Automatica 45 (2009) 1005–1011

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Perturbation analysis and condition numbers of symmetric algebraic Riccati equations[☆]

Liangmin Zhou[a,b], Yiqin Lin[c], Yimin Wei[a,b,*], Sanzheng Qiao[d]

[a] Institute of Mathematics, School of Mathematical Science, Fudan University, Shanghai, 200433, PR China

[b] Key Laboratory of Nonlinear Science (Fudan University), Education of Ministry, PR China

[c] Department of Mathematics and Computational Science, Hunan University of Science and Engineering, Yongzhou, 425100, PR China

[d] Department of Computing and Software, McMaster University, Hamilton, ON L8S 4K1, Canada

ARTICLE INFO

Article history:
Received 9 June 2008
Received in revised form
31 October 2008
Accepted 5 November 2008
Available online 21 January 2009

Keywords:
Symmetric algebraic Riccati equation
Perturbation analysis
Condition number

ABSTRACT

This paper is devoted to the perturbation analysis of symmetric algebraic Riccati equations. Based on our perturbation analysis, the upper bounds for the normwise, mixed and componentwise condition numbers are presented. The results are demonstrated by our preliminary numerical experiments.

1. Introduction

We consider the continuous-time algebraic Riccati equation (CARE)

$$C^{\mathrm{T}}C + A^{\mathrm{T}}X + XA - XBR^{-1}B^{\mathrm{T}}X = 0 \tag{1}$$

and the following discrete-time algebraic Riccati equation (DARE)

$$Y - A^{\mathrm{T}}YA + A^{\mathrm{T}}YB(R + B^{\mathrm{T}}YB)^{-1}B^{\mathrm{T}}YA - C^{\mathrm{T}}C = 0, \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{r\times n}$, $R \in \mathbb{R}^{m\times m}$ with R being symmetric and positive definite, and $X, Y \in \mathbb{R}^{n\times n}$ are unknown matrices. Let $G = BR^{-1}B^{\mathrm{T}}$ and $Q = C^{\mathrm{T}}C$, then G and Q are symmetric and positive semidefinite and the CARE (1)and DARE (2) can be respectively rewritten as

$$Q + A^{\mathrm{T}}X + XA - XGX = 0 \tag{3}$$

and

$$Y - A^{\mathrm{T}}Y(I + GY)^{-1}A - Q = 0. \tag{4}$$

We first introduce the following stability definitions, which play an important role in the study of algebraic Riccati equations. An $n \times n$ matrix M is said to be c-stable if all of its eigenvalues lie in the open left-half complex plane, and M is said to be d-stable if its spectral radius satisfies $\rho(M) < 1$. Then to ensure the existence and uniqueness of the solutions, we assume that (A, G) in the CARE (3) is a c-stabilizable pair, that is, there is a matrix $K \in \mathbb{R}^{n\times n}$ such that the matrix $A - GK$ is c-stable, and that (A, Q) is a c-detectable pair, that is, $(A^{\mathrm{T}}, Q^{\mathrm{T}})$ is c-stabilizable. It is known (Byers, 1985; Laub, 1979) that under these conditions there exists a unique symmetric positive semidefinite solution X to the CARE (3) and the matrix $A - GX$ is c-stable. Moreover, we also assume that (A, B) in the DARE (2) is a d-stabilizable pair, that is, if $\omega^{\mathrm{T}}B = 0$ and $\omega^{\mathrm{T}}A = \lambda\omega^{\mathrm{T}}$ hold for some constant λ, then $|\lambda| < 1$ or $\omega = 0$, and that (A, C) is a d-detectable pair, that is, $(A^{\mathrm{T}}, C^{\mathrm{T}})$ is d-stabilizable. It is known (Anderson & Moore, 1979; Gudmundsson, Kenney, & Laub, 1992; Konstantinov, Petkov, & Christov, 1993) that under these conditions there exists a unique symmetric positive semidefinite solution Y to the DARE (4), and the matrix $(I + GY)^{-1}A$ is d-stable.

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Maria Elena Valcher under the direction of Editor Roberto Tempo. Y. Wei is supported by the National Natural Science Foundation of China under grant 10871051, Shanghai Education Committee under grant 08SG01 and Shanghai Science & Technology Committee under grant 08DZ2271900. Y. Lin is supported by the National Natural Science Foundation of China under grant 10801048. S. Qiao is partially supported by the Shanghai Key Laboratory of Contemporary Applied Mathematics of Fudan University during his visit.

* Corresponding address: Fudan University, Institute of Mathematics, School of Mathematical Science, Key Laboratory of Nonlinear Science, (Fudan University), Han Dan Road 220, 200433 Shanghai, PR China. Tel.: +86 21 65642347; fax: +86 21 65646073.

E-mail addresses: 062018027@fudan.edu.cn (L. Zhou), yiqinlin@hotmail.com (Y. Lin), ymwei@fudan.edu.cn, yimin.wei@gmail.com (Y. Wei), qiao@cas.mcmaster.ca (S. Qiao).

0005-1098/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2008.11.010

The CARE (3) and the DARE (4) arise in linear control and system theory. For the theory, applications, and numerical solutions of the CARE (3) and the DARE (4), see, for example, Anderson and Moore (1979), Lancaster and Rodman (1991), Lancaster and Rodman (1995), Patel, Laub, and Van Dooren (1994), Sage and White (1977), Datta (2003), Ghavimi and Laub (1995) and Guo and Laub (1999).

Perturbation analysis (Stewart & Sun, 1990) is the study of the sensitivity of the solution to the perturbations in the data of a problem. A condition number (Higham, 2002) is a measurement of the sensitivity. In the area of the perturbation analysis of the CARE (3) and the DARE (4), Konstantinov, Mehrmann, Gu, and Petkov (2003), Lin and Xu (2006), Sun (1997, 1998a,b,c), Byers (1985), and Kenney and Hewer (1990) obtained the first-order perturbation bounds for the solution to the CARE (3), Kenney, Laub, and Wette (1990) derived residual error bounds associated with the Newton refinement of approximate solutions to the CARE (3), Gudmundsson et al. (1992) derived a condition number of the DARE (4) and a bound on the relative error of a computed solution, Konstantinov et al. (1993) obtained perturbation bounds that can determine the conditioning of the DARE (4), and Sun (2002) applied the theory of linear operators and derived explicit expressions of the normwise condition numbers of the CARE (3) and the DARE (4). In this paper, by using the Kronecker product Graham (1981), we give a simple presentation of the perturbation analysis and condition numbers for the CARE (3) and the DARE (4). We first present a normwise analysis, then a componentwise analysis. Two kinds of condition numbers, called mixed and componentwise defined in Gohberg and Koltracht (1993), are considered. To the best of our knowledge, this is the first study on the componentwise condition numbers for the symmetric algebraic Riccati equations.

We adopt the following notations: $\|X\|_2$ denotes the spectral norm of a matrix, given by the square root of the largest eigenvalue of X^TX; $\|X\|_F$ is the Frobenius norm given by $\|X\|_F = \sqrt{\sum_{i,j}|X_{ij}|^2}$; $\|X\|_{\max}$ is the max norm given by $\|X\|_{\max} = \max_{i,j}|X_{ij}|$; $\|X\|_\infty$ is the infinity norm given by $\|X\|_\infty = \max_i \sum_j |X_{ij}|$; X^T is the transpose of X; $|X|$ is the matrix whose elements are $|X_{ij}|$; $\operatorname{diag}(a)$ is the diagonal matrix whose diagonal is given by a vector a; $\|a\|_2$ is the Euclidean norm of a vector, given by $\|a\|_2 = \sqrt{\sum_i |a_i|^2}$; $\|a\|_\infty$ is the infinity norm of a vector, given by $\|a\|_\infty = \max_i |a_i|$; I_n is the $n\times n$ identity matrix; E_{ij} is the (i,j)th elementary matrix whose only nonzero (i,j)-entry equals 1; Π is an $n^2\times n^2$ permutation matrix given by $\Pi = \sum_{i,j} E_{ij}\otimes E_{ji}$. For matrices $X = [x_1, x_2, \ldots, x_n] = [X_{ij}]$ and Y, $X\otimes Y = [X_{ij}Y]$ is the Kronecker product of X and Y, and the linear operator $\operatorname{vec} : \mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ is defined by $\operatorname{vec}(X) = \left[x_1^T, x_2^T, \ldots, x_n^T\right]^T$ for all $X \in R^{m\times n}$. Note that vec is a homomorphism between $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{mn}$. For any X, $\|\operatorname{vec}(X)\|_\infty = \|X\|_{\max}$ and $\|\operatorname{vec}(X)\|_2 = \|X\|_F$. See Graham (1981) for the properties of the Kronecker product and the vec operation. In particular, for an $n\times n$ matrix A, $\operatorname{vec}(A^T) = \Pi\operatorname{vec}(A)$.

The rest of the paper is organized as follows. Section 2 is devoted to the perturbation analysis and explicit expressions of three kinds of normwise condition numbers for both the CARE (3) and the DARE (4). Section 3 presents the mixed and componentwise condition numbers. Our preliminary numerical experiments are demonstrated in Section 4. Finally, we make our conclusions in Section 5.

2. Normwise condition numbers

In this section, using the Kronecker product, we first present a perturbation analysis of the CARE (3) and derive its normwise condition numbers. Then, in a similar way, we give a perturbation analysis of the DARE (4) and obtain its normwise condition numbers.

For the CARE (3), we define the mapping

$$\varphi : (A, Q, G) \mapsto \operatorname{vec}(X),$$

where X is the unique symmetric and positive semidefinite solution to the CARE (3). Suppose we introduce perturbations ΔA, ΔQ, and ΔG to the data A, Q, and G respectively and the solution to the perturbed problem is $X+\Delta X$, then the perturbed CARE (3) is

$$(X+\Delta X)(G+\Delta G)(X+\Delta X) - (X+\Delta X)(A+\Delta A) - (A+\Delta A)^T(X+\Delta X) - (Q+\Delta Q) = 0. \tag{5}$$

Dropping the second and higher-order terms in (5) yields

$$(A^T - XG)\Delta X + \Delta X(A-GX) \approx X\Delta GX - X\Delta A - \Delta A^TX - \Delta Q,$$

where the CARE Eq. (3) is used. Applying the operator vec to both sides of the above relation, using the identity

$$\operatorname{vec}(UVW) = (W^T\otimes U)\operatorname{vec}(V), \tag{6}$$

which can be verified directly, and defining

$$Z = I_n\otimes(A^T - XG) + (A-GX)^T\otimes I_n, \tag{7}$$

we obtain

$$\begin{aligned} Z\operatorname{vec}(\Delta X) &\approx (X\otimes X)\operatorname{vec}(\Delta G) - (I_n\otimes X)\operatorname{vec}(\Delta A) \\ &\quad - (X\otimes I_n)\operatorname{vec}(\Delta A^T) - \operatorname{vec}(\Delta Q) \\ &= [-(I_n\otimes X) - (X\otimes I_n)\Pi \quad -I_{n^2} \quad X\otimes X] \\ &\quad \times \operatorname{vec}([\Delta A \quad \Delta Q \quad \Delta G]). \end{aligned} \tag{8}$$

Denoting $S_2 = [-(I_n\otimes X) - (X\otimes I_n)\Pi \;\; -I_{n^2} \; X\otimes X]$, we get

$$Z\operatorname{vec}(\Delta X) \approx S_2\operatorname{vec}([\Delta A \quad \Delta Q \quad \Delta G]). \tag{9}$$

The above relation gives a first-order perturbation ΔX in the solution corresponding to the perturbations ΔA, ΔQ, and ΔG. Based on this perturbation analysis of the mapping φ, we now investigate three kinds of normwise condition numbers defined by

$$\kappa_i(\varphi) = \lim_{\epsilon\to 0}\sup_{\Delta_i\le\epsilon}\frac{\|\Delta X\|_F}{\epsilon\|X\|_F}, \quad i = 1,2,3, \tag{10}$$

where

$$\begin{aligned} \Delta_1 &= \left\|\left[\frac{\|\Delta A\|_F}{\delta_1}, \frac{\|\Delta Q\|_F}{\delta_2}, \frac{\|\Delta G\|_F}{\delta_3}\right]\right\|_2, \\ \Delta_2 &= \max\left\{\frac{\|\Delta A\|_F}{\delta_1}, \frac{\|\Delta Q\|_F}{\delta_2}, \frac{\|\Delta G\|_F}{\delta_3}\right\}, \\ \Delta_3 &= \frac{\|[\|\Delta A\|_F, \|\Delta Q\|_F, \|\Delta G\|_F]\|_2}{\|[\|A\|_F, \|Q\|_F, \|G\|_F]\|_2}, \end{aligned} \tag{11}$$

here, the nonzero parameters δ_i, $i = 1,2,3$, provide three ways in which the perturbations approach zero. In general, they are often chosen to be functions of $\|A\|_F$, $\|Q\|_F$ and $\|G\|_F$, respectively. Among all these options, the most intriguing one is that $\delta_1 = \|A\|_F$, $\delta_2 = \|Q\|_F$ and $\delta_3 = \|G\|_F$.

Before deriving the explicit expressions and an upper bound for the three kinds of normwise condition numbers for the CARE (3), we state a lemma which will be very useful throughout our discussion; see Horn and Johnson (1985) for a proof.

Lemma 1 (*Horn & Johnson, 1985*). *Let $A\in\mathbb{R}^{m\times m}$, $B\in\mathbb{R}^{n\times n}$, and $\lambda_1, \lambda_2, \ldots, \lambda_m$ and $\mu_1, \mu_2, \ldots, \mu_n$ be the eigenvalues of A and B, respectively. Then the eigenvalues of $A\otimes I_n + I_m\otimes B$ can be expressed by $\lambda_i+\mu_j$ ($i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$).*

In our case, since X and G are symmetric, $(A-GX)^T = A^T - XG$. It then follows from Lemma 1 that the matrix Z defined in (7) is nonsingular, since $A-GX$ is c-stable.

The following theorem gives explicit expressions of $\kappa_1(\varphi)$ and $\kappa_3(\varphi)$ and an upper bound for $\kappa_2(\varphi)$.

Theorem 2. *Using the notations given above, the explicit expressions and an upper bound for the three kinds of normwise numbers of the CARE (3) are*

$$\kappa_1(\varphi) \approx \frac{\|Z^{-1}S_1\|_2}{\|X\|_F}, \tag{12}$$

$$\kappa_2(\varphi) \lesssim \min\left\{\sqrt{3}\,\kappa_1(\varphi),\, \beta_c/\|X\|_F\right\}, \tag{13}$$

$$\kappa_3(\varphi) \approx \frac{\|Z^{-1}S_2\|_2\sqrt{\|A\|_F^2 + \|Q\|_F^2 + \|G\|_F^2}}{\|X\|_F}, \tag{14}$$

where $S_1 = S_2\,\mathrm{diag}([\delta_1,\ \delta_2,\ \delta_3]^{\mathrm{T}})$ *and*

$$\beta_c = \delta_1\|Z^{-1}[I_n \otimes X + (X \otimes I_n)\Pi]\|_2 + \delta_2\|Z^{-1}\|_2 + \delta_3\|Z^{-1}(X \otimes X)\|_2.$$

Proof. Introducing nonzero parameters δ_1, δ_2, and δ_3 into (9), we get $Z\mathrm{vec}(\Delta X) \approx S_1 r_1$, where $r_1 = \mathrm{vec}([\Delta A/\delta_1\ \ \Delta Q/\delta_2\ \ \Delta G/\delta_3])$. Since Z is nonsingular, we see that

$$\mathrm{vec}(\Delta X) \approx Z^{-1}S_1 r_1. \tag{15}$$

By taking the norms of both sides of (15), we obtain

$$\|\Delta X\|_F = \|\mathrm{vec}(\Delta X)\|_2 \approx \|Z^{-1}S_1 r_1\|_2 \le \|Z^{-1}S_1\|_2\|r_1\|_2. \tag{16}$$

Noting definition (10), when $i = 1$, of the condition number $\kappa_1(\varphi)$, we obtain $\|r_1\|_2 = \Delta_1 \le \epsilon$. Hence (12) holds.

In particular, setting $\delta_1 = \delta_2 = \delta_3 = 1$ in (15), we obtain

$$\mathrm{vec}(\Delta X) \approx Z^{-1}S_2 r_2, \tag{17}$$

where $r_2 = \mathrm{vec}([\Delta A\ \ \Delta Q\ \ \Delta G])$. Thus we have $\|\Delta X\|_F = \|\mathrm{vec}(\Delta X)\|_2 \lesssim \|Z^{-1}S_2\|_2\|r_2\|_2$. Similarly, we obtain (14) by using $\|r_2\|_2 = \Delta_3\|[\|A\|_F,\ \|Q\|_F,\ \|G\|_F]\|_2 \le \epsilon\|[\|A\|_F,\ \|Q\|_F,\ \|G\|_F]\|_2$.

Let $\epsilon = \Delta_2$, then it follows from (16) that

$$\|\Delta X\|_F \lesssim \|Z^{-1}S_1\|_2\sqrt{\frac{\|\Delta A\|_F^2}{\delta_1^2} + \frac{\|\Delta Q\|_F^2}{\delta_2^2} + \frac{\|\Delta G\|_F^2}{\delta_3^2}} \le \sqrt{3}\epsilon\|X\|_F\kappa_1(\varphi). \tag{18}$$

On the other hand, we rewrite (15) as

$$\mathrm{vec}(\Delta X) \approx -\delta_1 Z^{-1}[(I_n \otimes X) + (X \otimes I_n)\Pi]\frac{\mathrm{vec}(\Delta A)}{\delta_1} - \delta_2 Z^{-1}\frac{\mathrm{vec}(\Delta Q)}{\delta_2} + \delta_3 Z^{-1}(X \otimes X)\frac{\mathrm{vec}(\Delta G)}{\delta_3},$$

from which it is easy to see that

$$\|\Delta X\|_F \lesssim \epsilon\beta_c. \tag{19}$$

Finally, by (18) and (19), we get (13). □

Analogous to the analysis of the CARE (3), for the DARE (4), we define the mapping

$$\psi : (A, Q, G) \mapsto \mathrm{vec}(Y),$$

where Y is the unique symmetric and positive semidefinite solution to the DARE (4). By introducing similar perturbations, the perturbed DARE (4) is

$$(Y + \Delta Y) - (Q + \Delta Q) - (A + \Delta A)^{\mathrm{T}}(Y + \Delta Y) \times [I_n + (G + \Delta G)(Y + \Delta Y)]^{-1}(A + \Delta A) = 0. \tag{20}$$

Dropping the second and higher-order terms in (20) and denoting $W = (I_n + GY)^{-1}$, we get

$$\Delta Y - A^{\mathrm{T}}\Delta YWA + A^{\mathrm{T}}YWG\Delta YWA \approx A^{\mathrm{T}}YW\Delta A + \Delta A^{\mathrm{T}}YWA - A^{\mathrm{T}}YW\Delta GYWA + \Delta Q,$$

where the DARE (4) is used. Similarly to (8), we have

$$L\mathrm{vec}(\Delta Y) \approx [I_n \otimes (A^{\mathrm{T}}YW) + ((A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes I_n)\Pi]\mathrm{vec}(\Delta A) - [(A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes (A^{\mathrm{T}}YW)]\mathrm{vec}(\Delta G) + \mathrm{vec}(\Delta Q),$$

where $L = I_{n^2} + (A^{\mathrm{T}}W^{\mathrm{T}}) \otimes (A^{\mathrm{T}}YWG - A^{\mathrm{T}})$. Noticing that $Y(I_n + GY)^{-1} = (I_n + YG)^{-1}Y$, we get

$$\begin{aligned} A^{\mathrm{T}}YWG - A^{\mathrm{T}} &= A^{\mathrm{T}}Y(I_n + GY)^{-1}G - A^{\mathrm{T}} \\ &= A^{\mathrm{T}}(I_n + YG)^{-1}[YG - (I_n + YG)] \\ &= -A^{\mathrm{T}}W^{\mathrm{T}}. \end{aligned}$$

Therefore, we obtain a simpler definition

$$L = I_{n^2} - (A^{\mathrm{T}}W^{\mathrm{T}}) \otimes (A^{\mathrm{T}}W^{\mathrm{T}}). \tag{21}$$

Finally, similar to (9), we have

$$L\,\mathrm{vec}(\Delta Y) \approx P_2\mathrm{vec}([\Delta A\quad \Delta Q\quad \Delta G]), \tag{22}$$

where $P_2 = [I_n \otimes (A^{\mathrm{T}}YW) + ((A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes I_n)\Pi, I_{n^2}, -(A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes (A^{\mathrm{T}}YW)]$.

The above relation (22) gives a first-order perturbation ΔY in the solution corresponding to the perturbations ΔA, ΔQ, and ΔG. Based on this perturbation analysis of the mapping ψ, we now investigate three kinds of normwise condition numbers defined by

$$\kappa_i(\psi) = \lim_{\epsilon \to 0}\sup_{\Delta_i \le \epsilon}\frac{\|\Delta Y\|_F}{\epsilon\|Y\|_F}, \quad i = 1, 2, 3, \tag{23}$$

where Δ_i, $i = 1, 2, 3$, are defined in (11).

Parallel to Lemma 1, the following lemma is useful throughout our discussion, see Horn and Johnson (1985) for a proof.

Lemma 3. *Let* $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{n\times n}$, $\lambda_1, \lambda_2, \ldots, \lambda_m$ *and* $\mu_1, \mu_2, \ldots, \mu_n$ *be eigenvalues of A and B, respectively. Then the eigenvalues of* $A \otimes B$ *can be expressed by* $\lambda_i\mu_j$ $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$.

Since Y and G are symmetric, $(A^{\mathrm{T}}W^{\mathrm{T}})^{\mathrm{T}} = [A^{\mathrm{T}}(I_n + GY)^{-T}]^{\mathrm{T}} = (I_n + GY)^{-1}A$. It then follows from Lemma 3 that L defined in (21) is nonsingular, since $(I_n + GY)^{-1}A$ is d-stable.

Similarly to the proof of Theorem 2, we can obtain explicit expressions of $\kappa_1(\psi)$ and $\kappa_3(\psi)$ and an upper bound for $\kappa_2(\psi)$ given in the following theorem.

Theorem 4. *Using the notations given above, the explicit expressions and an upper bound for the three kinds of normwise numbers of the DARE (4) are*

$$\kappa_1(\psi) \approx \frac{\|L^{-1}P_1\|_2}{\|Y\|_F}, \tag{24}$$

$$\kappa_2(\psi) \lesssim \min\left\{\sqrt{3}\kappa_1(\psi),\, \beta_d/\|Y\|_F\right\}, \tag{25}$$

$$\kappa_3(\psi) \approx \frac{\|L^{-1}P_2\|_2\sqrt{\|A\|_F^2 + \|Q\|_F^2 + \|G\|_F^2}}{\|Y\|_F}, \tag{26}$$

where $P_1 = [\delta_1(I_n \otimes (A^{\mathrm{T}}YW) + [(A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes I_n]\Pi),\ \delta_2 I_{n^2}, -\delta_3((A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes (A^{\mathrm{T}}YW))]$, *and*

$$\beta_d = \delta_1\|L^{-1}(I_n \otimes (A^{\mathrm{T}}YW) + [(A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes I_n]\Pi)\|_2 + \delta_2\|L^{-1}\|_2 + \delta_3\|L^{-1}((A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes (A^{\mathrm{T}}YW))\|_2.$$

The condition numbers given by Theorem 2 involve extensive computation of Kronecker products, which are impractical to compute. However, Theorem 2 indicates that when Z is ill-conditioned, $\|Z^{-1}S_1\|_2$ and $\|Z^{-1}S_2\|_2$ can be considerably large, consequently, the CARE (3) is ill-conditioned. Thus, a large condition number of Z, which can be easily estimated by using, for example, LAPACK (Anderson et al., 1999), is an indication of an ill-conditioned CARE (3). Similarly, from Theorem 4, an ill-conditioned L indicates that the DARE (4) is ill-conditioned.

3. Mixed and componentwise condition numbers

Componentwise analysis (Higham, 1994; Rohn, 1989; Skeel, 1979) is more informative than its normwise counterpart when the data are badly scaled or sparse. To define mixed and componentwise condition numbers, we introduce the following distance function. For any $a, b \in \mathbb{R}^n$, we define $a./b = [c_1, c_2, \dots, c_n]^{\mathrm{T}}$ with

$$c_i = \begin{cases} a_i/b_i, & \text{if } b_i \neq 0, \\ 0, & \text{if } a_i = b_i = 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Then we define the distance

$$d(a, b) = \|(a - b)./b\|_\infty = \max_i\{|a_i - b_i|/|b_i|\}.$$

Note that $d(a, b) = \min\{\gamma \geq 0 \mid |a_i - b_i| \leq \gamma |b_i|,\ i = 1, 2, \dots, n\}$ if $d(a, b) < \infty$. In the rest of this paper we assume $d(a, b) < \infty$ for any pair (a, b). We can extend the function d to matrices A and B in an obvious manner: $d(A, B) = d(\mathrm{vec}(A), \mathrm{vec}(B))$. For $\epsilon > 0$, we denote $B^0(a, \epsilon) = \{x \mid d(x, a) \leq \epsilon\}$. For a vector-valued function $F : \mathbb{R}^p \to \mathbb{R}^q$, we denote $\mathrm{Dom}(F)$ as the domain of the function F. The two kinds of condition numbers introduced by Gohberg and Koltracht (1993) are considered here. The first kind, called the mixed condition number, measures the output errors in norms while the input perturbations are componentwise. The second kind, called the componentwise condition number, measures both the output and the input perturbations componentwise. They are defined as follows.

Definition 5. Let $F : \mathbb{R}^p \to \mathbb{R}^q$ be a continuous mapping defined on an open set $\mathrm{Dom}(F) \subset \mathbb{R}^p$ such that $0 \notin \mathrm{Dom}(F)$ and $F(a) \neq 0$ for a given $a \in \mathbb{R}^p$.

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

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

(2) Suppose $F(a) = \left[f_1(a), f_2(a), \dots, f_q(a)\right]^{\mathrm{T}}$ such that $f_j(a) \neq 0$ for $j = 1, 2, \dots, q$. The componentwise condition number of F at a is defined by

$$c(F, a) = \lim_{\epsilon \to 0} \sup_{\substack{x \in B^0(a,\epsilon) \\ x \neq a}} \frac{d(F(x), F(a))}{d(x, a)}.$$

The explicit expressions of the mixed and componentwise condition numbers of F at a are given by the following lemma; see Gohberg and Koltracht (1993) or Cucker, Diao, and Wei (2007) for a proof.

Lemma 6. *Suppose F is Fréchet differentiable at a. We have:*

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

$$m(F, a) = \frac{\|F'(a)\,\mathrm{diag}(a)\|_\infty}{\|F(a)\|_\infty} = \frac{\||F'(a)||a|\|_\infty}{\|F(a)\|_\infty};$$

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

$$c(F, a) = \left\|\mathrm{diag}(F(a))^{-1} F'(a)\,\mathrm{diag}(a)\right\|_\infty = \|(|F'(a)||a|)./|F(a)|\|_\infty.$$

Note that the second equality for $m(F, a)$ in the above lemma, i.e., $\|F'(a)\,\mathrm{diag}(a)\|_\infty = \||F'(a)||a|\,\|_\infty$, can be derived by

$$\|F'(a)\,\mathrm{diag}(a)\|_\infty = \||\,|F'(a)|\mathrm{diag}(a)|e\|_\infty = \||F'(a)||a|\|_\infty,$$

where $e \in \mathbb{R}^n$ is a vector whose elements are all equal to one. Similarly, we can derive the second equality for $c(F, a)$ in Lemma 6.

In the following two theorems, we present the mixed and componentwise condition numbers of the CARE (3) and the DARE (4).

Theorem 7. *For the mixed and componentwise condition numbers of the CARE (3), we have*

$$m(\varphi) \approx \|w\|_\infty / \|X\|_{\max} \text{ and } c(\varphi) \approx \|w./|\mathrm{vec}(X)|\,\|_\infty,$$

where

$$\begin{aligned} w &= |Z^{-1}S_2||\mathrm{vec}([A : Q : G])| \\ &= |Z^{-1}(I_n \otimes X) + Z^{-1}(X \otimes I_n)\Pi|\mathrm{vec}(|A|) \\ &\quad + |Z^{-1}|\mathrm{vec}(|Q|) + |Z^{-1}(X \otimes X)|\mathrm{vec}(|G|). \end{aligned}$$

Furthermore, we have simpler upper bounds

$$m_U(\varphi) := \|X\|_{\max}^{-1}(\|Z^{-1}\|_\infty \|\,|X||A| + |A|^{\mathrm{T}}|X| + |Q| + |X||G||X|\,\|_{\max}) \gtrsim m(\varphi),$$

and

$$c_U(\varphi) := \|\mathrm{diag}^{-1}(\mathrm{vec}(X))Z^{-1}\|_\infty \|\,|X||A| + |A|^{\mathrm{T}}|X| + |Q| + |X||G||X|\,\|_{\max} \gtrsim c(\varphi).$$

Proof. It follows from (17) that $\varphi'(A, Q, G) \approx Z^{-1}S_2$. From the definition of v and (1) of Lemma 6, we obtain $m(\varphi) \approx \||Z^{-1}S_2||v|\|_\infty / \|\mathrm{vec}(X)\|_\infty = \|w\|_\infty / \|X\|_{\max}$. An upper bound for $\|w\|_\infty$ can be derived from $\|w\|_\infty \leq \||Z^{-1}||S_2||v|\|_\infty \leq \|Z^{-1}\|_\infty \||S_2||v|\|_\infty = \|Z^{-1}\|_\infty \||X||A| + |A|^{\mathrm{T}}|X| + |Q| + |X||G||X|\|_{\max}$, which leads to the upper bound $m_U(\varphi)$.

For the componentwise condition number $c(\varphi)$, it follows from (2) of Lemma 6 that $c(\varphi) \approx \left\||Z^{-1}S_2||v|./|\mathrm{vec}(X)|\right\|_\infty = \|w./|\mathrm{vec}(X)|\|_\infty$. Therefore, we easily obtain $c(\varphi) \lesssim \|\mathrm{diag}^{-1}(\mathrm{vec}(X))Z^{-1}\|_\infty \||X||A| + |A|^{\mathrm{T}}|X| + |Q| + |X||G||X|\|_{\max}$. □

Similarly to the proof of Theorem 7, we have the mixed and componentwise condition numbers for the DARE (4) given by the following theorem.

Theorem 8. *For the mixed and componentwise condition numbers of the DARE (4), we have*

$$m(\psi) \approx \|\eta\|_\infty / \|Y\|_{\max} \quad \text{and} \quad c(\psi) \approx \|\eta./|\mathrm{vec}(Y)|\,\|_\infty,$$

where

$$\begin{aligned} \eta &= |L^{-1}P_2||\mathrm{vec}([A\ Q\ G])| = |L^{-1}(I_n \otimes (A^{\mathrm{T}}YW) \\ &\quad + [(A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes I_n]\Pi)|\mathrm{vec}(|A|) + |L^{-1}|\mathrm{vec}(|Q|) \\ &\quad + |L^{-1}((A^{\mathrm{T}}W^{\mathrm{T}}Y) \otimes (A^{\mathrm{T}}YW))|\mathrm{vec}(|G|). \end{aligned}$$

Furthermore, we have simpler upper bounds

$$m_U(\psi) := \|Y\|_{\max}^{-1}(\|L^{-1}\|_\infty \|\,|A^{\mathrm{T}}YW||A| + |A|^{\mathrm{T}}|YWA| + |Q| + |A^{\mathrm{T}}YW\|G\|YWA|\,\|_{\max}) \gtrsim m(\psi)$$

and

$$c_U(\psi) := \|\mathrm{diag}^{-1}(\mathrm{vec}(Y))L^{-1}\|_\infty \|\,|A^{\mathrm{T}}YW||A| + |A|^{\mathrm{T}}|YWA| + |Q| + |A^{\mathrm{T}}YW\|G\|YWA|\,\|_{\max} \gtrsim c(\psi).$$

Theorems 7 and 8 show that an ill-conditioned Z indicates large condition numbers $m(\varphi)$ and $c(\varphi)$, whereas an ill-conditioned L indicates large condition numbers $m(\psi)$ and $c(\psi)$.

4. Numerical examples

In this section, we adopt the examples in Sun (2002) to illustrate the effectiveness of our results. All the experiments were performed using MATLAB 7.0.