# On the inverse of the Lyapunov operator


Tobias Damm
Zentrum Mathematik, TU München, Germany
damm@ma.tum.de



**Abstract**

We derive minimal representations for the inverses of Lyapunov and Sylvester operators.


## 1 Introduction

Let $A, B \in \mathbb{C}^{n\times n}$ and define the Sylvester operator $\mathcal{S}_{A,B} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ by

$$\mathcal{S}_{A,B}(X) \quad = \quad AX - XB \; .$$

If $B = -A^*$ then $\mathcal{S}_{A,B}$ is a Lyapunov-operator, and we write $\mathcal{S}_{A,B} = \mathcal{L}_A$.
It is assumed that $\mathcal{S}_{A,B}$ is nonsingular i.e. $0 \notin \sigma(A) + \sigma(B)$.

During the Oberwolfach meeting on 'Nonnegative matrices, M-matrices and their generalizations', V. Mehrmann raised a question about minimal representations for the inverse of the Lyapunov-operator:
Clearly, there exist matrices $B_i, C_i \in \mathbb{C}^{n\times n}$ such that for all $Y \in \mathbb{C}^{n\times n}$ we have

$$\mathcal{L}_A^{-1}(Y) = \sum_{i=1}^{N} B_i Y C_i \; . \tag{1}$$

But what is the minimal number $N$ of terms in this sum?

In the present paper we show, that this question can be answered straightforwardly by combining results from [3] and from [2]. More precisely, our aim is to prove the following:

**Theorem 1.1** *Let $\nu_A = \deg \mu_A$ and $\nu_B = \deg \mu_B$ denote the degrees of the minimal polynomials of $A$ and $B$, respectively, and $\nu = \min\{\nu_A, \nu_B\}$.*

*(i) The inverse* $\mathcal{S}_{A,B}^{-1} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ *has a representation of the form*

$$\mathcal{S}_{A,B}^{-1}(Y) \;=\; \sum_{i=1}^{\nu} B_i Y C_i, \qquad B_i, C_i \in \mathbb{C}^{n\times n} .$$

*(ii) If* $B = -A^*$*, i.e.* $\mathcal{S}_{A,B} = \mathcal{L}_A$*, there exists a symmetric representation*

$$\mathcal{L}_A^{-1}(Y) \;=\; \sum_{i=1}^{\nu_A} \varepsilon_i A_i Y A_i^*, \quad \varepsilon_i = \pm 1, \;\; A_i \in \mathbb{C}^{n\times n}$$

*with* $\varepsilon_i = 1$ *for all* $i$ *if and only if* $\sigma(A) \subset \mathbb{C}_+$*.*
*In the latter case* $\mathcal{L}_A$ *is said to be* completely positive *(compare [1]).*

*(iii) If an arbitrary representation*

$$\mathcal{S}_{A,B}^{-1}(Y) \;=\; \sum_{i=1}^{N} B_i Y C_i, \qquad B_i, C_i \in \mathbb{C}^{n\times n}$$

*is given, then necessarily* $N \geq \nu$*.*

We begin with some facts from representation theory for mappings between matrix spaces. These results are not new, but we take the chance of giving a concise self-contained presentation, providing all the details we need.

## 2 Representation of mappings between matrix spaces

We consider linear mappings $\mathcal{T} : \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{p\times q}$ and try to give a representation of the form $\mathcal{T}(X) = \sum_{i=1}^{N} B_i X C_i$, with $B_i \in \mathbb{C}^{p\times m}$, $C_i \in \mathbb{C}^{n\times q}$.
Obviously $\mathcal{T}$ can also be regarded as a linear mapping between the vector spaces $\mathbb{C}^{mn}$ and $\mathbb{C}^{pq}$. To make use of this observation, we recall the definition and some basic properties of the Kronecker product and the vec-operator.

**Definition 2.1** *Let* $V = (v_{ij}) = (v_1, \ldots, v_n) \in \mathbb{C}^{m\times n}$ *and* $U = (u_{ij}) \in \mathbb{C}^{p\times q}$*. Then*

$$V \otimes U \;=\; \Big(v_{ij} U\Big) \in \mathbb{C}^{mp\times nq} \quad \text{and} \quad \operatorname{vec} V \;=\; \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^{nm} .$$

**Lemma 2.2 (e.g. [4])** *Let* $U_1$*,* $U_2$*,* $V_1$*,* $V_2$*, and* $X$ *be matrices of appropriate sizes. Then*

$$(V_1 \otimes U_1)(V_2 \otimes U_2) \;=\; (V_1 V_2) \otimes (U_1 U_2) \quad \text{and} \quad \operatorname{vec}(U_1 X V_1) \;=\; (U_1^T \otimes V_1) \operatorname{vec} X .$$

In the following let $e_i^{(m)}$ denote the $i$-th canonical unit vector in $\mathbb{C}^m$ and $E_{ij}^{(mn)} = e_i^{(m)} e_j^{(n)*}$ the $m \times n$ matrix with the only nonzero entry 1 in the $i$-th row and $j$-th column. By $E^{mn}$ we denote the $m^2 \times n^2$-block matrix $\left(E_{ij}^{(mn)}\right)_{i,j=1}^{m,n}$ and with the mapping $\mathcal{T}$ we associate the $mp \times nq$ matrix $(I_{pq} \otimes \mathcal{T})(E^{(mn)}) = \left(\mathcal{T}(E_{ij}^{(mn)})\right)_{i,j=1}^{(m,n)}$.

The following simple identity plays a crucial role.

**Lemma 2.3** *If $V = (v_1, \ldots, v_m) \in \mathbb{C}^{p\times m}$ and $W = (w_1, \ldots, w_q) \in \mathbb{C}^{n\times q}$ then*

$$\left(V E_{ij}^{(mn)} W^*\right)_{i,j=1}^{m,n} = \operatorname{vec} V (\operatorname{vec} W)^*. \tag{2}$$

**Proof:** By Lemma 2.2 we have

$$\begin{aligned}
\operatorname{vec}(V E_{ij}^{(mn)} W^*) &= \bar{W} \otimes V \operatorname{vec} E_{ij}^{(mn)} \\
&= \left(\bar{w}_1 \otimes v_1, \bar{w}_1 \otimes v_2, \ldots, \bar{w}_q \otimes v_m\right) e_{(j-1)n+i}^{(nm)} \\
&= \bar{w}_j \otimes v_i \\
&= \operatorname{vec} v_i w_j^* \,.
\end{aligned}$$

Thus $V E_{ij}^{(mn)} W^* = v_i w_j^*$ and

$$\left(V E_{ij}^{(mn)} W^*\right)_{i,j=1}^{m,n} = \left(v_i w_j^*\right)_{i,j=1}^{m,n} = \operatorname{vec} V (\operatorname{vec} W)^*.$$

□

**Theorem 2.4** *Let $\mathcal{T} : \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{p\times q}$ be linear.*
*The matrices $B_1, \ldots, B_N \in \mathbb{C}^{p\times m}$, $C_1, \ldots, C_N \in \mathbb{C}^{n\times q}$ yield a representation of $\mathcal{T}$:*

$$\forall X \in \mathbb{C}^{m\times n} : \; \mathcal{T}(X) = \sum_{i=1}^{N} B_i X C_i \tag{3}$$

*if and only if*

$$(I_{pq} \otimes \mathcal{T})(E^{(mn)}) = \sum_{i=1}^{N} \operatorname{vec} B_i (\operatorname{vec} C_i)^* \,. \tag{4}$$

*In particular, the minimal number of summands is $\nu = \operatorname{rk}(I_{pq} \otimes \mathcal{T})(E^{(mn)})$, and one can choose a representation with matrices $B_1, \ldots, B_\nu \in \mathbb{C}^{p\times m}$, $C_1, \ldots, C_\nu \in \mathbb{C}^{n\times q}$, such that both sets $\{\operatorname{vec} B_1, \ldots, \operatorname{vec} B_\nu\}$ and $\{\operatorname{vec} C_1, \ldots, \operatorname{vec} C_\nu\}$ are orthogonal.*

**Proof:** By Lemma 2.3 the identity (3) clearly implies (4); vice versa (4) implies $\mathcal{T}(X) = \sum_{i=1}^N B_i X C_i$ for all $X = E_{ij}^{mn}$, and thus for all $X \in \mathbb{C}^{m\times n}$.
A representation with a minimal number of summands is given e.g. by a singular value decomposition of $(I_{pq} \otimes \mathcal{T})(E^{(mn)})$, in which case also the orthogonality property holds. □

Now we consider mappings between spaces of quadratic matrices.

**Definition 2.5** *Let* $\mathcal{T} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{m\times m}$ *be a linear map. Then* $\mathcal{T}$ *is called* Hermitian-preserving *if* $\mathcal{T}(\mathcal{H}^n) \subset \mathcal{H}^m$.

It is immediate to see, that the Lyapunov operator is Hermitian preserving. The following representation result can be found in [3]; the proof is adapted from [1].

**Theorem 2.6** *For a linear map* $\mathcal{T} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{m\times m}$ *the following are equivalent:*

1. $\mathcal{T}$ *is Hermitian-preserving.*
2. $\forall X \in \mathbb{C}^{n\times n}: \quad \mathcal{T}(X^*) = (\mathcal{T}(X))^*$
3. *The* $nm \times nm$ *matrix* $(I_{mm} \otimes \mathcal{T})(E^{(nn)}) = \left(\mathcal{T}(E_{ij}^{(nn)})\right)_{i,j=1}^n$ *is Hermitian.*
4. *There exist matrices* $V_1, \ldots, V_N \in \mathbb{C}^{m\times n}$ *and numbers* $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$, *such that* $\forall X \in \mathbb{C}^{n\times n} : \mathcal{T}(X) = \sum_{l=1}^{N} \varepsilon_i V_l X V_l^*$.
   *In particular one can choose* $N = \mathrm{rk}(I_{mm} \otimes \mathcal{T})(E^{(nn)})$.
   *One can choose* $\varepsilon_i = 1$ *for all* $i$, *if and only if* $(I_{mm} \otimes \mathcal{T})(E^{(nn)}) \geq 0$.

**Proof:** 1.⇒ 2.: For skew-Hermitian matrices $S$ (i.e. $iS$ Hermitian) we have on the one hand

$$\mathcal{T}(iS) = \mathcal{T}(iS)^* = -i\mathcal{T}(S)^*$$

and on the other

$$\mathcal{T}(iS) = \mathcal{T}((iS)^*) = \mathcal{T}(-iS^*) = -i\mathcal{T}(S^*)$$

which yields $\mathcal{T}(S)^* = \mathcal{T}(S^*)$.
As any matrix $A$ can be decomposed in $A = H + S$, where $H = \frac{1}{2}(A + A^*)$ is Hermitian and $S = \frac{1}{2}(A - A^*)$ is skew-Hermitian, the result follows.
2.⇒ 3.: This follows from

$$\begin{aligned}\left(\left(\mathcal{T}(E_{ij}^{(nn)})\right)_{i,j=1}^n\right)^* &= \left(\left(\mathcal{T}(E_{ij}^{(nn)})^*\right)_{j,i=1}^n\right) = \left(\left(\mathcal{T}(E_{ij}^{(nn)*})\right)_{j,i=1}^n\right)\\ &= \left(\left(\mathcal{T}(E_{ji}^{(nn)})\right)_{j,i=1}^n\right) = \left(\left(\mathcal{T}(E_{ij}^{(nn)})\right)_{i,j=1}^n\right).\end{aligned}$$

3.⇒ 4.: By Sylvester's Theorem there is a decomposition

$$\left(\mathcal{T}(E_{ij}^{(nn)})\right)_{i,j=1}^{n} = \sum_{l=1}^{N} \varepsilon_l \operatorname{vec} \bar{V}_l (\operatorname{vec} \bar{V}_l)^* \tag{5}$$

for appropriate $V_l$, where $N$ is the rank of the matrix on the left. By equation (2) we can also write

$$\left(\mathcal{T}(E_{ij}^{(nn)})\right)_{i,j=1}^{n} = \left(\sum_{l=1}^{N} \varepsilon_l V_l E_{ij}^{(nn)} V_l^*\right)_{i,j=1}^{n}$$

that means

$$\forall E_{ij}^{(nn)} : \quad \mathcal{T}(E_{ij}^{(nn)}) = \sum_{l=1}^{N} \varepsilon_l V_l E_{ij}^{(nn)} V_l^*$$

and thus the same holds with $E_{ij}^{(nn)}$ replaced by any $X \in \operatorname{span}\{E_{ij}^{(nn)} \mid i,j = 1,\ldots,n\} = \mathbb{C}^{n\times n}$.
Moreover it folows from Sylvester's Theorem, that one can choose $\varepsilon_i = 1$ for all $i$ if and only if $(I_{mm} \otimes \mathcal{T})(E^{(nn)}) \geq 0$.
4.⇒ 1.: Since for each $l$ the matrix $V_l^* X V_l$ is Hermitian if $X$ is Hermitian the sum is Hermitian, too. □

# 3 Proof of Theorem 1.1

In view of the Theorems 2.4 and 2.6 we need to determine the rank of the matrix

$$\mathcal{X} \;=\; (I_{nn} \otimes \mathcal{S}_{A,B}^{-1})(E^{(nn)}) \;=\; \left(\mathcal{S}_{A,B}^{-1}(E_{ij}^{(nn)})\right)_{i,j=1}^{n} . \tag{6}$$

To this end we apply the following result from [2] (Theorem 4):

**Theorem 3.1** *Let $A, B \in \mathbb{C}^{n\times n}$, $y \in \mathbb{C}^n$, and assume that the Sylvester equation*

$$AX - XB = yy^*$$

*has a unique solution $X$. Then*

$$\operatorname{rk} X \;=\; \min_{C\in\{A,B^*\}} \operatorname{rk}\left(y, Cy, \ldots, C^{n-1}y\right) .$$

*(The matrix on the right hand side is the controllability Gramian of the pair $(C, y)$.)*

Now, if in (6) we partition $\mathcal{X} = \left(\mathcal{X}_{ij}\right)_{i,j=1}^{n}$ conformably with $E^{(nn)}$, such that $\mathcal{X}_{ij} \in \mathbb{C}^{n\times n}$, then we can also write $A\mathcal{X}_{ij} - \mathcal{X}_{ij}B = E_{ij}^{(nn)}$ and thus

$$(I_n \otimes A)\mathcal{X} - \mathcal{X}(I_n \otimes B) \;=\; E^{(nn)} . \tag{7}$$

In other words $\mathcal{X}$ is a solution of the Sylvester equation (7) with the right hand side $E^{(nn)}$. By (2) (with $V = W = I_n$), we have $E^{(nn)} = \operatorname{vec} I_n (\operatorname{vec} I_n)^*$.
It follows therefore from Theorem 3.1, that

$$\begin{aligned}
\operatorname{rk}\mathcal{X} &= \min_{C\in\{A,B^*\}} \operatorname{rk}\left(\operatorname{vec} I_n, (I_n\otimes C)\operatorname{vec} I_n, \ldots, (I_n\otimes C)^{n^2}\operatorname{vec} I_n\right) \\
&= \min_{C\in\{A,B^*\}} \operatorname{rk}\begin{pmatrix} e_1^{(n)} & Ce_1^{(n)} & \ldots & C^{n^2}e_1^{(n)} \\ e_2^{(n)} & Ce_2^{(n)} & \ldots & C^{n^2}e_2^{(n)} \\ \vdots & \vdots & & \vdots \\ e_n^{(n)} & Ce_n^{(n)} & \ldots & C^{n^2}e_n^{(n)} \end{pmatrix} .
\end{aligned}$$

It is clear, that the first $\nu+1$ columns are linearly dependent, if $\nu = \deg \mu_C$ denotes the degree of the minimal polynomial of $C$; hence $\operatorname{rk}\mathcal{X} \le \nu$.
If we consider arbitrary linear combinations of the block rows, we find

$$\operatorname{rk}\mathcal{X} \;\ge\; \max_{x\in\mathbb{C}^n} \operatorname{rk}\left(x\,,\, Cx\,,\,\ldots\,,\, C^n x\right) = \nu\,.$$

Hence $\operatorname{rk}\mathcal{X} = \min\{\nu_A, \nu_B\}$, which proves (i), (iii) and the first part of (ii).
It remains to show that the $\varepsilon_i$ can be chosen positive if and only if $\sigma(A) \in \mathbb{C}_+$. If all $\varepsilon_i$ are positive, then $\mathcal{L}_A^{-1}$ is positive, i.e. $\mathcal{L}_A^{-1}(Y) \ge 0$ for all $Y \ge 0$ (the converse, however, is not true, see [1]). But it is well-known (e.g. [4]), that $\mathcal{L}_A^{-1}$ is positive if and only if $\sigma(A) \subset \mathbb{C}_+$. Hence $\varepsilon_i \ge 0$ for all $i$ implies $\sigma(A) \subset \mathbb{C}_+$.
To prove the converse, we need to show $\mathcal{X} \ge 0$. Let $\mathcal{L}_A = R_+ - R_-$ with

$$R_\pm(X) \;=\; \frac{1}{2}(A\pm I)X(A\pm I)^* .$$

One readily verifies that $\rho(R_+^{-1}R_-) < 1$ if $\sigma(A) \subset \mathbb{C}_+$, whence

$$\begin{aligned}
\mathcal{L}_A^{-1}(Y) &= \sum_{k=0}^{\infty} R_+^{-(k+1)} \circ R_-(Y) \\
&= \sum_{k=0}^{\infty} V_k Y V_k^* \quad \text{with } V_k = (A+I)^{-(k+1)}(A-I)^k .
\end{aligned}$$

Therefore

$$\mathcal{X} \;=\; \sum_{k=0}^{\infty} (I_n\otimes V_k)E^{(nn)}(I_n\otimes V_k)^* \ge 0\,,$$

which completes the proof. □

# References


[1] M.-D. Choi. Completely positive linear maps on complex matrices. *Lin. Alg. Appl.*, 10:285–290, 1975.

[2] E. de Souza and S. Bhattacharyya. Controllability, observability and the solution of $AX - XB = C$. *Lin. Alg. Appl.*, 39:167–188, 1981.

[3] R. D. Hill. Linear transformations which preserve Hermitian matrices. *Lin. Alg. Appl.*, 6:257–262, 1973.

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