SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

A study of Kronecker product and Lyapunov equations

av

Viktor Sas

2019 - No K11

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

A study of Kronecker product and Lyapunov equations

Viktor Sas

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2019

Abstract

In this thesis we study the Kronecker product and its applications in solving matrix equations. First we will give some preliminaries as a good tool to understand the calculations we will do with the Kronecker product. The preliminaries contain material from linear algebra and ordinary differential equations (ODE). We deal with the Kronecker product together with the vec – operator on matrix equations. The method is then applied to a special class of matrix equations, Lyapunov equations, in particular their relations to stability theory for linear dynamical systems are investigated. We will also study three different methods to solve the least square problem in the formulation of Kronecker product.

Acknowledgement

I would like to thank my supervisor Yishao Zhou for guiding and helping me on this subject.

Contents

1. Introduction

In this thesis we study the Kronecker product and its applications. In particular we investigate the preservation of matrix structures under such an operation. The thesis is organized as follows.

In Section 2 we gather some definitions and theorems from linear algebra and linear ordinary differential equations in matrix form for later use. The purpose is to make it easier for the reader to follow the rest of the text. In Section 3 we present the main topic in details. We begin by the definition of the Kronecker product and present its basic properties, especially the matrix structure preservation under the Kronecker product, such as LU-, QR-, Cholesky-factorization, and singular value decomposition. For the reader to feel comfortable with the properties of the Kronecker product we also present some examples. Then we study some linear matrix equations using the Kronecker product in order to get the "standard" form of linear equation "matrix-times-vector-equals-vector-form". To this end we introduce the vec-operator, Kronecker sum and their properties and how they can be used to reshape the matrix equations. Among these equations, the Sylvester and the Lyapunov equations brought up in Section 4 where we investigate the Lyapunov equation in great detail, from a brief introduction of the stability theory to different solution methods to unique solvability of positive definite solutions of the resulting Lyapunov equation. Finally in Section 5 we study how the least square problems in Kronecker product form can be solved using the information and solutions or factorizations from the smaller matrices involved and thus we can avoid working on large problems.

2. Preliminaries

In this section some elemental material relevant to this text is collected.
The material will be from linear algebra and ordinary differential equations.

2.1 Linear Algebra

Let us go through some basic definition from linear algebra focusing on factorization properties.

2.1.1 Basic definitions

We will though take as granted that the meaning of vector space, linear combinations, determinant is known for the reader. However, we repeat matrices and linear independence of vectors and some of their properites.

We use the notation $\mathbb{M}_{n,m}$ for the set of $n{\times}m$ matrices over some field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) and $\mathbb{M}_n$ if $m = n$.

Definition 2.1 (Invertible Matrix), see page 24 in [1]:
Let $A \in \mathbb{M}_n$ is called **invertible** if there exists an $B \in \mathbb{M}_n$ such that

$$AB = BA = I_n$$

where I_n denotes the $n{\times}n$ **identity matrix** and the multiplication used is the ordinary matrix multiplication. If A is invertible, then the matrix B is uniquely determined by A and is called the **inverse** of A, denoted A^{-1}. A simple property is that $\det(A) \neq 0$.

Definition 2.2 (Norm), see page 235 in [1]:

Let V be a complex or real vector space. A **norm** in V is a function

$$V \to \mathbb{R},$$

which for every vector v calculates $\|v\|$ satisfying follwing three properties:

- $\|v\| \geq 0$, and $\|v\| = 0 \Leftrightarrow v = 0$,
- $\|\alpha v\| = |\alpha| \cdot \|v\|$, for any $\alpha \in \mathbb{K}$, $\quad (\mathbb{K}\ is\ either\ \mathbb{R}\ or\ \mathbb{C})$,
- $\|v + u\| \leq \|v\| + \|u\|$ for any $u, v \in V$ (**triangle inequality**).

In this text we consider only finite dimensional vector spaces. More precisely $\dim(V) = n$.

Theorem 2.1
Let $v_1, \dots, v_n$ be vectors in V. Let A be a matrix, the columns of which, A_i are the coordinates of v_i in some basis $e_1, \dots, e_n$. Then the following conditions are equivalent:

- The vectors v_i are linearly independent,
- The vectors v_i generate V,
- The vectors v_i form a basis for V,
- $\det(A) \neq 0$,
- rank of A or $r(A) = n$,
- A is invertible or A is nonsingular,
- The system $AX = B$ has a unique solution for any column vector B,
- The system $AX = 0$ has only the trivial solution $X = 0$.

Proof: see page 75 in [1].

Definition 2.3 (Diagonalizable), see page 118 in [1]:
A square matrix A is **diagonalizable** if it is similar to a diagonal matrix $D = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$. In other words if there exists an invertible matrix S such that

$$S^{-1}AS = D = diag(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Definition 2.4 (Eigenvalues and eigenvectors), see page 118 in [1]:

A non-zero vector v is called an **eigenvector** for a linear operator $\mathcal{A}$ if

$$\mathcal{A}v = \lambda v$$

for some scalar λ. This scalar is called an **eigenvalue** for $\mathcal{A}$ associated with v.
Considering a matrix as an operator we can say the same:
A nonzero column vector v is called an **eigenvector** for A if

$$Av = \lambda v$$

and λ is called its **eigenvalue**.

Definition 2.5 (Positive definite), see page 429-430 in [2]:
A real symmetric matrix $A \in \mathbb{M}_n$ over $\mathbb{R}$ is **positive definite** if $x^T Ax > 0$ for all nonzero $x \in \mathbb{R}^n$. A matrix is **positive semidefinite** if $x^T Ax \geq 0$ for all nonzero $x \in \mathbb{R}^n$. Consequently $x^T Ax$ is *real* for all $x \in \mathbb{R}^n$. Conversely, if $A \in \mathbb{M}_n$ and $x^T Ax$ is *real* for all $x \in \mathbb{R}^n$, then A is symmetric, so assuming that A is symmetric in the preceding definitions, while customary, is actually superfluous. Of course, if A is positive definite, it is also positive semidefinite.

Theorem 2.2:
Any real symmetric matrix has only real eigenvalues and is always diagonalizable and the eigenvectors may be choosen orthogonal.
Furthermore $A = Q^T \Lambda Q$, where $Q^T Q = QQ^T = I$.

Proof: See page 266 in [1].

Definition 2.6 (Trace) see page 88 in [1]:
Assume $A \in \mathbb{M}_n$. The trace of A is $a_{11} + \cdots + a_{nn}$ (or $\sum_{i=1}^{n} a_{ii}$), the sum of the diagonal elements.

Definition 2.7 (Commutator) [3],
If $A, B \in \mathbb{M}_n$ then we can define the **commutator** of A and B to be

$$[A, B] = AB - BA.$$

2.1.2 Matrix factorization

In this subsection we only work on $\mathbb{K} = \mathbb{R}$ for simplicity of exposition.

Theorem 2.3 (LU-factorization/decomposition):
For every matrix A there exist a permutation matrix P, an matrix U which has the shape resulting from Gaussian elimination (echelon form) and a lower triangular matrix L with ones on the main diagonal such that

$$PA = LU$$

This is called the **LU-decomposition** (or LU-factorization) of A.

***Proof*:** See page 35 [1].
For more information about LU-decomposition see page 216 [2].

Theorem 2.4 (QR-decomposition), see page 262 in [1].
Every invertible matrix A has a unique ***QR-decomposition*** (or QR-factorization), namely

$$A = QR,$$

where Q is orhtogonal matrix and R is a upper triangular matrix with positive elements on the main diagonal. Generally, if $A \in \mathbb{M}_{n,m}$ we can find a decomposition $A = QR$ where Q is an orthogonal matrix and R has an echelon form with the positive pivot elements. But this decomposition is not unique if A is not invertible.

More concrete,
Let $A \in \mathbb{M}_{n,m}$ be given.

a) If $n \geq m$, there is a $Q \in \mathbb{M}_{n,m}$ with orthonormal columns and an upper triangular $R \in \mathbb{M}_m$ with nonnegative main diagonal entries such that $A = QR$.
b) If $n < m$, then the factors Q and R in a) are uniquely determined and the main diagonal entries of R are all positive.
c) If $n = m$, then the factor Q in a) is unitary.
d) There is a unitary $Q \in \mathbb{M}_n$ and an upper triangular $R \in \mathbb{M}_{n,m}$ with nonnegative diagonal entries such that $A = QR$.

***Proof*:** See the pages 262 in [1] and pages 89-90 in [2].

Theorem 2.5 (Cholesky factorization), see page 441 in [2]**:**
Let $A \in \mathbb{M}_n$ be symmetric. Then A is positive semidefinite (positive definite, respectively) if and only if there is a lower triangular matrix $L \in \mathbb{M}_n$ with nonnegative (respectively, positive) diagonal entries such that $A = LL^T$. If A is positive definite, L is unique. If A is real, L may be taken to be real.

Proof:
First we show that there exists a positive definite (semi-definite) matrix $A^{\frac{1}{2}}$ such that $A = A^{\frac{1}{2}} A^{\frac{1}{2}}$. Since A is positive definite (semi-definite) there exists a Q with $Q^TQ = QQ^T = I$ such that $A = Q^T\Lambda Q$ and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. By the **Theorem 2.2** in the previous subsection all $\lambda_i > 0$ $(or \geq 0)$, so $\Lambda = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}$, with $\Lambda^{\frac{1}{2}} = diag\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\right)$.
Thus $A = Q^T\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}Q = Q^T\Lambda^{\frac{1}{2}}QQ^T\Lambda^{\frac{1}{2}}Q = \left(Q^T\Lambda^{\frac{1}{2}}Q\right)\left(Q^T\Lambda^{\frac{1}{2}}Q\right) = A^{\frac{1}{2}}A^{\frac{1}{2}}$.
Let $A^{\frac{1}{2}} = QR$ be a QR factorization and let $L = R^T$. Then $A = A^{\frac{1}{2}}A^{\frac{1}{2}} = R^TQ^TQR = LL^T$.
The asserted properties of L follow from the properties of R. ■

Definition 2.8 (Rank), see page 37 in [1]:
The rank $r(A)$ of a matrix A is the rank of the matrix U in its *LU-decomposition* $PA = LU$ i.e. the number r of pivot elements in U.

Theorem 2.6 (Sinuglar value decomposition or SVD), see page 149-154 in [2]:
Let $A \in \mathbb{M}_{n,m}$ be given, let $q = \min\{m, n\}$, and suppose that $r(A) = r$.

(a) There are orthogonal matrices $V \in \mathbb{M}_n$ and $W \in \mathbb{M}_m$, and a square diagonal matrix

$$\Sigma_q = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_q \end{bmatrix}$$

such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_q$ and $A = V\Sigma W^T$ in which

- $\Sigma = \Sigma_q$ if $m = n$,
- $\Sigma = \begin{bmatrix} \Sigma_q & 0 \end{bmatrix} \in \mathbb{M}_{n,m}$ if $m > n$,
- $\Sigma = \begin{bmatrix} \Sigma_q \\ 0 \end{bmatrix} \in \mathbb{M}_{n,m}$ if $n > m$.

(b) The parameters $\sigma_1, \ldots, \sigma_r$ are the positive square roots of the decreasingly ordered nonzero eigenvalues of AA^T, which are the same as the decreasingly ordered nonzero eigenvalues of A^TA. They are called singular values of A.

The proof of the theorem and more can you find on page 150-151 in [2].

Remark:
It is very easy to think about SVD is just another matrix decomposition. However we want to point out something more fundamental. We recall the following theorem

The fundamental theorem of linear algebra [4].

Remember that linear transformation (or matrices A) are the important object in linear algebra. Associated with them are the four fundamental vector spaces,

I. Column space $\mathcal{C}(A)$ that spans by the columns in A, (or image/range space of A).
II. Row space (or coimage) $\mathcal{C}(A^T)$, spans by the rows in A.
III. Nullspace (or kernel) $\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$,
IV. Left null space (or cokernel) $\mathcal{N}(A^T)$ where A^T is the transpose of A.

They are related to each other by **The fundamental theorem of linear algebra:**

Let $A \in \mathbb{M}_{m,n}$ be real matrix.

1. The column space and the row space have the samme dimension r which is called rank of A. The nullspace $\mathcal{N}(A)$ have dimension $n - r$, and the left nullspace $\mathcal{N}(A^T)$ have dimension $m - r$.

More precisely:

- $\dim(\,\mathcal{C}(A)) = \dim(\mathcal{C}(A^T)) = r = rank(A)$
- $dim(\mathcal{N}(A)) = n - r$
- $\dim(\mathcal{N}(A^T)) = m - r$

2. $\mathcal{C}(A^T) = \mathcal{N}(A)^{\perp}$ or $\mathbb{R}^m = \mathcal{C}(A^T) \oplus \mathcal{N}(A)$
   $\mathcal{N}(A^T) = \mathcal{C}(A)^{\perp}$ or $\mathbb{R}^n = \mathcal{N}(A^T) \oplus \mathcal{C}(A)$

3. There exists orthogonal matrices $W \in \mathbb{M}_n$ respectively $V \in \mathbb{M}_m$ such that

$$AW = V\Sigma$$

   where Σ is the block matrix in the form

$$\Sigma = \begin{pmatrix} diag(\sigma_1, \dots, \sigma_r) & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$$

   and $diag(\sigma_1, \dots, \sigma_r)$ is the diagonal matrix with $\sigma_1, \dots, \sigma_r$ elements on the diagonal. Note that

$$AW = V\Sigma \;\; \Leftrightarrow A = V\Sigma W^T$$

   which is the wellknown *Singular value decomposition (SVD).*

Let columns in *V* and *W* be $v_1, \dots, v_m$ and $w_1, \dots, w_n$, respectively. Then $v_1, \dots, v_m$ are orthonormal and similar for $w_1, \dots, w_n$.

It's obvious that SVD isn't just a matrix decomposition. It creates ON-basis $v_1, \dots, v_m$ and $w_1, \dots, w_n$ for the four fundamental vector spaces (I. – IV.).

More precisely we have columnwise from 3. above:

- $Aw_i = 0, i = r+1, \dots, n.$ so $\mathcal{N}(A) = span\{w_{r+1}, \dots, w_n\}$,
- $Aw_i = v_i\sigma_i, i = 1, \dots, r, \sigma_i \neq 0 \Rightarrow v_i \in \mathcal{C}(A)$

  but $\mathcal{N}(A) = \mathcal{C}(A^T)^{\perp}$ and $\mathcal{N}(A^T) = \mathcal{C}(A)^{\perp}$ respectively. So $\{w_1, \dots, w_r\}$ is an ON-basis for the row space and $\{v_{r+1}, \dots, v_m\}$ is an ON-basis for the column space.

- If we consider *A* as a linear transformation from the row space to the column space, $Aw_i = \sigma_i v_i$ for $i = 1, \dots, r$ means that the matrix representation for *A* in the basis $\{w_1, \dots, w_r\}$ in the rowspace is a diagonal matrix in ON-basis $\{v_1, \dots, v_r\}$ (in the column space).

Remark:
We can use the singular value for the norm $\|A\|_2 = \sigma_{max}(A)$ (the maximal singular value of A).

2.2 Linear algebra and relations to linear systems ODEs

The material in this subsection is based on e.g. Sontag. See page 467 – 492 in [5].

2.2.1 System of first order equations and state form

In continuous time, a first order system is defined in terms of $x_1(t), x_2(t), \ldots, x_n(t)$ that are functions of the continuous variable t. These variables are related by a system of n first order differential equations of the following general form:

$$\begin{cases} \dot{x}_1(t) = f_1(x_1(t), \ldots, x_n(t), t) \\ \dot{x}_2(t) = f_2(x_1(t), \ldots, x_n(t), t) \\ \qquad\vdots \\ \dot{x}_n(t) = f_n(x_1(t), \ldots, x_n(t), t) \end{cases}$$

where $\dot{x}_i = \frac{dx_i}{dt}$ and f_i are continuous functions of $x_1(t), x_2(t), \ldots, x_n(t), t$. In matrix form

$$\dot{x}(t) = f(x(t), t),$$

where $x(t) = \left(x_1(t), \ldots, x_n(t)\right)^T$ and $f = (f_1, \ldots, f_n)^T$

If a continuous-time system is *linear* then it has the following form called state space form

$$\begin{cases} \dot{x}_1(t) = a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \cdots + a_{1n}(t)x_n(t) + u_1(t) \\ \dot{x}_2(t) = a_{21}(t)x_1(t) + a_{22}(t)x_2(t) + \cdots + a_{2n}(t)x_n(t) + u_2(t) \\ \qquad\vdots \\ \dot{x}_n(t) = a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \cdots + a_{nn}(t)x_n(t) + +u_n(t) \end{cases}$$

as before, the $x_i(t), i = 1,2, \ldots, n$ are state variables, the $a_{ij}(t)$ are coefficients, and $u_i(t), i = 1, \ldots, n$ are forcing terms. In order to guarantee existence and uniqueness of solution, the $a_{ij}(t)$ are assumed to be continuous in t. The linear system in matrix form is

$$\dot{x} = A(t)x(t) + u(t)$$

where the $x(t)$ is the $n \times 1$ vector, $u(t)$ is the $n \times 1$ forcing vector and $A(t)$ is the $n \times n$ matrix of coefficients referred to as the *system matrix*. If the matrix $A(t) = A$, independent of t, the system is said to be *time invariant*. [6]

2.2.2 Matrix exponential

Theorem 2.7 (Matrix exponential), see page 386 in [7]:
Let $A \in \mathbb{M}_n$.

a) There exist functions $\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)$, such that

$$e^{At} = \alpha_1(t)A^{n-1}t^{n-1} + \alpha_2(t)A^{n-2}t^{n-2} + \cdots + \alpha_{n-1}(t)At + \alpha_n I.$$

b) For the polynomial (in *r*)

$$p(r) = \alpha_1(t)r^{n-1} + \alpha_2(t)r^{n-2} + \cdots + \alpha_{n-1}(t)r + \alpha_n(t).$$

if λ is an eigenvalue of A, then $e^{\lambda} = p(\lambda)$, so that $e^{\lambda t} = p(\lambda t)$.

c) If λ is an eigenvalue of multiplicity k, then

$$e^{\lambda} = \left.\frac{dp(r)}{dr}\right|_{r=\lambda}, e^{\lambda} = \left.\frac{d^2p(r)}{dr^2}\right|_{r=\lambda}, \dots, e^{\lambda} = \left.\frac{d^{k-1}p(r)}{dr^{k-1}}\right|_{r=\lambda}.$$

2.2.3 Variation of constant formula, see page 21 in [7]

A way to solve for to solve linear equations via an integrating factor

1. Write the linear equation in the form $\frac{dy}{dx} + P(x)y = Q(x)$.
2. Calculate the integrating factor $e^{\int P(x)dx}$.
3. Evaluate the integral $\int Q(x)\, e^{\int P(x)dx}dx$ and then multiply this result by $e^{-\int P(x)dx}$.
4. The general solution to $\frac{dy}{dx} + P(x)y = Q(x)$ is

$$y = Ce^{-\int P(x)dx} + e^{-\int P(x)dx} \int Q(x)e^{\int P(x)dx}\, dx.$$

So if we consider first order system of the form

$$\dot{x} = Ax + F(t)$$

where F is a vector with continuous functions as its entries.
That system has solution

$$x(t) = e^{At}c + e^{At} \int e^{-At}F(t)dt.$$

It is easiest to understand the derivation of this solution in terms of a general fundamental matrix and the method of variation of parameters. If we also have an initial condition, the solution can be written to take this into account. Thus

$$\dot{x} = Ax + F(t), \qquad x(t_0) = x_0$$

has solution

$$x(t) = e^{A(t-t_0)}x_0 + e^{At}\int_{t_0}^{t} e^{-As}F(s)ds.$$

Lets us do one example with it without the initial condition in mind.
See page 396-397 in [7]. We will solve the following problem

$$\dot{x} = \begin{pmatrix} 1 & -4 \\ -2 & -1 \end{pmatrix} x + \begin{pmatrix} -\sin(t) \\ e^t \end{pmatrix}.$$

To solve this we can calculate the matrix exponential by **Theorem 2.6** and get

$$e^{At} = \frac{1}{3}\begin{pmatrix} e^{-3t} + 2e^{3t} & 2e^{-3t} - 2e^{3t} \\ e^{-3t} - e^{3t} & 2e^{-3t} + e^{3t} \end{pmatrix}$$

We also need to calculate $\int e^{-At}F(t)dt$. We note that e^{-At} is easily calculated from e^{At} to simply replacing t with $-t$. Then we have

$$\int e^{-At}F(t)dt = \int \frac{1}{3}\begin{pmatrix} e^{3t} + 2e^{-3t} & 2e^{3t} - 2e^{-3t} \\ e^{3t} - e^{-3t} & 2e^{3t} + e^{-3t} \end{pmatrix}\begin{pmatrix} -\sin(t) \\ e^t \end{pmatrix} dt =$$

$$= \int \frac{1}{3}\begin{pmatrix} -e^{3t}\sin(t) - 2e^{-3t}\sin(t) + 2e^{4t} - 2e^{-2t} \\ -e^{3t}\sin(t) + e^{-3t}\sin(t) + 2e^{4t} + e^{-2t} \end{pmatrix} dt =$$

$$= \frac{1}{30}\begin{pmatrix} 2e^{-3t}\cos(t) + 6e^{-3t}\sin(t) + e^{3t}\cos(t) - 3e^{3t}\sin(t) + 5e^{4t} + 10e^{-2t} \\ e^{3t}\cos(t) - 3e^{3t}\sin(t) - e^{-3t}\cos(t) - 3e^{-3t}\sin(t) - 5e^{-2t} + 5e^{4t} \end{pmatrix}.$$

We then left multiply by e^{At} and obtain, after a lot of simplifications,

$$\frac{1}{10}\begin{pmatrix} \cos(t) + \sin(t) + 5e^t \\ -2\sin(t) \end{pmatrix}.$$

We add this last vector to the product of e^{At} with an arbitrary vector. In the formula, we gave the arbitrary constant vector as c, but for comparison purposes, we let k be our arbitrary constant vector.

$$x(t) = \frac{1}{3}\begin{pmatrix} e^{-3t} + 2e^{3t} & 2e^{-3t} - 2e^{3t} \\ e^{-3t} - e^{3t} & 2e^{-3t} + e^{3t} \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} + \frac{1}{10}\begin{pmatrix} \cos(t) + \sin(t) + 5e^t \\ -2\sin(t) \end{pmatrix} =$$

$$= \frac{1}{3}\begin{pmatrix} (k_1 + 2k_2)e^{-3t} + (2k_1 - 2k_2)e^{3t} \\ (k_1 + 2k_2)e^{-3t} + (-k_1 + k_2)e^{3t} \end{pmatrix} + \frac{1}{10}\begin{pmatrix} \cos(t) + \sin(t) + 5e^t \\ -2\sin(t) \end{pmatrix}.$$

Now let $c_1 = \frac{(k_1+2k_2)}{3}$ and $c_2 = \frac{(-k_1+k_2)}{3}$ then our answer can be written as

$$x(t) = c_1 e^{-3t}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{3t}\begin{pmatrix} -2 \\ 1 \end{pmatrix} + e^t\begin{pmatrix} 1/2 \\ 1 \end{pmatrix} + \cos(t)\begin{pmatrix} 1/10 \\ 0 \end{pmatrix} + \sin(t)\begin{pmatrix} 1/10 \\ -1/5 \end{pmatrix}.$$

3. Kronecker Product

Now we introduce the definition of Kronecker product. We will do some examples with the properties of Kronecker product!

Definition 3.1 (Kronecker product), see page 11 in [8]:
Consider two matrices $A \in \mathbb{M}_{n,m}$ and $B \in \mathbb{M}_{p,q}$ we define the Kronecker product of A and B as follows

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix},$$

which is a matrix of size $np \times mq$.

3.1 Basic properties of Kronecker product [9]

$\boldsymbol{KP1}$) $Scalar -$ $For\ all\ elements\ \alpha\ in\ \mathbb{C}, A \in \mathbb{M}_{n,m}\ and\ B \in \mathbb{M}_{p,q}$:

$$(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B).$$

$\boldsymbol{KP2}$) $Right - distributive - Let\ A \in \mathbb{M}_{n,m}\ and\ B \in \mathbb{M}_{n,m}\ and\ C \in \mathbb{M}_{p,q}$:

$$(A + B) \otimes C = A \otimes C + B \otimes C.$$

$\boldsymbol{KP3}$) $Left - distributive - Let\ A \in \mathbb{M}_{n,m}\ and\ B \in \mathbb{M}_{p,q}\ and\ C \in \mathbb{M}_{p,q}$:

$$A \otimes (B + C) = A \otimes B + A \otimes C.$$

$\boldsymbol{KP4}$) *Combined distribution* (*see page* 12 *in* [8]):
If $(A+B) \in \mathbb{M}_{n,m}$ *and* $(C+D) \in \mathbb{M}_{p,q}$ *then*:

$$(A+B) \otimes (C+D) = (A \otimes C) + (B \otimes C) + (A \otimes D) + (B \otimes D)$$

Proof:
Let $E = (A+B)$ then we have

$$E \otimes (C+D) = [\boldsymbol{KP3})] = E \otimes C + E \otimes D = [put\ back\ (A+B)] =$$
$$= (A+B) \otimes C + (A+B) \otimes D = [\boldsymbol{KP2})] = (A \otimes C) + (B \otimes C) + (A \otimes D) + (B \otimes D).$$

Same result will be obtained if we choose instead $E = (C+D)$. ■

$\boldsymbol{KP5}$) *Associative* $-$ *Let* $A \in \mathbb{M}_{n,m}$, $B \in \mathbb{M}_{p,q}$ *and* $C \in \mathbb{M}_{r,s}$:

$$(A \otimes B) \otimes C = A \otimes (B \otimes C).$$

$\boldsymbol{KP6}$) *Transpose* $-$ *Let* $A \in \mathbb{M}_{n,m}$ *and* $B \in \mathbb{M}_{p,q}$:

$$(A \otimes B)^T = A^T \otimes B^T.$$

$\boldsymbol{KP7}$) *Conjugate* $-$ *Let* $A \in \mathbb{M}_{n,m}$ *and* $B \in \mathbb{M}_{p,q}$:

$$(A \otimes B)^* = A^* \otimes B^*.$$

$\boldsymbol{KP8}$) *Product* (*even called* ***mixed product property***)
Let $A \in \mathbb{M}_{n,m}, B \in \mathbb{M}_{p,q}, C \in \mathbb{M}_{r,s}$ *and* $D \in \mathbb{M}_{u,v}$:

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

$\boldsymbol{KP9}$) *Commutator* $-$ *Let* $A \in \mathbb{M}_{n,m}, B \in \mathbb{M}_{p,q}, C \in \mathbb{M}_{r,s}, D \in \mathbb{M}_{u,v}$:

$$[A \otimes B, C \otimes D] = AC \otimes BD - CA \otimes DB.$$

Proof:

$$[A \otimes B, C \otimes D] = [\boldsymbol{def.\ 2.6}] = (A \otimes B)(C \otimes D) - (C \otimes D)(A \otimes B) = [\boldsymbol{KP8})] =$$
$$= AC \otimes BD - CA \otimes DB.$$ ■

$\boldsymbol{KP10}$) *Trace* $-$ *Let* $A \in \mathbb{M}_n, B \in \mathbb{M}_m$:

$$Tr(A \otimes B) = Tr(B \otimes A) = Tr(A)Tr(B)$$

$\boldsymbol{KP11}$) $rank - with\ m{\times}m\ matrix\ A\ and\ n{\times}n\ matrix\ B$:

$$rank(A \otimes B) = rank(A)rank(B)$$

$\boldsymbol{KP12}$) $Determinant - Let\ A \in \mathbb{M}_n, B \in \mathbb{M}_m$:

$$\det(A \otimes B) = \det(B \otimes A) = (\det(A))^m (\det(B))^n$$

A consequence of this property is that $A \otimes B$ or $B \otimes A$ is nonsingular if and only if both A and B are nonsingular.

$\boldsymbol{KP13}$) $If\ A \in \mathbb{M}_n\ and\ B \in \mathbb{M}_m\ are\ nonsingular\ then$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

We get this property directly from $\boldsymbol{KP7}$) and $\boldsymbol{KP12}$)

$\boldsymbol{KP14}$) $Permutation - Let\ A \in \mathbb{M}_{n,m}\ and\ B \in \mathbb{M}_{p,q}$:

$$B \otimes A = S_{n,p}(A \otimes B)S_{m,q}^T,$$

$where$

$$S_{n,m} = \sum_{i=1}^{n}(e_i^T \otimes I_m \otimes e_i) = \sum_{j=1}^{m}\left(e_i \otimes I_n \otimes e_j^T\right)$$

$is\ called\ the\ \boldsymbol{perfect\ shuffle}\ permutation\ matrix.\ Although\ the\ Kronecker\ product\ is\ not\ commutative,\ we\ see\ here\ that\ it\ is\ so\ under\ permutation\ described\ in.$
$\boldsymbol{KP14}$) $In\ Section\ 3.3\ we\ provide\ an\ example\ to\ show\ its\ role\ in\ vectorized\ matrix\ A\ and\ vectorized\ matrix\ A^T$.

3.2 Properties of factorization [9]

$\boldsymbol{KP15}$) $\boldsymbol{LU - factorization}$ - $Let\ A \in \mathbb{M}_n\ and\ B \in \mathbb{M}_m\ be\ invertible$

$and\ let\ P_A, L_A, U_A, P_B, L_B, U_B\ be\ the\ matrices\ corresponding\ to\ their\ LU\ factorizations\ with\ partial\ pivoting. Then\ we\ have\ the\ LU\ factorization\ with\ partial\ pivoting\ of\ their\ Kronecker\ product$:

$$A \otimes B = (P_A \otimes P_B)^T (L_A \otimes L_B)(U_A \otimes U_B).$$

Proof:
$Let\ A = P_A^T L_A U_A,\ B = P_B^T L_B U_B, J_A = P_A^T L_A\ and\ J_B = P_B^T L_B\ we\ get$

$$A \otimes B = (P_A^T L_A U_A) \otimes (P_B^T L_B U_B) = (J_A U_A) \otimes (J_B U_B) = [\boldsymbol{KP8})] =$$
$$= (J_A \otimes J_B)(U_A \otimes U_B) = \left((P_A^T L_A) \otimes (P_B^T L_B)\right)(U_A \otimes U_B) = [\boldsymbol{KP8})] =$$
$$= (P_A^T \otimes P_B^T)(L_A \otimes L_B)(U_A \otimes U_B) = [\boldsymbol{KP6})] = (P_A \otimes P_B)^T (L_A \otimes L_B)(U_A \otimes U_B). \quad \blacksquare$$

$\boldsymbol{KP16}) - \boldsymbol{Cholesky\ factorization}$ – $Let\ A \in \mathbb{M}_n\ and\ B \in \mathbb{M}_m\ to\ be\ positive\ definite, and\ let\ L_A, L_B\ be\ the\ matrices\ corresponding\ to\ their\ Cholesky\ factorizations\ Then\ we\ can\ easily\ derive\ the\ Cholesky\ factorization\ of\ their\ Kronecker\ product;$

$$A \otimes B = (L_A \otimes L_B)(L_A \otimes L_B)^T.$$

Proof:

$Let\ A = L_A L_A^T, B = L_B L_B^T\ we\ get$

$$A \otimes B = (L_A L_A^T) \otimes (L_B L_B^T) = [\boldsymbol{KP8}) = (L_A \otimes L_B)(L_A^T \otimes L_B^T) = [\boldsymbol{KP6})] = \\ = (L_A \otimes L_B)(L_A \otimes L_B)^T. \qquad \blacksquare$$

The fact that $A \otimes B$ is positive (semi) definite follows from the eigenvalue theorem which we will establish later.

$\boldsymbol{KP17}) - \boldsymbol{QR\ factorization} - Let\ A \in \mathbb{M}_{n,m}, B \in \mathbb{M}_{p,q}, 1 \le m \le n, 1 \le q \le p,$ $be\ of\ full\ rank, and\ let\ Q_A, R_A, Q_B, R_B\ be\ the\ matrices\ corresponding\ to\ their\ QR - factorizations. Then\ we\ have\ the\ QR\ factorization\ of\ their\ Kronecker\ product.$

$$A \otimes B = (Q_A R_A) \otimes (Q_B R_B) = (Q_A \otimes Q_B)(R_A \otimes R_B).$$

$\boldsymbol{KP18}) - \boldsymbol{Schur\ factorization} - with\ A \in \mathbb{M}_n, B \in \mathbb{M}_m\ and\ let\ U_A, T_A, U_B, T_B$ $be\ the\ matrices\ corresponding\ to\ their\ Schur\ factorizations.$
$Then\ we\ have\ the\ Shur\ factorization\ of\ their\ Kronecker\ product:$

$$A \otimes B = (U_A \otimes U_B)(T_A \otimes T_B)(U_A \otimes U_B)^*.$$

***Proof*:**

$Let\ A = U_A T_A U_A^*, B = U_B T_B U_B^*,\ H_A = U_A T_A$ and $H_B = U_B T_B, we\ will\ have$

$$A \otimes B = (U_A T_A U_A^*) \otimes (U_B T_B U_B^*) = (H_A U_A^*) \otimes (H_B U_B^*) = [\boldsymbol{KP8})] = \\ = (H_A \otimes H_B)(U_A^* \otimes U_B^*) = (U_A T_A \otimes U_B T_B)(U_A^* \otimes U_B^*) = [\boldsymbol{KP8})] = \\ = (U_A \otimes U_B)(T_A \otimes T_B)(U_A^* \otimes U_B^*) = [\boldsymbol{KP7})] = (U_A \otimes U_B)(T_A \otimes T_B)(U_A \otimes U_B)^*. \quad \blacksquare$$

$\boldsymbol{KP19}) - \boldsymbol{Singular\ value\ decomposition\ or\ SVD} -$
$Let\ A \in \mathbb{M}_{n,m}, B \in \mathbb{M}_{p,q}\ have\ rank\ r_A\ and\ r_b\ and\ let\ V_A, W_A, \Sigma_A, V_B, W_B, \Sigma_B\ be\ the\ matrices\ corresponding\ to\ their\ SVDs.$
$Then\ we\ have\ the\ SVD\ of\ their\ Kronecker\ product:$

$$A \otimes B = (V_A \otimes V_B)(\Sigma_A \otimes \Sigma_B)(W_A \otimes W_B)^T.$$

Proof:
$Let\ A = V_A\Sigma_A W_A^T, B = V_B\Sigma_B W_B^T,\ Q_A = V_A\Sigma_A\ and\ Q_B = V_B\Sigma_B, we\ will\ get$

$$A \otimes B = (V_A\Sigma_A W_A^T) \otimes (V_B\Sigma_B W_B^T) = (Q_A W_A^T) \otimes (Q_B W_B^T) = [\boldsymbol{KP8}] =$$
$$= (Q_A \otimes Q_B)(W_A^T \otimes W_B^T) = \big((V_A\Sigma_A) \otimes (V_B\Sigma_B)\big)(W_A^T \otimes W_B^T) = [\boldsymbol{KP8}] =$$
$$= (V_A \otimes V_B)(\Sigma_A \otimes \Sigma_B)(W_A^T \otimes W_B^T) = [\boldsymbol{KP6}] = (V_A \otimes V_B)(\Sigma_A \otimes \Sigma_B)(W_A \otimes W_B)^T. \ \blacksquare$$

3.3 Examples on Kronecker product properties

In this subsection we illustrate some properties by working through some examples.

$\boldsymbol{EX\_KP2}$:

Let $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}\ and\ \ C = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$

And do first $(A + B) \otimes C$:

$$\left(\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}\right) \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 3\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \end{pmatrix} =$$

$$= \begin{pmatrix} 5 & 10 & 1 & 2 \\ 10 & 15 & 2 & 3 \\ 1 & 2 & 3 & 6 \\ 2 & 3 & 6 & 9 \end{pmatrix}.$$

And now the other way :

$$A \otimes C + B \otimes C = \begin{pmatrix} 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \end{pmatrix} + \begin{pmatrix} 3\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \end{pmatrix} =$$

$$= \begin{pmatrix} 2 & 4 & 1 & 2 \\ 4 & 6 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 1 & 2 \\ 10 & 15 & 2 & 3 \\ 1 & 2 & 3 & 6 \\ 2 & 3 & 6 & 9 \end{pmatrix}.$$

So $(A + B) \otimes C = A \otimes C + B \otimes C$.
We will get similar result if we did the same with $\boldsymbol{KP3}$)

EX_KP4:

Let us use same A, B and C to check the next property – $(A \otimes B) \otimes C = A \otimes (B \otimes C)$:

$$(A \otimes B) \otimes C = \left(\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}\right) \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} =$$

$$= \begin{pmatrix} 2\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} & 1\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} \\ 0\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} & 1\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 3 & 0 \\ 2 & 4 & 1 & 2 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} =$$

$$= \begin{pmatrix} 6\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 3\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 4\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 3\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \end{pmatrix} =$$

$$= \begin{pmatrix} 6 & 12 & 0 & 0 & 3 & 6 & 0 & 0 \\ 12 & 18 & 0 & 0 & 6 & 9 & 0 & 0 \\ 2 & 4 & 4 & 8 & 1 & 2 & 2 & 4 \\ 4 & 6 & 8 & 12 & 2 & 3 & 4 & 6 \\ 0 & 0 & 0 & 0 & 3 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 & 2 & 3 & 4 & 6 \end{pmatrix}.$$

$$A \otimes (B \otimes C) = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \otimes \left(\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\right) =$$

$$= \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 3\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 0\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \\ 1\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & 2\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} =$$

$$= \begin{pmatrix} 2\begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} & 1\begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} \\ 0\begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} & 1\begin{pmatrix} 3 & 6 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{pmatrix} \end{pmatrix} =$$

$$= \begin{pmatrix} 6 & 12 & 0 & 0 & 3 & 6 & 0 & 0 \\ 12 & 18 & 0 & 0 & 6 & 9 & 0 & 0 \\ 2 & 4 & 4 & 8 & 1 & 2 & 2 & 4 \\ 4 & 6 & 8 & 12 & 2 & 3 & 4 & 6 \\ 0 & 0 & 0 & 0 & 3 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 & 2 & 3 & 4 & 6 \end{pmatrix}.$$

$(A \otimes B) \otimes C$ and $A \otimes (B \otimes C)$ are the same.

***EX_KP*9**:
Let $A, B \in \mathbb{M}_n$

$$[A \otimes I_n, I_n \otimes B] = (A \otimes I_n)(I_n \otimes B) - (I_n \otimes B)(A \otimes I_n) =$$
$$= (A \cdot I_n) \otimes (I_n \cdot B) - (I_n \cdot A) \otimes (B \cdot I_n) = A_n \otimes B_n - A_n \otimes B_n = 0$$

***EX_KP*14**:
Here is an example of (2,3)-perfect shuffle matrix

$$S_{2,3} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Note that $A_1 \otimes A_2 \neq A_2 \otimes A_1$. However if $A_1 \in \mathbb{M}_{n,m}, A_2 \in \mathbb{M}_{k,l}$ then

$$S_{n,k}(A_1 \otimes A_2)S_{m,l}^T = A_2 \otimes A_1$$

The perfect shuffle is also "behind the scenes" when the transpose of a matrix is taken, e.g.,

$$S_{2,3}vec(A) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{32} \end{pmatrix} = vec(A^T)$$

where $vec(A)$ is stacking the columns on top of each other from the first column to the last column. The $vec(A)$ is studied in next section.

3.4 Vec-operator and Kronecker product

Definition 3.2 (Vec-operator), see page 2 and 4 in [10]:
For any matrix $A \in \mathbb{M}_{n,m}$ the **vec-operator** is defined as

$$vec(A) = (a_{11}, \dots, a_{n1}, a_{12}, \dots, \ _{n2}, \dots, a_{1m}, \dots, a_{nm})^T,$$

i.e. the entries of A are stacked columnwise forming a vector of length nm.
A property of this definition is with the square matrices $m \times m \ \ A \ and \ B$:

$$trace(A^T B) = vec(A)^T vec(B),$$

Theorem 3.1
Vec-operator is linear. Linearity holds for the vec-operator.

***Proof*:**
To be a linear operator the following properties must hold:

- $vec(A + B) = vec(A) + vec(B) \ for \ \ A, B \in \mathbb{M}_n.$
- $vec(\alpha A) = \alpha \cdot vec(A) \ for \ A \ to \ be \ n \ x \ n \ matrix. and \ \alpha \ to \ be \ any \ real \ number.$

Let us have two vec-operator

$$vec(A) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{nn} \end{pmatrix}, vec(B) = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{nn} \end{pmatrix}.$$

Proof of the first property:

$$vec(A + B) = vec(a_{11} + b_{11}, a_{21} + b_{21}, \dots, a_{nn} + b_{nn}) =$$
$$= \begin{pmatrix} a_{11} + b_{11} \\ a_{21} + b_{21} \\ \vdots \\ a_{nn} + b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{nn} \end{pmatrix} + \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{nn} \end{pmatrix} = vec(A) + vec(B).$$

and the second property:

$$vec(\alpha A) = \begin{pmatrix} \alpha a_{11} \\ \alpha a_{21} \\ \vdots \\ \alpha a_{nn} \end{pmatrix} = \alpha \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{nn} \end{pmatrix} = \alpha vec(A). \qquad \blacksquare$$

Definition 3.3 (Kronecker Sum), see page 268 in [11]:
Let $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$, the **Kronecker sum** of A and B is defined as

$$A \oplus B = (I_m \otimes A) + (B \otimes I_n),$$

which will be used later.

3.5 Matrix equations

Consider matrix equations

I. $AX = B,$ $(Linear\ equation\ 1)$
II. $AX + XB = C,$ $(Sylvester\ equation)$
III. $AXB = C,$ $(Linear\ equation\ 2)$
IV. $A^T X + XA = C$ $(Lyapunov - equation)$

They can be formulated as system of linear equations in the form of matrix times a vector using Kronecker product and vec-operator.

i. $(I \otimes A)vec(X) = vec(B)$
ii. $[(I \otimes A) + (B^T \otimes I)]vec(X) = [A \oplus B^T]vec(X) = vec(C)$
iii. $(B^T \otimes A)vec(X) = vec(C)$
iv. $(I \otimes A^T)vec(X) + (A^T \otimes I)vec(X) = [A^T \oplus A^T]vec(X) = vec(C)$

Let us prove 2 of them:

***Proof* (i.)**, see page 2-3 in [10]:

Note that $AX = AXI$.

$$AX = B \Rightarrow vec(AX) = vec(B) \Rightarrow vec(AXI) = vec(B) \Rightarrow$$
$$\Rightarrow (I \otimes A)vec(X) = vec(B). \blacksquare$$

***Proof* (ii.)**

$$AX + XB = C \Rightarrow vec(AX + XB) = vec(C) \Rightarrow by\ linearity \Rightarrow$$
$$\Rightarrow vec(AX) + vec(XB) = vec(C) \Rightarrow like\ above \Rightarrow vec(AXI) + vec(IXB) = vec(C) \Rightarrow$$
$$\Rightarrow (I \otimes A)vec(X) + (B^T \otimes I)vec(X) = vec(C) \Rightarrow [\boldsymbol{Definition\ 3.3}] \Rightarrow$$
$$\Rightarrow (A \oplus B^T)vec(X) = vec(C). \quad \blacksquare$$

Theorem 3.2 (eigenvalues and eigenvectors for Kronecker product),
see page 13-14 in [8].

Let $A \in \mathbb{M}_m$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_m$ and the corresponding eigenvectors $u_1, u_2, u_3, \ldots, u_m$. Let $B \in \mathbb{M}_n$ with eigenvalues $\mu_1, \mu_2, \mu_3, \ldots, \mu_n$ and the corresponding eigenvectors $v_1, v_2, v_3, \ldots, v_n$. Then the matrix $A \otimes B$ has the eigenvalues $\lambda_j \mu_k$ with the corresponding eigenvectors $u_j \otimes v_k$, where $1 \leq j \leq m$ and $1 \leq k \leq n$.

Proof:
We will use the eigenvalue equations to solve this theorem.
The equations look like following:

$$Au_j = \lambda_j u_j \quad where \ \ 1 \leq j \leq m \tag{1}$$

$$Bv_k = \mu_k v_k \quad where \ \ 1 \leq k \leq n \tag{2}$$

Lets do the Kronecker product of Au_j with Bv_k. We will get to results
1]

$$\left(Au_j\right) \otimes (Bv_k) = [use\ (1)\ and(\ 2)] = \left(\lambda_j u_j\right) \otimes (\mu_k v_k) = \left[\boldsymbol{KP1}\right)\ with\ \lambda_j\ \Big] =$$
$$= \lambda_j \left(u_j \otimes (\mu_k v_k)\right) = [\boldsymbol{KP1})\ with\ \mu_k\ ] = \lambda_j \mu_k \left(u_j \otimes v_k\right)$$

2]

$$\left(Au_j\right) \otimes (Bv_k) = [\boldsymbol{KP8})] = (A \otimes B)\left(u_j \otimes v_k\right).$$

We can now combine 1] and 2] and get

$$(A \otimes B)\left(u_j \otimes v_k\right) = \lambda_j \mu_k \left(u_j \otimes v_k\right)$$

which is exaclty of the form of eigenvalue equation. Here the eigenvalues of $(A \otimes B)$ will be $\lambda_j \mu_k$ and its corresponding eigenvectors will be $u_j \otimes v_k$ which we were looking for.

■

There is an other theorem and will be important later:

Theorem 3.3 (Eigenvalues and eigenvectors for Kronecker sum), see page 14-15 in [8]: Assume the same condition on A and B as in **Theorem 3.2** then the eigenvalues and eigenvectors of $A \oplus B = A \otimes I_n + I_m \otimes B$ will be $\lambda_j + \mu_k$ and $u_j \otimes v_k$ respectively.

Proof:
Our goal is to get the eigenvalue equation for

$$(A \otimes I_n + I_m \otimes B)(u_j \otimes v_k)$$

Let us calculate it and see what happens:

$$\begin{aligned}
&(A \otimes I_n + I_m \otimes B)(u_j \otimes v_k) = [\boldsymbol{KP2})] = \\
&= (A \otimes I_n)(u_j \otimes v_k) + (I_m \otimes B)(u_j \otimes v_k) = [\boldsymbol{KP8})] = \\
&= (Au_j) \otimes (I_n v_k) + (I_m u_j) \otimes (Bv_k) = [use\ (1)\ and (\ 2)] = \\
&= (\lambda_j u_j) \otimes (I_n v_k) + (I_m u_j) \otimes (\mu_k v_k) = \\
&= (\lambda_j u_j) \otimes v_k + u_j \otimes (\mu_k v_k) = [\boldsymbol{KP1}) - scalar\ on\ \mu_k\ and\ \lambda_j] = \\
&= \lambda_j (u_j \otimes v_k) + \mu_k (u_j \otimes v_k) = (\lambda_j + \mu_k)(u_j \otimes v_k).
\end{aligned}$$

Now the expression $(A \otimes I_n + I_m \otimes B)(u_j \otimes v_k) = (\lambda_j + \mu_k)(u_j \otimes v_k)$
looks exactly like the eigenvalue equations which we wanted. ■

4. Lyapunov equation

In this section we will use our knowledge of Kronecker product to solve the Lyapunov equation,

$$A^T X + XA = -Q.$$

where A is known square matrix and Q is symmetric and known. X is symmetric and unknown.

This equation plays very important role in Lyapunov stability and optimal control theory. In Section 4.1 we give a brief derivation of Lyapunov equation for linear systems. Then in Section 4.2 we study the existence and uniqueness of solution to this equation, and we give in Section 4.3 the closed form of solution and conditions for it to be positive definite. Finally in Section 4.4 we show that this closed form solution is exactly the same solution obtained using Kronecker product.

4.1 Lyapunov theory for linear systems

Lyapunov, in his original 1892 work, proposed two methods for demonstrating stability. Roughly speaking stability is about convergence of solutions a dynamical to its equilibria, which is called (asymptotical) stability. The first method developed the solution in a series which was then proved convergent within limits. The second method, which is now referred to as the Lyapunov stability criterion or the direct method, makes use of a *Lyapunov function V(x)* which has an analogy to the potential function of classical dynamics. It is introduced as follows for a system $\dot{x} = f(x)$ (mentioned in Section 2.2) having an equilibrium at $x = 0$. It is locally (asymptotically) stable if there is a differentiable function $V: \mathbb{R}^n \to \mathbb{R}$ such that, in a neighborhood U of $x = 0$, $V(x) > 0$ for all $x \neq 0, V(0) = 0$ and

$$\frac{dV}{dt} = \frac{\partial V}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial V}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial V}{\partial x_n}\frac{dx_n}{dt} = \frac{\partial V}{\partial x_1}f_1(x) + \frac{\partial V}{\partial x_2}f_2(x) + \cdots + \frac{\partial V}{\partial x_n}f_n(x) =$$

$$= \nabla V \cdot f(x) < 0 \text{ for all } x \neq 0.$$

It is easy to prove that this local property is global for the linear system $\dot{x} = Ax$. In other words, $x(t) \to 0$ as $t \to \infty$. A Lyapunov candidate $V(x)$ for this system could be $V(x) = x^T P x$ where P is a symmetric positive definite matrix. Clearly $V(0) = 0$ and $V(x) > 0$ for all $x \neq 0$. Next we derive the condition for P to fullfil the last condtion $\dot{V} = \frac{dV}{dt} < 0$ (See [12] and see page 218-233 in [5]).

The derivative of *V* is:

$$\dot{V} = \dot{x}^T P x + x^T P \dot{x}$$

Since $\dot{x} = Ax$ we have

$$\dot{V} = (Ax)^T P x + x^T P (Ax) = x^T A^T P x + x^T P A x =$$
$$= x^T (A^T P x + P A x) = x^T\left((A^T P + PA)x\right) = x^T (A^T P + PA) x$$

Hence $\dot{V} < 0$ if there exists a positive definite matrix Q and P satisfies the matrix equation

$$A^T P + PA = -Q$$

because

$$x^T (A^T P + PA) x = -x^T Q x < 0$$

since Q is positive definite.

Now we can see an expression, $A^T P + PA = -Q$ which is the Lyapunov equation!

4.2 Solution of Lyapunov equation using Kronecker product

There are a few ways to find the solution P of the Lyapunov equation. In this subsection we make use of Kronecker product. Vectorize the equation in the following steps.

We rewrite it in Vec-operator as

$$vec(A^TP + PA) = -vec(Q) \Leftrightarrow vec(A^TPI) + vec(IPA) = -vec(Q) \Leftrightarrow$$
$$\Leftrightarrow (I \otimes A^T)vec(P) + (A^T \otimes I)vec(P) = -vec(Q) \Leftrightarrow$$
$$\Leftrightarrow (A^T \oplus A^T)vec(P) = -vec(Q).$$

Our expression $(A^T \oplus A^T)vec(P) = -vec(Q)$ looks like a familiar one which is $Ax = b$. (a linear equation). One way for us to solve such expression is by taking the left inverse of A and get

$$Ax = b \Leftrightarrow A^{-1}Ax = A^{-1}b \Leftrightarrow x = A^{-1}b.$$

To implement this on our original expression we would get:

$$(A^T \oplus A^T)vec(P) = -vec(Q) \Leftrightarrow vec(P) = -(A^T \oplus A^T)^{-1}vec(Q)$$

if $(A^T \oplus A^T)$ is invertible which we'll prove now.

To solve that kind of equation on our original expression we would begin by finding the eigenvalue of $(A^T \oplus A^T)$ and the eigenvectors. By **Theorem 3.3**,
the eigenvalue of $(A^T \oplus A^T)$ is $\lambda_i + \lambda_j$ and eigenvectors $v_i \otimes v_j$. Note that $\lambda_i + \lambda_j \neq 0$, for $i \neq j \Leftrightarrow \det(A^T \oplus A^T) \neq 0$ will show that the solution is unique (**Theorem 2.1**).
In other words, A^T and $-A^T$, or equivalently A and $-A$, have no common eigenvalues.
However this does not guarantee that the solution is positive definite.

4.3 Positive definite solution of Lyapunov equation

In order to find the unique positive definite solution we will show that there is a closed form of P which is

$$P = \int_0^\infty e^{tA^T}Qe^{tA}dt.$$

It is clearly well defined if $Re\big(\lambda_i(A)\big) < 0$ for all $i = 1, \dots, n$.
Moreover P satisfies the Lyapunov equations because

$$A^T P + PA = \int_0^\infty \left(A^T e^{tA^T} Q e^{tA} + e^{tA^T} Q e^{tA} A\right) dt =$$

$$\int_0^\infty \left(\frac{d}{dt} e^{tA^T} Q e^{tA}\right) dt = \left[e^{tA^T} Q e^{tA}\right]_0^\infty = \lim_{R\to\infty} e^{RA^T} Q e^{RA} - Q =$$

$$= [By\ having\ Re(A) < 0] = 0 - Q = -Q.$$

Since Q is positive definite, there is an invertible matrix C (e.g. a Cholesky factor) such that $Q = C^T C$. Next we show that P defined above is positive semi-definite. It is obvious that for any vector z,

$$z^T P z = z^T \left(\int_0^\infty e^{tA^T} Q e^{tA} dt\right) z = \int_0^\infty \left(z^T e^{tA^T} C^T\right)(C e^{tA} z) dt = \int_0^\infty \|C e^{tA} z\|^2 dt \geq 0$$

So P is positive semidefinite. Finally we see that P is positive definite since C and e^{tA^T} are invertible. Note that $Re(\lambda_i(A)) < 0$ for all $i = 1, \dots, n$ guarantees that A and $-A$ have no common eigenvalues which is required in the previous section.

4.4 Uniqueness of solutions

It remains to show that P obtained in Section 4.2 and 4.3 are the same. In this text we take an alternative approach (which we couldn't find in the literature).

We have already shown that the vectorized Lyapunov equation has a unique solution if $\lambda_i + \lambda_j \neq 0$, for $i \neq j$ which holds since now we have $Re(\lambda_i(A)) < 0$ for all $i \neq j$. Hence we have to prove that

The LHS is

$$\int_0^\infty vec\left(e^{tA^T} Q e^{tA}\right) dt = \int_0^\infty \left(e^{tA^T} \otimes e^{tA^T}\right) dt\, vec(Q)$$

Thus it remains to show that

$$\int_0^\infty \left(e^{tA^T} \otimes e^{tA^T}\right) dt = -(A^T \oplus A^T)^{-1}$$

or equivalently

$$(A^T \oplus A^T) \int_0^\infty \left(e^{tA^T} \otimes e^{tA^T}\right) dt = -I.$$

We get

$$(A^T \oplus A^T)\int_0^\infty \left(e^{tA^T} \otimes e^{tA^T}\right)dt = \int_0^\infty (A^T \otimes I + I \otimes A^T)\left(e^{tA^T} \otimes e^{tA^T}\right)dt.$$

Now let $X = A^T, Y = I, V = e^{tA^T}$, our expression will now look,

$$\int_0^\infty (X \otimes Y + Y \otimes X)(V \otimes V)dt.$$

We will substitute again where $N = X \otimes Y$, $M = Y \otimes X$ and $R = (V \otimes V)$ and have

$$\int_0^\infty (X \otimes Y + Y \otimes X)(V \otimes V)dt = \int_0^\infty (N + M)R\,dt = \int_0^\infty (NR + MR)dt =$$

$$= \int_0^\infty NR\,dt + \int_0^\infty MR\,dt.$$

which look easier. Let us now substitute back and will get

$$\int_0^\infty NR\,dt + \int_0^\infty MR\,dt = \int_0^\infty (X \otimes Y)(V \otimes V)\,dt + \int_0^\infty (Y \otimes X)(V \otimes V)\,dt =$$

$$= \int_0^\infty (X \otimes Y)(V \otimes V)\,dt + \int_0^\infty (Y \otimes X)(V \otimes V)\,dt = [\boldsymbol{KP8})] =$$

$$= \int_0^\infty (XV \otimes YV)\,dt + \int_0^\infty (YV \otimes XV)\,dt = \int_0^\infty \left(A^T e^{tA^T} \otimes I e^{tA^T}\right)dt + \int_0^\infty \left(I e^{tA^T} \otimes A^T e^{tA^T}\right)dt =$$

$$= \int_0^\infty \left(\left(A^T e^{tA^T}\right) \otimes e^{tA^T}\right)\,dt + \int_0^\infty \left(e^{tA^T} \otimes \left(A^T e^{tA^T}\right)\right)\,dt.$$

.

Since $A^T e^{tA^T} = \frac{d}{dt}\left(e^{tA^T}\right)$ we use partial integration to evaluate the integral i.e.

$$\int_0^\infty \left(\left(A^T e^{tA^T}\right) \otimes e^{tA^T}\right)\,dt = \int_0^\infty \left(\frac{d}{dt}\left(e^{tA^T}\right) \otimes e^{tA^T}\right)\,dt =$$

$$= \left[e^{tA^T} \otimes e^{tA^T}\right]_0^\infty - \int_0^\infty \left(e^{tA^T} \otimes \frac{d}{dt}\left(e^{tA^T}\right)\right)dt + \int_0^\infty \left(e^{tA^T} \otimes \frac{d}{dt}\left(e^{tA^T}\right)\right)dt =$$

$$= \left[e^{tA^T} \otimes e^{tA^T}\right]_0^\infty = -I,$$

where all eigenvalue of A have negative real parts. ■

5. Least square problem

In many applications we have to solve the following minimization problem

$$\min_x \|(B \otimes C)x - b\|$$

where $\|.\|$ is the 2-norm (Euclidean norm), can be efficiently solved by computing the QR factorizations (or SVDs) of B and C. Barrlund's book *Efficient solution of constrained least square problems with Kronecker product structure (1998, SIAM), page 154-160* shows how to minimize $\|(A_1 \otimes A_2)x - f\|$ subject to the constraint that $(B_1 \otimes B_2)x = g$, a problem that comes up in surface fitting with certain kinds of splines. [13]

5.1 Least square – primer

First we give a short review on how to solve minimization problem

$$\min_x \|Ax - b\|^2,$$

where $A \in \mathbb{M}_{m,n}$ and $m > n$ and $rank(A) = n$.

5.1.1 Normal equation

Using standar calculus the necessary condition for the minimum is

$$\text{grad}(\|b - Ax\|^2) = 2(A^T Ax - A^T b) = 0.$$

The last equation is equivalent to

$$A^T Ax = A^T b \ (normal\ equation),$$

Since $A^T A$ is positive definite the function $\|Ax - b\|^2$ is convex and the normal equation has unique solution. Therefore this unique solution is the global minimum.
In other words to solve the optimization problem above is equivalent to find the solution to the over-determined system $Ax = b$,which can be done as follows

$$A^T Ax = A^T b \Leftrightarrow x = (A^T A)^{-1} A^T b$$

The expression $(A^T A)^{-1} A^T = A^\dagger$ is called (Moore-Penrose) pseudoinverse.
It would now look like $x = A^\dagger\, b$.

In practise we don't invert matrix but we use *Cholesky-decomposition*

$$A^T A = R^T R$$

where *R* is overtriangular matrix.

5.1.2 QR-decomposition

Another alternative is the use of *QR-decomposition*:

$$A = QR = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1$$

where $Q \in \mathbb{M}_m$ is orthogonal ($Q^T Q = I$) and *R* is overtriangular.
The columns in $Q_1 \in \mathbb{M}_{m,n}$ form an *ON-basis* for the range space $\mathcal{V}(A)$ and the columns in Q_2 spans the orthogonal complement. The decomposition $A = Q_1 R_1$ is called sometimes "economy" *QR – decomposition*.

Let $r := b - Ax$, be called the residue, then

$$\|r\|_2^2 = \|QQ^T r\|_2^2 = \|Q^T r\|_2^2 = \left\| \begin{pmatrix} R_1 \\ 0 \end{pmatrix} x - Q^T b \right\|_2^2 = \|R_1 x - Q_1^T b\|_2^2 + \|Q_2^T b\|_2^2.$$

We can't do so much with the term $\|Q_2^T b\|_2^2$ but we can choose a proper value of x such that it affects $\|R_1 x - Q_1^T b\|_2^2$. Because we minimize $\|r\|$ we choose x such that we can solve the equation

$$R_1 x - Q_1^T b = 0 \Leftrightarrow x = R_1^{-1} Q_1^T b.$$

5.1.3 Singular value decomposition

A third method to solve the overdetermined system of equation is with the *Singular Value Decomposition (SVD).*

For $A \in \mathbb{M}_{m,n}$, we have

$$A = V \Sigma W^T = \begin{pmatrix} V_1 & V_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} W^T = V_1 \Sigma_1 W^T,$$

where $V \in \mathbb{M}_m$and $W \in \mathbb{M}_n$ is orthogonal matrices and $\Sigma \in \mathbb{M}_{m,n}$ is diagonal with the diagonalelement

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0 \; (singular\ value\ of\ A)$$

$V_1 \in \mathbb{M}_{m,n} \Sigma_1 \in \mathbb{M}_n$ and $W \in \mathbb{M}_n$.

Let the columns in V be $v_1, \dots, v_m$ and in W be $w_1, \dots, w_m$. Then we have

$$\|Ax - b\|_2^2 = \|(VV^T)(A(WW^T)x - b)\|_2^2 = \|V^T(AWW^Tx - b)\|_2^2 =$$

$$= \|V^TAWW^Tx - V^Tb\|_2^2 = \|(V^TAW)W^Tx - V^Tb\|_2^2 = \|\Sigma W^Tx - V^Tb\|_2^2 =$$

$$= \|\Sigma z - U^Tb\|_2^2 = \sum_{i=1}^{n}(\sigma_i z_i - u_i^T b)^2 + \sum_{i=n+1}^{m}(u_i^T b)^2.$$

The second term on the last line is known and we can't do anything about. But the first one we can choose such that it can be 0. Then we get a solution for z and therefore $x = Vz$. We will notice that 2-norm of z and x are the same:

$$z_i = \frac{u_i^T b}{\sigma_i}, \quad i = 1, \dots, r \text{ and } z_i \text{ arbitrarily for } i = n+1, \dots, m$$

and

$$\min_x \|Ax - b\|_2^2 = \sum_{i=n+1}^{m}(u_i^T b)^2.$$

5.2 Least Square - Kronecker Product

Now we solve the problem $\min\|(B \otimes C) - b\|$. Assuming $B \otimes C$ has full column rank, i.e B and C have full column ranks. Since $rank(B \otimes C)$ is equal to $rank(B)rank(C)$.

Alternative 1 (Using normal equation):
We need to solve

$$(B \otimes C)^T(B \otimes C)x = (B \otimes C)^T b \Leftrightarrow \big((B^TB) \otimes (C^TC)\big)x = (B^T \otimes C^T)b.$$

Since B, C have full column ranks, B^TB and C^TC is invertible. Thus $\big((B^TB) \otimes (C^TC)\big)$ is invertible. Then

$$x = \big((B^TB) \otimes (C^TC)\big)^{-1}(B^T \otimes C^T)b \Leftrightarrow [\boldsymbol{KP}\mathbf{13})] \Leftrightarrow$$
$$\Leftrightarrow ((B^TB)^{-1} \otimes (C^TC)^{-1})(B^T \otimes C^T)b \Leftrightarrow [\boldsymbol{KP}\mathbf{8})] \Leftrightarrow ((B^TB)^{-1}B^T) \otimes ((C^TC)^{-1}C^T)b.$$

From this we see that we only have to find the inverse for the matrices B and C, respectively. That is, it is not necessary to invert the big matrix $(B \otimes C)^T(B \otimes C)$.

Alternative 2 (using QR):
Assume that we have the economy QR-factorization for B and C:

$$B = Q_B R_B, \qquad C = Q_C R_C$$

where R_B and R_C are triangular. Then

$$B \otimes C = (Q_B R_B) \otimes (Q_C R_C) = (Q_B \otimes Q_C)(R_B \otimes R_C)$$

According to the solution in the previous subsection we have

$$x = (R_B \otimes R_C)^{-1}(Q_B \otimes Q_C)^T b \Leftrightarrow [\boldsymbol{KP13}) \; and \; \boldsymbol{KP6})] \Leftrightarrow (R_B^{-1} \otimes R_C^{-1})(Q_B^T \otimes Q_C^T)b \Leftrightarrow$$
$$\Leftrightarrow [\boldsymbol{KP8})] \Leftrightarrow \left(\left(R_B^{-1} Q_B^T\right) \otimes \left(R_C^{-1} Q_C^T\right)\right) b.$$

Again we only have to carry out matrix computaions for the smaller matrices B and C.

Alternative 3 (using SVD):
Let $B = U_B \Sigma_B V_B^T$, $C = U_C \Sigma_C V_C^T$, U_B, V_B and U_C, V_C have orthonormal columns and the factorizations are in economy form. Then

$$B \otimes C = (U_B \Sigma_B V_B^T) \otimes (U_C \Sigma_C V_C^T) \Leftrightarrow [\boldsymbol{KP8}) \; and \; \boldsymbol{KP6})] \Leftrightarrow (U_B \otimes U_C)(\Sigma_B \otimes \Sigma_C)(V_B \otimes V_C)^T$$

So all the calculations are boiled down to the smaller matrices B and C as follows::
$x = (V_B \otimes V_C)z$, $z = (z_1^T, z_2^T)$ with $z_1 \in \mathbb{R}^{rs}$ and

$$(\Sigma_B \otimes \Sigma_C)z_1 = (U_B \otimes U_C)^T b = [\boldsymbol{KP6})] = (U_B^T \otimes U_C^T)b \Leftrightarrow z_1 = (\Sigma_B \otimes \Sigma_C)^{-1}(U_B^T \otimes U_C^T)b =$$
$$= [\boldsymbol{KP13})] = (\Sigma_B^{-1} \otimes \Sigma_C^{-1})(U_B^T \otimes U_C^T)b = [\boldsymbol{KP8})] = \left((\Sigma_B^{-1} U_B^T) \otimes (\Sigma_C^{-1} U_C^T)\right)b$$

and z_2 arbitrary.

6. References

[1] Anders Holst, Victor Ufnarovski – Matrix Theory – LTH – 2013.
[2] Roger A. Horn, Charles R. Johnson – Matrix Analysis (2013. Cambrdge University Press).
[3] http://vergil.chemistry.gatech.edu/notes/quantrev/node13.html#subsub:BasicOps
[4] Lecture notes in Numerical Analysis, Department of Mathematics, Stockholm University.
[5] E. Sontag, Mathematical Control Theory, Deterministic finite dimensional systems, 2nd edition, Springer.
[6] Lecture notes in Dynamical system and control theory, Department of Mathematics, Stockholm University.
[7] Stephen A. Wirkus, Rabdall J. Swift – A course in Ordinary differential equations – 2nd edition, CRC Press.
[8] K. Kowalski, W. Hans Steeb - Nonlinear dynamical systems and Carleman linearization – (1991, World Scientific).
[9] K. Schäcke – On the Kronecker Product – 2013-08-01 – University of Waterloo
[10] https://www.ime.unicamp.br/~cnaber/Kronecker.pdf
[11] Roger A. Horn, Charles R. Johnson - Topics in matrix analysis (1991, Cambridge University Press).
[12] https://en.wikipedia.org/wiki/Lyapunov_stability
[13] Charles F. Van Loan –Journal of computational and Applied Mathematics 123 (2000) - The ubiquitous Kronecker product – Cornell University.