# Study Notes of Matrix and Tensor

Pei Zhong

March 15, 2024

# Contents

# Preface

The notes mainly refer to:

- From Algebraic Structures to Tensors
- Matrix and Tensor Decompositions in Signal Processing

# Chapter 1

# Matrix Algebra

## 1.1 Notations and definitions

Scalars, column vectors, matrices, and hypermatrices/tensors of order higher than two will be denoted by lowercase letters $(a, b, \ldots)$, bold lowercase letters $(\mathbf{a}, \mathbf{b}, \ldots)$, bold uppercase letters $(\mathbf{A}, \mathbf{B}, \ldots)$, and calligraphic letters $(\mathcal{A}, \mathcal{B}, \ldots)$ respectively.

A matrix $\mathbf{A}$ of dimensions $I \times J$, with $I$ and $J \in \mathbb{N}^*$, denoted by $\mathbf{A}(I, J)$, is an array of $IJ$ elements stored in $I$ rows and $J$ columns; the elements belong to a field $\mathbb{K}$. Its $i$th row and $j$th column, denoted by $A_{i.}$ and $A_{.j}$, respectively, are called $i$th row vector and $j$th column vector. The element located at the intersection of $A_{i.}$ and $A_{.j}$ is designated by $a_{ij}$. We will use the notation $\mathbf{A} = (a_{ij})$, with $a_{ij} \in \mathbb{K}$, $i \in \langle I \rangle = \{1, 2, \ldots, I\}$ and $j \in \langle J \rangle = \{1, 2, \ldots, J\}$.

A matrix $A \in \mathbb{K}^{I \times J}$ is written in the form:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1J} \\ a_{12} & a_{22} & \ldots & a_{2J} \\ \ldots & \ldots & \ldots & \ldots \\ a_{I1} & a_{I2} & \ldots & a_{IJ} \end{pmatrix}$$

The special cases $I = 1$ and $J = 1$ correspond respectively to row vectors of dimension $J$ and to column vectors of dimension $I$:

$$\mathbf{v} = \begin{pmatrix} v_1 & v_2 & \ldots & v_J \end{pmatrix} \in \mathbb{K}^{1 \times J}, \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_I \end{pmatrix} \in \mathbb{K}^{I \times 1}.$$

In the following, for column vectors, $\mathbb{K}^I$ will be used instead of $\mathbb{K}^{I \times 1}$.

$\mathbf{e}_i^{(I)}$ is the column vector of dimension $I$, in which element is equal to $1$ at position $i$ and $0s$ elsewhere. $\mathbf{E}_{ij}^{I \times J}$ is the matrix of dimension $I \times J$ in which element is equal to $1$ at position $(i, j)$ and $0s$ elsewhere.

## 1.2 Transposition and conjugate transposition

**Definition 1.1: transpose and the conjugate transpose of a column vector**

The transpose and the conjugate transpose (also called transconjugate) of a column vector $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_I \end{pmatrix} \in \mathbb{C}^I$, denoted by $\mathbf{u}^T$ and $\mathbf{u}^H$, respectively, are the row vectors defined as:

$$\mathbf{u}^T = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_I \end{pmatrix} \text{ and } \mathbf{u}^H = \begin{pmatrix} u_1^* & u_2^* & \dots & u_I^* \end{pmatrix},$$

where $u_i^*$ is the conjugate of $u_i$ also denoted by $\overline{u}_i$.

**Definition 1.2: transpose and the conjugate transpose of a matrix**

The transpose of $\mathbf{A} \in \mathbb{K}^{I\times J}$ is the matrix denoted by $\mathbf{A}^T$, of dimensions $J \times I$, such that $A^T = (a_{ji})$, with $i \in \langle I \rangle$ and $j \in \langle J \rangle$. In the case of a complex matrix, the conjugate transpose, also known as Hermitian transpose and denoted by $\mathbf{A}^H$, is defined as: $\mathbf{A}^H = (\mathbf{A}^*)^T = (\mathbf{A}^T)^* = (a_{ji}^*)$, where $\mathbf{A}^* = (a_{ij}^*)$ is the conjugate of $\mathbf{A}$.

**<u>Remark.</u>** In mathematics, the complex conjugate of a complex number is the number with an equal real part and an imaginary part equal in magnitude but opposite in sign. That is, if $a$ and $b$ are real numbers then the complex conjugate of $a + ib$ is $a - ib$. The complex conjugate of $z$ is often denoted as $\bar{z}$ or $z^*$.

**Proposition 1.1**

The operations of transposition and conjugate transposition satisfy:

$$(\mathbf{A}^T)^T = \mathbf{A}, (\mathbf{A}^H)^H = \mathbf{A}$$
$$(\mathbf{A}+\mathbf{B}) = \mathbf{A}^T + \mathbf{B}^T, (\mathbf{A}+\mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H,$$
$$(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T, (\alpha\mathbf{A})^H = \alpha^*\mathbf{A}^H,$$

for any matrix $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{I\times J}$ and any scalar $\alpha \in \mathbb{C}$.

**<u>Remark.</u>** By decomposing $\mathbf{A}$ using its real and imaginary parts, we have:

$$\mathbf{A} = \mathrm{Re}(\mathbf{A}) + \mathrm{i}\mathrm{Im}(\mathbf{A}) \Rightarrow \begin{cases} \mathbf{A}^T = (\mathrm{Re}(\mathbf{A}))^T + \mathrm{i}(\mathrm{Im}(A))^T \\ \mathbf{A}^H = (\mathrm{Re}(\mathbf{A}))^H - \mathrm{i}(\mathrm{Im}(A))^H \end{cases}$$

.

# 1.3 Vector outer product and vectorization

## 1.3.1 Vector outer product

The outer product of two vectors $\mathbf{u} \in \mathbb{K}^I$ and $\mathbf{v} \in \mathbb{K}^J$, denoted $\mathbf{u} \circ \mathbf{v}$, gives a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ such that $a_{ij} = (\mathbf{u} \circ \mathbf{v})_{ij} = u_i v_j$, and therefore, $\mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^T = (u_i v_j)$, with $i \in \langle I \rangle, j \in \langle J \rangle$.

**Example 1.1**

For $I = 2, J = 3$, we have:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \circ \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} = \begin{pmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \end{pmatrix}.$$

## 1.3.2 Vectorization

A very widely used operation in matrix computation is vectorization which consists of stacking the columns of a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ on top of each other to form a column vector of dimension $JI$:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{.1} & \mathbf{A}_{.2} & \dots & \mathbf{A}_{.J} \end{pmatrix} \in \mathbb{K}^{I \times J} \Rightarrow \text{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{A}_{.1} \\ \mathbf{A}_{.2} \\ \dots \\ \mathbf{A}_{.\mathbf{J}} \end{pmatrix} \in \mathbb{K}^{JI}.$$

This operation defines an isomorphism between the space $\mathbb{K}^{JI}$ of vectors of dimension $JI$ and the space $\mathbb{K}^{I \times J}$ of matrices $I \times J$. Indeed, the canonical basis of $\mathbb{K}^{JI}$, denoted by $\{\mathbf{e}^{(JI)}_{(j-1)I+i}\}$, allows us to write vec$(\mathbf{A})$ as:

$$\mathbf{A} = \sum_{i=1}^{I} \sum_{j=1}^{J} a_{ij} \mathbf{e}_i^{(I)} \circ \mathbf{e}_j^{(J)} \Rightarrow \text{vec}(\mathbf{A}) = \sum_{i=1}^{I} \sum_{j=1}^{J} a_{ij} \mathbf{e}^{(JI)}_{(j-1)I+i},$$

with $\mathbf{e}^{(JI)}_{(j-1)I+i} = \text{vec}(\mathbf{e}_i^{(I)} \circ \mathbf{e}_j^{(J)}) = \text{vec}(\mathbf{e}_i^{(I)} (\mathbf{e}_j^{(J)})^T)$.

**Remark.** Since the operator vec satisfies $\text{vec}(\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha\text{vec}(\mathbf{A}) + \beta\text{vec}(\mathbf{B})$ for all $\alpha, \beta \in \mathbb{K}$, it is linear.

# 1.4 Vector inner product, norm and orthogonality

## 1.4.1 Inner product

In this section, we recall the definition of the inner product(also called dot product) of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{K}^I$.

Definition 1.3

If $\mathbb{K} = \mathbb{R}$, the inner product is defined as:

$$\langle \cdot, \cdot \rangle : \mathbb{R}^I \times \mathbb{R}^I \to \mathbb{R}$$
$$(a, b) \mapsto \langle a, b \rangle = a^T b = \sum_{i=1}^{I} a_i b_i.$$

In $\mathbb{C}^I$, the definition of the inner product is given by:

$$\langle \cdot, \cdot \rangle : \mathbb{C}^I \times \mathbb{C}^I \to \mathbb{C}$$
$$(a, b) \mapsto \langle a, b \rangle = a^H b = \sum_{i=1}^{I} a_i^* b_i.$$

### 1.4.2 Euclidean/Hermitian norm

Definition 1.4

The Euclidean (Hermitian) norm of a vector $\mathbf{a}$, denoted $||\mathbf{a}||$, associates to $a \in \mathbb{R}^I$ ($a \in \mathbb{C}^I$) a non-negative real number according to the following definition:

$$|| \cdot ||_2 : \mathbb{K}^I \to \mathbb{R}^+$$
$$\mathbf{a} \mapsto ||\mathbf{a}||_2 = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$$

.

### 1.4.3 Orthogonality

Definition 1.5

Two vectors $\mathbf{a}$ and $\mathbf{b}$ of $\mathbb{K}^I$ are said to be orthogonal if and only if $\langle a, b \rangle = 0$.

## 1.5 Matrix multiplication

### 1.5.1 Definition and properties

given matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{J \times K}$, the product of $\mathbf{A}$ by $\mathbf{B}$ gives a matrix $\mathbf{C} = \mathbf{AB} \in \mathbb{K}^{I \times K}$ such that $c_{ik} = \sum_{i=1}^{I} a_{ij} b_{jk}$, for $j \in \langle J \rangle$; $k \in \langle K \rangle$.

This product can be written in terms of the outer products of column vectors of

## 1.6 Matrix trace, inner product and Frobenius norm

### 1.6.1 Definition and properties of the trace

Definition 1.6

The trace of a square matrix $\mathbf{A}$ of order $I$ is defined as the sum of its diagonal elements:

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{I} a_{ii}.$$

Proposition 1.2

The trace satisfies the following properties:

$$\begin{aligned} \text{tr}(\alpha\mathbf{A} + \beta\mathbf{B}) &= \alpha\text{tr}(\mathbf{A}) + \beta\text{tr}(\mathbf{B}), \\ \text{tr}(\mathbf{A}^T) &= \text{tr}(\mathbf{A}), \\ \text{tr}(\mathbf{A}^*) &= \text{tr}(\mathbf{A}^H) = (\text{tr}(\mathbf{A}))^*, \end{aligned}$$

## 1.7 Subspaces associated with a matrix

## 1.8 Matrix rank

## 1.9 Determinant, inverses and generalized inverses

## 1.10 Eigenvalues and eigenvectors

## 1.11 Reference

- From Algebraic Structures to Tensors ch4 matrix algebra

# Chapter 2

# Matrix Decompositions

# Chapter 3

# Hadamard, Kronecker and Khatri–Rao Products

## 3.1 Partitioned matrices

Let $\{\alpha_{m_1}, \ldots, \alpha_{m_R}\}$ and $\{\beta_{n_1}, \ldots, \beta_{n_S}\}$ be partitions of the sets $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$, respectively, with $m_r \in \langle m \rangle$ and $n_s \in \langle n \rangle$, such that $\sum_{r=1}^{R} m_r = m$ and $\sum_{s=1}^{S} n_s = n$. It is said that matrices $\mathbf{A}_{rs}$ of dimensions $(m_r, n_s)$ form a partition of the matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ into $(R, S)$ blocks, or that $\mathbf{A}$ is partitioned into $(R, S)$ blocks, if $\mathbf{A}$ can be written as:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \ldots & \mathbf{A}_{1S} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \ldots & \mathbf{A}_{2S} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{A}_{R1} & \mathbf{A}_{R2} & \ldots & \mathbf{A}_{RS} \end{pmatrix} = (\mathbf{A}_{rs}), r \in \langle R \rangle, s \in \langle S \rangle.$$

All submatrices of the same row-block $(r)$ contain the same number $(m_r)$ of rows. Similarly, all submatrices of the same column-block $(s)$ contain the same number $(n_s)$ of columns, that is:

$$\begin{pmatrix} \mathbf{A}_{r1} & \mathbf{A}_{r2} & \ldots & \mathbf{A}_{rS} \end{pmatrix} \in \mathbb{K}^{m_r \times n}, \begin{pmatrix} \mathbf{A}_{1s} & \mathbf{A}_{2s} & \ldots & \mathbf{A}_{Rs} \end{pmatrix} \in \mathbb{K}^{m \times n_s}.$$

It is then said that the submatrices $\mathbf{A}_{rs}$ are of compatible dimensions.

In the particular case where $n = 1$, the partitioned matrix becomes a block-column vector:

$$\mathbf{a} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_R \end{pmatrix} \in \mathbb{K}^{m \times 1}, \mathbf{a}_r \in \mathbb{K}^{m_r \times 1}, r \in \langle R \rangle.$$

Similarly, when $m = 1$, the partitioned matrix becomes a block-row vector:

$$\mathbf{a}^T = \begin{pmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T & \ldots & \mathbf{a}_S^T \end{pmatrix} \in \mathbb{K}^{1 \times n}, \mathbf{a}_s \in \mathbb{K}^{n_s \times 1}, s \in \langle S \rangle.$$

## 3.2 Notation

We will write $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$, $\mathbf{A}_{i\cdot}$, $\mathbf{A}_{\cdot j}$, $r(\mathbf{A})$ and $\det(\mathbf{A})$, for the conjugate, the transpose, the transconjugate (also known as conjugate transpose or Hermitian transpose), the Moore-Penrose pseudo-inverse, the $i$th row, the $j$th column, the rank and the determinant of $\mathbf{A} \in \mathbb{K}^{I\times J}$, respectively. In the following literature, the Hadamard product is denoted by $*$. The symbols $\otimes$ and $\odot$ are used for the Kronecker and Khatri-Rao product, respectively.

The symbol $\mathbf{1}_I$ denotes a column vector of size $I$ whose elements are all equal to $1$. The elements of the matrices $\mathbf{0}_{I\times J}$ and $\mathbf{1}_{I\times J}$ of size $(I \times J)$ are all equal to $0$ and $1$, respectively. The symbols $\mathbf{I}_N$ and $\mathbf{e}_n^{(N)}$ denote the identity matrix of order $N$ and the $n$th vector of the canonical basis of the vector space $\mathbb{R}^N$, respectively.

## 3.3 Hadamard product

### 3.3.1 Definition and identities

**Definition 3.1**

Let $\mathbf{A}$ and $\mathbf{B} \in \mathbb{K}^{I\times J}$ be two matrices of the same size. The Hadamard product of $\mathbf{A}$ and $\mathbf{B}$ is the matrix $\mathbf{C} \in \mathbb{K}^{I\times J}$ defined as follows:

$$\mathbf{C} = \mathbf{A} * \mathbf{B} = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2J}b_{2J} \\ \dots & \dots & \dots & \dots \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \dots & a_{IJ}b_{IJ} \end{pmatrix},$$

i.e. $c_{ij} = a_{ij}b_{ij}$, and therefore $\mathbf{C} = (a_{ij}b_{ij})$, with $i \in \langle I \rangle$, $j \in \langle J \rangle$.

## 3.4 Kronecker product

### 3.4.1 Kronecker product of vectors

#### Definition

Definition 3.2

(a) For $\mathbf{u} \in \mathbb{K}^I$ and $\mathbf{v} \in \mathbb{K}^J$, we have:

$$\mathbf{x} = \mathbf{u} \otimes \mathbf{v} \in \mathbb{K}^{IJ} \Leftrightarrow x_{j+(i-1)J} = u_i v_j$$

or equivalently:

$$\mathbf{u} \otimes \mathbf{v} = \begin{pmatrix} u_1 v_1 & \dots & u_1 v_J & u_2 v_1 & \dots & u_2 v_J & \dots & u_I v_1 & \dots & u_I v_J \end{pmatrix}^T = \begin{pmatrix} u_1 \mathbf{v} & \dots & u_I \mathbf{v} \end{pmatrix}.$$

(b) Similarly, for $\mathbf{u} \in \mathbb{K}^I$, $\mathbf{v} \in \mathbb{K}^J$, and $\mathbf{w} \in \mathbb{K}^K$, we have:

$$\mathbf{x} = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} \in \mathbb{K}^{IJK} \Leftrightarrow x_{k+(j-1)K+(i-1)JK} = u_i v_j w_k.$$

**Remark.** By convention, the order of the dimensions in a product $IJK$ follows the order of variation of the corresponding indices $(i, j, k)$. For example, $\mathbb{K}^{IJK}$ means that the index $i$ varies more slowly than $j$, which itself varies more slowly than $k$.

### 3.4.2 Kronecker product of matrices

#### Definitions and identities

Definition 3.3

Given two matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{M \times N}$ of arbitrary size, the right Kronecker product of $\mathbf{A}$ by $\mathbf{B}$ is the matrix $\mathbf{C} \in \mathbb{K}^{IM \times JN}$ defined as follows:

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2J}\mathbf{B} \\ \dots & \dots & \dots & \dots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \dots & a_{IJ}\mathbf{B} \end{pmatrix} = (a_{ij}\mathbf{B}).$$

**Remark.** The Kronecker product is a matrix partitioned into $(I, J)$ blocks, where the block $(i, j)$ is given by the matrix $a_{ij}\mathbf{B} \in \mathbb{K}^{M \times N}$. The element $a_{ij} b_{mn}$ is located at the position $((i-1)M + m, (j-1)N + n)$ in $\mathbf{A} \otimes \mathbf{B}$.

## 3.5 Reference

- From Algebraic Structures to Tensors ch5
- Matrix and Tensor Decompositions in Signal Processing ch2

# Chapter 4

# Tensor Operations

## 4.1 Notation

Let $\chi$

## 4.2 Notion of slice

## 4.3 Mode combination

## 4.4 Matricization

## 4.5 Multiplication operations