TECHNISCHE UNIVERSITÄT
CHEMNITZ

# On Properties of General Sylvester and Lyapunov Operators

M. Konstantinov, V. Mehrmann, P. Petkov

Preprint 98-12

Fakultät für Mathematik

# ON PROPERTIES OF GENERAL SYLVESTER AND LYAPUNOV OPERATORS*

M. Konstantinov† V. Mehrmann‡ P. Petkov§

**Abstract**

General Sylvester and Lyapunov operators in real and complex matrix spaces are studied, which include as a particular case the operators arising in the theory of linear time-invariant descriptor systems. For linear matrix operators an index which characterizes the operator is introduced and determined for general linear matrix operators. The problem of representing such an operator as a sum of elementary operators is posed and solved. The dimensions of the spaces of Lyapunov operators are determined and the concept of symmetrised singular values of a Lyapunov operator is introduced. The application of symmetric singular values to the perturbation and error analysis of Lyapunov equations is discussed.

**Keywords:** Linear matrix operators, Sylvester operators, Lyapunov operators, singular values, symmetrised singular values, perturbation analysis and error analysis.

**AMS Subject Classification:** 15 A 24, 93 B 35, 93 C 73.

## 1 Introduction and notations

Linear matrix equations and linear matrix operators have been studied, since the pioneering work of Sylvester and Kronecker [14, 23, 22], see also [26, 18, 17] and [1]. Now there are hundreds of papers, surveys and many books, e.g., [3, 21, 2, 10, 11, 25, 7] devoted to the analysis, existence, uniqueness and representation of the solution and also to the numerical algorithms and software to solve various types of linear matrix equations. Most of the existing results, however, are connected with particular classes of such

---

*Work supported by grant 5/72 764 of VW Stiftung.

†University of Architecture and Civil Engineering, 1 Hr. Smirnenski Blv., 1421 Sofia, Bulgaria, E-mail: mmk_fte@uacg.acad.bg

‡Fakultät für Mathematik, TU Chemnitz, D-09107 Chemnitz, Germany, E-mail: mehrmann@mathematik.tu-chemnitz.de

§Department of Automatics, TU Sofia, 1756 Sofia, Bulgaria, E-mail: php@mbox.digsys.bg

matrix equations. In particular the problem of representing a general linear matrix operator as a sum of elementary operators seems to be not completely settled so far.

An important class of linear matrix equations are the Lyapunov equations. Since the fundamental work of Lyapunov on stability of motion, these matrix equations have been widely used in stability theory of differential equations [27], in the theory of linear-quadratic optimisation and filtering [16], in the perturbation analysis of linear and non-linear matrix equations [9, 6, 12, 13] and other fields of pure and applied mathematics. This has motivated a continuous interest to both the theory and numerical treatment of Lyapunov operators and equations [5, 24, 8, 19, 20, 4] and also recently in the context of the analysis and numerical simulation of descriptor systems via generalized Lyapunov equations [15]. Some general properties of finite-dimensional Lyapunov operators, however, have not been studied to a sufficient extent. In particular, the notion of the minimal singular value of a Lyapunov operator is sometimes misused. Introducing the new concept of symmetric singular values of a Lyapunov operator, some well-known estimates in the sensitivity theory of matrix equations may be substantially improved.

In this paper we first investigate a general class of linear matrix operators, the Sylvester operators, and introduce the index of a linear matrix operator as the minimum number of terms in which it can be represented as a sum of elementary Sylvester operators. We then give an explicit expression for the index and derive a procedure for determining the representation of a general Sylvester operator as a sum of elementary Sylvester operators.

Furthermore we study the general class of Lyapunov operators and determine the dimensions of the spaces of real and complex Lyapunov operators. We then introduce the concept of symmetric singular values of Lyapunov operators and show their application to the perturbation and a posteriori error analysis of Lyapunov equations.

We use the following notation.

- $\overline{m,n}$ – the set of integers $m, m+1, , \ldots, n$, where $m \leq n$;
- $\mathcal{R}$ and $\mathcal{C}$ – the sets of real and complex numbers, $\jmath = \sqrt{-1}$, $\mathcal{R}_+ = [0, \infty)$;
- $\mathcal{F}^{m\times n}$ – the space of $m \times n$ matrices over $\mathcal{F}$, $\mathcal{F}^n = \mathcal{F}^{n\times 1}$, where $\mathcal{F}$ is $\mathcal{R}$ or $\mathcal{C}$;
- $A^{\mathrm{T}}$, $\overline{A}$ and $A^{\mathrm{H}} = \overline{A}^{\mathrm{T}}$ - the transpose, the complex conjugate and the complex conjugate transpose of a matrix $A$;
- $\mathrm{Rg}(A)$ and $\mathrm{Ker}(A)$ – the image and the kernel of the matrix $A$;
- $I_n$ – the unit $n \times n$ matrix;

- $E_{ij}(m,n) \in \mathcal{R}^{m\times n}$ – a matrix with a single non-zero entry equal to 1, in position $(i,j)$;

- $\mathrm{tr}(A)$ and $\mathrm{rank}(A)$ – the trace and rank of a matrix $A$;

- $\|A\|_2 = \sigma_{\max}(A)$ – the spectral norm of $A$, where $\sigma_{\max}(A)$ is the maximum singular value of $A$;

- $\|A\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(A^{\mathrm{H}}A)}$ – the Frobenius norm of $A$ (we use the same notation for the Frobenius norm of a linear operator);

- $\mathrm{vec}[A] \in \mathcal{F}^{mn}$ – the column-wise vector representation of $A \in \mathcal{F}^{m\times n}$;

- $\Pi(m,n) \in \mathcal{R}^{mn\times mn}$ – the vec-permutation matrix such that $\mathrm{vec}[M^{\top}] = \Pi(m,n)\mathrm{vec}[M]$, $M \in \mathcal{F}^{m\times n}$, $\Pi_n = \Pi(n,n)$;

- $\Omega^{\mathcal{F}}(n) \subset \mathcal{F}^{n^2\times n^2}$ – the set of all matrices $M \in \mathcal{F}^{n^2\times n^2}$ such that $M\Pi_n = \Pi_n\overline{M}$; the set $\Omega(n) = \Omega^{\mathcal{R}}(n)$ is the subspace of all $M \in \mathcal{R}^{n^2\times n^2}$, such that $M\Pi_n = \Pi_n M$;

- $A \otimes B$ – the Kronecker (tensor) product of the matrices $A$ and $B$;

- $\lambda(A) = \{\lambda_1(A), \ldots, \lambda_n(A)\} \subset \mathcal{C}$ – the spectrum of $A \in \mathcal{F}^{n\times n}$, where the eigenvalues $\lambda_i(A)$ of $A$ are counted according to their algebraic multiplicities;

- $\sigma(A) = \{\sigma_1(A), \ldots, \sigma_r(A)\} \subset \mathcal{R}_+$ – the set of singular values $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$ of $A \in \mathcal{F}^{m\times n}$, $\sigma_i(A) = \sqrt{\lambda_i(A^{\mathrm{H}}A)}$, counted according to their algebraic multiplicities, where $r = \min\{m,n\}$;

- $\mathcal{GL}(n) \subset \mathcal{F}^{n\times n}$ – the group of non-singular matrices; $\mathcal{U}(n) \subset \mathcal{GL}(n)$ – the group of unitary matrices; $\mathbf{Lin}^{\mathcal{F}}(n) \simeq \mathcal{F}^{n^4}$ – the space of Sylvester operators $\mathcal{M} : \mathcal{F}^{n\times n} \to \mathcal{F}^{n\times n}$,

  $$\mathcal{M}[X] = \sum_{k=1}^{r} A_k X B_k$$

  where $A_k, B_k \in \mathcal{F}^{n\times n}$ are given matrices, $\mathbf{Lin}(n) = \mathbf{Lin}^{\mathcal{R}}(n)$;

- $\mathrm{Mat}(\mathcal{M}) = M \in \mathcal{F}^{n^2\times n^2}$ – the matrix representation of $\mathcal{M} \in \mathbf{Lin}^{\mathcal{F}}(n)$, i.e., $\mathrm{vec}[\mathcal{M}[X]] = M\mathrm{vec}[X]$ and hence

  $$M = \sum_{k=1}^{r} B_k^{\top} \otimes A_k.$$

The singular values of an operator $\mathcal{M} \in \mathbf{Lin}^{\mathcal{F}}(n)$ are the singular values $\sigma_1(M) \geq \cdots \geq \sigma_{n^2}(M) \geq 0$ of its matrix $M$ and we write $\sigma_i(\mathcal{M}) = \sigma_i(M)$.

The abbreviation ":=" stands for "equal by definition".

# 2 Linear matrix operators

## 2.1 Basic concepts

Denote by $\mathbf{Lin} = \mathbf{Lin}(p, m, n, q)$ the linear space of linear matrix operators $\mathcal{M} : \mathcal{F}^{m\times n} \to \mathcal{F}^{p\times q}$, i.e., $\mathcal{M}[X] \in \mathcal{F}^{p\times q}$, $X \in \mathcal{F}^{m\times n}$. In what follows an operator will often depend on a collection of $2r$ matrices

$$C = (A_1, B_1, \ldots, A_r, B_r) \in \Gamma_r := (\mathcal{F}^{p\times m} \times \mathcal{F}^{n\times q})^r, \tag{1}$$

where $A_k \in \mathcal{F}^{p\times m}$ and $B_k \in \mathcal{F}^{n\times q}$. To emphasize this dependence we also write $\mathcal{L}(C) \in \mathbf{Lin}$ for the operator itself and $\mathcal{L}(C)[X] \in \mathcal{R}^{p\times q}$ for its matrix value. Thus we have a family of operators $\{\mathcal{L}(C)\}_{C\in\Sigma_r}$ and $\mathcal{L}$ may be considered as a mapping $\mathcal{L} : \Sigma_r \times \mathcal{F}^{m\times n} \to \mathcal{F}^{p\times q}$.

A general linear matrix operator can be defined as follows. Let $pq$ vectors $m_{i,j} \in \mathcal{F}^{mn}$, $i = 1, ldots, p$, $j = 1, ldots, q$ be given. For every $X \in \mathcal{F}^{m\times n}$ let $\mathcal{M}_{i,j}[X] := m_{i,j}^{\top}\mathrm{vec}[X] \in \mathcal{F}$. The operator $\mathcal{M} : \mathcal{F}^{m\times n} \to \mathcal{F}^{p\times q}$, defined from $\mathcal{M}[X] = (\mathcal{M}_{i,j}[X])_{i,j=1}^{p,q}$, is a linear matrix operator. The matrix $M :=$ Mat$[\mathcal{M}] \in \mathcal{F}^{pq\times mn}$ associated with $\mathcal{M}$ is defined via $\mathrm{vec}[\mathcal{L}[X]] = M\mathrm{vec}[X]$ and hence

$$M = [m_{1,1}, \ldots, m_{p,1}, \ldots, m_{1,q}, \ldots, m_{p,q}]^{\top}.$$

In this formulation a linear matrix operator has no particular structure and may be identified with its matrix $M \in \mathcal{F}^{pq\times mn}$ according to the commutative diagram

$$\begin{array}{ccc} \mathcal{F}^{m\times n} & \xrightarrow{\mathcal{M}} & \mathcal{F}^{p\times q} \\ \downarrow \mathrm{vec} & & \downarrow \mathrm{vec} \\ \mathcal{F}^{mn} & \xrightarrow{M} & \mathcal{F}^{pq} \end{array}.$$

At the same time any linear matrix operator may be expressed directly in terms of matrix products. In this framework the specific structure of the operator may be revealed as an alternative to its representation as a general $pq \times mn$ matrix. This special structure is encoded in the mapping $\mathcal{L}(\cdot) : \Sigma_r \to \mathbf{Lin}$.

**Definition 1** *The operator $\mathcal{E}(A_1, B_1) \in \mathbf{Lin}$, such that $\mathcal{E}(A_1, B_1)[X] = A_1XB_1$ for $X \in \mathcal{F}^{m\times n}$, where $A_1 \in \mathcal{F}^{p\times m}$ and $B_1 \in \mathcal{F}^{n\times q}$, is called an elementary Sylvester operator with generating matrices $(A_1, B_1)$.*

The zero operator $0_{(p,m,n,q)} \in \mathbf{Lin}(p, m, n, q)$ and the identity operator $1_{(m,m,n,n)} \in \mathbf{Lin}(m, m, n, n)$ are elementary Sylvester operators with generating matrices $(A_1, 0_{n\times q})$ (or $(0_{p\times m}, B_1)$) and $(I_m, I_n)$, respectively, where $A_1 \in \mathcal{F}^{p\times m}$ (or $B_1 \in \mathcal{F}^{n\times q}$) is arbitrary.

Let $r \geq 1$ and let a matrix $2r$-tuple as in (1) be given. Consider an operator $\mathcal{L}(C) \in \mathbf{Lin}$, which is represented as a sum of $r$ non-zero elementary Sylvester operators $\mathcal{E}(A_k, B_k)$, i.e. ,

$$\mathcal{L}(C)[X] = \sum_{k=1}^{r} \mathcal{E}(A_k, B_k)[X] = \sum_{k=1}^{r} A_k X B_k \,, \ X \in \mathcal{F}^{m\times n}. \tag{2}$$

Operators of the form (2) are called *Sylvester operators.*

Each $\mathcal{M} \in \mathbf{Lin}$ may be represented in the form (2), i.e. $\mathcal{M} = \mathcal{L}(C)$ for some $r$ and $C$, although this is not a trivial task as shown below.

Applying the vec operation to the expression for $\mathcal{L}(C)[X]$ we get

$$\text{vec}[\mathcal{L}(C)[X]] = L(C)\text{vec}[X], \tag{3}$$

where

$$L = L(C) := \text{Mat}[\mathcal{L}(C)] = \sum_{k=1}^{r} B_k^{\top} \otimes A_k \in \mathcal{F}^{pq\times mn} \tag{4}$$

is the matrix associated with $\mathcal{L}(C)$.

Every collection $C$ determines a unique Sylvester operator $\mathcal{L}(C)$ through (2) but the converse is of course not true. Using the inverse of the vec operator $\text{vec}^{-1} : \mathcal{F}^{pq} \to \mathcal{F}^{p\times q}$, any operator $\mathcal{M} \in \mathbf{Lin}$ and its associated matrix $M \in \mathcal{F}^{pq\times mn}$ are related via

$$\mathcal{M}[X] = \text{vec}^{-1}(p,q)[M\text{vec}[X]], \ X \in \mathcal{F}^{m\times n}.$$

There exist different integers $r$ and infinitely many collections $C \in \Sigma_r$, such that $\mathcal{M}$ has a representation of type (2), i.e. $\mathcal{M} = \mathcal{L}(C)$, where $C$ satisfies the non-linear equation

$$\sum_{k=1}^{r} B_k^{\top} \otimes A_k = M. \tag{5}$$

Obviously pairs $(A_k, B_k)$ and $(\lambda_k A_k, \mu_k B_k)$ with $\lambda_k\mu_k = 1$ give rise to the same Sylvester operator. Another possibility to get different representations of the same operator is when the matrix $M_k := B_k^{\top} \otimes A_k$ of some elementary Sylvester operator $\mathcal{E}(A_k, B_k)$ is a linear combination of the matrices associated with other Sylvester operators.

**Example 1** Given $A_1 \in \mathcal{F}^{m\times m}$ and $B_1 \in \mathcal{F}^{n\times n}$ such that

$$\begin{aligned}\mathcal{M}[X] &= A_1XB_1 + A_1X + XB_1 = (A_1 + I_m)X(B_1 + I_n) - X \\ &= A_1X(B_1 + I_n) + XB_1 = A_1X + (A_1 + I_m)XB_1\end{aligned}$$

we have that the operator $\mathcal{M} := \mathcal{L}(A_1, B_1, A_1, I_n, I_m, B) \in \mathbf{Lin}$ may be represented in at least three more ways $\mathcal{L}(A_1+I_m, B_1+I_n, I_m, -I_n)$, $\mathcal{L}(A_1, B_1+I_n, I_m, B_1$, $\mathcal{L}(A_1, I_n, A_1 + I_m, B_1)$.

These observations lead to the problem of representing an operator $\mathcal{M} \in \mathbf{Lin}$ as a sum of a minimum number of elementary Sylvester operators.

**Definition 2** *The minimum number $\ell \geq 1$, such that the operator $\mathcal{M} \in \mathbf{Lin}(p, m, n, q)$ may be represented as a sum of $\ell$ elementary Sylvester operators, is said to be the Sylvester-index of $\mathcal{M}$ and is denoted by* $\mathrm{ind}(p, m)[\mathcal{M}]$. *Any representation of $\mathcal{M}$ as a sum of minimum number of elementary operators is called a condensed representation.*

It immediately follows from the definition that a Sylvester operator is elementary if and only it has Sylvester-index 1.

We have explicitely indicated the dependence of the Sylvester index of $\mathcal{M}$ on $p$ and $m$ in order to have a universal definition if a general operator $\mathcal{M}$ is given. Indeed, the matrix $M$ of $\mathcal{M}$ is $(pq \times mn)$. Let $p'$, $m'$, $n'$ and $q'$ be any integers such that $p'q' = pq$ and $m'n' = mn$. Then $\mathcal{M}$ may be represented as a sum of (a minimum number) $\ell$ of elementary operators from $\mathbf{Lin}(p, m, n, q)$ and as a sum of (minimum number) $\ell'$ of elementary operators from $\mathbf{Lin}(p', m', n', q')$. In general $\ell \neq \ell'$ and the index of $\mathcal{M}$ depends on the representation of the dimensions $pq$ and $mn$ as products of two factors. In particular $\mathrm{ind}(1, 1)[\mathcal{M}] = 1$. As may be expected, the Sylvester-index is symmetric in the sense that $\mathrm{ind}(p, m) = \mathrm{ind}(q, n)$ (see Proposition 3 below).

Definition 2 applies also to operators $\mathcal{L}(C)$ in the form (2) and here we write simply $\mathrm{ind}[\mathcal{L}(C)]$. The index of $\mathcal{L}(C)$ in (2) is at most $r$ but may be much less.

**Example 2** The Sylvester-index of the operator

$$\mathcal{M} := \mathcal{L}(A, B, A, I_n, I_m, B, I_n, I_m)$$

is at most 4, but in fact it is equal to 1, since $\mathcal{M}$ is in fact the elementary operator $\mathcal{E}(A + I_m, B + I_n)$.

In Examples 1 and 2 some of the elementary Sylvester operators were linear combinations of other elementary Sylvester operators in the representation (2). Such elementary operators may be removed from the representation of a general Sylvester operator according to the following proposition.

**Proposition 1** *Let an operator $\mathcal{L}(C)$ as in (2) be given. Then*

$$\mathrm{ind}[\mathcal{L}(C)] \leq r_1 := \mathrm{rank}\left[\mathrm{vec}\left[B_1^\top \otimes A_1\right], \ldots, \mathrm{vec}\left[B_r^\top \otimes A_r\right]\right]. \qquad (6)$$

*Proof.* Suppose that $r_1 < r$ (if $r_1 = r$ there is nothing to prove, since $\mathrm{ind}[L(C)] \leq r$). Let $Z_j := \mathrm{vec}\left[B_j^\top \otimes A_j\right]$ and assume w.l.o.g. that

$Z_1, \ldots, Z_{r_1}$ are linearly independent. Then every $Z_k$ with $k > r_1$ may be expressed as

$$Z_k = \sum_{j=1}^{r_1} \lambda_{k,j} Z_j, \; \lambda_{k,j} \in \mathcal{F}.$$

Hence, for $k > r_1$

$$\left(B_k^{\top} \otimes A_k\right) \mathrm{vec}[X] = \sum_{j=1}^{r_1} \lambda_{k,j} \left(B_j^{\top} \otimes A_j\right) \mathrm{vec}[X]$$

and

$$A_k X B_k = \sum_{j=1}^{r_1} \lambda_{k,j} A_j X B_j.$$

Substituting this expression in (2) we obtain

$$\mathcal{L}(C)[X] = \sum_{k=1}^{r_1} \alpha_k A_k X B_k,$$

where $\alpha_k := \sum_{i=r_1+1}^{r} \lambda_{i,k}$ and hence $\mathrm{ind}[\mathcal{L}(C)] \leq r_1$ as claimed. □

It follows that we may assume that the representation (2) of a Sylvester operator is condensed, i.e. $r = \mathrm{ind}[\mathcal{L}(C)]$. For Lyapunov operators, however, a non-condensed but symmetric representation may also be useful.

For $\mathcal{M} \in \mathbf{Lin}$ we have

$$\|\mathcal{M}[X]\|_{\mathrm{F}} = \|\mathrm{vec}[\mathcal{M}[X]]\|_2 \leq \|M\|_2 \|\mathrm{vec}[X]\|_2 = \|M\|_2 \|X\|_{\mathrm{F}}$$

with equality holding if $\mathrm{vec}[X]$ is the right singular vector of the matrix $M$, corresponding to its maximum singular value $\|M\|_2$. Hence, we may define a norm in **Lin** as follows.

**Definition 3** *The (Frobenius) norm of* $\mathcal{M} \in \mathbf{Lin}$ *is defined as*

$$\|\mathcal{M}\|_{\mathrm{F}} := \max\{\|\mathcal{M}[X]\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1\} = \|M\|_2.$$

Other norms as

$$\|\mathcal{M}\|_{pq} := \max\{\|\mathcal{M}[X]\|_p : \|X\|_q = 1\}; \; p, q \geq 1$$

where $\|\cdot\|_p$ and $\|\cdot\|_q$ are Hölder norms, may also be used. Here convenient expressions for $\|\cdot\|_{pq}$ are known only for $p = q = 2$ when $\mathcal{M}$ is the standard Lyapuov operator of Sylvester-index 2, see e.g. [9, 6].

## 2.2 Representation of a linear matrix operator as a sum of elementary Sylvester operators

Consider the problem of representing a general linear matrix operator $\mathcal{M}$ with associated matrix $M$ in the form (2). The dimension (real or complex) of **Lin** $\simeq \mathcal{F}^{pq\times mn} \simeq \mathcal{F}^{pmnq}$ is $pmnq$. In particular, for each $M \in \mathcal{F}^{pq\times mn}$ there exist $C \in \Sigma_r$, $r = \text{ind}(p,m)[\mathcal{M}]$, and an operator $\mathcal{L}(C) \in$ **Lin**, such that the associated matrix of $\mathcal{L}(C)$ is $M$, i.e., $L(C) = M$. This equation in $C$, of the form (5), may be consistent or not depending on $r$. If it is consistent, then it is also underdetermined and has a multi-parameter family of solutions.

A simple solution is obtained as follows. Partition the matrix $M \in \mathcal{F}^{pq\times mn}$ as $M = [M_{i,j}]$; $i = 1,\ldots,q$, $j = 1,\ldots,n$, where $M_{i,j} \in \mathcal{F}^{p\times m}$. Then

$$M = \sum_{i,j=1}^{q,n} E_{i,j}(q,n) \otimes M_{i,j}.$$

Hence, in view of (5), a possible solution for $C$ is

$$A_k = M_{i,j},\ B_k = E_{j,i}(n,q),\ k = k(i,j) := i + (j-1)q,$$

in which there are at most $nq$ non-zero pairs $(A_k, B_k)$, i.e. the resulting operator $\mathcal{L}(C)$ and hence $\mathcal{M}$ has Sylvester-index at most $nq$. A similar argument for the transposed operator shows that $\text{ind}[\mathcal{L}(C)] \le pm$. Thus we have proved the following Proposition.

**Proposition 2** *The Sylvester-index of an operator $\mathcal{M} \in$ **Lin** satisfies*

$$\text{ind}(m,n)[\mathcal{M}] \le \min\{pm, nq\}.$$

Using the described construction, we may calculate the Sylvster-index of a linear operator and construct a representation of type (2) in the following way. Suppose that $M \in \mathcal{F}^{pq\times mn}$ is the matrix associated with $\mathcal{M} \in$ **Lin**. Partition $M = [M_{i,j}]$ as before and let

$$M^* := \Pi(p,q) M \Pi(n,m) = [M^*_{k,\ell}],$$

with $M^*_{k,\ell} \in \mathcal{F}^{q\times n}$ for $k = 1, ldots, p$ and $\ell = 1, ldots, m$. Introducing

$$\widehat{M} = [\text{vec}[M_{1,1}],\ldots,\text{vec}[M_{q,1}],\ldots,\text{vec}[M_{1,n}],\ldots,\text{vec}[M_{q,n}]] \in \mathcal{F}^{pm\times qn}$$

and

$$\widehat{M}^* = \left[\text{vec}[M^*_{1,1}],\ldots,\text{vec}[M^*_{p,1}],\ldots,\text{vec}[M^*_{1,m}],\ldots,\text{vec}[M^*_{q,n}]\right] \in \mathcal{F}^{qn\times pm}.$$

we can determine the Sylvester-index of an arbitrary operator $\mathcal{M} \in$ **Lin** and construct a matrix collection $C$ such that $\mathcal{M} = \mathcal{L}(C)$.

**Proposition 3** *Let $M$ be the matrix associated with $\mathcal{M} \in$ **Lin**. Then*

$$\mathrm{ind}(p, m)[\mathcal{M}] = \mathrm{ind}(q, n)[\mathcal{M}] = \max\{1, \nu(M)\}$$

*where*

$$\nu(M) := \mathrm{rank}[\widehat{M}] = \mathrm{rank}[\widehat{M}^*].$$

*Proof.* For a given $r \geq 1$ equation (5) may be written as a bilinear equation

$$AB = \widehat{M} \tag{7}$$

in the unknown matrices

$$A := [\mathrm{vec}[A_1], \mathrm{vec}[A_2], \ldots, \mathrm{vec}[A_r]] \in \mathcal{F}^{pm \times r}$$

$$B := [\mathrm{vec}[B_1], \mathrm{vec}[B_2], \ldots, \mathrm{vec}[B_r]]^\top \Pi(q, n) = \begin{bmatrix} \mathrm{vec}^\top[B_1^\top] \\ \mathrm{vec}^\top[B_2^\top] \\ \vdots \\ \mathrm{vec}^\top[B_r^\top] \end{bmatrix} \in \mathcal{F}^{r \times nq}.$$

Let $\Theta_r(M) \subset \mathcal{F}^{pm \times r} \times \mathcal{F}^{r \times nq}$ be the set of solutions of (7). We show that $\Theta_r(M) \neq \emptyset$ if and only if $r \geq \nu(M)$ and hence equation (7) is solvable for $r = \nu(M)$. The proof is constructive and we give explicit expressions for $\Theta_{\nu(M)}(M)$.

In the trivial case $\mathcal{M} = 0$ we have $r = 1$ and the solution may be taken as $(A, 0)$ or $(0, B)$ with $\max\{pm, nq\}$ free parameters. Hence $\Theta_1(0)$ is either $\{0\} \times \mathcal{F}^{1.nq}$ or $\mathcal{F}^{pm} \times \{0\}$.

Consider the general case $\mathcal{M} \neq 0$. It follows from (7) that

$$\nu = \nu(M) \leq \min\{\mathrm{rank}[A], \mathrm{rank}[B]\} \leq r.$$

We show now that if $r = \nu$, then (7) is easily solved.

If $r = \nu = pm \leq nq$ then the solution set is

$$\Theta_r(M) = \{(P, P^{-1}\widehat{M}) : P \in \mathcal{GL}(pm)\},$$

while for $r = \nu = nq < mn$ it is

$$\Theta_r(M) = \{(\widehat{M}P^{-1}, P) : P \in \mathcal{GL}(nq)\}.$$

If $r = \nu < \min\{pm, nq\}$, then let

$$\widehat{M} = USV^{\mathrm{H}} = \begin{bmatrix} M_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_1 \in \mathcal{GL}(r)$$

be the singular value decomposition of $\widehat{M}$, where $U \in \mathcal{U}(pm)$, $V \in \mathcal{U}(nq)$. Then the solution set of (7) is

$$\Theta_r(M) = \left\{ \left( U^\top \begin{bmatrix} P \\ 0 \end{bmatrix}, \left[P^{-1}M_1, 0\right] V \right) : P \in \mathcal{LG}(r) \right\}.$$

Similar arguments hold true for the transposed operator with a matrix $M^*$, showing that $\mathrm{ind}(p,m) = \mathrm{ind}(q,n)$. Note that $\widehat{M^*} = \widehat{M}^\top$. □

We see that in all cases with $\mathcal{M} \neq 0$ the solution set is $\Theta_r(M)$ of (7) with $r = \nu(M)$ is parametrized via the $r^2$ free elements of the matrix $P \in \mathcal{GL}(r)$. Note that equation (7) is of the form $\pi(\theta) = 0$, where the entries of $\pi : \Sigma_r \simeq \mathcal{F}^{r(pm+nq)} \to \mathcal{F}^{pmnq}$ are second order polynomials, i.e., we have $pmnq$ scalar quadratic equations in $r(pm+nq)$ scalar unknowns (the elements of $A$ and $B$). Hence we may expect that generically the solution set $\Theta_r(M)$ is an $\ell$-parameter family, where $\ell := r(pm+nq) - pmnq$. Since $r^2 - \ell = (pm-r)(nq-r) \geq 0$, we see that in the generic case $\nu = \min\{pm, nq\}$ this is true. In the non-generic cases $\nu < \min\{pm, nq\}$ there are in general more free parameters in $\Theta_r(M)$ than the differnce between the number of unknowns and equations.

**Example 3** Consider the operator $\mathcal{P}_n \in \mathbf{Lin}(n,n,n,n)$ acting as $\mathcal{P}_n[X] = X^\top$. The matrix associated with $\mathcal{P}_n$ is $\Pi_n$. Since $\mathrm{rank}[\Pi_n] = n^2$ we see easily that $\mathrm{ind}(n,n)[\Pi_n] = n^2$. In particular we have

$$X^\top = \sum_{i,j=1}^{n} E_{i,j}(n,n) X E_{i,j}(n,n).$$

Consider finally the case when $mn = pq$ and the operator $\mathcal{M}$ is invertible, i.e. its associated matrix $M$ is non-singular. For some classes of invertible operators it may be shown that $\mathrm{ind}(p,m)[\mathcal{M}] = \mathrm{ind}(m,p)[\mathcal{M}^{-1}]$. Whether it is true in general is an open question.

# 3 Lyapunov operators

## 3.1 Real Lyapunov operators

An important class of linear operators are the Lyapunov operators, which are automorphisms in $\mathcal{F}^{n\times n}$. In this section we consider the class of real Lyapunov operators in $\mathbf{Lin}(n) := \mathbf{Lin}(n,n,n,n)$.

**Definition 4** *An operator $\mathcal{L} \in \mathbf{Lin}(n)$ is called a real Lyapunov operator if $\mathcal{L}^\top[X] = \mathcal{L}[X^\top]$ for all $X \in \mathcal{R}^{n\times n}$. We denote by $\mathbf{Lyap}(n) \subset \mathbf{Lin}(n)$ the set of Lyapunov operators.*

It follows from Definition 4 that

$$\begin{aligned} X = X^\top &\;\Rightarrow\; \mathcal{L}[X] = \mathcal{L}^\top[X] \\ X = -X^\top &\;\Rightarrow\; \mathcal{L}[X] = -L^\top[X] \end{aligned}$$

provided $\mathcal{L} \in \mathbf{Lyap}(n)$. Hence the subspaces of symmetric and anti-symmetric $n \times n$ real matrices are invariant subspaces for real Lyapunov operators, see also [4].

Obviously $\mathbf{Lyap}(n)$ is a linear subspace of $\mathbf{Lin}(n)$, which may be characterised by the next proposition.

**Proposition 4** *The following statements are equivalent:*

**(i)** $\mathcal{L} \in \mathbf{Lyap}(n)$.

**(ii)** There exists $\mathcal{M} \in \mathbf{Lin}(n)$ such that $\mathcal{L}[X] = \mathcal{M}[X] + \mathcal{M}^{\top}[X^{\top}]$ for all $X \in \mathcal{F}^{n\times n}$, i.e.,

$$\mathcal{L}[X] = \sum_{k=1}^{r} \left( A_k X B_k + B_k^{\top} X A_k^{\top} \right)$$

or equivalently

$$L := \mathrm{Mat}(\mathcal{L}) = \sum_{k=1}^{r} \left( B_k^{\top} \otimes A_k + A_k \otimes B_k^{\top} \right),$$

where $r \geq 1$ and $A_k, B_k \in \mathcal{R}^{n\times n}$ are given matrices.

**(iii)** $L \in \Omega(n)$, i.e. $\Pi_n L = L \Pi_n$.

*Proof.* The proof follows from the definitions. □

The representation of $\mathcal{L} \in \mathbf{Lyap}(n)$ as in Proposition 4(i) is not unique. As in the case of a general Sylvester operator $\mathcal{M} \in \mathbf{Lin}(n)$, the Lyapunov operator $\mathcal{L}$ may be represented in a condensed form as a sum of $\ell := \mathrm{ind}[\mathcal{L}]$ elementary linear operators but in this case the symmetry in Proposition 4(ii) may be lost. However, for each $\mathcal{L} \in \mathbf{Lyap}(n)$ there exist two integers $\ell_1, \ell_2 \geq 0$, such that $\ell \leq \ell_1 + 2\ell_2$ and

$$\mathcal{L}[X] = \sum_{k=1}^{\ell_1} \varepsilon_k C_k X C_k^{\top} + \sum_{k=1}^{\ell_2} (A_k X B_k + B_k^{\top} X A_k^{\top}),\ \varepsilon_k = \pm 1.$$

Note that the strict inequality $\ell < \ell_1 + 2\ell_2$ is possible.

**Example 4** Let the operator $\mathcal{L}$ be defined by the symmetric expression

$$\mathcal{L}[X] = A^{\top} X A + A^{\top} X + X A$$

i.e., $\ell_1 = \ell_2 = 1$. At the same time the Sylvester index of $\mathcal{L}$ is at most 2, since $\mathcal{L}[X] = A^{\top} X (A + I_n) + XA$. Hence $\ell \leq 2 < 3 = \ell_1 + 2\ell_2$.

According to parts (i) and (iii) of Proposition 4 a matrix $L \in \mathcal{R}^{n^2\times n^2}$ is the matrix associated with a Lyapunov operator if and only if it has the symmetry property $\Pi_n L = L\Pi_n$, or, equivalently, $L = \Pi_n L \Pi_n$.

**Proposition 5** *The set $\Omega(n)$ of matrices associated with real Lyapunov operators is isomorphic to the subspace*

$$\mathrm{Ker}\,(I_{n^2} \otimes \Pi_n - \Pi_n \otimes I_{n^2}) = \mathrm{Ker}(\Pi_n \otimes \Pi_n - I_{n^4}) \subset \mathcal{R}^{n^4}. \qquad (8)$$

*Proof.* Taking the vec operation on both sides of the characteristic equation $\mathcal{L}^\top[X] = \mathcal{L}[X^\top]$ of a Lyapunov operator $\mathcal{L}$ with associated matrix $L$ we get

$$\begin{aligned} \mathrm{vec}[\mathcal{L}^\top[X]] &= \mathrm{vec}[\mathcal{L}[X^\top]] \\ \Pi_n \mathrm{vec}[L[X]] &= L\mathrm{vec}[X^\top] \\ \Pi_n L\mathrm{vec}[X] &= L\Pi_n \mathrm{vec}[X] \end{aligned}$$

and hence $\Pi_n L = L\Pi_n$. Multiplying the last equation with $\Pi_n$ and taking into consideration that $\Pi_n^2 = I_{n^2}$ we also get $L = \Pi_n L \Pi_n$. The characterisation of $\Omega(n)$ by the subspace (8) is obtained by taking the vec operation on both sides of $\Pi_n L - L\Pi_n = 0$, namely $(I_{n^2} \otimes \Pi_n - \Pi_n \otimes I_{n^2})\mathrm{vec}[L] = 0$. □

**Proposition 6** *The (real) dimension of* **Lyap**$(n)$ *is* $n^2(n^2+1)/2$.

*Proof.* We can give the proof via (8), but an alternative proof is as follows. The matrix equation $R(L) := \Pi_n L - L\Pi_n = 0$ for $L = [l_{ij}]$ is equivalent to $n^4$ scalar equations of the type $r_{ij}(L) = l_{pq} - l_{st} = 0$ for the elements of $L$. The equations, corresponding to $i = kn + k + 1,\ k \le n-1$, and/or $j = \ell n + \ell + 1,\ \ell \le n-1$, are zero identities $l_{pq} - l_{pq} = 0$ and there are $n^2$ of these identities. For each one of the remaining $n^4 - n^2$ equations of the form $l_{pq} - l_{st} = 0$ there is a corresponding equation $l_{st} - l_{pq} = 0$. Thus the number of linearly independent scalar equations is $(n^4 - n^2)/2$. Hence the number of free parameters in $L$ are $n^4 - (n^4 - n^2)/2 = n^2(n^2+1)/2$ as claimed. □

**Example 5** For $n = 2$ and $n = 3$ the sets $\Omega(2)$ and $\Omega(3)$ are 10- and 45-dimensional real spaces with patterns $\Lambda_2$ and $\Lambda_3$ of the free parameters as follows:

$$\Lambda_2 = \begin{bmatrix} \underline{1} & 2 & 2 & \underline{3} \\ 4 & 5 & 6 & 7 \\ 4 & 6 & 5 & 7 \\ \underline{8} & 9 & 9 & \underline{10} \end{bmatrix}, \quad \Lambda_3 = \begin{bmatrix} \underline{1} & 2 & 3 & 2 & \underline{4} & 5 & 3 & 5 & \underline{6} \\ 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\ 7 & 10 & 13 & 8 & 11 & 14 & 9 & 12 & 15 \\ \underline{25} & 26 & 27 & 26 & \underline{28} & 29 & 27 & 29 & \underline{30} \\ 31 & 32 & 33 & 34 & 35 & 36 & 37 & 38 & 39 \\ 16 & 19 & 22 & 17 & 20 & 23 & 18 & 21 & 24 \\ 31 & 34 & 37 & 32 & 35 & 38 & 33 & 36 & 39 \\ \underline{40} & 41 & 42 & 41 & \underline{43} & 44 & 42 & 44 & \underline{45} \end{bmatrix}.$$

In both examples, the underlined elements are in the positions corresponding to the zero identities in the equation $R(L) = 0$.

If $\mathcal{M} \in \mathbf{Lin}(n)$ is a general Sylvester operator, then according to Definition 3 we have

$$\|\mathcal{M}\|_{\mathrm{F}} := \sigma_{\max}(\mathcal{M}) := \sigma_1(\mathrm{Mat}(\mathcal{M})) = \max\{\|\mathcal{M}(X)\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1\}.$$

Similarly

$$\sigma_{\min}(\mathcal{M}) := \sigma_{n^2}(\mathrm{Mat}(\mathcal{M})) = \min\{\|\mathcal{M}(X)\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1\}$$

and if $\mathcal{M} \in \mathbf{Lin}(n)$ is invertible, then $\|\mathcal{M}^{-1}\|_{\mathrm{F}} = 1/\sigma_{\min}(\mathcal{M})$.

For Lyapunov operators $\mathcal{L} \in \mathbf{Lyap}(n)$, however, in addition to $\sigma_{\max}(\mathcal{L})$ and $\sigma_{\min}(\mathcal{L})$, we may also define the symmetrised values

$$\|\mathcal{L}\|_{\mathrm{F}}^* := \sigma_{\max}^*(\mathcal{L}) := \max\{\|\mathcal{L}[X]\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1,\ X = X^\top\}$$

and

$$\sigma_{\min}^*(\mathcal{L}) := \min\{\|\mathcal{L}[X]\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1,\ X = X^\top\},$$

and if $\mathcal{L}$ is invertible, then

$$\|\mathcal{L}^{-1}\|_{\mathrm{F}}^* = 1/\sigma_{\min}^*(\mathcal{L}).$$

Obviously

$$\sigma_{\min}(\mathcal{L}) \le \sigma_{\min}^*(\mathcal{L}) \le \sigma_{\max}^*(\mathcal{L}) \le \sigma_{\max}(\mathcal{L}).$$

Each of these inequalities may be strict, i.e., $\sigma_{\min}(\mathcal{L}) < \sigma_{\min}^*(\mathcal{L})$ and $\sigma_{\max}^*(\mathcal{L}) < \sigma_{\max}(\mathcal{L})$ is possible. Moreover, as we show below, the differences $\sigma_{\min}^*(\mathcal{L}) - \sigma_{\min}(\mathcal{L})$ and $\sigma_{\max}(\mathcal{L}) - \sigma_{\max}^*(\mathcal{L})$ may be arbitrarily large, see Example 7.

Let $A \in \mathcal{R}^{n\times n}$ and $a := \mathrm{vec}[A] \in \mathcal{R}^{n^2}$. The map $\mathrm{vec} : \mathcal{R}^{n\times n} \to \mathcal{R}^{n^2}$ is an isomorphism and its inverse $\mathrm{vec}^{-1} : \mathcal{R}^{n^2} \to \mathcal{R}^{n\times n}$ is well defined. If we use the notation $\mathrm{vec}^{-\top}[a] = (\mathrm{vec}^{-1}[a])^\top$, then it follows that the set

$$Z(n) := \left\{a \in \mathcal{R}^{n^2} : \mathrm{vec}^{-1}[a] = \mathrm{vec}^{-\top}[a]\right\}$$

is an $n(n+1)/2$-dimensional subspace of $\mathcal{R}^{n^2}$. Moreover, we will show that

$$Z(n) = \mathrm{Rg}(I_{n^2} + \Pi_n) = \mathrm{Rg}(P_n),$$

where

$$P_n = [P_{n,ij}] \in \mathcal{R}^{n^2\times n(n+1)/2},\ i,j \in \{1,\dots,n\},$$

is an block upper-triangular matrix. The blocks $P_{n,ij} \in \mathcal{R}^{n\times j}$ are defined via

$$P_{n,ij} = \begin{cases} 0_{n\times j} & \text{if } i > j, \\ \begin{bmatrix} I_i \\ 0_{(n-i)\times i} \end{bmatrix} & \text{if } i = j, \\ E_{ji}(n,j) & \text{if } i < j. \end{cases}$$

If $L$ is the matrix associated with the Lyapunov operator $\mathcal{L}$, then we can rewrite the expression for $\sigma^*_{\max}$ in the equivalent form

$$\begin{aligned}
\sigma^*_{\max}(\mathcal{L}) &= \max\left\{\frac{\|La\|_2}{\|a\|_2} : 0 \neq a \in Z(n)\right\} \\
&= \max\left\{\frac{\|LP_nb\|_2}{\|P_nb\|_2} : 0 \neq b \in \mathcal{R}^{n(n+1)/2}\right\} \\
&= \|LQ_n\|_2 = \sigma_{\max}(LQ_n),
\end{aligned}$$

where

$$Q_n = P_n(P_n^\top P_n)^{-1} = [Q_{n,ij}] \in \mathcal{R}^{n^2\times n(n+1)/2};\ i,j = 1,\ldots,n$$

is a block upper-triangular projector ($Q_n^\top Q_n = I_{n(n+1)/2}$). The blocks $Q_{n,ij} \in \mathcal{R}^{n\times j}$ are given by $Q_{n,ij} = 0$ if $i > j$, $Q_{n,11} = [1,0,\ldots,0]^\top \in \mathcal{R}^n$, $Q_{n,kk} = [\mathrm{diag}(qI_{k-1},1),0]^\top$ and $Q_{n,ij} = qE_{ji}(n,j)$ if $i < j$, where $q := 1/\sqrt{2}$.

The matrices $P_n$ and $Q_n$ have the same sign-patterns, the only difference being that the non-zero elements of $P_n$ are equal to 1, while the non-zero elements of $Q_n$ are equal to 1 or $q$.

**Example 6** *The matrices $Q_2, Q_3, Q_4$ are*

$$Q_2 = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & q & 0 \\ \hline 0 & q & 0 \\ 0 & 0 & 1 \end{array}\right], Q_3 = \left[\begin{array}{c|cc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q & 0 & 0 \\ \hline 0 & q & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 0 \\ \hline 0 & 0 & 0 & q & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$
Q_4 = \left[\begin{array}{c|cc|ccc|cccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\
\hline
0 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 \\
\hline
0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right].
$$

Similarly, we have for the minimal singular value

$$\sigma^*_{\min}(\mathcal{L}) = \sigma_{\min}(LQ_n).$$

**Definition 5** *The singular values of the matrix $LQ_n$ are called symmetrised singular values of $\mathcal{L}$ and the set of symmetrised singular values of the Lyapunov operator $\mathcal{L}$ is denoted as*

$$\sigma^*(\mathcal{L}) := \sigma(LQ_n).$$

We immediately obtain that

$$\sigma^*(\mathcal{L}) = \sigma(LQ_n) = \sigma(L^\top Q_n) = \sigma(Q_n^\top LQ_n).$$

To compare the classical and symmetrised maximal and minimal singular values, consider the following example.

**Example 7** Let $n = 2$ and let

$$
\begin{aligned}
\mathcal{L}_1[X] &:= E_{11}XE_{22} + E_{22}XE_{11} - E_{12}XE_{12} - E_{21}XE_{21}, \\
\mathcal{L}_2[X] &:= X + \beta L_1[X],
\end{aligned}
$$

where $E_{ij} := E_{ij}(2,2)$, $L_k := \mathrm{Mat}(\mathcal{L}_k)$ and $\beta > -1/2$. Then we have the associated matrices

$$
L_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1+\beta & -\beta & 0 \\ 0 & -\beta & 1+\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

Since $\sigma_{\max}(\mathcal{L}_1) = 2$ and $L_1 Q_2 = 0$, the maximum singular value $\sigma_{\max}(\beta\mathcal{L}_1) = 2|\beta|$ of the operator $\beta\mathcal{L}_1$ may be arbitrarily larger than its maximum symmetrised singular value $\sigma^*_{\max}(\beta\mathcal{L}_1) = 0$. Furthermore we have $\sigma(\mathcal{L}_2) = \{2\beta+1, 1, 1, 1\}$ and since $L_2 Q_2 = Q_2$ we obtain $\sigma^*(\mathcal{L}_2) = \{1, 1, 1\}$. Then for large $\beta$ the maximum singular value $\sigma_{\max}(\mathcal{L}_2) = 2\beta + 1$ of $\mathcal{L}_2$ is arbitrarily larger than its maximum symmetrised singular value $\sigma^*_{\max}(\mathcal{L}_2) = 1$. Finally, let $\beta = -1/2 + \varepsilon/2$, where $\varepsilon > 0$ is a small parameter. Then the minimum singular value $\sigma_{\min}(\mathcal{L}_2) = \varepsilon$ of $\mathcal{L}_2$ may be arbitrarily smaller than its minimum symmetrised singular value, which is equal to 1.

Currently, it is not clear what the exact relationship between the set of standard and symmetrised singular values is, but, based on several numerical experiments, we **conjecture** that for $\mathcal{L} \in \mathbf{Lyap}(n)$ and the associated matrix $L = \mathrm{Mat}(\mathcal{L})$ we have that

$$\sigma^*(\mathcal{L}) \subset \sigma(\mathcal{L}). \tag{9}$$

It is also interesting to define the class of Lyapunov $\mathcal{L}$ with Sylvester index $\mathrm{ind}[\mathcal{L}] \le 2$ such that

$$\sigma_{\min}(\mathcal{L}) = \sigma^*_{\min}(\mathcal{L}), \sigma_{\max}(\mathcal{L}) = \sigma^*_{\max}(\mathcal{L}). \tag{10}$$

A staightforward calculation shows that (9) holds for $n = 2$. As Example 7 shows for for $\mathrm{ind}[\mathcal{L}] \ge 4$ it is possible that $\sigma_{\min}(\mathcal{L}) < \sigma^*_{\min}(\mathcal{L})$ and/or $\sigma_{\max}(\mathcal{L}) > \sigma^*_{\max}(\mathcal{L})$. It is shown in [4] that for $n = 3$ and $\mathrm{ind}(\mathcal{L}) = 2$ relation 10 is not valid. The case $\mathrm{ind}[\mathcal{L}] = 3$ when, e.g.

$$\mathcal{L}[X] = AXB + B^\top XA^\top + C^\top XC$$

is also not completely analyzed yet.

If 10 holds, then for the corresponding Lyapunov operators that are most used in practice, i.e., for

$$\mathcal{L}_c[X] = A^\top XE + E^\top XA, \ \mathcal{L}_d[X] = A^\top XA - E^\top XE$$

it is justified to use the minimum singular value instead of the minimum symmetrised singular value, since they would be equal. Of course, for general Lyapunov operators one must use the symmetrised singular values.

Note that if (9) holds, then it is not necessarily true that $\mathcal{L} \in \mathbf{Lyap}(n)$, as is demonstrated in the following example.

**Example 8** Let

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

then $\sigma(MQ_2) = \{\sqrt{10}, 1, 1\} \subset \sigma(M) = \{\sqrt{10}, 1, 1, 0\}$, but $M \notin \Omega(2)$.

### 3.2 Complex Lyapunov operators

The above results for real Lyapunov operators can be extended to Lyapunov operator in complex linear spaces as we demonstrate now. In this section the superscript $^{\mathcal{C}}$ means that the corresponding linear space is over $\mathcal{C}$. In particular $\mathbf{Lin}^{\mathcal{C}}(n)$ is the space of linear matrix opertors $\mathcal{C}^{n\times n} \to \mathcal{C}^{n\times n}$.

An operator $\mathcal{M} \in \mathbf{Lin}^{\mathcal{C}}(n)$ is represented in the form (2), where $A_k, B_k \in \mathcal{C}^{n\times n}$. Definition 3 is directly applicable to such operators and Proposition 3 holds as well. Definition 4 is modified as follows:

**Definition 6** *The operator $\mathcal{L} \in \mathbf{Lin}^{\mathcal{C}}(n)$ is said to be a Lyapunov operator if $\mathcal{L}^{\mathrm{H}}[X] = \mathcal{L}[X^{\mathrm{H}}]$ for $X \in \mathcal{C}^{n\times n}$.*

In the complex case, due to the non-linearity of the complex conjugation, the set $\mathbf{Lyap}^{\mathcal{C}}(n) \subset \mathbf{Lin}^{\mathcal{C}}(n)$ of Lyapunov operators is not a subspace of $\mathbf{Lin}^{\mathcal{C}}(n)$ and the set $\Omega^{\mathcal{C}}(n) \subset \mathcal{C}^{n^2\times n^2}$ is not a subspace of $\mathcal{C}^{n^2\times n^2}$ (these sets become subspaces if we consider linear spaces of complex matrices with $\mathcal{R}$ as a field of scalars or if we pass to the representation $\mathcal{C}^{n^2\times n^2} \simeq \mathcal{R}^{2n^2\times 2n^2}$).

We obtain the following modification of Proposition 4.

**Proposition 7** *The following statements are equivalent:*

**(i)** $\mathcal{L} \in \mathbf{Lyap}^{\mathcal{C}}(n)$.

**(ii)** *There exists $\mathcal{M} \in \mathbf{Lin}^{\mathcal{C}}(n)$ such that $\mathcal{L}[X] = \mathcal{M}[X] + \mathcal{M}^{\mathrm{H}}[X^{\mathrm{H}}]$ for $X \in \mathcal{C}^{n\times n}$, i.e.,*

$$\mathcal{L}[X] = \sum_{k=1}^{r} \left( A_k X B_k + B_k^{\mathrm{H}} X A_k^{\mathrm{H}} \right)$$

*and*

$$L := \mathrm{Mat}(\mathcal{L}) = \sum_{k=1}^{r} \left( B_k^{\mathrm{T}} \otimes A_k + \overline{A}_k \otimes B_k^{\mathrm{H}} \right)$$

*where $r \geq 1$ and $A_k, B_k \in \mathcal{C}^{n\times n}$ are given matrices.*

**(iii)** $L \in \Omega^{\mathcal{C}}(n)$, *i.e.,* $\Pi_n L = \overline{L}\Pi_n$.

If we represent $L$ as $L = L_0 + \jmath L_1$, where $L_0, L_1 \in \mathcal{R}^{n^2\times n^2}$, then Proposition 7(iii) yields

$$L_0\Pi_n - \Pi_n L_0 = 0 \qquad (11)$$

$$L_1\Pi_n + \Pi_n L_1 = 0. \qquad (12)$$

Hence we come to the following analogues of Propositions 5 and 6.

**Proposition 8** *The set $\Omega^{\mathcal{C}}(n)$ of matrices associated with complex Lyapunov operators is isomorphic to the subspace*

$$\mathrm{Ker}(\mathrm{diag}(U_-(n), U_+(n))) = \mathrm{Ker}(\mathrm{diag}(V_-(n), V_+(n))) \subset \mathcal{R}^{2n^4},$$

*where*

$$U_\pm(n) := I_{n^2} \otimes \Pi_n \pm \Pi_n \otimes I_{n^2}, \; V_\pm(n) := \Pi_n \otimes \Pi_n \pm I_{n^4}.$$

*Proof.* The proof follows directly from (11) and (12). □

**Proposition 9** *The real dimension of* $\mathbf{Lyap}^{\mathcal{C}}(n) \simeq \Omega^{\mathcal{C}}(n)$ *is* $n^4$.

*Proof.* Based on equation (11) we have $n^2(n^2+1)/2$ free (real) parameters in the matrix $L_0$ according to Proposition 8. The number of free parameters in $L_1$ is obtained as follows. Equation (12) for the matrix $L_1$ is equivalent to $n^4$ scalar equations $l_{pq} + l_{st} = 0$ for its elements $l_{ij}$. In the $n^2$ positions of the zero identities in equation (11) (see the proof of Proposition 4) the corresponding scalar equations in (12) are of the form $2l_{pq} = 0$ and are linearly independent. For each of the remaining $n^4 - n^2$ equations $l_{pq} + l_{st} = 0$, there is an equivalent equation $l_{st} + l_{pq} = 0$. Thus the number of linearly independent scalar equations in (12) is $n^2 + (n^4 - n^2)/2 = (n^4 + n^2)/2$ and the number of free elements in $L_1$ becomes $n^4 - (n^4 + n^2)/2 = n^2(n^2 - 1)/2$. Adding this number to the number $n^2(n^2 + 1)/2$ of free elements in $L_0$ we obtain that the number of free real scalars in $L$ is $n^4$ as claimed. □

The maximum and minimum symmetrised singular values of the operator $\mathcal{L} \in \mathbf{Lyap}^{\mathcal{C}}(n)$ are defined as

$$\begin{aligned} \sigma^*_{\max}(\mathcal{L}) &:= \max\{\|\mathcal{L}[X]\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1, \; X = X^{\mathrm{H}}\}, & (13)\\ \sigma^*_{\min}(\mathcal{L}) &:= \min\{\|\mathcal{L}[X]\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1, \; X = X^{\mathrm{H}}\}, & (14) \end{aligned}$$

respectively.

The symmetrised singular values for a complex Lyapunov operator $\mathcal{L}$ with matrix $L = L_0 + \jmath L_1$, $L_i \in \mathcal{R}^{n\times n}$, are determined as follows. Let $X = X_0 + \jmath X_1$, $X_i \in \mathcal{R}^{n\times n}$. Then the restriction $X = X^{\mathrm{H}}$ in (13) gives $X_0^{\mathrm{T}} = X_0$, $X_1^{\mathrm{T}} = -X_1$. Hence we may take $X_0 = Y_0 + Y_0^{\mathrm{T}}$, $X_1 = Y_1 - Y_1^{\mathrm{T}}$, where the matrices $Y_i \in \mathcal{R}^{n\times n}$ are arbitrary. Thus

$$\begin{aligned} \mathrm{vec}[X_0] &\in Z(n) = \mathrm{Rg}(I_{n^2} + \Pi_n),\\ \mathrm{vec}[X_1] &\in Z'(n) := \mathrm{Rg}(I_{n^2} - \Pi_n), \end{aligned}$$

where $Z'(n)$ is an $n(n-1)/2$-dimensional subspace of $\mathcal{R}^{n^2}$. As in the real case, we get

$$\sigma^*_{\max}(\mathcal{L}) = \|\widehat{L}_n\|_2$$

where

$$\widehat{L}_n := \begin{bmatrix} L_0 Q_n & -L_1 \widehat{Q}_n \\ L_1 Q_n & L_0 \widehat{Q}_n \end{bmatrix} \in \mathcal{R}^{2n^2 \times n}.$$

The matrix $\widehat{Q}_n \in \mathcal{R}^{n^2 \times n(n-1)/2}$ is obtained from $Q_n$ by deleting the columns contained ones which are numbered as $k(k+1)/2$, $k = 1, \ldots, n$, and by changing the sign of each second element $q$ in each column of the reduced matrix. Formally this procedure is described as follows. Let

$$\Delta_n = (\Delta_n)_{ij} := [\delta_{i(i+1)/2,j}] \in \mathcal{R}^{n(n+1)/2 \times n(n-1)/2}$$

where $\delta_{ij}$ is the Kronecker $\delta$, and

$$J := \{(kn+l, k(k-1)/2+l) : k = 1, \ldots, n-1,\ l = 1, ldots, k\}.$$

Then

$$(\widehat{Q}_n)_{ij} = \begin{cases} (Q_n \Delta_n)_{ij} & \text{if } (i,j) \notin J \\ -(Q_n \Delta_n)_{ij} & \text{if } (i,j) \in J \end{cases}.$$

**Definition 7** *The symmetrised singular values of the complex Lyapunov operator $\mathcal{L}$ with associated matrix $L$ are the singular values of $\widehat{L}_n$:*

$$\sigma^*(\mathcal{L}) := \sigma(\widehat{L}_n).$$

A similar conjecture as in the real case can be stated for complex Lyapunov operators $\mathcal{L}$.

So far we have made some formal analysis of generalized Sylvester and Lyapunov operators. In the following section, we discuss the application of these formal results to the sensitivity and perturbation analysis of Lyapunov equations.

# 4 Sensitivity and perturbation analysis of Lyapunov equations

Consider the Hermitian Lyapunov equation

$$\mathcal{L}[X] = Q,\ Q^{\mathrm{H}} = Q \neq 0 \tag{15}$$

with an invertible Lyapunov operator $\mathcal{L}$. The minimum symmetrised singular value $\sigma^*_{\min}(\mathcal{L})$ of $\mathcal{L}$ is a relevant measure for the sensitivity of the Lyapunov equation (15) relative to perturbations in the coefficient matrices of $\mathcal{L}$ and Hermitian perturbations $\Delta Q = \Delta Q^{\mathrm{H}}$ in the matrix $Q$.

Denote by $P = P^{\mathrm{H}} = \mathcal{L}^{-1}[Q]$ the solution of (15) and let $X = P + \Delta P$ be the solution to the perturbed Lyapunov equation $\mathcal{L}(X) = Q + \Delta Q$. We have $\Delta P = \mathcal{L}^{-1}[\Delta Q]$ and hence

$$\|\Delta P\|_{\mathrm{F}} \leq \|\mathcal{L}^{-1}\|^*_{\mathrm{F}} \|\Delta Q\|_{\mathrm{F}} = \frac{1}{\sigma^*_{\min}(\mathcal{L})} \|\Delta Q\|_{\mathrm{F}}.$$

In terms of relative perturbations it is fulfilled that

$$\delta_P \le \kappa^* \delta_Q, \ \kappa^* := \frac{1}{\sigma^*_{\min}(\mathcal{L})} \frac{\|Q\|_{\mathrm{F}}}{\|P\|_{\mathrm{F}}},$$

where $\delta_Z := \|\Delta Z\|_{\mathrm{F}}/\|Z\|_{\mathrm{F}}$ and $\kappa^*$ is the *relative condition number* of the Lyapunov equation (15) with respect to Hermitian perturbations in $Q$. Note that usually $Q = C^{\mathrm{H}}C$ and if the matrix $C$ is perturbed then the perturbation $\Delta Q = \Delta C^{\mathrm{H}}C + C^{\mathrm{H}}\Delta C + \Delta C^{\mathrm{H}}\Delta C$ in $Q$ is Hermitian.

Most of the perturbation bounds in the literature [9, 6] are based on $\sigma_{\min}(\mathcal{L})$ instead on $\sigma^*_{\min}(\mathcal{L})$, e.g. the condition number is taken as $\kappa := \|Q\|_{\mathrm{F}}/(\|P\|_{\mathrm{F}}\sigma_{\min}(\mathcal{L}))$. Since $\kappa \ge \kappa*$ may be much larger than $\kappa^*$, it is clear that in case of Hermitian perturbations one should use the relevant sensitivity estimates based on symmetrised singular values instead on usual singular values of Lyapunov operators. At the same time sensitivity estimates, based on the usual singular values, should be used in case of non-Hermitian perturbations.

Consider now the a posteriori error analysis of equation (15). Suppose that $\widehat{P}$ is an approximate solution of equation (15), e.g. the solution produced by a numerical method in floating-point computing environment. Then it is important to have a sharp computable bound on the actual relative error

$$\delta_{\widehat{P}} := \frac{\|\widehat{P} - P\|_{\mathrm{F}}}{\|P\|_{\mathrm{F}}}.$$

Such a tight bound may be derived using the symmetrised singular values of $\mathcal{L}$ and in particular the symmetrised relative condition number of $\mathcal{L}$, defined below.

Denote $\widehat{Q} := \mathcal{L}[\widehat{P}]$. We have $\mathcal{L}[\widehat{P} - P] = \widehat{Q} - Q$ which gives $\widehat{P} - P = \mathcal{L}^{-1}[\widehat{Q} - Q]$ and

$$\|\widehat{P} - P\|_{\mathrm{F}} \le \frac{\|\widehat{Q} - Q\|_{\mathrm{F}}}{\sigma^*_{\min}(\mathcal{L})}. \tag{16}$$

Since $\|Q\|_{\mathrm{F}} \le \sigma^*_{\max}(\mathcal{L})\|P\|_{\mathrm{F}}$ we have

$$\frac{1}{\|P\|_{\mathrm{F}}} \le \frac{\sigma^*_{\max}(\mathcal{L})}{\|Q\|_{\mathrm{F}}}. \tag{17}$$

Combining (16) and (17) we get the desired estimate

$$\delta_{\widehat{P}} \le \mathrm{cond}^*_2(\mathcal{L}) \frac{\|\widehat{Q} - Q\|_{\mathrm{F}}}{\|Q\|_{\mathrm{F}}},$$

where

$$\mathrm{cond}^*_2(\mathcal{L}) := \frac{\sigma^*_{\max}(\mathcal{L})}{\sigma^*_{\min}(\mathcal{L})}$$

is the *symmetrised relative condition number* of $\mathcal{L}$ with respect to inversion.

## References


[1] J.B. Barlow, M.M. Monahemi, and D.P. O'Leary. Constrained matrix Sylvester equations. *SIAM J. Matrix Anal. Appl.*, 13:1-9, 1992.

[2] S. Barnett. *Introduction to Mathematical Control Theory.* Oxford University Press, Oxford, 1975.

[3] S. Barnett and C. Storey. *Matrix Methods in Stability Theory.* Nelson, London, 1970.

[4] R. Byers and S. Nash. On the singular "vectors" of the Lyapunov operator. *SIAM J. Algebraic Discrete Methods*, 8:59-66, 1987.

[5] D. Carlson and R. Loewy. On ranges of Lyapunov transformations. *Linear Algebra Appl.*, 8:237-248, 1974.

[6] P.M. Gahinet, A.J. Laub, C.S. Kenney, and G.A. Hewer. Sensitivity of the stable discrete-time Lyapunov equation. *IEEE Trans. Automat. Control*, 35:1209-1217, 1990.

[7] Z. Gajić and M.T.J. Qureshi. *Lyapunov Matrix Equation in System Stability and Control.* Academic Press, San Diego, CA, 1995.

[8] W. Givens. Elementary divisors and some properties of the Lyapunov mapping $X \to AX + XA^*$. Technical Report Rep. ANL-6456, Argonne Nat. Lab., 1961.

[9] G. Hewer and C. Kenney. The sensitivity of the stable Lyapunov equation. *SIAM J. Cont. Optim.*, 26:321-344, 1988.

[10] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, UK, 1985.

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

[12] C. Kenney and G. Hewer. The sensitivity of the algebraic and differential Riccati equations. *SIAM J. Cont. Optim.*, 28:50-69, 1990.

[13] M.M. Konstantinov, P.Hr. Petkov, and N.D. Christov. Perturbation analysis of the discrete Riccati equation. *Kybernetica* (Prague), 29:18-29, 1993.

[14] L. Kronecker. Algebraische Reduction der Schaaren bilineare Formen. *S.B. Akad. Berlin*, pages 1225-1237, 1880.

[15] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[16] P.Hr. Petkov, N.D. Christov, and M.M. Konstantinov. *Computational Methods for Linear Control Systems.* Prentice-Hall, Hemel Hempstead, Herts, UK, 1991.

[17] W.E. Roth. The equations $AX - YB = C$ and $AX - XB = C$ in matrices. *Proc. Amer. Math. Soc.*, 3:392–396, 1952.

[18] D.E. Rutherford. On the solution of the matrix equation $AX+XB = C$. *Nederl. Akad. Wetensch. Proc. Ser. A*, 35:53–59, 1932.

[19] P. Stein. On the ranges of two functions of positive definite matrices. *J. Algebra*, 2:350–353, 1965.

[20] P. Stein and A. Pfeffer. On the ranges of two functions of positive definite matrices ii. *ICC Bull.*, 6:81–86, 1967.

[21] G.W. Stewart. *Introduction to Matrix Computations.* Academic Press, New York, 1973.

[22] J.J. Sylvester. Sur la solution du cas le plus général des équations linéaires en quantités binaires, c'est-à-dire en quaternions ou en matrices du second ordre. *C.R. Acad. Sci. Paris*, 99:117–118, 1884.

[23] J.J. Sylvester. Sur l'équation en matrices $px = xq$. *C.R. Acad. Sci. Paris*, 99:67–71, 1884.

[24] O. Taussky and H. Wielandt. On the matrix function $AX + X'A'$. *Arch. Rat. Mech. Anal.*, 9:93–96, 1962.

[25] D.S. Watkins. *Fundamentals of Matrix Computations.* John Wiley and Sons Inc., 1991.

[26] J.H.M. Wedderburn. Note on the linear matrix equation. *Proc. Edinburgh Math. Soc.*, 22:49–53, 1904.

[27] W.M. Wonham. *Linear Multivariable Control: A Geometric Approach.* Springer-Verlag, New York, second edition, 1979.