Linear Algebra and its Applications 469 (2015) 114–129

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Simple spectral bounds for sums of certain Kronecker products

S.V. Lototsky [1]

Department of Mathematics, USC, Los Angeles, CA 90089, USA

ARTICLE INFO

Article history:
Received 4 April 2014
Accepted 20 November 2014
Available online 2 December 2014
Submitted by Y. Wei

MSC:
primary 15A18
secondary 15A69

Keywords:
Covariance matrix
Mean-square stability
Spectral abscissa
Spectral radius

ABSTRACT

New bounds are derived for the eigenvalues of sums of Kronecker products of square matrices by relating the corresponding matrix expressions to the covariance structure of suitable bi-linear stochastic systems in discrete and continuous time.

1. Introduction

Kronecker product reduces a matrix–matrix equation to an equivalent matrix–vector form ([1] or [4, Chapter 4]). For example, consider a matrix equation $BXA^{\top} = C$ with known n-by-n matrices A, B, C, and the unknown n-by-n matrix X. To cover the most general setting, all matrices are assumed to have complex-valued entries. Introduce a column vector $\mathrm{vec}(X) = \boldsymbol{X} \in \mathbb{C}^{n^2}$ by stacking together the columns of X, left-to-right:

E-mail address: lototsky@usc.edu.
URL: http://www-bcf.usc.edu/~lototsky.
[1] Tel.: (+1) 213 740 2389; fax: (+1) 213 740 2424.

http://dx.doi.org/10.1016/j.laa.2014.11.026
0024-3795/

$$\mathrm{vec}(X) = \boldsymbol{X} = (X_{11}, \dots, X_{n1}, X_{12}, \dots, X_{n2}, \dots, X_{1n}, \dots, X_{nn})^\top. \tag{1.1}$$

Then direct computations show that the matrix equation $AXB^\top = C$ can be written in the matrix-vector form for the unknown vector $\boldsymbol{X}$ as

$$(A \otimes B)\boldsymbol{X} = \boldsymbol{C}, \quad \boldsymbol{C} = \mathrm{vec}(C), \tag{1.2}$$

where $A \otimes B$ is the `Kronecker product` of matrices A and B, that is, an n^2-by-n^2 block matrix with blocks $A_{ij}B$. In other words, (1.2) means

$$\mathrm{vec}\big(BXA^\top\big) = (A \otimes B)\,\mathrm{vec}(X), \tag{1.3}$$

with $\mathrm{vec}(\cdot)$ operation defined in (1.1).

In what follows, an n-dimensional column vector will be denoted by a lower-case bold Latin letter, e.g. $\boldsymbol{h}$, whereas upper-case regular Latin letter, e.g. A, will mean an n-by-n matrix. Then $|\boldsymbol{h}|$ is the Euclidean norm of $\boldsymbol{h}$ and $|A|$ is the induced matrix norm

$$|A| = \max\big\{|A\boldsymbol{h}| : |\boldsymbol{h}| = 1\big\}.$$

For a matrix $A \in \mathbb{C}^{n\times n}$, $\overline{A}$ is the matrix with complex conjugate entries, $A^\top$ means transposition, and A^* denotes the conjugate transpose: $A^* = \overline{A^\top} = \overline{A}^\top$. The same notations, $\overline{}$, $^\top$, and *, will also be used for column vectors in $\mathbb{C}^n$. The identity matrix is I.

For a square matrix A, define the following numbers:

$$\boldsymbol{\rho}(A) = \max\big\{\big|\lambda(A)\big| : \lambda(A) \text{ is an eigenvalue of } A\big\} \quad (\texttt{spectral radius} \text{ of } A);$$
$$\boldsymbol{\alpha}(A) = \max\big\{\Re\lambda(A) : \lambda(A) \text{ is an eigenvalue of } A\big\} \quad (\texttt{spectral abscissa} \text{ of } A);$$
$$\boldsymbol{\varrho}(A) = \min\big\{\Re\lambda(A) : \lambda(A) \text{ is an eigenvalue of } A\big\}.$$

If H is a Hermitian matrix, then

$$\boldsymbol{\varrho}(H) = \lambda_{\min}(H), \qquad \boldsymbol{\alpha}(H) = \boldsymbol{\rho}(H) = \lambda_{\max}(H), \tag{1.4}$$

$$\boldsymbol{\varrho}(H)|\boldsymbol{x}|^2 \le \boldsymbol{x}^* H \boldsymbol{x} \le \boldsymbol{\alpha}(H)|\boldsymbol{x}|^2. \tag{1.5}$$

While eigenvalues of the matrices $A \otimes B$ and $A \otimes I + I \otimes B$ can be easily expressed in terms of the eigenvalues of the matrices A and B [4, Theorems 4.2.12 and 4.4.5], there is, in general, no easy way to bound the eigenvalues of the matrices

$$C_{A,B} = \overline{A} \otimes I + I \otimes A + \sum_{k=1}^{m} \overline{B}_k \otimes B_k \quad \text{and} \tag{1.6}$$

$$D_{A,B} = \overline{A} \otimes A + \sum_{k=1}^{m} \overline{B}_k \otimes B_k. \tag{1.7}$$

Matrices $C_{A,B}$ and $D_{A,B}$ appear, for example, in the study of bi-linear stochastic systems, both finite-dimensional [2] and infinite-dimensional [6]. Paper [3] presents one of the first investigations of the spectral properties of (1.7) and (1.6). The main result of the current paper provides another contribution to the subject by establishing explicit upper and lower bounds on $\boldsymbol{\alpha}(C_{A,B})$ and $\boldsymbol{\rho}(D_{A,B})$ in terms of the eigenvalues of other matrices that are Hermitian and of smaller size.

The matrix expressions $A \otimes B$ and $A \otimes I + I \otimes B$ have designated names (Kronecker product and Kronecker sum), but there is no established terminology for (1.6) and (1.7). In what follows, (1.6) will be referred to as the continuous-time stochastic Kronecker sum and (1.7) will be referred to as the discrete-time stochastic Kronecker sum. The reason for this choice of names is motivated by the type of problems in which the corresponding matrix expressions appear.

The paper is organized as follows. Section 2 outlines the well-known approach to the analysis of matrices $C_{A,B}$ and $D_{A,B}$ using, respectively, the generalized Lyapunov and Stein operators, and illustrates the reasons why an alternative approach, based on the analysis of the covariance structure of a suitable stochastic system, is necessary. Section 3 explains how matrices of the type (1.6) appear in the analysis of continuous-time bi-linear stochastic systems and establishes upper and lower bounds on the spectral abscissa of the matrix $C_{A,B}$ (Theorem 3.3). Section 4 explains how matrices of the type (1.7) appear in the analysis of discrete-time bi-linear stochastic systems and establishes upper and lower bounds on the spectral radius of the matrix $D_{A,B}$ (Theorem 4.2). Theorems 3.3 and 4.2 are the main results of the paper. The connection with stochastic systems also illustrates why it is indeed natural to bound spectral abscissa for matrices of the type (1.6) and the spectral radius for matrices of the type (1.7). Section 5 provides several examples illustrating Theorems 3.3 and 4.2 and discusses the particular case when all the matrices A and B_k are Hermitian.

2. General background

Given matrices $A, B_1, \ldots, B_m$ in $\mathbb{C}^{n\times n}$, define the generalized Lyapunov operator $\mathcal{L}_{A,B}$ and the generalized Stein operator $\mathcal{S}_{0;A,B}$, each acting from $\mathbb{C}^{n\times n}$ to $\mathbb{C}^{n\times n}$, by

$$\mathcal{L}_{A,B}(X) = A^*X + XA + \sum_{k=1}^{m} B_k^* X B_k,$$

$$\mathcal{S}_{0;A,B}(X) = A^*XA + \sum_{k=1}^{m} B_k^* X B_k, \quad X \in \mathbb{C}^{n\times n}.$$

The adjoint operators $\mathcal{L}^*_{A,B}$ and $\mathcal{S}^*_{0;A,B}$ are defined using the matrix inner product

$$\langle X, Y\rangle = \operatorname{tr}\bigl(Y^*X\bigr) = \bigl(\operatorname{vec}(X)\bigr)^* \operatorname{vec}(Y), \quad X, Y \in \mathbb{C}^{n\times n},$$

so that

$$\big\langle Y, \mathcal{L}_{A,B}(X)\big\rangle = \big\langle \mathcal{L}^*_{A,B}(Y), X\big\rangle, \qquad \big\langle Y, \mathcal{S}_{0;A,B}(X)\big\rangle = \big\langle \mathcal{S}^*_{0;A,B}(Y), X\big\rangle.$$

The following equalities are verified by direct computation (cf. [2, Chapter 3]):

$$\big(\mathcal{L}_{A,B}(X)\big)^* = \mathcal{L}_{A,B}\big(X^*\big), \qquad \big(\mathcal{S}_{0;A,B}(X)\big)^* = \mathcal{S}_{0;A,B}\big(X^*\big); \tag{2.1}$$

$$\mathcal{L}^*_{A,B} = \mathcal{L}_{A^*,B^*}, \qquad \mathcal{S}^*_{0;A,B} = \mathcal{S}_{0;A^*,B^*}; \tag{2.2}$$

$$C_{A,B}\operatorname{vec}(X) = \operatorname{vec}\big(\mathcal{L}^*_{A,B}(X)\big), \qquad D_{A,B}\operatorname{vec}(X) = \operatorname{vec}\big(\mathcal{S}^*_{0;A,B}(X)\big). \tag{2.3}$$

Equalities (2.3) imply that the matrix $C_{A,B}$ has the same eigenvalues as the operator $\mathcal{L}^*_{A,B}$, and the matrix $D_{A,B}$ has the same eigenvalues as the operator $\mathcal{S}^*_{0;A,B}$. If the operators $\mathcal{L}_{A,B}$ and $\mathcal{S}_{0;A,B}$ are Hermitian (which, by (2.2), happens if $A, B_1, \dots, B_m$ are all Hermitian matrices), then one can use variational characterization of the eigenvalues of $\mathcal{L}_{A,B}$ (resp. $\mathcal{S}_{0;A,B}$) to bound eigenvalues of $C_{A,B}$ (resp. $D_{A,B}$). For example,

$$\lambda_{\max}(C_{A,B}) = \sup_{X\neq 0} \frac{\langle X, \mathcal{L}_{A,B}(X)\rangle}{\operatorname{tr}(X^*X)} \geq \frac{\langle I, \mathcal{L}_{A,B}(I)\rangle}{\operatorname{tr}(I^*I)} = \frac{\operatorname{tr}(M_{A,B})}{n}, \tag{2.4}$$

where

$$M_{A,B} = \mathcal{L}_{A,B}(I) = A + A^* + \sum_{k=1}^{m} B_k^* B_k. \tag{2.5}$$

Similarly,

$$\lambda_{\max}(D_{A,B}) = \sup_{X\neq 0} \frac{\langle X, \mathcal{S}_{0;A,B}(X)\rangle}{\operatorname{tr}(X^*X)} \geq \frac{\langle I, \mathcal{S}_{0;A,B}(I)\rangle}{\operatorname{tr}(I^*I)} = \frac{\operatorname{tr}(N_{A,B})}{n}, \tag{2.6}$$

where

$$N_{A,B} = \mathcal{S}_{0;A,B}(I) = A^*A + \sum_{k=1}^{m} B_k^* B_k. \tag{2.7}$$

The matrices $M_{A,B}$ and $N_{A,B}$ will be of central importance in the following sections, and this is why these matrices are written in the general form; when A, B_k are Hermitian, then, of course, $M_{A,B} = 2A + \sum_{k=1}^m B_k^2$ and $N_{A,B} = A^2 + \sum_{k=1}^m B_k^2$.

If $A, B_1, \dots, B_m$ are all Hermitian, then both $C_{A,B}$ and $D_{A,B}$ are sums of Hermitian matrices, and the eigenvalues of each matrix in the sums are explicitly expressed in terms of the eigenvalues of A and B_k. This leads to other bounds on the eigenvalues of $C_{A,B}$ and $D_{A,B}$, such as Weyl's inequality, and Section 5 discusses some of those bounds.

In the general case, when A and B_k are not Hermitian, equalities (2.1)–(2.3) can connect various properties of the spectra of $C_{A,B}$ and $D_{A,B}$ with those of $\mathcal{L}_{A,B}$ and $\mathcal{S}_{0;A,B}$

(see [3]), but there are no analogues of (2.4) or (2.6). Indeed, for a non-Hermitian matrix Y, the number $\mathbf{r}(Y) = \sup_{|\boldsymbol{x}|=1} |\boldsymbol{x}^* Y \boldsymbol{x}|$ provides an *upper* bound on both the spectral radius $\boldsymbol{\rho}(Y)$ (cf. [4, Property 1.2.6]) and the spectral abscissa $\boldsymbol{\alpha}(Y) \le \boldsymbol{\rho}(Y)$. Because of that, a *lower* bound on $\mathbf{r}$ provides no information about either $\boldsymbol{\alpha}$ or $\boldsymbol{\rho}$.

Nonetheless, with an alternative approach, it is still possible to derive upper and lower bounds on $\boldsymbol{\alpha}(C_{A,B})$ (resp. $\boldsymbol{\rho}(D_{A,B})$) using the matrix $M_{A,B}$ from (2.5) (resp. $N_{A,B}$ from (2.7)) when the matrices $A, B_1, \dots, B_m$ are not necessarily Hermitian. This alternative approach is based on the analysis of the covariance structure of a suitably constructed stochastic system.

Given two $\mathbb{C}^n$-valued random column-vectors

$$\boldsymbol{x} = (x_1, \dots, x_n)^\top \quad \text{and} \quad \boldsymbol{y} = (y_1, \dots, y_n)^\top,$$

define the matrix $U_{x,y} = \mathbb{E}(\boldsymbol{x}\boldsymbol{y}^*)$, where $\mathbb{E}$ denotes the expected value. If $\mathbb{E}\boldsymbol{x} = \mathbb{E}\boldsymbol{y} = 0$, then $U_{x,y}$ is the covariance matrix of $\boldsymbol{x}$ and $\boldsymbol{y}$. With the notation $\boldsymbol{U}_{x,y} = \mathrm{vec}(U_{x,y})$, we find $|\boldsymbol{U}_{x,y}|^2 = \sum_{i,j=1}^n |\mathbb{E}x_i y_j^*|^2$, and the Cauchy–Schwarz inequality $|\mathbb{E}x_i y_j^*|^2 \le \mathbb{E}|x_i|^2 \mathbb{E}|y_j|^2$ leads to an upper bound on $|\boldsymbol{U}_{x,y}|$:

$$|\boldsymbol{U}_{x,y}|^2 \le r_x r_y, \tag{2.8}$$

where $r_x = \sum_{i=1}^n \mathbb{E}|x_i|^2$. In the particular case $\boldsymbol{x} = \boldsymbol{y}$,

$$\begin{aligned} n|\boldsymbol{U}_{x,x}|^2 &= n\sum_{i,j=1}^n \left|\mathbb{E}x_i x_j^*\right|^2 = n\sum_{i=1}^n \left(\mathbb{E}|x_i|^2\right)^2 + n\sum_{i\ne j}\left|\mathbb{E}x_i x_j^*\right|^2 \\ &\ge n\sum_{i=1}^n \left(\mathbb{E}|x_i|^2\right)^2 \ge \left(\sum_{i=1}^n \mathbb{E}|x_i|^2\right)^2, \end{aligned}$$

leading to a lower bound:

$$|\boldsymbol{U}_{x,x}| \ge n^{-1/2} r_x. \tag{2.9}$$

In Section 3, we construct a continuous-time stochastic system for which the matrix $C_{A,B}$ describes time evolution of $U_{x,y}$ and the matrix $M_{A,B}$ describes time evolution of r_x; then (2.8) and (2.9) will lead to upper and lower bounds on $\boldsymbol{\alpha}(C_{A,B})$.

In Section 4, we construct a discrete-time stochastic system for which the matrix $D_{A,B}$ describes time evolution of $U_{x,y}$ and the matrix $N_{A,B}$ describes time evolution of r_x; then (2.8) and (2.9) will lead to upper and lower bounds on $\boldsymbol{\rho}(D_{A,B})$.

3. Continuous-time stochastic Kronecker sum

Given matrices $A, B_1, \dots, B_m \in \mathbb{C}^{n\times n}$, consider two $\mathbb{C}^n$-valued stochastic processes $\boldsymbol{x}(t) = (x_1(t), \dots, x_n(t))^\top$ and $\boldsymbol{y}(t) = (y_1(t), \dots, y_n(t))^\top$, $t \ge 0$, defined by the Itô integral equations

$$\boldsymbol{x}(t) = \boldsymbol{u} + \int_0^t A\boldsymbol{x}(s)ds + \sum_{k=1}^m \int_0^t B_k\boldsymbol{x}(s)dw_k(s),$$

$$\boldsymbol{y}(t) = \boldsymbol{v} + \int_0^t A\boldsymbol{y}(s)ds + \sum_{k=1}^m \int_0^t B_k\boldsymbol{y}(s)dw_k(s). \tag{3.1}$$

Both equations in (3.1) are driven by independent standard Brownian motions $w_1, \ldots, w_m$. The initial conditions $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ are non-random. The processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ satisfy the same equation and differ only in the initial conditions. Existence and uniqueness of solution of (3.1) are well-known [7, Theorem 5.2.1], and then $\boldsymbol{u} = \boldsymbol{v}$ implies $\boldsymbol{x}(t) = \boldsymbol{y}(t)$ for all $t > 0$. The terms $dw_k(t)$ can be considered as continuous-time analogues of a discrete-time white noise sequence. The term *bi-linear* in connection with (3.1) reflects the fact that the noise process enters the system in a multiplicative, as opposed to additive, way.

The differential form

$$d\boldsymbol{x}(t) = A\boldsymbol{x}(t)dt + \sum_{k=1}^m B_k\boldsymbol{x}(t)dw_k(t), \qquad d\boldsymbol{y}(t) = A\boldsymbol{y}(t)dt + \sum_{k=1}^m B_k\boldsymbol{y}(t)dw_k(t)$$

is a more compact, and less formal, way to write (3.1).

The peculiar behavior of white noise in continuous time, often written informally as $(dw(t))^2 = dt$, makes it necessary to modify the usual product rule for the derivatives. The result is known as the `Itô formula`; its one-dimensional version is presented below for reader's convenience.

Proposition 3.1. *If a, b, σ, and μ are globally Lipschits continuous functions and x, y are non-random numbers, then*

(a) *There are unique continuous random processes f and g such that*

$$f(t) = x + \int_0^t a\big(f(s)\big)ds + \int_0^t \sigma\big(f(s)\big)dw(s),$$

$$g(t) = y + \int_0^t b\big(g(s)\big)ds + \int_0^t \mu\big(g(s)\big)dw(s);$$

(b) *The following equality holds:*

$$\mathbb{E}\big(f(t)g(t)\big) = xy + \int_0^t \mathbb{E}\Big(f(s)b\big(g(s)\big) + g(s)a\big(f(s)\big) + \sigma\big(f(s)\big)\mu\big(g(s)\big)\Big)ds. \tag{3.2}$$

Proof. In differential form,

$$d(fg) = f dg + g df + \sigma\mu\, dt,$$

where the first two terms on the right come from the usual product rule and the third term, known as the Itô correction, is a consequence of $(dw(t))^2 = dt$. The expected value of the Itô stochastic integral is zero:

$$\mathbb{E}\int_0^t f(s)\mu\big(g(s)\big)dw(s) = \mathbb{E}\int_0^t g(s)\sigma\big(f(s)\big)dw(s) = 0,$$

and then (3.2) follows. See [7, Chapter 4] for more details and [2, Sections 1.1–1.3] for a more detailed summary. □

Proposition 3.2. *Given the random vectors* $\boldsymbol{x}(t)$ *and* $\boldsymbol{y}(t)$ *from* (3.1), *define the matrix* $V(t) = \mathbb{E}(\boldsymbol{x}(t)\boldsymbol{y}^*(t))$ *and the number* $r_x(t) = \mathbb{E}(\boldsymbol{x}^*(t)\boldsymbol{x}(t))$. *Then*

(1) *The vector*

$$\boldsymbol{U}(t) = \mathrm{vec}\big(V(t)\big)$$

satisfies

$$\boldsymbol{U}(t) = e^{tC_{A,B}}\boldsymbol{U}(0), \tag{3.3}$$

with the matrix $C_{A,B}$ *from* (1.6);

(2) *The number* $r_x(t)$ *satisfies*

$$|\boldsymbol{u}|^2 e^{\gamma t} \le r_x(t) \le |\boldsymbol{u}|^2 e^{\beta t}, \tag{3.4}$$

where $\gamma = \lambda_{\min}(M_{A,B})$, $\beta = \lambda_{\max}(M_{A,B})$, *and* $M_{A,B}$ *is defined in* (2.5).

Proof. While equality (3.3) is well-known (see, for example, [2, Theorem 1.4.3]), there is no standard reference for inequality (3.4). For reader's convenience, below is an outline of the computations leading to (3.3) and (3.4).

In differential form,

$$d\boldsymbol{x}(t) = A\boldsymbol{x}(t)dt + \sum_{k=1}^m B_k\boldsymbol{x}(t)dw_k(t), \qquad d\boldsymbol{y}^*(t) = \boldsymbol{y}^*(t)A^*dt + \sum_{k=1}^m \boldsymbol{y}^*(t)B_k^*dw_k(t).$$

By the Itô formula,

$$V(t) = V(0) + \int_0^t \left(AV(s) + V(s)A^* + \sum_{k=1}^m B_k V(s) B_k^* \right) ds,$$

and (3.3) follows from (1.3).

Similarly,

$$r_x(t) = r_x(0) + \int_0^t \mathbb{E}\big(\boldsymbol{x}^*(s) M_{A,B} \boldsymbol{x}(s)\big) ds,$$

and then, for every real number a,

$$r_x(t) = r_x(0) + \int_0^t a r_x(s) ds + \int_0^t f_a(s) ds,$$

where

$$f_a(s) = \int_0^t \mathbb{E}\big(\boldsymbol{x}^*(s) M_{A,B} \boldsymbol{x}(s) - a\boldsymbol{x}^*(s)\boldsymbol{x}(s)\big) ds.$$

In other words,

$$r_x(t) = |\boldsymbol{u}|^2 e^{at} + \int_0^t e^{a(t-s)} f_a(s) ds.$$

If $a = \gamma$ (the smallest eigenvalue of $M_{A,B}$), then $f_a(s) \geq 0$ and the lower bound in (3.4) follows; if $a = \beta$ (the largest eigenvalue of $M_{A,B}$), then $f_a(s) \leq 0$, and the upper bound in (3.4) follows. $\square$

Given the origin of Eq. (3.3), the matrix $C_{A,B}$ is natural to call the `continuous-time stochastic Kronecker sum` of the matrices A and B_k.

It is known [8, Theorem 15.3] that the spectral abscissa $\boldsymbol{\alpha}(A)$ of a matrix $A \in \mathbb{C}^{n\times n}$ satisfies

$$\boldsymbol{\alpha}(A) = \lim_{t\to+\infty} \frac{1}{t} \ln\left|e^{tA}\right|. \tag{3.5}$$

Together with Proposition 3.2, equality (3.5) leads to a two-sided bound on the spectral abscissa of $C_{A,B}$, which is the main result of this section.

Theorem 3.3. *For every matrices $A, B_1, \ldots, B_m \in \mathbb{C}^{n\times n}$,*

$$\lambda_{\min}(M_{A,B}) \leq \boldsymbol{\alpha}(C_{A,B}) \leq \lambda_{\max}(M_{A,B}). \tag{3.6}$$

Proof. As in Proposition 3.2, we write $\beta = \lambda_{\max}(M_{A,B})$, $\gamma = \lambda_{\min}(M_{A,B})$. It follows from (3.3) that $|\boldsymbol{U}(t)| = |e^{tC_{A,B}}\boldsymbol{U}(0)|$. By (2.8), $|\boldsymbol{U}(t)| \le \sqrt{r_x(t)r_y(t)}$, and then (3.4) implies

$$\left|e^{tC_{A,B}}\boldsymbol{U}(0)\right| \le |\boldsymbol{u}||\boldsymbol{v}|e^{\beta t}. \tag{3.7}$$

Recall that $\boldsymbol{U}(0) = \mathrm{vec}(\boldsymbol{u}\boldsymbol{v}^*) = \boldsymbol{v} \otimes \boldsymbol{u}$. Let $\boldsymbol{u}_i$, $i = 1, \ldots, n$, be a unit basis in $\mathbb{C}^n$. Then, for every $\boldsymbol{U} \in \mathbb{C}^{n^2}$, $\boldsymbol{U} = \sum_{i,j=1}^n c_{ij}\boldsymbol{u}_i \otimes \boldsymbol{u}_j$ with $c_{ij} \in \mathbb{C}$, and $|\boldsymbol{U}|^2 = \sum_{i,j=1}^n |c_{ij}|^2$. By (3.7), followed by the triangle and Cauchy–Schwarz inequalities,

$$\left|e^{tC_{A,B}}\boldsymbol{U}\right| \le \left(\sum_{i,j=1}^n |c_{ij}|\right)e^{\beta t} \le n|\boldsymbol{U}|e^{\beta t},$$

that is,

$$\left|e^{tC_{A,B}}\right| \le ne^{\beta t}. \tag{3.8}$$

Then the upper bound in (3.6) follows from (3.8) and (3.5).

To derive the lower bound, take $\boldsymbol{u} = \boldsymbol{v}$ with $|\boldsymbol{u}| = 1$, so that $\boldsymbol{x}(t) = \boldsymbol{y}(t)$ for all $t \ge 0$. Then (2.9) and (4.11) imply

$$n^{-1/2}e^{\gamma t} \le \left|\boldsymbol{U}(n)\right| \le \left|e^{tC_{A,B}}\right|,$$

and the lower bound in (3.6) follows from (3.5). □

Examples of applications of (3.6) are in Section 5.

4. Discrete-time stochastic Kronecker sum

Given matrices $A, B_1, \ldots, B_m \in \mathbb{C}^{n\times n}$, consider two $\mathbb{C}^n$-valued random sequences $\boldsymbol{x}(\ell) = (x_1(\ell), \ldots, x_n(\ell))^\top$ and $\boldsymbol{y}(\ell) = (y_1(\ell), \ldots, y_n(\ell))^\top$, $\ell = 0, 1, 2, \ldots$, defined by

$$\begin{aligned}\boldsymbol{x}(\ell+1) &= A\boldsymbol{x}(\ell) + \sum_{k=1}^m B_k\boldsymbol{x}(\ell)\xi_{\ell+1,k}, \quad \boldsymbol{x}(0) = \boldsymbol{u},\\ \boldsymbol{y}(\ell+1) &= A\boldsymbol{y}(\ell) + \sum_{k=1}^m B_k\boldsymbol{y}(\ell)\xi_{\ell+1,k}, \quad \boldsymbol{y}(0) = \boldsymbol{v}.\end{aligned} \tag{4.1}$$

Both equations in (3.1) are driven by a `white noise` sequence $\xi_{\ell,k}$, $\ell \ge 1$, $k = 1, \ldots, m$ of independent, for all ℓ and k, random variables, all with zero mean and unit variance:

$$\mathbb{E}\xi_{\ell,k} = 0, \qquad \mathbb{E}\xi_{\ell,k}^2 = 1, \qquad \mathbb{E}(\xi_{\ell,k}\xi_{p,q}) = 0 \quad \text{if } \ell \ne p \text{ or } k \ne q. \tag{4.2}$$

The initial conditions $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ are non-random. Note that the sequences $\boldsymbol{x}(\ell)$ and $\boldsymbol{y}(\ell)$ satisfy the same equation and differ only in the initial conditions. In particular, $\boldsymbol{u} = \boldsymbol{v}$ implies $\boldsymbol{x}(\ell) = \boldsymbol{y}(\ell)$ for all $\ell \geq 0$. The term *bi-linear* in connection with (4.1) reflects the fact that the noise sequence enters the system in a multiplicative, as opposed to additive, way.

Proposition 4.1. *Given the random vectors $\boldsymbol{x}(\ell)$ and $\boldsymbol{y}(\ell)$ from (4.1), define the matrix $V(\ell) = \mathbb{E}(\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell))$ and the number $r_x(\ell) = \mathbb{E}(\boldsymbol{x}^*(\ell)\boldsymbol{x}(\ell)) = \mathbb{E}|\boldsymbol{x}(\ell)|^2$. Then*

(1) *The vector $\boldsymbol{U}(\ell) = \mathrm{vec}(V(\ell))$ satisfies*

$$\boldsymbol{U}(\ell) = (D_{A,B})^{\ell}\boldsymbol{U}(0), \tag{4.3}$$

with the matrix $D_{A,B}$ from (1.7);

(2) *The number $r_x(\ell)$ satisfies*

$$|\boldsymbol{u}|^2\gamma^{\ell} \leq r_x(\ell) \leq |\boldsymbol{u}|^2\beta^{\ell}, \tag{4.4}$$

where $\gamma = \lambda_{\max}(N_{A,B})$, $\beta = \lambda_{\min}(N_{A,B})$, and the matrix $N_{A,B}$ is defined in (2.7).

Proof. By (4.1),

$$\boldsymbol{x}(\ell+1) = A\boldsymbol{x}(\ell) + \sum_{k=1}^{m} B_k\boldsymbol{x}(\ell)\xi_{\ell+1,k}, \qquad \boldsymbol{y}^*(\ell+1) = \boldsymbol{y}^*(\ell)A^* + \sum_{k=1}^{m} \boldsymbol{y}^*(\ell)B_k^*\xi_{\ell+1,k},$$

so that

$$\boldsymbol{x}(\ell+1)\boldsymbol{y}^*(\ell+1) = A\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)A^* + \sum_{k,p=1}^{m} B_k\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_p^*\xi_{\ell+1,k}\xi_{\ell+1,p} \tag{4.5}$$

$$+ \sum_{k=1}^{m} A\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_k^*\xi_{\ell+1,k} + \sum_{k=1}^{m} B_k\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)A^*\xi_{\ell+1,k}. \tag{4.6}$$

The vectors $\boldsymbol{x}(\ell)$ and $\boldsymbol{y}(\ell)$ are independent of every $\xi_{\ell+1,k}$. Therefore, using (4.2),

$$\mathbb{E}\big(A\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_k^*\xi_{\ell+1,k}\big) = \mathbb{E}\big(A\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_k^*\big)\mathbb{E}\xi_{\ell+1,k} = 0, \tag{4.7}$$

$$\sum_{k,p=1}^{m} \mathbb{E}\big(B_k\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_p^*\xi_{\ell+1,k}\xi_{\ell+1,p}\big) = \sum_{k,p=1}^{m} \mathbb{E}\big(B_k\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)B_p^*\big)\mathbb{E}(\xi_{\ell+1,k}\xi_{\ell+1,p})$$

$$= \sum_{k=1}^{m} B_k\mathbb{E}\big(\boldsymbol{x}(\ell)\boldsymbol{y}^*(\ell)\big)B_k^* = \sum_{k=1}^{m} B_kV(\ell)B_k^*. \tag{4.8}$$

As a result,

$$V(\ell+1) = AV(\ell)A^* + \sum_{k=1}^{m} B_k V(\ell) B_k^*,$$

and (4.3) follows from (1.3).

Similarly,

$$r_x(\ell+1) = \mathbb{E}\boldsymbol{x}^*(\ell)\left(A^*A + \sum_{k=1}^{m} B_k^* B_k\right)\boldsymbol{x}(\ell) = \mathbb{E}\big(\boldsymbol{x}^*(\ell) N_{A,B} \boldsymbol{x}(\ell)\big).$$

Then (1.5) implies $\gamma r_x(\ell) \le r_x(\ell+1) \le \beta r_x(\ell)$, and (4.4) follows. $\square$

Given the origin of Eq. (4.3), the matrix $D_{A,B}$ from (1.7) is natural to call the `discrete-time stochastic Kronecker sum` of the matrices A and B_k.

It is really very well known that the spectral radius $\boldsymbol{\rho}(A)$ of a matrix $A \in \mathbb{C}^{n\times n}$ satisfies

$$\boldsymbol{\rho}(A) = \lim_{\ell\to+\infty} \left|A^\ell\right|^{1/\ell}. \tag{4.9}$$

Together with Proposition 4.1, equality (4.9) leads to a two-sided bound on the spectral radius of $D_{A,B}$, which is the main result of this section.

Theorem 4.2. *For every matrices $A, B_1, \ldots, B_m \in \mathbb{C}^{n\times n}$,*

$$\lambda_{\min}(N_{A,B}) \le \boldsymbol{\rho}(D_{A,B}) \le \lambda_{\max}(N_{A,B}). \tag{4.10}$$

Proof. Similar to Proposition 4.1, write $\gamma = \lambda_{\min}(N_{A,B})$ and $\beta = \lambda_{\max}(N_{A,B})$. To derive the upper bound in (4.10), note that (2.8) and (4.4) imply

$$\left|\boldsymbol{U}(\ell)\right| \le \sqrt{r_x(\ell) r_y(\ell)} \le |\boldsymbol{u}||\boldsymbol{v}|\beta^\ell. \tag{4.11}$$

Combining (4.3) and (4.11) leads to

$$\left|(D_{A,B})^\ell \boldsymbol{U}(0)\right| \le |\boldsymbol{u}||\boldsymbol{v}|\beta^\ell. \tag{4.12}$$

Since $\boldsymbol{U}(0) = \mathrm{vec}(\boldsymbol{u}\boldsymbol{v}^*) = \boldsymbol{v} \otimes \boldsymbol{u}$, and $\boldsymbol{u}$ and $\boldsymbol{v}$ are arbitrary vectors in $\mathbb{C}^n$, the same arguments as in the continuous-time case show that (4.12) implies

$$\left|(D_{A,B})^\ell\right| \le n\beta^\ell. \tag{4.13}$$

Then the upper bound in (4.10) follows from (4.13) and (4.9).

To derive the lower bound, take $\boldsymbol{u} = \boldsymbol{v}$ with $|\boldsymbol{u}| = 1$ so that $\boldsymbol{x}(\ell) = \boldsymbol{y}(\ell)$ for all $\ell \geq 0$. Then (2.9) and (4.3) imply

$$n^{-1/2}\gamma^{\ell} \leq \big|\boldsymbol{U}(\ell)\big| \leq \big|(D_{A,B})^{\ell}\big|,$$

and the lower bound in (4.10) follows from (4.9). □

Examples of applications of (4.10) are in Section 5.

5. Examples and further discussions

One reason (3.6) and (4.10) are potentially useful is that the matrices $M_{A,B}$ and $N_{A,B}$ are Hermitian and have size n-by-n, whereas the matrices $C_{A,B}$ and $D_{A,B}$ are in general not Hermitian or even normal and have a much bigger size n^2-by-n^2. In particular, if the matrix $M_{A,B}$ is scalar, that is, $M_{A,B} = \beta I$, then $\boldsymbol{\alpha}(C_{A,B}) = \beta$; if $N_{A,B} = \beta I$, then $\boldsymbol{\rho}(C_{A,B}) = \beta$.

For example, with $m = 1$,

(1) If $A = aI + S$ for a real number a and a skew-symmetric matrix S, and B is orthogonal, then $M_{A,B} = (2a+1)I$ and $\boldsymbol{\alpha}(C_{A,B}) = 2a + 1$;
(2) If matrices A and B are orthogonal, then $N_{A,B} = 2I$ and $\boldsymbol{\rho}(D_{A,B}) = 2$.

Without additional information about the matrices A and B, it is not possible to know how tight the bounds in (3.6) and (4.10) will be. As an example, consider two real matrices

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ \sigma & 0 \end{pmatrix}.$$

The corresponding stochastic systems are

$$dx_1(t) = ax_1(t)dt, \qquad dx_2(t) = bx_2(t)dt + \sigma x_1(t)dw(t)$$

in continuous time and

$$x_1(\ell+1) = ax_1(\ell), \qquad x_2(\ell+1) = bx_2(\ell) + \sigma x_1(\ell)\xi_{\ell+1}$$

in discrete time. Then

$$C_{A,B} = A \otimes I + I \otimes A + B \otimes B = \begin{pmatrix} 2a & 0 & 0 & 0 \\ 0 & a+b & 0 & 0 \\ 0 & 0 & a+b & 0 \\ \sigma^2 & 0 & 0 & 2b \end{pmatrix},$$

$$M_{A,B} = A^\top + A + B^\top B = \begin{pmatrix} 2a+\sigma^2 & 0 \\ 0 & 2b \end{pmatrix},$$

$$D_{A,B} = A \otimes A + B \otimes B = \begin{pmatrix} a^2 & 0 & 0 & 0 \\ 0 & ab & 0 & 0 \\ 0 & 0 & ab & 0 \\ \sigma^2 & 0 & 0 & b^2 \end{pmatrix},$$

$$N_{A,B} = A^\top A + B^\top B = \begin{pmatrix} a^2+\sigma^2 & 0 \\ 0 & b^2 \end{pmatrix}.$$

In particular, both $\boldsymbol{\alpha}(C_{A,B})$ and $\boldsymbol{\rho}(D_{A,B})$ do not depend on σ:

$$\boldsymbol{\alpha}(C_{A,B}) = \max(2a, 2b), \qquad \boldsymbol{\rho}(D_{A,B}) = \max\bigl(a^2, b^2\bigr)$$

whereas

$$\lambda_{\max}(M_{A,B}) = \max\bigl(2a+\sigma^2, 2b\bigr) \quad \text{and} \quad \lambda_{\max}(N_{A,B}) = \max\bigl(a^2+\sigma^2, b^2\bigr)$$

can be arbitrarily large.

An important question in the study of stochastic systems is whether the matrices $C_{A,B}$ and $D_{A,B}$ are stable, that is, $\boldsymbol{\alpha}(C_{A,B}) < 0$ and $\boldsymbol{\rho}(D_{A,B}) < 1$. One consequence of Propositions 3.2 and 4.1 is that stability of the stochastic Kronecker sum matrix is equivalent to the mean-square asymptotic stability of the corresponding stochastic system (see also [2, Theorem 1.5.3]):

$$\boldsymbol{\alpha}(C_{A,B}) < 1 \quad \Leftrightarrow \quad \lim_{t\to+\infty} \mathbb{E}\bigl|\boldsymbol{x}(t)\bigr|^2 = 0,$$

$$\boldsymbol{\rho}(D_{A,B}) < 1 \quad \Leftrightarrow \quad \lim_{\ell\to\infty} \mathbb{E}\bigl|\boldsymbol{x}(\ell)\bigr|^2 = 0.$$

The example shows that it is possible to have this stability even when the matrices $M_{A,B}$ and $N_{A,B}$ are not stable: $C_{A,B}$ is stable if (and only if) $\max(a,b) < 0$, and $D_{A,B}$ is stable if (and only if) $\max(|a|,|b|) < 1$; this is also clear by looking directly at the corresponding stochastic system.

One can always use the lower bounds in (3.6) and (4.10) to check if the matrices $C_{A,B}$ and $D_{A,B}$ (and hence the corresponding systems) are not stable. In the above example, if

$$\lambda_{\min}(M_{A,B}) = \min\bigl(2a+\sigma^2, 2b\bigr) > 0,$$

then $b > 0$ and $C_{A,B}$ is certainly not stable. Similarly, if

$$\lambda_{\min}(N_{A,B}) = \min\bigl(a^2+\sigma^2, b^2\bigr) > 1,$$

then $|b| > 1$ and $D_{A,B}$ is certainly not stable.

To conclude this section, assume that the matrices $A, B_1, \ldots, B_m$ are all Hermitian. Then $C_{A,B}$ and $D_{A,B}$ are also Hermitian, and, with (1.4) in mind, inequalities (3.6) and (4.10) become, respectively,

$$\lambda_{\min}(M_{A,B}) \le \lambda_{\max}(C_{A,B}) \le \lambda_{\max}(M_{A,B}), \tag{5.1}$$

$$\lambda_{\min}(N_{A,B}) \le \lambda_{\max}(D_{A,B}) \le \lambda_{\max}(N_{A,B}). \tag{5.2}$$

Let us compare (5.1) and (5.2) with some other bounds that can be derived in this Hermitian case using the following approaches:

(1) Variation characterization of the eigenvalues of the operators $\mathcal{L}_{A,B}$ and $\mathcal{S}_{0;A,B}$ (cf. (2.4) and (2.6));
(2) Weyl's inequality for the eigenvalues of the sum of two Hermitian matrices [5, Theorem 4.3.1].

Both approaches provide bounds on individual eigenvalues of $C_{A,B}$ and $D_{A,B}$. To stay within the scope of the paper, let us restrict our discussion to the bounds on the largest eigenvalues.

Inequality (2.4) is an improvement of the lower bound in (5.1). Indeed, (a) (2.4) is sharper because the average of the eigenvalues of $M_{A,B}$ is bigger than or equal to the smallest eigenvalue; (b) (2.4) is explicit in terms of the entries of the matrices $A = (A_{ij})$ and $B = (B_{ij})$:

$$\lambda_{\max}(C_{A,B}) \ge \frac{2\sum_{i=1}^{n} A_{ii} + \sum_{i,j=1}^{n} |B_{ij}|^2}{n}. \tag{5.3}$$

Similarly, (2.6) is an improvement of the lower bound in (5.2):

$$\lambda_{\max}(D_{A,B}) \ge \frac{\sum_{i,j=1}^{n}(|A_{ij}|^2 + |B_{ij}|^2)}{n}. \tag{5.4}$$

Upper bounds on $\lambda_{\max}(C_{A,B})$ and $\lambda_{\max}(D_{A,B})$ are not readily available with this approach.

Next, let us bound $\lambda_{\max}(C_{A,B})$ and $\lambda_{\max}(D_{A,B})$ directly from (1.6) and (1.7) using Weyl's inequality for the eigenvalues of the sum $H + J$ of two Hermitian matrices:

$$\begin{aligned}\lambda_{\min}(H) + \lambda_{\min}(J) \le \lambda_{\min}(H+J) &\le \lambda_{\max}(H) + \lambda_{\min}(J) \\ &\le \lambda_{\max}(H+J) \le \lambda_{\max}(H) + \lambda_{\max}(J);\end{aligned} \tag{5.5}$$

also, recall that eigenvalues of $A \otimes B$ are of the form $\lambda(A)\lambda(B)$ [4, Theorem 4.2.12], whereas the eigenvalues of $A \otimes I + I \otimes B$ are of the form $\lambda(A) + \lambda(B)$ [4, Theorem 4.4.5]. Combining (5.5) with (1.6) yields

$$\max\big(2\lambda_{\max}(A)+\lambda^2_{\min}(B), 2\lambda_{\min}(A)+\lambda^2_{\max}(B)\big) \le \lambda_{\max}(C_{A,B}) \le 2\lambda_{\max}(A)+\lambda^2_{\max}(B). \quad (5.6)$$

For the matrix $M_{A,B} = 2A + B^2$, Weyl's inequality yields

$$\lambda_{\min}(M_{A,B}) \le \max\big(2\lambda_{\max}(A)+\lambda^2_{\min}(B), 2\lambda_{\min}(A)+\lambda^2_{\max}(B)\big) \le \lambda_{\max}(M_{A,B}) \le 2\lambda_{\max}(A)+\lambda^2_{\max}(B),$$

leading to the following comparison of (5.1) and (5.6):

(1) (5.6) is more explicit than (5.1);
(2) (5.6) provides a sharper lower bound;
(3) (5.1) provides a sharper upper bound.

Comparison between (5.3) and the lower bound in (5.6) depends on the particular matrices A and B.

In the case of the matrix $D_{A,B}$, the same analysis produces the following upper and lower bounds:

$$\max\big(\lambda^2_{\max}(A)+\lambda^2_{\min}(B), \lambda^2_{\min}(A)+\lambda^2_{\max}(B)\big) \le \lambda_{\max}(D_{A,B}) \le \lambda^2_{\max}(A)+\lambda^2_{\max}(B), \quad (5.7)$$

and we conclude that

(1) (5.7) is more explicit than (5.2);
(2) (5.7) provides a sharper lower bound;
(3) (5.2) provides a sharper upper bound.

Comparison between (5.4) and the lower bound in (5.7) depends on the particular matrices A and B.

To conclude, in Hermitian case, there are several immediate ways to improve the lower bounds in (5.1) and (5.2), but no immediate ways to improve the upper bounds.

Acknowledgements

The author is grateful to the referee for very helpful and insightful comments on an earlier version of the paper.

References

[1] J.W. Brewer, Kronecker products and matrix calculus in system theory, IEEE Trans. Circuits Syst. 25 (9) (1978) 772–781.

[2] T. Damm, Rational Matrix Equations in Stochastic Control, Lecture Notes in Control and Inform. Sci., vol. 297, Springer-Verlag, Berlin, 2004.
[3] J. Feng, J. Lam, Y. Wei, Spectral properties of sums of certain Kronecker products, Linear Algebra Appl. 431 (9) (2009) 1691–1701.
[4] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1994.
[5] R.A. Horn, C.R. Johnson, Matrix Analysis, 2nd ed., Cambridge University Press, Cambridge, 2013.
[6] S. Lototsky, J. Zhong, Stochastic evolution systems with constant coefficients, Stoch. Partial Differ. Equ. Anal. Comput. 1 (4) (2013) 687–711.
[7] B. Øksendal, Stochastic Differential Equations, sixth ed., Universitext, Springer-Verlag, Berlin, 2003.
[8] L.N. Trefethen, M. Embree, Spectra and Pseudospectra, Princeton University Press, Princeton, NJ, 2005.