Chapter 1
Multi-dimensional Data

Abstract Roughly speaking, the classical matrix theory can mainly deal with one- or two-dimensional data. The main purpose of semi-tensor product of matrices is to use matrix tools to deal with higher-dimensional data. Hence the multi-dimensional data become the main objective of this book. This chapter considers how to arrange a set of higher-dimensional data into a vector or a matrix. First, the ordered multi-index is introduced to arrange a set of data into a properly ordered form. Then we briefly introduce some other matrix products, including the Kronecker Product (also called the tensor product) of matrices, which is a fundamental tool in this book; the Hadamard product, which is also used in the sequel; etc. Tensor and Nash equilibrium are two useful examples for multi-dimensional data. They are introduced in this chapter and will be used in the sequel. Symmetric group is another useful tool and it is also introduced here. Finally, we propose a special matrix, called the swap matrix. Its certain properties are discussed. It will be used largely to overcome the non-commutativity of matrix product.

1.1 Multi-dimensional Data

In scientific research we have to deal with various kinds of data. First, we would like to clarify what do we mean the dimension of data. A set of data may depend on k factors. Assume each factor can have n_j levels, $j = 1, \cdots, k$. Then to label a data, we may need k indices $i_1, i_2, \cdots, i_k$, and allow i_j runs from 1 to n_j. Then we can have a finite set of data as:

$$D := \{d_{i_1,\cdots,i_k} \mid 1 \le i_j \le n_j \; j = 1, \cdots, k\}. \tag{1.1}$$

Throughout this book only finite sets of data are considered, unless elsewhere stated.

We say that the dimension of D is k, i.e., $\dim(D) = k$. Roughly speaking, the dimension of a set of data is the number of indices of the data.

Example 1.1. 1. Consider a vector $X \in \mathbb{R}^n$. It can be expressed as $X = (x_1, x_2, \cdots, x_n)^T$. Hence a vector can be considered as a set of one-dimensional data.

2. Consider a matrix $A \in \mathcal{M}_{m\times n}$. It can be expressed as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Hence in a natural way, a matrix can be considered as a set of two-dimensional data.

3. Let $y = f(x_1, x_2, \cdots, x_n)$ be a function of n variables. To get numerical expression of f we let x_i take values $x_i^1, x_i^2, \cdots, x_i^{n_i}$. Then we have a set of n-dimensional data as
$$Y := \left\{ y_{i_1,\cdots,i_n} = f(x_1^{i_1}, \cdots, x_n^{i_n}) \,\middle|\, 1 \le i_j \le n_j;\ j = 1, \cdots, n \right\}.$$

To arrange a multi-dimensional data, or equivalently, to decide the order of a data, we introduce an ordered multi-index, briefly, multi-index.

Definition 1.1. A set of k-dimensional data (1.1) is said to be arranged by the (ordered) multi-index $\mathrm{id}(i_1, \cdots, i_k; n_1, \cdots, n_k)$, if the data are labeled by indices $i_1, \cdots, i_k$, and arranged in the order that $d_{p_1,\cdots,p_k} \prec d_{q_1,\cdots,q_k}$, iff there exists a $1 \le r \le k$, such that $p_i = q_i$, $i < r$ and $p_r < q_r$.

Example 1.2. Let
$$D = \{x_{i,j,k} | 1 \le i \le 2;\ 1 \le j \le 3;\ 1 \le k \le 4\}.$$

(i) If we arrange it by the multi-index $\mathrm{id}(i, j, k; 2, 3, 4)$, we have

$$[x_{111}, x_{112}, x_{113}, x_{114}, x_{121}, x_{122}, x_{123}, x_{124}, x_{131}, x_{132}, x_{133}, x_{134},$$
$$x_{211}, x_{212}, x_{213}, x_{214}, x_{221}, x_{222}, x_{223}, x_{224}, x_{231}, x_{232}, x_{233}, x_{234}].$$

(ii) If we arrange it by the multi-index $\mathrm{id}(j, i, k; 3, 2, 4)$, we have

$$[x_{111}, x_{112}, x_{113}, x_{114}, x_{211}, x_{212}, x_{213}, x_{214}, x_{121}, x_{122}, x_{123}, x_{124},$$
$$x_{221}, x_{222}, x_{223}, x_{224}, x_{131}, x_{132}, x_{133}, x_{134}, x_{231}, x_{232}, x_{233}, x_{234}].$$

(iii) If we arrange it by the multi-index $\mathrm{id}(k, j, i; 4, 3, 2)$, we have

$$[x_{111}, x_{211}, x_{121}, x_{221}, x_{131}, x_{231}, x_{112}, x_{212}, x_{122}, x_{222}, x_{132}, x_{232},$$
$$x_{113}, x_{213}, x_{123}, x_{223}, x_{133}, x_{233}, x_{114}, x_{214}, x_{124}, x_{224}, x_{134}, x_{234}].$$

Assume we have a set of n data, where $n = \prod_{i=1}^k n_i$. Then we can either use a single index to label the data as

$$D = [x_1, x_2, \cdots, x_n], \tag{1.2}$$

or use multi-index $\mathrm{id}(i_1, \cdots, i_k; n_i, \cdots, n_k)$ to arrange the data as

$$D = [x_{11\cdots 1}, x_{11\cdots 2}, \cdots, x_{11\cdots n_k}, \cdots, x_{n_1 n_2 \cdots n_k}]. \tag{1.3}$$

Then we need to find formulas to convert the single index to multi-index and vise versa. In the following we deduce the formulas. For notational ease, we introduce some notations in the following.

- Let $a \in \mathbb{Z}$, $b \in \mathbb{Z}_+$. As in C-language, the $a\%b$ is used for the remaining of a/b.
- Denote by $[t]$ the largest integer, which is less than or equal to t.

For instance,

$$\begin{gathered} 100\%3 = 1, \quad 100\%7 = 2, \quad (-7)\%3 = 2; \\ \left[\frac{7}{3}\right] = 2, \quad [-1.25] = -2. \end{gathered}$$

It is easy to see that

$$a = \left[\frac{a}{b}\right] b + a\%b. \tag{1.4}$$

Next, we give the converting formulas between single index and multi-index of a set of data. We leave the proves to the reader.

Proposition 1.1. *Let D be a set of $n = \prod_{i=1}^{k} n_i$ data. It has been labeled by single index as in (1.2) and by multi-index as in (1.3). An element $x \in D$ is labeled by single index p and multi-index $\mu_1 \cdots \mu_k$. That is, $x \in D$ is expressed as*

$$x = x_p = x_{\mu_1 \cdots \mu_k}.$$

Then we have the following converting formulas:

1. *Set $p_k := p - 1$, then $(\mu_1, \cdots, \mu_k)$ can be calculated inductively by*

$$\begin{cases} \mu_k = p_k \% n_k + 1, \\ p_j = \left[\frac{p_{j+1}}{n_{j+1}}\right], \quad \mu_j = p_j \% n_j + 1, \quad j = k-1, \cdots, 1. \end{cases} \tag{1.5}$$

2. *Conversely, from single index to multi-index we have:*

$$p = \sum_{j=1}^{k-1} (\mu_j - 1) n_{j+1} n_{j+2} \cdots n_k + \mu_k. \tag{1.6}$$

The following example shows the conversions.

Example 1.3. A set of data

$$D = \{d_1, d_2, \cdots, d_{100}\}.$$

1. Given a number $p = 35$, what is the label of d_p under multi-index $\mathrm{id}(i_1, i_2, i_3; 4, 5, 5)$? Using (1.5), we have

$$
\begin{aligned}
p_3 &= p - 1 = 34,\\
\mu_3 &= p_3 \% n_3 + 1 = 34\%5 + 1 = 4 + 1 = 5,\\
p_2 &= [p_3/n_3] = [34/5] = 6,\\
\mu_2 &= p_2 \% n_2 + 1 = 6\%5 + 1 = 1 + 1 = 2,\\
p_1 &= [p_2/n_2] = [6/4] = 1,\\
\mu_1 &= p_1 \% n_1 + 1 = 1\%4 + 1 = 1 + 1 = 2.
\end{aligned}
$$

Thus the label of p is $(2,2,5)$.

2. Assume the label of $d_q \in D$ under multi-index $\mathrm{id}(i_1, i_2, i_3; 5, 2, 10)$ is $(3,2,8)$. What is the single index q?
   Using (1.6), we have

$$q = (\mu_1 - 1)n_2 n_3 + (\mu_2 - 1)n_3 + \mu_3 = 2 \cdot 2 \cdot 10 + 10 + 8 = 58.$$

1.2 Arrangement of Data

Let D be a k-dimensional data as in (1.1) with $n = \prod_{i=1}^{k} n_i$. When $k = 1$, $n = n_1$ and the data can be arranged into a vector as

$$V_D = (d_1, d_2, \cdots, d_n)^T. \tag{1.7}$$

When $k = 2$, the data can naturally be arranged into a matrix as

$$M_D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n_2} \\ d_{21} & d_{22} & \cdots & d_{2n_2} \\ \vdots & & & \\ d_{n_1 1} & d_{n_1 2} & \cdots & d_{n_1 n_2} \end{bmatrix}. \tag{1.8}$$

Vector and matrix are two major objects for matrix theory or linear algebra. Roughly speaking, matrix theory is a theory for one- or two-dimensional data. Now when $k \geq 3$, then what can we do? Let us first consider the case of $k = 3$. Say, we have a set of data D as

$$D = \{d_{ijk} \,|\, i = 1, \cdots, p; j = 1, \cdots, m; k = 1, \cdots, n\}. \tag{1.9}$$

Then how can we arrange the data? It was proposed by some researchers that the data are arranged into a cube, which consists of p layers and each layer is an $m \times n$ matrix [1, 7]. Such a compounded matrix is called a cubic matrix (Fig. 1.1). Now a natural question is: when $k > 3$ what can we do? It seems that this is not a proper way to arrange the higher-dimensional data.

As we know that a set of higher-dimensional data can easily be stored in a computer memory, where they are not arranged into a cubic or even higher-dimensional cube. In fact, the data are arranged into a line regardless the dimension of the data.

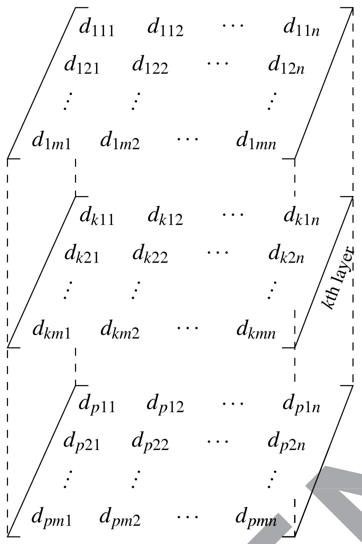

Fig. 1.1 A cubic matrix

Then how the computers to find the hierarchy structure of the data? they use some marks. Say, in C-language the pointer, pointer to pointer, pointer to pointer to pointer etc. are used to indicate the data structure.

Motivated by the computer technology, we propose to arrange a set of data into either a vector or a matrix. One may ask that why not just use vector only. It might be more convenient? In fact, to use tools developed in matrix theory, both vector and matrix forms are necessary.

To arrange a set of multi-dimensional data into a vector is rather easy. Formula (1.6) provides the single index label for each data. We are particularly interested in arrange the entries of a matrix into a vector.

Definition 1.2. Consider the matrix M_D in (1.8).

1. Its row stacking form, denoted by $V_r(M_D)$, is defined as

$$V_r(M_D) = (d_{11}, d_{12}, \cdots, d_{1n_2}, d_{21}, d_{22}, \cdots, d_{2n_2}, \cdots d_{n_1 1}, d_{n_1 2}, \cdots, d_{n_1 n_2})^T. \tag{1.10}$$

2. Its column stacking form, denoted by $V_c(M_D)$, is defined as

$$V_c(M_D) = (d_{11}, d_{21}, \cdots, d_{n_1 1}, d_{12}, d_{22}, \cdots, d_{n_1 2}, \cdots, d_{1n_2}, d_{2n_2}, \cdots, d_{n_1 n_2})^T. \tag{1.11}$$

By definition, it is obvious that for any matrix A

$$V_r(A) = V_c(A^T); \quad V_c(A) = V_r(A^T). \tag{1.12}$$

Next, we consider how to arrange a set of multi-dimensional data into a matrix form. Let $\mathrm{id}(i_1,\cdots,i_k;n_1,\cdots,n_k)$ be a multi-index.

$$\{i_{j_1},\cdots,i_{j_p}\}\subset\{i_1,\cdots,i_k\}.$$

Then $\mathrm{id}\left(i_{j_1},\cdots,i_{j_p};n_{j_1},\cdots,n_{j_p}\right)$ is called a sub-index of $\mathrm{id}(i_1,\cdots,i_k;n_1,\cdots,n_k)$.

Definition 1.3. Let D be a k-dimensional data as in (1.1). Assume there is a partition of the indices as

$$\{i_1,i_2,\cdots,i_k\}=\{i_{r_1},i_{r_2},\cdots,i_{r_p}\}\cup\{i_{c_1},i_{c_2},\cdots,i_{c_q}\},$$

where $p\geq 1$, $q\geq 1$ and $p+q=k$. Then the multi-indexed matrix M_D, corresponding to this index partition, is defined as follows:

(i) $M_D\in\mathscr{M}_{n_r\times n_c}$, where $n_r=\prod_{i=1}^{p}n_{r_i}$ and $n_c=\prod_{i=1}^{q}n_{c_i}$.
(ii) The rows of M_D is labeled by multi-index $\mathrm{id}\left(i_{r_1},i_{r_2},\cdots,i_{r_p};n_{r_1},n_{r_2},\cdots,n_{r_p}\right)$ and its columns is labeled by multi-index $\mathrm{id}\left(i_{c_1},i_{c_2},\cdots,i_{c_q};n_{c_1},n_{c_2},\cdots,n_{c_q}\right)$.
(iii) the $((\alpha_1,\cdots,\alpha_p),(\beta_1,\cdots,\beta_q))$th element of M_D is $d_{i_1,\cdots,i_k}$ where $i_{r_s}=\alpha_s$, $s=1,\cdots,p$, and $i_{c_s}=\beta_s$, $s=1,\cdots,q$.

The set of multi-indexed matrices, which have row and column multi-indexes as $\mathrm{id}(i_1,i_2,\cdots,i_p;m_1,m_2,\cdots,m_p)$ and $\mathrm{id}(i_1,i_2,\cdots,i_q;n_1,n_2,\cdots,n_q)$ respectively, is in $\mathscr{M}_{\prod_{i=1}^{p}m_i\times\prod_{i=1}^{q}n_i}$. To emphasize its multi-index structure, it can alternatively be denoted by $\mathscr{M}_{(m_1,\cdots,m_p)\times(n_1,\cdots,n_q)}$.

Example 1.4. Given a 4-dimensional data

$$D=\{d_{i,j,k,r}|i=1,2;j=1,2,3;k=1,2,3,4;r=1,2\}.$$

1. Partition the index set as $\{i,j,k,r\}=\{i,j\}\cup\{k,r\}$, and let the corresponding matrix M_D be row-indexed by $\mathrm{id}(i,j;2,3)$ and the column-indexed by $\mathrm{id}(k,r;4,2)$. Then we have the matrix as

$$M_D=\begin{bmatrix}
d_{1111} & d_{1112} & d_{1121} & d_{1122} & d_{1131} & d_{1132} & d_{1141} & d_{1142}\\
d_{1211} & d_{1212} & d_{1221} & d_{1222} & d_{1231} & d_{1232} & d_{1241} & d_{1242}\\
d_{1311} & d_{1312} & d_{1321} & d_{1322} & d_{1331} & d_{1332} & d_{1341} & d_{1342}\\
d_{2111} & d_{2112} & d_{2121} & d_{2122} & d_{2131} & d_{2132} & d_{2141} & d_{2142}\\
d_{2211} & d_{2212} & d_{2221} & d_{2222} & d_{2231} & d_{2232} & d_{2241} & d_{2242}\\
d_{2311} & d_{2312} & d_{2321} & d_{2322} & d_{2331} & d_{2332} & d_{2341} & d_{2342}
\end{bmatrix}\in\mathscr{M}_{(2,3)\times(4,2)}.$$

2. Partition the index set as $\{i,j,k,r\}=\{i,k\}\cup\{j,r\}$, and let the corresponding matrix M_D be row-indexed by $\mathrm{id}(i,k;2,4)$ and the column-indexed by $\mathrm{id}(j,r;3,2)$. Then we have the matrix as

$$
M_D = \begin{bmatrix}
d_{1111} & d_{1112} & d_{1211} & d_{1212} & d_{1311} & d_{1312} \\
d_{1121} & d_{1122} & d_{1221} & d_{1222} & d_{1321} & d_{1322} \\
d_{1131} & d_{1132} & d_{1231} & d_{1232} & d_{1331} & d_{1332} \\
d_{1141} & d_{1142} & d_{1241} & d_{1242} & d_{1341} & d_{1342} \\
d_{2111} & d_{2112} & d_{2211} & d_{2212} & d_{2311} & d_{2312} \\
d_{2121} & d_{2122} & d_{2221} & d_{2222} & d_{2321} & d_{2322} \\
d_{2131} & d_{2132} & d_{2231} & d_{2232} & d_{2331} & d_{2332} \\
d_{2141} & d_{2142} & d_{2241} & d_{2242} & d_{2341} & d_{2342}
\end{bmatrix} \in \mathcal{M}_{(2,4)\times(3,2)}.
$$

1.3 Matrix Products

In addition to conventional matrix product, there are some other matrix products. This section gives a brief survey on their definitions and basic properties without proves. We refer to many standard references of matrix theory for details.

1.3.1 Kronecker Product of Matrices

The Kronecker product of matrices is also called the tensor product of matrices. This product is applicable to any two matrices. It will be used from time to time throughout this book. We refer to [4] for a complete discussion.

Definition 1.4. Let $A = [a_{ij}] \in \mathcal{M}_{m\times n}$ and $B = [b_{ij}] \in \mathcal{M}_{p\times q}$. The Kronecker product of A and B is defined as

$$
A \otimes B = \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{11}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & & & \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{bmatrix} \in \mathcal{M}_{mp\times nq}. \tag{1.13}
$$

Next, we introduce some basic properties of Kronecker product:

Proposition 1.2. *1. (Associative Law)*

$$
A \otimes (B \otimes C) = (A \otimes B) \otimes C. \tag{1.14}
$$

2. (Distributive Law)

$$
(\alpha A + \beta B) \otimes C = \alpha(A \otimes C) + \beta(B \otimes C), \tag{1.15}
$$

$$
A \otimes (\alpha B + \beta C) = \alpha(A \otimes B) + \beta(A \otimes C), \quad \alpha, \beta \in \mathbb{R}. \tag{1.16}
$$

Proposition 1.3. *1.*

$$
(A \otimes B)^T = A^T \otimes B^T. \tag{1.17}
$$

2. Assume A and B are invertible. Then

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}. \tag{1.18}$$

3.

$$\mathrm{rank}(A \otimes B) = \mathrm{rank}(A)\,\mathrm{rank}(B). \tag{1.19}$$

4. Let $A \in \mathscr{M}_{m\times m}$ and $B \in \mathscr{M}_{n\times n}$. Then

$$\det(A \otimes B) = (\det(A))^n (\det(B))^m. \tag{1.20}$$
$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B). \tag{1.21}$$

The next proposition is very useful.

Proposition 1.4. *Let $A \in \mathscr{M}_{m\times n}$, $B \in \mathscr{M}_{p\times q}$, $C \in \mathscr{M}_{n\times r}$, and $D \in \mathscr{M}_{q\times s}$. Then*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \tag{1.22}$$

Particularly, we have

$$A \otimes B = (A \otimes I_p)(I_n \otimes B). \tag{1.23}$$

The next proposition is about the vector form of matrices.

Proposition 1.5. *1. Let $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^n$ be two column vectors. Then*

$$V_c(XY^T) = Y \otimes X. \tag{1.24}$$

2. Let $A \in \mathscr{M}_{m\times p}$, $B \in \mathscr{M}_{p\times q}$, and $C \in \mathscr{M}_{q\times n}$. Then

$$V_c(ABC) = (C^T \otimes A) V_c(B). \tag{1.25}$$

1.3.2 Hadamard Product

Hadamard product of matrices is another useful product in certain problems. It will also used in the sequel. The reader is referred to [4, 8] for more about it.

Definition 1.5. Let $A = [a_{i,j}], B = [b_{i,j}] \in \mathscr{M}_{m\times n}$. The Hadamard product of A and B is defined as

$$A \odot B = [a_{i,j} b_{i,j}] \in \mathscr{M}_{m\times n}. \tag{1.26}$$

Hadamard product has some important properties as

Proposition 1.6. *1. (Commutativity) For any two matrices $A, B \in \mathscr{M}_{m\times n}$*

$$A \odot B = B \odot A. \tag{1.27}$$

2. *(Associative Law) Let $A, B, C \in \mathcal{M}_{m\times n}$. Then*

$$(A \odot B) \odot C = A \odot (B \odot C). \tag{1.28}$$

3. *(Distributive Law) Let $A, B, C \in \mathcal{M}_{m\times n}$. Then*

$$(\alpha A + \beta B) \odot C = \alpha(A \odot C) + \beta(B \odot C), \quad \alpha, \beta \in \mathbb{R}. \tag{1.29}$$

Proposition 1.7. *1.*

$$(A \odot B)^T = A^T \odot B^T. \tag{1.30}$$

2. *Let $A \in \mathcal{M}_n$ and $E = \mathbf{1}_n$, i.e., $E = (\underbrace{1\ 1\ \cdots\ 1}_{n})^T$. Then*

$$A \odot (EE^T) = A = (EE^T) \odot A. \tag{1.31}$$

3. *Let $X, Y \in \mathbb{R}^n$ be two column vectors. Then*

$$(XX^T) \odot (YY^T) = (X \odot Y)(X \odot Y)^T. \tag{1.32}$$

Define

$$H_n = \mathrm{diag}(\delta_n^1, \cdots, \delta_n^n).$$

Then we have

Proposition 1.8. *Let $A, B \in \mathcal{M}_{m\times n}$. Then*

$$A \odot B = H_m^T (A \otimes B) H_n. \tag{1.33}$$

Proposition 1.9 (Schur's Theorem). *Let $A, B \in \mathcal{M}_n$ be symmetric.*

(i) If $A \geq 0$ and $B \geq 0$, then $A \circ B \geq 0$;
(ii) If $A > 0$ and $B > 0$, then $A \circ B > 0$.

Proposition 1.10 (Oppenbeim's Theorem). *Let $A, B \in \mathcal{M}_n$ be symmetric. If $A \geq 0$ and $B \geq 0$, then*

$$\det(A \odot B) \geq \det(A)\det(B). \tag{1.34}$$

1.3.3 Khatri-Rao Product

We refer to [6] or [8] for details of Khatri-Rao product of matrices.

Definition 1.6. Let $A \in \mathscr{M}_{m\times r}$ and $B \in \mathscr{M}_{n\times r}$. The Khatri-Rao product of A and B is defined as

$$A * B = [\mathrm{Col}_1(A) \otimes \mathrm{Col}_1(B), \mathrm{Col}_2(A) \otimes \mathrm{Col}_2(B), \cdots, \mathrm{Col}_r(A) \otimes \mathrm{Col}_r(B)]. \quad (1.35)$$

Proposition 1.11. *1. (Associative Law) Let $A \in \mathscr{M}_{m\times r}$, $B \in \mathscr{M}_{n\times r}$, and $C \in \mathscr{M}_{p\times r}$. Then*

$$(A * B) * C = A * (B * C). \quad (1.36)$$

2. (Distributive Law) Let $A, B \in \mathscr{M}_{m\times r}$ and $C \in \mathscr{M}_{n\times r}$. Then

$$(aA + bB) * C = a(A * C) + b(B * C), \quad a, b \in \mathbb{R}. \quad (1.37)$$
$$C * (aA + bB) = a(C * A) + b(C * B), \quad a, b \in \mathbb{R}. \quad (1.38)$$

The following example is useful in the sequel.

Example 1.5. A matrix $A \in \mathscr{M}_{m\times r}$ is called a logical matrix if all its columns are of the form δ_m^i, $1 \le i \le m$. The set of $m \times r$ logical matrices is denoted by $\mathscr{L}_{m\times r}$.

Assume $A \in \mathscr{L}_{m\times r}$ and $B \in \mathscr{L}_{n\times r}$. Then

$$A * B \in \mathscr{L}_{mn\times r}.$$

Remark 1.1. In addition to conventional matrix product, we have introduced Kronecker product, Hadamard product, and Khatri-Rao product of matrices. One sees easily that the associativity and distributivity are two common properties. These two properties may be considered as two fundamental requirements for any matrix products.

1.4 Tensor

Tensor is a typical multi-linear mapping. This section is a brief introduction. We refer to [2] for details.

Let V be an n-dimensional vector space with a basis $\{d_1, \cdots, d_n\}$. Denote by V^* the dual space of V, that is V^* is the set of linear functions on V. Let $\{e_1, \cdots, e_n\} \subset V^*$ be a basis of V^*, dual to $\{d_1, \cdots, d_n\}$. That is,

$$d_i(d_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Then $X = \sum_{i=1}^{n} x_i d_i \in V$ can be expressed as a column vector $X = (x_1, \cdots, x_n)^T$, and $\omega = \sum_{i=1}^{n} \omega_i e_i \in V^*$ as a row vector $\omega = (\omega_1, \cdots, \omega_n)$.

Definition 1.7. 1. Let $f : V^s \to \mathbb{R}$ be an s-linear mapping, and

$$f(d_{i_1},\cdots,d_{i_s}) = \mu_{i_1 i_2 \cdots i_s}, \quad 1 \le i_p \le n,\ p = 1,\cdots,s.$$

Arrange $\{\mu_{i_1 i_2 \cdots i_s} | 1 \le i_j \le n,\ j = 1,\cdots,s\}$ into a row by using multi-index $\mathrm{id}(i_1 i_2 \cdots i_s; n,\cdots,n)$ as

$$M_f = [\mu_{11\cdots1}\ \cdots\ \mu_{11\cdots n}\ \cdots\ \mu_{nn\cdots n}]. \tag{1.39}$$

M_f is called the structure matrix of f. By the linearity, it is easy to check that for $X_1,\cdots,X_s \in V$

$$f(X_1,\cdots,X_s) = M_f(X_1 \otimes \cdots \otimes X_s). \tag{1.40}$$

f is called a tensor of covariant order s. The set of tensors on V of covariant order s is denoted by $\mathscr{T}^s$.

2. Let $f : (V^*)^t \to \mathbb{R}$ be a t-linear mapping, and

$$f(e_{j_1},\cdots,e_{j_t}) = \mu^{j_1 j_2 \cdots j_t}, \quad 1 \le j_q \le n,\ q = 1,\cdots,t.$$

Arrange $\{\mu^{j_1 j_2 \cdots j_t} | 1 \le j_q \le n,\ q = 1,\cdots,t\}$ into a column by using multi-index $\mathrm{id}(j_1 j_2 \cdots j_t; n,\cdots,n)$ as

$$M_f = [\mu^{11\cdots1}\ \cdots\ \mu^{11\cdots n}\ \cdots\ \mu^{nn\cdots n}]^T. \tag{1.41}$$

M_f is called the structure matrix of f. By the linearity, it is easy to check that for $\omega_1,\cdots,\omega_t \in V^*$

$$f(\omega_1,\cdots,\omega_t) = (\omega_1 \otimes \cdots \otimes \omega_t)M_f. \tag{1.42}$$

f is called a tensor of contravariant order t. The set of tensors on V of contravariant order t is denoted by $\mathscr{T}_t$.

3. Let $f : V^s \times (V^*)^t \to \mathbb{R}$ be an $s+t$-linear mapping, and

$$f(d_{i_1},\cdots,d_{i_s},e_{j_1},\cdots,e_{j_t}) = \mu_{i_1 i_2 \cdots i_s}^{j_1 j_2 \cdots j_t},$$

$$1 \le i_p \le n,\ p = 1,\cdots,s;\ 1 \le j_q \le n,\ q = 1,\cdots,t.$$

Arrange $\left\{\mu_{i_1 i_2 \cdots i_s}^{j_1 j_2 \cdots j_t} | 1 \le i_p \le n,\ p = 1,\cdots,s;\ 1 \le j_q \le n,\ q = 1,\cdots,t\right\}$ into a matrix, whose columns are labeled by multi-index $\mathrm{id}(i_1 i_2 \cdots i_s; n,\cdots,n)$ and rows are labeled by multi-index $\mathrm{id}(j_1 j_2 \cdots j_t; n,\cdots,n)$. Then we have

$$M_f = \begin{bmatrix} \mu_{11\cdots 1}^{11\cdots 1} & \mu_{11\cdots 2}^{11\cdots 1} & \cdots & \mu_{11\cdots n}^{11\cdots 1} & \cdots & \mu_{nn\cdots n}^{11\cdots 1} \\ \mu_{11\cdots 1}^{11\cdots 2} & \mu_{11\cdots 2}^{11\cdots 2} & \cdots & \mu_{11\cdots n}^{11\cdots 2} & \cdots & \mu_{nn\cdots n}^{11\cdots 2} \\ \vdots & & & & & \\ \mu_{11\cdots 1}^{11\cdots n} & \mu_{11\cdots 2}^{11\cdots n} & \cdots & \mu_{11\cdots n}^{11\cdots n} & \cdots & \mu_{nn\cdots n}^{11\cdots n} \\ \vdots & & & & & \\ \mu_{11\cdots 1}^{nn\cdots n} & \mu_{11\cdots 2}^{nn\cdots n} & \cdots & \mu_{11\cdots n}^{nn\cdots n} & \cdots & \mu_{nn\cdots n}^{nn\cdots n} \end{bmatrix} \tag{1.43}$$

By the linearity, it is also easy to check that for $X_1, \cdots, X_s \in V$ and $\omega, \cdots, \omega_t \in V^*$

$$f(X_1, \cdots, X_s, \omega_1, \cdots, \omega_t) = (\omega_1 \otimes \cdots \otimes \omega_t) M_f (X_1 \otimes \cdots \otimes X_s). \tag{1.44}$$

f is called a tensor of covariant order s and contravariant order t. The set of tensors on V of covariant order s and contravariant t is denoted by $\mathscr{T}_t^s$.

In the following we assume contravariant order $t = 0$.

Definition 1.8. A tensor $f \in \mathscr{T}^s$ is symmetric if for any $i \neq j$

$$f(X_1, \cdots, X_i, \cdots, X_j \cdots, X_s) = f(X_1, \cdots, X_j, \cdots, X_i \cdots, X_s), \quad X_1, \cdots, X_s \in V. \tag{1.45}$$

A tensor $f \in \mathscr{T}^s$ is sky-symmetric if for any $i \neq j$

$$f(X_1, \cdots, X_i, \cdots, X_j \cdots, X_s) = -f(X_1, \cdots, X_j, \cdots, X_i \cdots, X_s), \quad X_1, \cdots, X_s \in V. \tag{1.46}$$

From high school algebra we know that in $\mathbb{R}^3$ two products were defined: (i) inner product; (ii) cross product. Fix a basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ as $\mathbf{i} = (1,0,0)^T$, $\mathbf{j} = (0,1,0)^T$, and $\mathbf{k} = (0,0,1)^T$. Let $X = (x_1, x_2, x_3)^T$ and $Y = (y_1, y_2, y_3)^T$. The inner product is defined as

$$\langle X, Y \rangle := x_1 y_1 + x_2 y_2 + x_3 y_3. \tag{1.47}$$

The cross product, denoted by $\times_c$, is defined as

$$X \times_c Y = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

Example 1.6. 1. The inner product on $\mathbb{R}^3$ is a tensor of covariant order $r = 2$. It is symmetric. (This result is also true for $\mathbb{R}^n$.)

2. Let $X, Y, Z \in \mathbb{R}^3$.

$$v := \langle X, Y \times_c Z \rangle.$$

We leave to the reader to check that v is a tensor of covariant order $r = 3$. It is sky-symmetric. In fact, v is the volume of the parallelepiped (as X, Y, Z satisfy the right hand rule, otherwise, it is the negative volume).

1.5 Nash Equilibrium

This section gives a very limited introduction to game theory, including strategic form and Nash equilibrium. We refer to [3] for details.

Assume a game has n players, denoted by $P_1, \cdots, P_n$, and each player P_i can have $n_i > 0$ possible actions, called his strategies, denoted by $\{s^i_1, s^i_2, \cdots, s^i_{n_i}\}$. Let

$$f^i(s^1_{k_1}, s^2_{k_2}, \cdots, s^n_{k_n}), \quad i = 1, \cdots n \tag{1.48}$$

be the payoff of the ith player, which means what the player i obtains from the game when P_j takes his strategy $s^j_{k_j}$, $j = 1, \cdots, n$. For compactness, denote by

$$\mu^i_{k_1,k_2,\cdots,k_n} := f^i(s^1_{k_1}, s^2_{k_2}, \cdots, s^n_{k_n}), \quad i = 1, \cdots, n;\ k_j = 1, \cdots, n_j,\ j = 1, \cdots, n. \tag{1.49}$$

It is reasonable to assume that each player is pursuing his maximum payoff.

Definition 1.9. A set of strategies $\{s^1_*, s^2_*, \cdots, s^n_*\}$ is called a Nash equilibrium, if

$$f^i(s^1_*, s^2_*, \cdots, s^i_*, \cdots, s^n_*) \geq f^i(s^1_*, s^2_*, \cdots, s^i, \cdots, s^n_*), \quad \forall s^i,\ i = 1, \cdots, n. \tag{1.50}$$

Nash equilibrium is extremely important because once it is reached, each player intends to stick on this strategy forever.

Now let us see how to find the Nash equilibrium. The following procedure comes from definition directly. For each i we can put the data

$$D_i = \{\mu^i_{k_1,k_2,\cdots,k_n} \,|\, k_j = 1, \cdots, n_j;\ j = 1, \cdots, n\}$$

into a matrix M_i, which has $\mathrm{id}(k_i; n_i)$ as its row index and use all other indexes to label columns, that is, $\mathrm{id}(k_1, \cdots, k_{i-1}, k_{i+1}, \cdots, k_n; n_1, \cdots, n_{i-1}, n_{i+1}, \cdots, n_n;)$ is used as its column index. Then for each column of M_i we can find at least one n-index $(k^*_1, \cdots, k^*_n)$, which corresponds to the largest value of μ^i. Denote by K_i the set of n-indexed found from each columns of M_i. Then

$$E_N := \cap^n_{i=1} K_i$$

is the set of Nash equilibriums.

We give some examples to depict this.

Example 1.7. Two suspects are arrested by the police. The police have insufficient evidence for a conviction, and, having separated the prisoners, visit each of them to offer the same deal. If one testifies for the prosecution against the other (defects) and the other remains silent (cooperates), the defector goes free and the silent accomplice receives the full 10-year sentence. If both remain silent, both prisoners are sentenced to only 1 year in jail for a minor charge. If each betrays the other, each receives a 5-year sentence. Each prisoner must choose to betray the other or to remain silent. Each one is assured that the other would not know about the betrayal before the end of the investigation. How should the prisoners act?

The payoff bi-matrix is given in Table 1.1.

Table 1.1 Payoff of Prisoner's Dilemma

$P_1\backslash P_2$	C	D
C	$-1\ \ -1$	$-10\ \ 0$
D	$0\ \ -10$	$-5\ \ -5$

Now we have M_1 and M_2 as

$$M_1 = \begin{bmatrix} -1 & -10 \\ \underline{0} & \underline{-5} \end{bmatrix}; \quad M_2 = \begin{bmatrix} -1 & -10 \\ \underline{0} & \underline{-5} \end{bmatrix},$$

where the underline elements are the column maximal elements.

Note that in M_1 the row index is k_1 and the column index is k_2 while in M_2 the row index is k_2 and the column index is k_1, hence we have the K_i set as

$$K_1 = \{(2,1),(2,2)\}; \quad K_2 = \{(1,2),(2,2)\}.$$

It follows that the set of Nash equilibrium(s) is

$$E_N = K_1 \cap K_2 = \{(2,2)\},$$

which means (D,D) is the only Nash equilibrium.

Example 1.8. Assume a game has three players. Their strategies are: $S_1 = \{s_1^1, s_2^1, s_3^1\}$, $S_2 = \{s_1^2, s_2^2\}$, $S_3 = \{s_1^3, s_2^3\}$. And the payoffs are shown in Table 1.2.

Table 1.2 Payoffs

P_2	s_1^2		s_2^2	
$P_1\backslash P_3$	s_1^3	s_2^3	s_1^3	s_2^3
s_1^1	3 1 2	−1 2 1	2 −1 3	2 0 −1
s_2^1	2 2 0	1 −1 2	1 1 2	0 1 2
s_3^1	1 3 −1	−2 1 −1	0 2 −1	2 4 3

Then we have $M_i, i = 1,2,3$ as

$$M_1 = \begin{bmatrix} \underline{3} & -1 & \underline{2} & \underline{2} \\ 2 & \underline{1} & 1 & 0 \\ 1 & -2 & 0 & \underline{2} \end{bmatrix}, M_2 = \begin{bmatrix} \underline{1} & \underline{2} & \underline{2} & -1 & \underline{3} & 1 \\ -1 & 0 & 1 & \underline{1} & 2 & \underline{4} \end{bmatrix}, M_3 = \begin{bmatrix} \underline{2} & \underline{3} & 0 & \underline{2} & -1 & -1 \\ 1 & -1 & \underline{2} & \underline{2} & \underline{1} & \underline{3} \end{bmatrix}$$

Note the in M_1 the row index is k_1 and the column index is $Id(k_2,k_3;2,2)$, it follows that

$$K_1 = \{(1,1,1),(2,1,2),(1,2,1),(1,2,2),(3,2,2)\}.$$

In M_2 the row index is k_2 and the column index is $Id(k_1,k_3;3,2)$, it follows that

$$K_2 = \{(1,1,1),(1,1,2),(2,1,1),(2,2,2),(3,1,1),(3,2,2)\}.$$

In M_3 the row index is k_3 and the column index is $Id(k_1,k_2;3,2)$, hence

$$K_3 = \{(1,1,1),(1,2,1),(2,1,2),(2,2,1),(2,2,2),(3,1,2),(3,2,2)\}.$$

Therefore,

$$E_N = K_1 \cap K_2 \cap K_3 = \{(1,1,1),(3,2,2)\}.$$

That is, (s_1^1, s_1^2, s_1^3) and (s_3^1, s_2^2, s_2^3) are two Nash equilibriums.

1.6 Symmetric Group

Let $S = \{1,2,\cdots,k\}$. A permutation σ on S is a one-to-one mapping from S onto S. All the permutations on S with the product as the combination of two permutations as their product form a group called the symmetric group on k letters, or kth order symmetric group, denoted by $\mathbf{S}_k$ [5].

We use some numerical examples to depict it. For instance, let $k = 5$. A $\sigma \in \mathbf{S}_5$ may be expressed as

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 5 & 1 & 4 & 3 \end{pmatrix}.$$

Let another permutation $\tau \in \mathbf{S}_5$ be expressed as

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 3 & 4 & 1 & 2 \end{pmatrix}.$$

The product on $\mathbf{S}_5$ is defined as

$$\tau\sigma = \begin{pmatrix} 1\ 2\ 3\ 4\ 5 \\ \downarrow\downarrow\downarrow\downarrow\downarrow \\ 2\ 5\ 1\ 4\ 3 \\ \downarrow\downarrow\downarrow\downarrow\downarrow \\ 3\ 2\ 5\ 1\ 4 \end{pmatrix} = \begin{pmatrix} 1\ 2\ 3\ 4\ 5 \\ \downarrow\downarrow\downarrow\downarrow\downarrow \\ 3\ 2\ 5\ 1\ 4 \end{pmatrix}.$$

An alternative expression of an element in $\mathbf{S}_k$ is expressing it as a product of cycles. For instance, we can express $\sigma = (1\ 2\ 5\ 3)$, $\tau = (1\ 5\ 2\ 3\ 4)$, and $\tau\sigma = (1\ 3\ 5\ 4)$. Let

$$\mu = \begin{pmatrix} 1\ 2\ 3\ 4\ 5\ 6 \\ \downarrow\downarrow\downarrow\downarrow\downarrow\downarrow \\ 2\ 1\ 4\ 5\ 6\ 3 \end{pmatrix} \in \mathbf{S}_6.$$

Then it can be expressed as $\mu = (1\ 2)(3\ 4\ 5\ 6)$.

It is easy to check that the cardinal number $|\mathbf{S}_k| = k!$.

A cycle of two elements, such as $(a,b) \in \mathbf{S}_k$, is called a transposition.

Proposition 1.12. *Every permutation can be expressed as a product of transpositions.*

Proof. We have only to prove that each cycle can be expressed as a product of transpositions. Assume the length of a cycle is 1: we have $(r_1) = (r_1\ r_2)(r_2\ r_1)$. Assume the length of a cycle is greater than 1, then we have $(r_1\ r_2\ \cdots\ r_k) = (r_1\ r_k)(r_1\ r_{k-1})\cdots(r_1\ r_2)$. □

Note that a permutation σ can have different products of transpositions, but the number of transpositions can either be even or odd, but not both [5]. When the number of the transpositions is even we say $\mathrm{sgn}(\sigma) = 1$, otherwise, $\mathrm{sgn}(\sigma) = -1$.

For a $\sigma \in \mathbf{S}_k$ define a matrix M_σ as

$$M_\sigma = \delta_k[\sigma(1)\ \sigma(2)\ \cdots\ \sigma(k)].$$

Then M_σ can realize the permutation as

$$(\sigma(1)\ \sigma(2)\ \cdots\ \sigma(k)) = (1\ 2\ \cdots\ k)M_\sigma.$$

Moreover,

$$\mathrm{sgn}(\sigma) = \det(M_\sigma). \tag{1.51}$$

Note that sometimes to label a set of data the index order and the index arrange order do not coincide. Say the data are labeled by multi-index $(i_1, \cdots, i_k)$ and the multi-index may be ordered by id $\left(i_{\sigma(1)}, \cdots, i_{\sigma(k)}; n_{\sigma(1)}, \cdots, n_{\sigma(k)}\right)$. See the following example.

Example 1.9. Consider a set of data

$$D = \{d_{i_1,i_2,i_3} \,|\, i_1 = 1,2; i_2 = 1,2,3; i_3 = 1,2,3\}.$$

Let $\sigma = (1,3,2) \in \mathbf{S}_3$. Assume D is required to be arranged in the order of $\mathrm{id}(i_{\sigma(1)}, i_{\sigma(2)}, i_{\sigma(3)}; n_{\sigma(1)}, n_{\sigma(2)}, n_{\sigma(3)})$. Since $\sigma(1) = 3$, $\sigma(2) = 1$, and $\sigma(3) = 2$, the data are arranged in the order of $\mathrm{id}(i_3, i_1, i_2; 3, 2, 3)$. Hence we have

$$\begin{array}{l} d_{111}\ d_{121}\ d_{131}\ d_{211}\ d_{221}\ d_{231}\ d_{112}\ d_{122}\ d_{132} \\ d_{212}\ d_{222}\ d_{232}\ d_{113}\ d_{123}\ d_{133}\ d_{213}\ d_{223}\ d_{233} \end{array}$$

1.7 Swap Matrix

In this section we define a special matrix, called the swap matrix. It is very useful in overcoming the non-commutativity of the matrix product. Swap matrix was firstly introduced in [4], where it is called commutation matrix.

Definition 1.10. A swap matrix $W_{[m,n]} \in \mathscr{M}_{mn \times mn}$ is constructed in the following way:

Step 1. Label its columns by index (i, j) in the order of $\mathrm{id}(i, j; m, n)$ and its rows by index (I, J) in the order of $\mathrm{id}(J, I; n, m)$.

Step 2. The entry at row (I, J) and column (i, j), denoted by $w_{(I,J),(i,j)}$, is assigned as

$$w_{(I,J),(i,j)} = \begin{cases} 1, & I = i, \text{ and } J = j \\ 0, & \text{otherwise.} \end{cases} \tag{1.52}$$

We give some examples to depict swap matrices.

Example 1.10. 1. Consider $W_{[2,3]}$. Labeling its columns by (i, j) in the order of $\mathrm{id}(i, j; 2, 3)$ and its rows by (I, J) in the order of $\mathrm{id}(J, I; 3, 2)$. Then the swap matrix can be constructed as

$$W_{[2,3]} = \begin{array}{c} \begin{array}{cccccc} (11) & (12) & (13) & (21) & (22) & (23) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ (11) \\ (21) \\ (12) \\ (22) \\ (13) \\ (23) \end{array} .$$

2. Consider $W_{[3,2]}$. Labeling its columns by (i, j) in the order of $\mathrm{id}(i, j; 3, 2)$ and its rows by (I, J) in the order of $\mathrm{id}(J, I; 2, 3)$. Then the swap matrix can be constructed as

$$
W_{[3,2]} = \begin{array}{c}
\begin{array}{cccccc} (11) & (12) & (21) & (22) & (31) & (32) \end{array} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
\begin{array}{c} \\ (11) \\ (21) \\ (31) \\ (12) \\ (22) \\ (32) \end{array} .
$$

According to the construction of the swap matrix, the following two propositions are immediate consequences.

Proposition 1.13. *Let $A \in \mathscr{M}_{m\times n}$. Then*

$$W_{[m,n]}V_r(A) = V_c(A); \quad W_{[n,m]}V_c(A) = V_r(A). \tag{1.53}$$

Proposition 1.14. *1. Let $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ be two column vectors. Then*

$$W_{[m,n]}(X \otimes Y) = Y \otimes X. \tag{1.54}$$

2. Let $\omega \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^n$ be two row vectors. Then

$$(\omega \otimes \mu)W_{[n,m]} = \mu \otimes \omega. \tag{1.55}$$

This proposition has an equivalent statement.

Corollary 1.1. *Let $D = \{x_{ij} \mid i = 1,\cdots,m, j = 1,\cdots,n\}$ be a set of data. X is a column vector of the elements of D, labeled by (i,j) and arranged in the order of $\mathrm{id}(i,j;m,n)$, and Y is a column vector of the elements of D, labeled by (i,j) and arranged in the order of $\mathrm{id}(j,i;n,m)$. Then*

$$W_{[m,n]}X = Y; \quad W_{[n,m]}Y = X. \tag{1.56}$$

Swap matrix has some special properties, which follow from its definition immediately.

Proposition 1.15. *1. A swap matrix is an orthogonal matrix. It satisfies*

$$W^{\mathrm{T}}_{[m,n]} = W^{-1}_{[m,n]} = W_{[n,m]}. \tag{1.57}$$

2. When $m = n$, (1.55) becomes

$$W_{[n,n]} = W^{\mathrm{T}}_{[n,n]} = W^{-1}_{[n,n]}. \tag{1.58}$$

3.

$$W_{[1,n]} = W_{[n,1]} = I_n. \tag{1.59}$$

$m = n$ is a particularly useful in the sequel. To simplify the notation, we define

$$W_{[n]} := W_{[n,n]}.$$

Exercise 1

1. Prove the formulas (1.5) and (1.6).
2. A k-dimensional data D as in (1.1). Assume there is a partition of the indices as

$$\{i_1, i_2, \cdots, i_k\} = \{i_{r_1}, i_{r_2}, \cdots, i_{r_p}\} \cup \{i_{c_1}, i_{c_2}, \cdots, i_{c_q}\},$$

   such that M_D is row-indexed by $\mathrm{id}\left(i_{r_1}, i_{r_2}, \cdots, i_{r_p}; n_{r_1}, n_{r_2}, \cdots, n_{r_p}\right)$ and column-indexed by $\mathrm{id}\left(i_{c_1}, i_{c_2}, \cdots, i_{c_q}; n_{c_1}, n_{c_2}, \cdots, n_{c_q}\right)$.
   Find the pth element in $V_r(M_D)$ and the pth element in $V_c(M_D)$.
3. Let $Z \in \mathscr{L}_{mn\times k}$. Show that there exist unique $X \in \mathscr{L}_{m\times k}$ and $Y \in \mathscr{L}_{n\times k}$ such that Z is the Khatri-Rao product of X and Y. That is,

$$Z = X * Y.$$

4. Let ξ be an eigenvector of A with respect to the eigenvalue $\lambda \in \sigma(A)$ and η be an eigenvector of B with respect to the eigenvalue $\mu \in \sigma(B)$. Prove that $\xi \otimes \eta$ is an eigenvector of AB with respect to the eigenvalue $\lambda\mu \in \sigma(AB)$.
5. Check the v defined in Example 1.6 is a sky-symmetric tensor of covariant order $r = 3$.
6. Consider the game illustrated in Table 1.3.

Table 1.3 Payoffs

P_3	A						B					
$P_1 \backslash P_2$	L			R			L			R		
U	0	0	10	−5	−5	0	−2	−2	0	−5	−5	0
D	−5	−5	0	1	1	−5	−5	−5	0	−1	−1	5

   a. Find Nash equilibrium of this game.
   b. If Player 1 and Player 2 form a coalition, the coalition's payoff is the sum of their payoffs, then, do the Nash equilibriums found in *a* remain equilibriums?

7. Prove $\mathbf{S}_k$ is a group.
8. A set of data D is arranged by $\mathrm{id}(i_1, \cdots, i_k; n_1, \cdots, n_k)$, and under this order a data $d_p \in D$ is the pth element. $\sigma \in \mathbf{S}_k$ is a known permutation. Find the multi-index of d_p in the order of $\mathrm{id}(i_{\sigma(1)}, \cdots, i_(\sigma(k)); n_1, \cdots, n_k)$.
9. Let V be an n-dimensional vector space with a basis $\{d_1, \cdots, d_n\}$, which has dual basis $\{e_1, \cdots, e_n\}$. Consider a tensor $\omega \in T^r_s(V)$. The structure matrix of ω under these bases is M_ω. Let $\{\tilde{d}_1, \cdots, \tilde{d}_n\}$ be another basis of V, with dual bases $\{\tilde{e}_1, \cdots, \tilde{e}_n\}$. Moreover,

$$\begin{bmatrix} \tilde{d}_1 \\ \vdots \\ \tilde{d}_n \end{bmatrix} = A \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}.$$

Find the structure matrix of ω under new bases.

10. Proof the following two alternative expressions of the swap matrix.

$$W_{[m,n]} = \begin{bmatrix} I_m \otimes (\delta_n^1)^T \\ I_m \otimes (\delta_n^2)^T \\ \vdots \\ I_m \otimes (\delta_n^n)^T \end{bmatrix}. \tag{1.60}$$

$$W_{[m,n]} = \left[I_n \otimes \delta_m^1 \; I_n \otimes \delta_m^2 \; \cdots \; I_n \otimes \delta_m^m \right]. \tag{1.61}$$

11. Let $X \in \mathbb{R}^m$, $Y \in \mathbb{R}^n$, $Z \in \mathbb{R}^p$ be columns. Prove the following equations:

$$Y \otimes X \otimes Z = \left(W_{[m,n]} \otimes I_p\right) X \otimes Y \otimes Z. \tag{1.62}$$

$$X \otimes Z \otimes Y = \left(I_m \otimes W_{[n,p]}\right) X \otimes Y \otimes Z. \tag{1.63}$$

$$Z \otimes Y \otimes X = \left(W_{[n,p]} \otimes I_m\right) \left(I_n \otimes W_{[m,p]}\right) \left(W_{[m,n]} \otimes I_p\right) X \otimes Y \otimes Z. \tag{1.64}$$

12. Let $A \in \mathscr{M}_{m\times n}$. Prove the following equations.

$$V_c(A^T) = W_{[n,m]} V_c(A).$$
$$V_r(A^T) = W_{[m,mn]} V_r(A).$$

References

1. Bates, D., Watts, D.: Relative curvature measures of nonlinearity. Journal of the Royal Statistical Society. Series B (Methodological) **42**, 1–25 (1980)
2. Boothby, W.: An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd edn. Academic Press, Inc., Orlando (1986)
3. Fudenberg, D., Tirole, J.: Game Theory. MIT Press, Cambridge (1991)
4. Horn, R., Johnson, C.: Topics in Matrix Analysis, vol. 1, 2. Cambridge Univ. Press, Cambridge (1991)
5. Hungerford, T.: Algebra. Springer-Verlag, New York (1974)
6. Ljung, L., Söderström, T.: Theory and Practice of Recursive Identification. MIT Press (1982)
7. Tsai, C.: Contributions to the Design and Analysis of Nonlinear Models. Ph.D. thesis, Univ. of Minisota (1983)
8. Zhang, X.: Matrix Analysis and Applications. Tsinghua Univ. Press & Springer, Beijing (2004)