Study Notes of Matrix and Tensor

Pei Zhong

April 1, 2024

Contents

1 Matrix Algebra **3**

1.1 Notations and definitions . . . 3

1.2 Transposition and conjugate transposition . . . 4

1.3 Vector outer product and vectorization . . . 5

1.3.1 Vector outer product . . . 5

1.3.2 Vectorization . . . 5

1.4 Vector inner product and orthogonality . . . 5

1.4.1 Inner product . . . 5

1.4.2 Orthogonality . . . 6

1.5 Vector Norms . . . 6

1.5.1 Vector 2-norm . . . 6

1.5.2 Vector 1-norm . . . 7

1.5.3 Vector ∞-norm . . . 7

1.5.4 Vector p-norm . . . 7

1.6 Matrix Norms . . . 8

1.6.1 Frobenius norm . . . 8

1.6.2 Induced matrix norms . . . 8

1.7 Matrix multiplication . . . 9

1.8 Matrix trace, Matrix inner product . . . 10

1.8.1 Definition and properties of the trace . . . 10

1.9 Eigenvalues and eigenvectors . . . 10

1.10 Generalized inverses . . . 14

1.11 Reference . . . 14

2 Matrix Decompositions **16**

2.1 SVD . . . 16

2.1.1 Singular value decomposition (SVD) . . . 16

2.1.2 Geometric interpretation of SVD . . . 18

2.1.3 Applications of Singular-Value Decomposition . . . 18

2.2 Reference 18

3 Hadamard, Kronecker and Khatri-Rao Products 19

3.1 Partitioned matrices 19
3.2 Notation 20
3.3 Hadamard product 20
3.3.1 Definition and identities 20
3.4 Kronecker product 21
3.4.1 Kronecker product of vectors 21
3.4.2 Kronecker product of matrices 21
3.5 Reference 21

4 Tensor Operations 23

4.1 Notation 23
4.2 Notion of slice 23
4.3 Mode combination 23
4.4 Matricization 23
4.5 Multiplication operations 23

5 Tensor Decomposition 24

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 ith row and jth column, denoted by $A_{i\cdot}$ and $A_{\cdot j}$, respectively, are called ith row vector and jth column vector. The element located at the intersection of $A_{i\cdot}$ and $A_{\cdot j}$ 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 $\mathbb{K}^{I \times 1}$.

Identity matrix (i.e., a matrix with all 1's on the diagonal and 0's everywhere else) is denoted by $\mathbf{E}$. $\mathbf{e}_i^{(I)}$ 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 $\mathbf{A}$.

Remark. 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 $\alpha \in \mathbb{C}$.

Remark. 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:

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

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 $\text{vec}(\mathbf{A})$ 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 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 vector inner product is defined as:

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

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

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

Remark. Whether $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, $\langle \mathbf{a}, \mathbf{a} \rangle \in \mathbb{R}$.

1.4.2 Orthogonality

Definition 1.4

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

1.5 Vector Norms

Definition 1.5

Let $v : \mathbb{C}^n \to \mathbb{R}$. Then v is a norm if for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

- $\mathbf{x} \neq 0 \Rightarrow v(\mathbf{x}) > 0$,
- $v(\alpha \mathbf{x}) = |\alpha| v(\mathbf{x})$, and
- $v(\mathbf{x} + \mathbf{y}) \leqslant v(\mathbf{x}) + v(\mathbf{y})$

Remark. often we will use $||\cdot||$ to denote a vector norm.

1.5.1 Vector 2-norm

Definition 1.6

The vector 2-norm $||\cdot||_2 : \mathbb{C}^n \to \mathbb{R}$ is defined by

$$||\mathbf{x}||_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\mathbf{x}^H \mathbf{x}} = \sqrt{x_1^* x_1 + x_2^* x_2 + \ldots + x_n^* x_n} = \sqrt{|x_1|^2 + \ldots + |x_n|^2}.$$

Theorem 1.1

Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Then $|\langle \mathbf{x}, \mathbf{y} \rangle| = |\mathbf{x}^H \mathbf{y}| \leqslant ||\mathbf{x}||_2 \cdot ||\mathbf{y}||_2$.

Proposition 1.2

The vector 2-norm is a norm.

1.5.2 Vector 1-norm

Definition 1.7

The vector 1-norm $|| \cdot ||_1 : \mathbb{C}^n \to \mathbb{R}$ is defined by

$$||\mathbf{x}||_1 = |x_1| + |x_2| + \ldots + |x_n|.$$

Proposition 1.3

The vector 1-norm is a norm.

1.5.3 Vector ∞-norm

Definition 1.8

The vector ∞-norm $|| \cdot ||_\infty : \mathbb{C}^n \to \mathbb{R}$ is defined by $||\mathbf{x}||_\infty = \max_i |x_i|$.

Proposition 1.4

The vector ∞-norm is a norm.

1.5.4 Vector p-norm

Definition 1.9

The vector p-norm $|| \cdot ||_p : \mathbb{C}^n \to \mathbb{R}$ is defined by

$$||\mathbf{x}||_p = \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_n|^p}.$$

Proposition 1.5

The vector p-norm is a norm.

1.6 Matrix Norms

It is not hard to see that vector norms are all measures of how "big" the vectors are. Similarly, we want to have measures for how "big" matrices are. We will start with one that are somewhat artificial and then move on to the important class of induced matrix norms.

1.6.1 Frobenius norm

Definition 1.10

The Frobenius norm $||\cdot||_F : \mathbb{C}^{m\times n} \to \mathbb{R}$ is defined by

$$||\mathbf{A}||_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}.$$

Remark. $||\mathbf{A}||_F = ||\text{vec}(\mathbf{A})||_2$.

Proposition 1.6

The Frobenius norm is a norm.

1.6.2 Induced matrix norms

Definition 1.11

Let $||\cdot||_\mu : \mathbb{C}^m \to \mathbb{R}$ and $||\cdot||_v : \mathbb{C}^n \to \mathbb{R}$ be vector norms. Define $||\cdot||_{\mu,v} : \mathbb{C}^{m\times n} \to \mathbb{R}$ by

$$||\mathbf{A}||_{\mu,v} = \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_\mu}{||\mathbf{x}||_v}.$$

Remark. How "big" $\mathbf{A}$ is, as measured by $||\mathbf{A}||_{\mu,v}$, is defined as the most that $\mathbf{A}$ magnifies the length of nonzero vectors, where the length of the vector $\mathbf{x}$ is measured with norm $||\cdot||_v$ and the length of the transformed vector $\mathbf{A}\mathbf{x}$ is measured with norm $||\cdot||_\mu$.

Proposition 1.7

Let $||\cdot||_\mu : \mathbb{C}^m \to \mathbb{R}$ and $||\cdot||_v : \mathbb{C}^n \to \mathbb{R}$ be vector norms.

$$\begin{aligned}||\mathbf{A}||_{\mu,v} &= \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_\mu}{||\mathbf{x}||_v} \\ &= \max_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_\mu}{||\mathbf{x}||_v} \\ &= \max_{||\mathbf{x}||_v = 1} ||\mathbf{A}\mathbf{x}||_\mu.\end{aligned}$$

Proposition 1.8

$||\cdot||_{\mu,v} : \mathbb{C}^{m\times n} \to \mathbb{R}$ is a norm.

Definition 1.12

Define $||\cdot||_p : \mathbb{C}^{m\times n} \to \mathbb{R}$ by

$$\begin{aligned} ||\mathbf{A}||_p &= \max_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_p}{||\mathbf{x}||_p} \\ &= \max_{||\mathbf{x}||_p=1} ||\mathbf{A}\mathbf{x}||_p. \end{aligned}$$

Proposition 1.9

For all $\mathbf{A} \in \mathbb{C}^{m\times n}, \mathbf{x} \in \mathbb{C}^n$,

$$||\mathbf{A}\mathbf{x}|| \leqslant ||\mathbf{A}||_p \cdot ||\mathbf{x}||_p.$$

Proof. By the definition of $||\mathbf{A}||_p$,

$$\frac{||\mathbf{A}||_p}{||\mathbf{x}||_p} \leqslant ||\mathbf{A}||_p.$$

□

Proposition 1.10

For any $\mathbf{A} \in \mathbb{C}^{m\times k}$ and $\mathbf{B} \in \mathbb{C}^{k\times n}$,

$$\begin{aligned} ||\mathbf{AB}||_p &\leqslant ||\mathbf{A}||_p||\mathbf{B}||_p \\ ||\mathbf{AB}||_F &\leqslant ||\mathbf{A}||_F||\mathbf{B}||_F. \end{aligned}$$

1.7 Matrix multiplication

Suppose $\mathbf{A} \in \mathbb{R}^{m\times r}$ and $\mathbf{B} \in \mathbb{R}^{r\times n}$, the matrix multiplication $\mathbf{C} = \mathbf{AB}$ can be viewed from three different perspectives as follows:

Dot Product Matrix Multiply. Every element c_{ij} of $\mathbf{C}$ is the dot product of row vector $\mathbf{A}_{i.}$ and column vector $\mathbf{B}_{\cdot j}$.

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1\cdot}^T \\ \dots \\ \mathbf{A}_{m\cdot}^T \end{pmatrix}, \mathbf{A}_{k\cdot} \in \mathbb{R}^r$$
$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{\cdot 1} & \dots \mathbf{B}_{\cdot n} \end{pmatrix}, \mathbf{B}_{\cdot k} \in \mathbb{R}^r$$
$$\mathbf{C} = (c_{ij}), c_{ij} = \mathbf{A}_{i\cdot}^T \mathbf{B}_{\cdot j} = \sum_{k=1}^{r} a_{ik} b_{kj}.$$

Column Combination Matrix Multiply. Every column $\mathbf{C}_{\cdot j}$ of $\mathbf{C}$ is a linear combination of column vector $\mathbf{A}_{\cdot k}$ of $\mathbf{A}$ with columns b_{kj} as the weight coefficients.

1.8 Matrix trace, Matrix inner product

1.8.1 Definition and properties of the trace

Definition 1.13

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 1.11

The trace satisfies the following properties:

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

1.9 Eigenvalues and eigenvectors

Definition 1.14

A real matrix $\mathbf{A}$ is a symmetric matrix if it equals to its own transpose, that is $\mathbf{A} = \mathbf{A}^T$.

Definition 1.15

A complex matrix $\mathbf{A}$ is a hermitian matrix if it equals to its own complex conjugate transpose, that is $\mathbf{A} = \mathbf{A}^H$.

Definition 1.16

A real matrix $\mathbf{Q}$ is an orthogonal matrix if the inverse of $\mathbf{Q}$ equals to the transpose of $\mathbf{Q}$, $\mathbf{Q}^{-1} = \mathbf{Q}^T$, that is $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = I$.

Definition 1.17

A complex matrix $\mathbf{U}$ is a unitary matrix if the inverse of $\mathbf{U}$ equals the complex conjugate transpose of $\mathbf{U}$, $\mathbf{U}^{-1} = \mathbf{U}^H$, that is $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$.

Definition 1.18

A hermitian matrix $\mathbf{Q}$ is positive semidefinite (abbreviated SPSD and denoted by $\mathbf{Q} \succeq 0$) if

$$\mathbf{x}^H\mathbf{Q}\mathbf{x} \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{C}^n.$$

Definition 1.19

A hermitian matrix $\mathbf{Q}$ is positive semidefinite (abbreviated SPD and denoted by $\mathbf{Q} \succ 0$) if

$$\mathbf{x}^H\mathbf{Q}\mathbf{x} > 0 \text{ for all } \mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq 0.$$

A number $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{M}$ if there exists a vector $\bar{\mathbf{x}} \neq 0$ such that $\mathbf{M}\bar{\mathbf{x}} = \lambda\bar{\mathbf{x}}$. $\bar{\mathbf{x}}$ is called an eigenvector of $\mathbf{M}$ (and is called an eigenvector corresponding to λ). Note that λ is an eigenvalue of $\mathbf{M}$ if and only if there exists $\bar{\mathbf{x}} \neq 0$ such that $(\mathbf{M} - \lambda\mathbf{E})\bar{\mathbf{x}} = 0$ or, equivalently, if and only if $\det(\mathbf{M} - \lambda\mathbf{E}) = 0$.

Let $g(\lambda) = \det(\mathbf{M} - \lambda\mathbf{E})$. Then $g(\lambda)$ is a polynomial of degree n, and so will have n roots that will solve the equation $g(\lambda) = \det(\mathbf{M} - \lambda\mathbf{E}) = 0$, including multiplicities. These roots are the eigenvalues of $\mathbf{M}$.

Proposition 1.12

If $\mathbf{Q}$ is a real symmetric matrix, all of its eigenvalues are real numbers.

Proposition 1.13

If $\mathbf{Q}$ is a complex hermitian matrix, all of its eigenvalues are real numbers.

Proposition 1.14

If $\mathbf{Q}$ is a hermitian matrix, its eigenvectors corresponding to different eigenvalues are orthogonal.

Proposition 1.15

If $\mathbf{Q}$ is SPD, the eigenvalues of $\mathbf{Q}$ are nonnegative.

Theorem 1.2

If $\mathbf{A}$ is a real symmetric matrix, then $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$, where $\mathbf{Q}$ is an orthonormal matrix, the columns of $\mathbf{Q}$ are an orthonormal basis of eigenvectors of $\mathbf{A}$, and $\mathbf{\Lambda}$ is a diagonal matrix of the corresponding eigenvalues of $\mathbf{A}$.

Theorem 1.3

If $\mathbf{A}$ is a complex hermitian matrix, then $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is an unitary matrix, the columns of $\mathbf{U}$ are an orthonormal basis of eigenvectors of $\mathbf{A}$, and $\mathbf{\Lambda}$ is a diagonal matrix of the corresponding eigenvalues of $\mathbf{A}$.

Definition 1.20

The square root of a matrix $\mathbf{A}$ is a matrix $\mathbf{A}^{1/2}$ such that

$$\mathbf{A} = \mathbf{A}^{1/2}\mathbf{A}^{1/2}.$$

Proposition 1.16

If $\mathbf{A}$ is SPSD, the $\mathbf{A} = \mathbf{M}^H\mathbf{M}$ for some matrix $\mathbf{M}$.

Proof. $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = (\mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^H)(\mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^H) = \mathbf{M}^H\mathbf{M}$. □

Definition 1.21

The Rayleigh quotient of the matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ at the nonzero vector $\mathbf{x} \in \mathbb{C}^n$ is the scalar

$$\frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} \in \mathbb{C}.$$

Remark. If $(\lambda, \mathbf{u})$ is an eigenpair for $\mathbf{A}$, then notice that

$$\frac{\mathbf{u}^H\mathbf{A}\mathbf{u}}{\mathbf{u}^H\mathbf{u}} = \frac{\mathbf{u}^H(\lambda\mathbf{u})}{\mathbf{u}^H\mathbf{u}} = \lambda,$$

so Rayleigh quotients generalize eigenvalues.

Remark. For Hermitian $\mathbf{A} \in \mathbb{C}^{n\times n}$, then there exists unitary matrix $\mathbf{U}$ such that $A = \mathbf{U}\Lambda\mathbf{U}^H$, for any $\mathbf{x} \neq \mathbf{0}$, it can be represented by $\mathbf{x} = \mathbf{U}\mathbf{c} = \sum_{j=1}^{n} c_j\mathbf{u}_j$, then

$$\frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} = \frac{\mathbf{c}^H\mathbf{U}^H\mathbf{U}\Lambda\mathbf{U}^H\mathbf{U}\mathbf{c}}{\mathbf{c}^H\mathbf{U}^H\mathbf{U}\mathbf{c}} = \frac{c^H\Lambda c}{c^H c}$$

The diagonal structure of Λ allows for an illuminating refinement,

$$\frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} = \frac{\lambda_1|c_1|^2 + \ldots + \lambda_n|c_n|^2}{|c_1|^2 + \ldots + |c_n|^2}.$$

As the numerator and denominator are both real, notice that the Rayleigh quotients for a Hermitian matrix is always real. We can say more: if the eigenvalues are ordered, $\lambda_1 \leqslant \ldots \leqslant \lambda_n$,

$$\frac{\mathbf{x}^H \mathbf{A}\mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \frac{\lambda_1 |c_1|^2 + \ldots + \lambda_n |c_n|^2}{|c_1|^2 + \ldots + |c_n|^2} \geqslant \frac{\lambda_1 (|c_1|^2 + \ldots + |c_n|^2)}{|c_1|^2 + \ldots + |c_n|^2} = \lambda_1,$$

and similarly,

$$\frac{\mathbf{x}^H \mathbf{A}\mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \frac{\lambda_1 |c_1|^2 + \ldots + \lambda_n |c_n|^2}{|c_1|^2 + \ldots + |c_n|^2} \leqslant \frac{\lambda_n (|c_1|^2 + \ldots + |c_n|^2)}{|c_1|^2 + \ldots + |c_n|^2} = \lambda_n.$$

Proposition 1.17

For a Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ with eigenvalues $\lambda_1 \leqslant \ldots \leqslant \lambda_n$, the Rayleigh quotient for nonzero $\mathbf{x} \in \mathbb{C}^{n \times n}$ satisfies

$$\frac{\mathbf{x}^H \mathbf{A}\mathbf{x}}{\mathbf{x}^H \mathbf{x}} \in [\lambda_1, \lambda_n].$$

Proposition 1.18

If $\mathbf{A} \in \mathbb{C}^{m \times n}$ with $m < n$ and $\mathbf{A}$ has rank m, then

$$||\mathbf{A}||_2 = \sqrt{\lambda_{\max}(\mathbf{A}^H \mathbf{A})} = \sqrt{\lambda_{\max}(\mathbf{A}\mathbf{A}^H)},$$

where $\lambda_{\max}(\mathbf{M})$ denotes the largest eigenvalue of a matrix $\mathbf{M}$.

Proof.

$$||\mathbf{A}||_2^2 = \sup_{\mathbf{x} \in \mathbb{C}^n, ||\mathbf{x}|| \neq 0} \frac{||\mathbf{A}\mathbf{x}||_2^2}{||\mathbf{x}||_2^2} = \sup_{\mathbf{x} \in \mathbb{C}^n, ||\mathbf{x}|| \neq 0} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} = \sup_{\mathbf{x} \in \mathbb{C}^n, ||\mathbf{x}|| \neq 0} \frac{\mathbf{x}\mathbf{A}^H \mathbf{A}\mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \lambda_{\max}(\mathbf{A}^H \mathbf{A}).$$

□

Proposition 1.19

If $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a hermitian matrix, then

$$||\mathbf{A}||_2 = |\lambda_{\max}(\mathbf{A})|.$$

Proof. Since $\mathbf{A}$ is a hermitian matrix, we have $\mathbf{A} = \mathbf{U}^H \Lambda \mathbf{U}$ where $\mathbf{U}$ is an unitary matrix and Λ is a diagonal matrix containing the eigenvalue of $\mathbf{A}$. Let $\mathbf{y} = \mathbf{U}\mathbf{x}$, then $||\mathbf{y}||_2 = ||\mathbf{U}\mathbf{x}||_2$. Hence

$$||\mathbf{A}||_2^2 = \max_{||\mathbf{x}||_2 = 1} ||\mathbf{A}\mathbf{x}||_2^2 = \max_{||\mathbf{x}||_2 = 1} \langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{x} \rangle = \max_{||\mathbf{x}||_2 = 1} \mathbf{x}^H \mathbf{A}^H \mathbf{A}\mathbf{x} = \max_{||\mathbf{y}||_2 = 1} \mathbf{y}^H \Lambda \mathbf{y} = (\lambda_{\max}(\mathbf{A}))^2.$$

□

Proposition 1.20

Suppose that $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a SPSD. Then the following are equivalent:

(a) $h > 0$ satisfies $||\mathbf{A}^{-1}||_2 \leqslant \frac{1}{h}$.

(b) $h > 0$ satisfies $||\mathbf{A}\mathbf{x}||_2 \geqslant h \cdot ||\mathbf{x}||_2$ for any vector $\mathbf{x}$

(c) $h > 0$ satisfies $|\lambda_i(\mathbf{A})| \geqslant h$ for every eigenvalue $\lambda_i(\mathbf{A})$ of $\mathbf{A}$, $i = 1, ..., m$.

Proof. By proposition 1.17, we have for all $\mathbf{x} \neq 0$,

$$\frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2} \in [\lambda_{\min}(\mathbf{A}), \lambda_{\max}(\mathbf{A})].$$

Firstly, we claim that

$$||\mathbf{A}^{-1}||_2 = \frac{1}{\lambda_{\min}(\mathbf{A})}.$$

In fact,

$$\begin{aligned} ||\mathbf{A}^{-1}||_2 &= \max_{\mathbf{x} \neq 0} \frac{||\mathbf{A}^{-1}\mathbf{x}||_2}{||\mathbf{x}||_2} = \max_{\mathbf{A}\mathbf{x} \neq 0} \frac{||\mathbf{A}^{-1}\mathbf{A}x||_2}{||\mathbf{A}\mathbf{x}||_2} \\ &= \max_{\mathbf{A}\mathbf{x} \neq 0} \frac{||\mathbf{x}||_2}{||\mathbf{A}\mathbf{x}||_2} = \max_{\mathbf{x} \neq 0} \frac{||\mathbf{x}||_2}{||\mathbf{A}\mathbf{x}||_2} \\ &= \frac{1}{\min_{\mathbf{x} \neq 0} \frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2}} = \frac{1}{\lambda_{\min}(\mathbf{A})}, \end{aligned}$$

the second and forth equality follows from the fact that $\mathbf{x} \neq 0$ if and only if $\mathbf{A}\mathbf{x} \neq 0$ since $\mathbf{A}$ is nonsingular.

(a) $\Rightarrow$ (b) $\quad \frac{1}{\lambda_{\min}(\mathbf{A})} = ||\mathbf{A}^{-1}||_2 \leqslant \frac{1}{h} \Rightarrow \frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2} \geqslant \lambda_{\min}(\mathbf{A}) \geqslant h, \forall \mathbf{x} \neq 0$.

(b) $\Rightarrow$ (a) $\quad ||\mathbf{A}\mathbf{x}||_2 \geqslant h \cdot ||x||_2, \forall \mathbf{x} \Rightarrow \lambda_{\min}(\mathbf{A}) \geqslant h \Rightarrow \frac{1}{\lambda_{\min}(\mathbf{A})} = ||\mathbf{A}^{-1}||_2 \leqslant \frac{1}{h}$.

(b) $\Leftrightarrow$ (c) $\quad ||\mathbf{A}\mathbf{x}||_2 \geqslant h \cdot ||x||_2, \forall \mathbf{x} \Leftrightarrow \lambda_{\min}(\mathbf{A}) \geqslant h \Leftrightarrow \lambda_i(\mathbf{A}) \geqslant h, i = 1, ..., m$.

□

1.10 Generalized inverses

1.11 Reference

- From Algebraic Structures to Tensors ch4 matrix algebra
- Notes on Vector and Matrix Norms
- symmetric, positive definted, eigenvalues and eigenvectors
- Rayleigh quotient

- The Principal Axis Theorem
- linear algebra and geometry note
- eigenvalues application: ellipses
- Generalized inverses

Chapter 2

Matrix Decompositions

2.1 SVD

We have seen that hermitian matrices are always diagonalizable. What about general rectangular matrices? The following theorem shows that one can always find two matrix $\mathbf{U}$ and $\mathbf{V}$ so that the matrix $\mathbf{U}^H\mathbf{A}\mathbf{V}$ will be diagonal. This can be written as $\mathbf{U}^H\mathbf{A}\mathbf{V} = \mathbf{\Sigma}$ or equivalently, $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$.

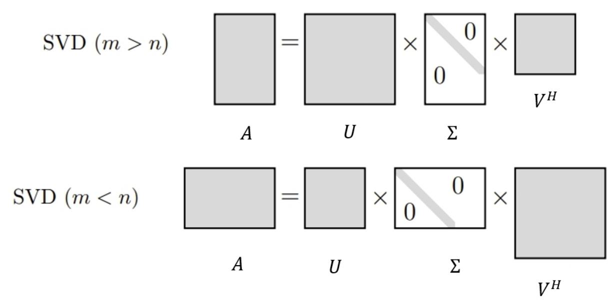

Figure 2.1

Remark. $\Sigma \in \mathbb{C}^{m\times n}$ is said to be diagonal if $\Sigma_{ij} = 0$ for $i \neq j$. It has exactly $p = \min\{m, n\}$ diagonal entries and can be denoted by $\Sigma = \text{diag}\{d_1, ...d_p\}$.

2.1.1 Singular value decomposition (SVD)

Proposition 2.1

For any matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{A}^H\mathbf{A}$ and $\mathbf{A}\mathbf{A}^H$ have non-negative eigenvalues.

Proof. $\mathbf{A}^H\mathbf{A}$ and $\mathbf{A}\mathbf{A}^H$ are hermitian and positive semi-definite. □

Proposition 2.2

For any matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$, $r(\mathbf{A}) = r(\mathbf{A}^H\mathbf{A}) = r(\mathbf{A}\mathbf{A}^H)$.

Proof. For $\mathbf{x}$, $\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{A}^H\mathbf{Ax} = 0$. For $\mathbf{x}$, $\mathbf{A}^H\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x}^H\mathbf{A}^H\mathbf{Ax} = \mathbf{0} \Rightarrow ||\mathbf{Ax}||_2^2 = 0 \Rightarrow \mathbf{Ax} = \mathbf{0}$. Hence, dim(ker $\mathbf{A}$) = dim(ker $\mathbf{A}^H\mathbf{A}$) and so $r(\mathbf{A}) = r(\mathbf{A}^H\mathbf{A})$. The result follows since $r(\mathbf{AA}^H) = r(\mathbf{A}^H\mathbf{A})$. $\square$

Theorem 2.1

Let $\mathbf{A} \in \mathbb{C}^{m\times n}$ and denote $p = \min\{m, n\}$. Denote the rank of $\mathbf{A}$ by r ($0 \leqslant r \leqslant p$). Then there are unitary matrices $\mathbf{U} \in \mathbb{C}^{m\times m}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ and a real diagonal matrix $\Sigma = \operatorname{diag}\{\sigma_1, ..., \sigma_p\} \in \mathbb{R}^{m\times n}$ such that

$$\mathbf{A} = \mathbf{U\Sigma V}^H$$

and

$$\sigma_1 \geqslant ... \geqslant \sigma_r \geqslant \sigma_{r+1} = ... = \sigma_p = 0,$$

the matrix Σ is uniquely determined by $\mathbf{A}$.

Proof Strategy: Assume $m > n$, if $\mathbf{A} = \mathbf{U\Sigma V}^H$, then

$$\begin{aligned}
\mathbf{A}^H\mathbf{A} &= \mathbf{V}(\mathbf{\Sigma}^H\mathbf{\Sigma})\mathbf{V}^H \\
&= \mathbf{V}(\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix})\mathbf{V}^H \qquad (\sigma_i^* = \sigma_i \text{ since } \sigma_i \in \mathbb{R}) \\
&= \mathbf{V}\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \\ & & \end{bmatrix}\mathbf{V}^H
\end{aligned}$$

Then we can let $\mathbf{V}$ be matrix containing the eigenvectors of $\mathbf{A}^H\mathbf{A}$ and $\mathbf{\Sigma}$ be matrix containing square roots of eigenvalues of $\mathbf{A}^H\mathbf{A}$. After we have found both $\mathbf{V}$ and $\mathbf{\Sigma}$, rewrite the matrix equation as

$$\mathbf{AV} = \mathbf{U\Sigma},$$

or in columns,

$$\mathbf{A}\begin{pmatrix}\mathbf{v}_1 & ... & \mathbf{v}_r & \mathbf{v}_{r+1} & ... & \mathbf{v}_n\end{pmatrix} = \begin{pmatrix}\mathbf{u}_1 & ... & \mathbf{u}_r & \mathbf{u}_{r+1} & ... & \mathbf{u}_n\end{pmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix}$$

By comparing columns, we obtain

$$\mathbf{X}\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i, & 1 \leqslant i \leqslant r(\text{nonzero singular values}) \\ \mathbf{0}, & r < i \leqslant d. \end{cases}$$

This tells us how to find the matrix $\mathbf{U} : \mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i$ for $1 \leqslant i \leqslant r$.

Proof. Let $\mathbf{C} = \mathbf{A}^H\mathbf{A} \in \mathbb{C}^{n\times n}$. Then $\mathbf{C}$ is square, hermitian, and positive semidefinite. Therefore, $\mathbf{C} = \mathbf{V}\Lambda\mathbf{V}^H$ for an unitary $\mathbf{V} \in \mathbb{C}^{n\times n}$ and diagonal $\Lambda = \mathrm{diag}\{\lambda_1, \ldots\lambda_n\} \in \mathbb{R}^{n\times n}$ with $\lambda_1 \geqslant \ldots \geqslant \lambda_r > 0 = \lambda_{r+1} = \ldots = \lambda_n$. Let $\sigma_i = \sqrt{\lambda_i}$ and correspondingly form the matrix $\Sigma \in \mathbb{R}^{m\times n}$:

$$\mathbf{\Sigma} = \begin{pmatrix} \mathrm{diag}(\sigma_1, \ldots, \sigma_r) & \mathbf{O}_{r\times(n-r)} \\ \mathbf{O}_{(m-r)\times r} & \mathbf{O}_{(m-r)\times(n-r)} \end{pmatrix}$$

Define also

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i \in \mathbb{C}^n, \text{ for each } 1 \leqslant i \leqslant r.$$

Then $\mathbf{u}_1, \ldots, \mathbf{u}_r$ are orthonormal vectors. To see this,

$$\begin{aligned} \mathbf{u}_i^H\mathbf{u}_j &= (\frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i)^H(\frac{1}{\sigma_j}\mathbf{A}\mathbf{v}_j) = \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^H\mathbf{A}^H\mathbf{A}\mathbf{v}_j \\ &= \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^H(\lambda_j v) \end{aligned}$$

□

2.1.2 Geometric interpretation of SVD

2.1.3 Applications of Singular-Value Decomposition

2.2 Reference

- Matrix Decomposition
- SVD notes
- Mathematical Methods for Data Visualization
- SVD notes XMUT
- Singular Value Decomposition
- least squares svd
- svd application

Chapter 3

Hadamard, Kronecker and Khatri-Rao Products

3.1 Partitioned matrices

Let $\{\alpha_{m_1}, ..., \alpha_{m_R}\}$ and $\{\beta_{n_1}, ..., \beta_{n_S}\}$ be partitions of the sets $\{1, ..., m\}$ and $\{1, ..., n\}$, 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:

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

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:

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

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:

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

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

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

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 ith row, the jth 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 nth 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

- Kronecker Product and the vec Operator
- On Hadamard and Kronecker Products Over Matrix of Matrices
- HADAMARD, KHATRI-RAO, KRONECKER AND OTHER MATRIX PRODUCTS

Chapter 4

Tensor Operations

4.1 Notation

Let $\chi \in \mathbb{K}^{I_1}$

4.2 Notion of slice

4.3 Mode combination

4.4 Matricization

4.5 Multiplication operations

Chapter 5

Tensor Decomposition