\|_\infty \cdot \left\| |I| + 2\left|A^*\right| \left|e^X\right| |A| \right\|_{\max} \geq c(\Psi, X). \tag{50}$$

This completes our proof. □

4. Numerical Experiments

In this section, we provide some numerical examples and results. Our tests were carried out in MATLAB mark 22.0 on an Intel(R) Core(TM)i3-4005u CPU@1.7GHz 1.70GHz with 64-bit operating system. Four examples are considered. In Example 1, we evaluate normwise, mixed, and componentwise condition numbers for different tridiagonal matrix sizes. In Example 2, two cases are considered, in the first case, a badly scaled input matrix A is considered and the three kinds of condition numbers are computed. In the second case, a well-known doubly stochastic matrix of different sizes is considered and the computed condition numbers are compared. In Example 3, we make some small random perturbation in matrices I and A and evaluate the local upper perturbation bounds for the computed condition numbers. In Example 4, we consider a symmetric matrix A with some specified perturbations in the matrices I and A and evaluate condition numbers and their local upper perturbation bounds.

In each example a comparison table for the computed condition numbers is provided and a general remark is provided for all results.

For $\epsilon > 0$, we define $\epsilon_0 = \max\{\epsilon : |\Delta A| \leq \epsilon|A|, |\Delta I| \leq \epsilon|I|\}$, $\widetilde{A} = A + \Delta A$, $\widetilde{I} = I + \Delta I$, and X, $\widetilde{X}$ are solutions of (1) and (6), respectively. The solutions X and $\widetilde{X}$ are computed by a fixed point algorithm. We obtain the local normwise, mixed, and componentwise condition numbers as follows. $\zeta_N = \|\Delta X\|_F / \|X\|_F$, $\zeta_M = \|\Delta X\|_{\max} / \|X\|_{\max}$, and $\zeta_C = \|\operatorname{vec}(\Delta X)./\operatorname{vec}(X)\|_{\max}$. Then, we also define the relative mixed and componentwise local upper perturbation bounds as $\|\Delta X\|_{\max}/\|X\|_{\max} \leq \epsilon_0 m_{up}(\Psi, X)$ and $\|\operatorname{vec}(\Delta X)./\operatorname{vec}(X)\|_{\max} \leq \epsilon_0 c_{up}(\Psi, X)$, respectively.

TABLE 2: Normwise and mixed condition numbers for different p-values using an ill-conditioned input matrix A.

p	K^1_{rel}	K^2_{rel}	$m(\Psi, X)$
2	34.6169	22.7843	1.6223
4	6.5804	4.4013	0.8919
6	3.5024	2.3663	0.6400
9	2.0393	1.3890	0.4498

Here, we propose a fixed point algorithm to compute the solutions X and $\widetilde{X}$.

Fixed Point Algorithm

(I) Input an $n\times n$ matrix with $\rho(A) < 1/e$, tolerance error tol=$n\times$eps, and an initial guess $X_0 \geq I$, where n is the size of matrix A and eps $\simeq 1.1 \times 10^{-16}$ is the standard machine precision.

(II) For $i = 0, 1, 2, \cdots$, compute $X_{i+1} = (I + A^* e^X A)^{1/p}$ and relative residual

$$\text{res} = \frac{\left\|X^p - \left(A^* e^X A + I\right)\right\|_F}{\|X^p\|_F + \left\|A^* e^X A + I\right\|_F}. \tag{51}$$

(III) Exit the loop if res $\leq$ *tol*. Otherwise, go to step (II).

(IV) Display the solution X.

Example 1. We consider the tridiagonal matrix $A = W_n/25$, where $W_n = \operatorname{diag}(1, 2, 1)$.

The matrix W_n is generated by a MATLAB function "full (gallery ('tridiag', n, 1,2,1))", where n denotes the size of matrix A. For $p = 2$ and using fixed point algorithm we evaluate the solution of (1) and compute relative normwise, mixed, and componentwise condition numbers. The summary of results is recorded in Table 1.

Example 2.

(I) We consider (1) with $A = (2/5)\begin{bmatrix} 0 & 0 & 10^{-17} \\ 2 & 0 & 3 \\ 1 & 10^{-14} & 0.05 \end{bmatrix}$. Then, we evaluate the normwise mixed and condition numbers for different p-values. We employ our proposed fixed point method with $X0$=eye(n). A summary of results is displayed in Table 2.

Table 3: Normwise and mixed condition numbers evaluated at different doubly stochastic input matrix A.

(p,n)	$\mathrm{K}_{\mathrm{rel}}^{1}$	$\mathrm{K}_{\mathrm{rel}}^{2}$	$m(\Psi,X)$	$c(\Psi,X)$
(3,3)	8.7952	5.2523	3.4629	12.6425
(4,4)	3.4727	2.0392	1.4709	9.7038
(5,5)	2.2008	1.2491	0.8735	8.1702

Table 4: Condition numbers and relative upper perturbation bounds for Example 3 evaluated at $h = \{6, 8, 10, 12\}$ and $n = 3$.

h	6	8	10	12
$\mathrm{K}_{\mathrm{rel}}^{1}\Delta_1$	1.2816×10^{-7}	1.3364×10^{-9}	1.2573×10^{-11}	1.2935×10^{-13}
ζ_N	3.0635×10^{-7}	2.8668×10^{-9}	2.9382×10^{-11}	2.7542×10^{-13}
ζ_M	5.6886×10^{-7}	5.0152×10^{-9}	5.7033×10^{-11}	5.9011×10^{-13}
ζ_C	5.6991×10^{-7}	5.0237×10^{-9}	5.7129×10^{-11}	5.9093×10^{-13}
$\mathrm{K}_{\mathrm{rel}}^{2}\Delta_2$	4.8743×10^{-6}	5.6170×10^{-8}	4.8345×10^{-10}	5.2817×10^{-12}
$\epsilon_0 m_{\mathrm{up}}(\Psi,X)$	5.5744×10^{-6}	5.9432×10^{-8}	5.2362×10^{-10}	5.7641×10^{-12}
$\epsilon_0 c_{\mathrm{up}}(\Psi,X)$	5.5987×10^{-6}	5.9691×10^{-8}	5.2590×10^{-10}	5.7893×10^{-12}
$\epsilon_0 \mathrm{cond}(X)$	1.0861×10^{-5}	1.1579×10^{-7}	1.0202×10^{-9}	1.1230×10^{-11}

(II) We consider (1) with a well-known doubly stochastic matrix which has applications in communication theory and graph theory [15]. We generate a doubly stochastic matrix $A = (a_{ij}) = H/N$ such that $\sum_i a_{ij} = \sum_j a_{ij} = 1$, where H =magic(n), N =sum($H(1,:)$), and n is the size of A. Then, we apply our proposed fixed point method with $X0$=eye(n). A summary of results is recorded in Table 3

Remark. In Table 2, the numerical result indicates that the mixed condition number is much smaller than the normwise condition number.

Equation (1) converges to a Hermitian positive definite solution if $\rho(A) < 1/e$ as considered in Theorem 7 and in our proposed fixed point method. In our several trials such matrices did not satisfy componentwise condition numbers to be less than the normwise condition number. Only mixed condition number yielded the best result.

From Table 3, all the computed condition numbers decrease as the size of doubly stochastic matrix and p-values are increased.

Example 3. In this example, we use the same matrix A as in Example 1 for $n = 3$ and $p = 2$. We also set $\Delta A = 10^{-h} \times (\mathrm{rand}(\mathrm{size}(A))/4)$ and $\Delta I = 2\Delta A$ as the perturbations in the matrices A and I, respectively. Δ_1 and Δ_2 are used as they were previously defined and h is a positive integer.

The proposed fixed point algorithm is used to obtain the solutions of (1) and (6), then a summary of results is recorded in Table 4.

Example 4. In this example, we consider (1) and (6) in which $p = 2$ and

$$A = \begin{bmatrix} 0.0382 & 0.0157 & 0.0395 \\ 0.0157 & 0 & 0.0478 \\ 0.03395 & 0.0478 & 0.1065 \end{bmatrix}. \tag{52}$$

We also suppose that the perturbations in A and I are

$$\Delta A = \begin{bmatrix} -0.2 & -0.3 & 0.1 \\ 0.1 & -0.1 & 0.1 \\ -0.1 & 0.1 & 0.2 \end{bmatrix} \times 10^{-h}$$
$$\text{and } \Delta I = \begin{bmatrix} -0.3 & 0.2 & 0.1 \\ 0.1 & -0.2 & 0.3 \\ -0.1 & 0.1 & -0.3 \end{bmatrix} \times 10^{h}, \tag{53}$$

where h is a positive integer. Then, using the proposed fixed point algorithm we evaluate solutions X and $\widetilde{X}$ for (1) and (6), respectively. A summary of results is recorded in Table 5.

Remarks

(I) In Table 1, both mixed and normwise condition numbers indicate that our nonlinear matrix equation is well-conditioned since the computed condition number is very small for different matrix sizes used in the experiment. However, the componentwise condition number shows that our equation is ill-conditioned. Moreover, the componentwise and normwise condition numbers increase as the matrix size increases whereas the results for the mixed condition number remains constant.

(II) In Tables 4 and 5, all the computed condition numbers are fairly sharp and the local upper perturbation bounds for the mixed and componentwise condition numbers exist as it was expected.

(III) In Tables 4 and 5, all condition numbers decrease as the perturbation in the matrices I and A are decreased.

TABLE 5: Condition numbers and relative upper perturbation bounds for Example 4 evaluated at $h = \{6, 8, 10, 12\}$.

h	6	8	10	12
ζ_N	1.6657×10^{-7}	1.6657×10^{-9}	1.6656×10^{-11}	1.6601×10^{-13}
$\mathrm{K}_{\mathrm{rel}}^{1}\Delta_1$	1.8385×10^{-7}	1.8385×10^{-9}	1.8385×10^{-11}	1.8385×10^{-13}
ζ_M	2.9229×10^{-7}	2.9229×10^{-9}	2.9228×10^{-11}	2.9087×10^{-13}
ζ_C	3.0312×10^{-7}	3.0312×10^{-9}	3.0310×10^{-11}	3.0164×10^{-13}
$\mathrm{K}_{\mathrm{rel}}^{2}\Delta_2$	1.8081×10^{-6}	1.8081×10^{-8}	1.8083×10^{-10}	1.8081×10^{-12}
$\epsilon_0 m_{\mathrm{up}}(\Psi, X)$	1.9060×10^{-6}	1.9060×10^{-8}	1.9060×10^{-10}	1.9060×10^{-12}
$\epsilon_0 c_{\mathrm{up}}(\Psi, X)$	1.9376×10^{-6}	1.9376×10^{-8}	1.9376×10^{-10}	1.9376×10^{-12}
$\epsilon_0 \mathrm{cond}(X)$	3.3083×10^{-6}	3.3083×10^{-8}	3.3083×10^{-10}	3.3083×10^{-12}

5. Conclusion

In this paper, we studied the normwise, mixed, and componentwise condition numbers of (1). Also, we derived the explicit expressions normwise mixed and componentwise condition numbers. Local upper bounds for the mixed and componentwise condition numbers exists as it was expected. A comparative analysis for the studied condition numbers is carried out. Componentwise condition number showed the worst results among the computed condition numbers as shown in Tables 1 and 3. In general, (1) seems to be well-conditioned since the computed condition numbers are relatively small. Results in Tables 1, 2, 3, 4, and 5 indicate that the mixed condition number can reveal the true sensitivity of the problem when its input data are badly scaled or sparse.

Data Availability

Due to the nature of our subject, all the necessary steps are included in the submitted manuscript. However, if more details will be required we will provide them immediately.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank Professor Hyun-Min Kim for his remarkable guidance during this research work. This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP) (NRF-2017R1A5A1015722).

References

[1] R. Loxton, K. L. Teo, and V. Rehbock, "Positive definite solutions of the nonlinear matrix equation $X + A^{*}X^{q}A = Q(q > 0)$," *Applied Mathematics and Computation*, vol. 217, no. 14, pp. 9182–9188, 2011.

[2] A. C. Ran and M. C. Reurings, "On the nonlinear matrix equation $X+A^{*}F(X)A = Q$: solutions and perturbation theory," *Linear Algebra and Its Applications*, vol. 346, pp. 15–26, 2002.

[3] D. Gao, "On Hermitian positive definite solutions of the nonlinear matrix equation $X-A^{*}e^{X}A = I$," *Applied Mathematics and Computation*, vol. 50, no. 1, pp. 109–116, 2016.

[4] I. P. Popchev, M. M. Konstantinov, P. H. Petkov, and V. A. Angelova, "Norm-wise, mixed and component-wise condition numbers of matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i^{*} X^{pi} A_i = 0, \sigma_i = \pm 1, k \geq 2$," *International Journal of Applied and Computational Mathematics*, vol. 13, no. 1, pp. 18–30, 2014.

[5] M. Konstantinov, P. Petkov, I. Popchev, and V. Angelova, "Perturbation bounds for the matrix $C+\sum_{i=1}^{r} \sigma_i A_i X B_i + DX^{s}E = 0$," *Comptes Rendus de l'Académie Bulgare des Sciences*, vol. 61, no. 9, pp. 1111–1120, 2008.

[6] I. G. Ivanov, "Perturbation analysis for solutions of $X \pm A^{*}X^{-2}A = Q$," *Linear Algebra and its Applications*, vol. 395, pp. 313–331, 2005.

[7] S. Wang, H. Yang, and H. Li, "Condition numbers for the nonlinear matrix equation and their statistical estimation," *Linear Algebra and its Applications*, vol. 482, pp. 221–240, 2015.

[8] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge press, Cambridge, UK, 1991.

[10] B. J. Broxson, "The Kronecker Product," UNF, Theses and Dissertations, Paper 25, http://digitalcommons.unf.edu/etd/25, 2006.

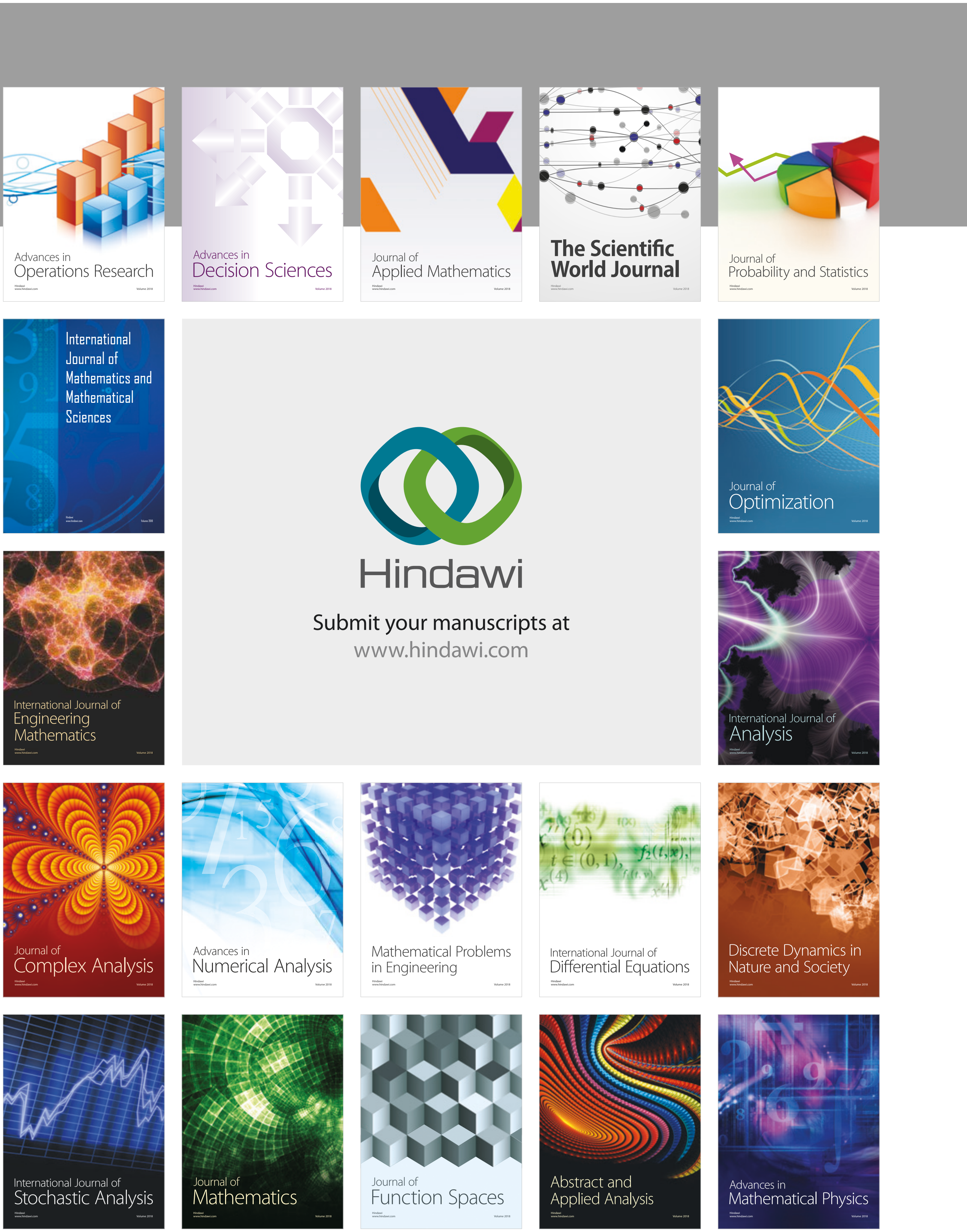
[11] J.-H. Seo and H.-M. Kim, "Convergence of pure and relaxed Newton methods for solving a matrix polynomial equation arising in stochastic models," *Linear Algebra and Its Applications*, vol. 440, pp. 34–49, 2014.

[12] J. R. Rice, "A theory of condition," *SIAM Journal on Numerical Analysis*, vol. 3, pp. 287–310, 1966.

[13] I. Gohberg and I. Koltracht, "Mixed, componentwise, and structured condition numbers," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 3, pp. 688–704, 1993.

[14] F. Cucker, H. Diao, and Y. Wei, "On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems," *Mathematics of Computation*, vol. 76, no. 258, pp. 947–963, 2007.

[15] R. A. Brualdi, "Some applications of doubly stochastic matrices," *Linear Algebra and Its Applications*, vol. 107, pp. 77–100, 1988.



Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
The Scientific
World Journal
Journal of
Probability and Statistics
International
Journal of
Mathematics and
Mathematical
Sciences
Hindawi
Submit your manuscripts at
www.hindawi.com
Journal of
Optimization
International Journal of
Engineering
Mathematics
International Journal of
Analysis
Journal of
Complex Analysis
Advances in
Numerical Analysis
Mathematical Problems
in Engineering
International Journal of
Differential Equations
Discrete Dynamics in
Nature and Society
International Journal of
Stochastic Analysis
Journal of
Mathematics
Journal of
Function Spaces
Abstract and
Applied Analysis
Advances in
Mathematical Physics