# Chapter 15
# Multi-variable Polynomials

This chapter considers the matrix expression of multi-variable polynomials via semi-tensor product. Under this expression the differential of functional matrices is introduced and its calculation formulas are obtained. Then we consider the expression under two different generators and the conversion formula is presented. As an application the Taylor expansion of multi-variable functions is investigated. Finally, the formula is presented for calculating Lie derivatives. This expression provides a convenient tool to solve nonlinear problems via matrices, and then it is ready to use computers.

## 15.1 Matrix Expression of Multi-variable Polynomials

Let $x = (x_1, \cdots, x_n)$ be a set of coordinate variables on $\mathbb{R}^n$. Denote the set of $k$-th degree homogeneous polynomials by $B_n^k$. Denote by $B_n^0 = \mathbb{R}$, it represents the zero-degree polynomials, that is, constants.

It is easy to see that the components of $x^k$ form a generator of $B_n^k$, that is, this set contains a basis of $B_n^k$. Precisely speaking, let $f(x) \in B_n^k$. Then there exists an matrix $F \in M_{1\times n^k}$ such that

$$f(x) = F \ltimes x^k. \tag{15.1}$$

But this generator is not a basis, because it contains some redundant elements. In other words, the elements in this generator are not linearly independent. It follows that the $F$ in (15.1) is not unique.

*Example 15.1.* Let $f(x) \in B_2^3$ be

$$f(x) = x_1^3 + x_1^2 x_2 - 2x_1 x_2^2 - x_2^3. \tag{15.2}$$

Note that

$$x = (x_1,\, x_2)^{\mathrm{T}}, \quad x^2 = (x_1^2,\, x_1x_2,\, x_2x_1,\, x_2^2)^{\mathrm{T}},$$
$$x^3 = (x_3,\, x_1^2x_2,\, x_1x_2x_1,\, x_1x_2^2,\, x_2x_1^2,\, x_2x_1x_2,\, x_2^2x_1,\, x_2^3)^{\mathrm{T}}.$$

Then $f(x)$ can be expressed as

$$f(x) = (1\ 1\ 0\ -2\ 0\ 0\ 0\ -1) \ltimes x^3. \tag{15.3}$$

Since the generator has redundant elements, the coefficient matrix is not unique. For instance, an alternative expression of $f(x)$ can be

$$f(x) = (1\ 1\ 0\ -1\ 0\ -1\ 0\ -1) \ltimes x^3. \tag{15.4}$$

An obvious advantage of the above semi-tensor product based matrix expression of multi-variable polynomial is that the semi-tensor product has many nice properties such as associativity etc., which provide us a convenient way to manipulate the polynomials. For instance, we may factorize the polynomials in the way of single-variable polynomials. Let us see the following example.

*Example 15.2.* Assume $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$, and the corresponding coefficient matrices have proper dimensions.

1. $$F_3 \ltimes x^3 + F_5 \ltimes x^5 = (F_3 + F_5 \ltimes x^2) \ltimes x^3.$$
2. $$(A \ltimes x)^2 - 1 = (A \ltimes x + 1)(A \ltimes x - 1).$$

Since $x^k$ has redundant elements, which makes the coefficient matrix $F$ ununique, we may try to pose some restrictions on $F$. Recall that in Linear Algebra a quadratic function $f(x) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} a_{i,j}x^ix^j$ can be expressed as

$$f(x) = x^{\mathrm{T}}Ax,$$

where $A = (a_{i,j})$ is not unique. But if we require $A$ being symmetric, then it is unique. This is a convention in Linear Algebra. Similarly, we can define a symmetric expression of (15.1).

**Definition 15.1.** Let $f(x) \in B_n^k$ be expressed as in (15.1) with coefficient matrix $F$. Assume the elements of $F$ are labeled by multi-index $Id(i_1, \cdots, i_k; n, \cdots, n)$. $F$ is said to be a symmetric coefficient matrix, if

$$F_{i_{\sigma(1)},\cdots,i_{\sigma(k)}} = F_{i_1,\cdots,i_k}, \quad \forall\, \sigma \in \mathbf{S}_k,$$

where $\mathbf{S}_k$ is the $k$-th order symmetric group.

*Remark 15.1.* 1. If the $F$ in (15.1) is symmetric, then it is easy to verify that it is unique.

2. If $f(x)$ is a quadratic form and $F$ is symmetric, i.e., $x \in \mathbb{R}^n$ and

$$f(x) = Fx^2 = (F_{11}, F_{12}, \cdots, F_{1n}, \cdots, F_{n1}, F_{n2}, \cdots, F_{nn})x^2,$$

where $F_{ij} = F_{ji}$, then $V_c^{-1}(F) = V_r^{-1}(F)$ is a symmetric matrix. That is, the symmetric coefficient matrix is a generalization of symmetric matrix.
3. In fact, we may consider $f(x)$ as a tensor as $f \in T^k(\mathbb{R}^n)$. That is,

$$f(z_1, \cdots, z_k) := F \ltimes z_1 \ltimes \cdots \ltimes z_k,$$

where $z_1 = \cdots = z_k = x$. Hence, a homogeneous multi-variable polynomial has tensor structure. The advantage of using semi-tensor product to tensor calculation has already been discussed in Chapter 2.

In the following discussion symmetric group is a fundamental tool. Symmetric group has been discussed in Chapter 1. Briefly speaking, $\mathbf{S}_n$ consists of all the permutations of $n$ elements. Hence $|S_n| = n!$. For instance, $|\mathbf{S}_5| = 5! = 120$. In the following we give an example for the permutation and the index permutation.

*Example 15.3.* 1. Let $\sigma_1, \sigma_2 \in \mathbf{S}_6$ be

$$\sigma_1 = (123)(456), \quad \sigma_2 = (15642).$$

Then the product

$$\sigma_2\sigma_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 & 5 & 6 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 3 & 5 & 6 & 4 & 2 \end{bmatrix}.$$

It can be briefly denoted as $\sigma_2\sigma_1 = (23546)$.
2. Let $F_{i_1,i_2,i_3,i_4,i_5,i_6}$ is labeled by multi-index $Id(i_1, i_2, i_3, i_4, i_5, i_6; n, n, n, n, n, n)$. For instance, $F_{i_1,i_2,i_3,i_4,i_5,i_6} = F_{1,3,2,3,1,1}$, that is, $(i_1, i_2, i_3, i_4, i_5, i_6) = (1, 3, 2, 3, 1, 1)$, Then we use $F_\sigma$ to represent the elements after subscript index permutation $F_{i_{\sigma(1)},\cdots,i_{\sigma(k)}}$. For instance, consider $a = F_{1,3,2,3,1,1}$ and take $\sigma = \sigma_1$. Since $\sigma_1(1) = 2$, and $i_2 = 3$, then correspondingly, $a$ has the first index "3"; similarly, since $\sigma_1(2) = 3$, and $i_3 = 2$, the second index is "2", and so on. Finally, we have $a = F_{\sigma_1} = F_{3,2,1,1,1,3}$. Similarly, we have $a = F_{\sigma_2} = F_{1,1,2,3,1,3}$, $a = F_{\sigma_2\sigma_1} = F_{1,2,1,1,1,3}$.

Next, we consider how to convert the $f(x)$ in (15.1) to symmetric form. A simple way is to collect terms by hand. But to deal with theoretical analysis or to use computer to carry out the conversion we need a general formula. To this end, we need some new concepts and notations.

**Definition 15.2.** Let $I = (I_1, I_2, \cdots, I_k) \in Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k})$. Then

1. the permutating set of $I$, denoted by $P_I$, is defined as

$$P_I = \{J \in Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k}) \mid (J_1, \cdots, J_k) = (I_{\sigma(1)}, \cdots, I_{\sigma(k)}),\ \sigma \in \mathbf{S}_k\};$$

2. the index frequency of $I$, denoted by $C_I = (c_1, \cdots, c_n) \in \mathbb{Z}_+^n$, is defined as follows: $c_j$ is the times of $j$ appearing into $\{I_1, I_2, \cdots, I_k\}$. Note that $1 \le j \le n$.

Denote the cardinality (size) of $P_I$ by $|P_I|$, and set

$$C_I! = \prod_{j=1}^{n} c_j!.$$

Then we have

$$|P_I| = \frac{k!}{C_I!}. \tag{15.5}$$

*Example 15.4.* Let $k = 4$, $n = 5$, $I = (2,5,2,3) \in Id(i_1, i_2, i_3, i_4; n^4)$. Then

$$P_I = \left\{ \begin{matrix} (2,2,3,5) & (2,2,5,3) & (2,3,2,5) & (2,3,5,2) & (2,5,2,3) & (2,5,3,2) \\ (3,2,2,5) & (3,2,5,2) & (3,5,2,2) & (5,2,2,3) & (5,2,3,2) & (5,3,2,2) \end{matrix} \right\},$$

Since there is no 1 in $I$, we have $c_1 = 0$. Similarly, we have $c_2 = 2$, $c_3 = 1$, $c_4 = 0$, $c_5 = 1$. Finally, we have $C_I = (0,2,1,0,1)$, and hence

$$|P_I| = \frac{k!}{C_I!} = \frac{4!}{0!2!1!0!1!} = 12.$$

The following result comes from the above definitions and notations.

**Proposition 15.1.** *Let $P(x) = Fx^k$, $x \in \mathbb{R}^n$ be a $k$-th homogeneous polynomial with its symmetric form as $P(x) = \tilde{F}x^k$. Assume both $F$ and $\tilde{F}$ are labeled by $Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k})$, then for $I = (i_1, \cdots, i_k) \in Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k})$ we have*

$$\tilde{F}_I = \frac{1}{|P_I|} \sum_{J \in P_I} F_J. \tag{15.6}$$

**Definition 15.3.** Let $F = \{F_{i_1, \cdots, i_k}\} \in \mathbb{R}^{n^k}$ be a set of $n^k$ numbers, labeled by the multi-index $Id(i_1, \cdots, i_k; \underbrace{n, \cdots, n}_{k})$. The equation (15.6) defines a mapping $\tilde{F} = \psi_n^k(F) : \mathbb{R}^{n^k} \to \mathbb{R}^{n^k}$. Then the mapping $\psi$ is called a symmetrization on $\mathbb{R}^{n^k}$. .

Note that it is easy to verify by definition that both $\psi_n^1$ and $\psi_n^0$ are identity mapping.

The following proposition is an immediate consequence of the definition.

**Proposition 15.2.** *Let*

$$P(x) = P_0 + P_1 x + P_2 x^2 + \cdots + P_k x^k, \quad x \in \mathbb{R}^n$$

*be a k-th degree polynomial.* $P(x) \equiv 0$*, if and only if,*

$$\psi_n^i(P_i) = 0, \quad i = 0, 1, \cdots, k.$$

In fact, $\psi$ is a linear mapping, hence it can be expressed by a matrix. When $s \geq 2$, we define an $n^s \times n^s$ matrix $\Psi_n^s$ as following: using indices $\{i_1, \cdots, i_s\}$ to label its rows and columns in the order of $Id(i_1, \cdots, i_s; n, \cdots, n)$. Let $J = (J_1, \cdots, J_s)$ be its row index and $I = (I_1, \cdots, I_s)$ its column index. Then we assign the elements of $\Psi_n^s = (\psi_{J,I})$ by

$$\psi_{J,I} = \begin{cases} 0, & J \notin P_I, \\ \frac{1}{|P_I|}, & J \in P_I. \end{cases}$$

As for $s \leq 1$ we set $\Psi_n^0 = 1$ and $\Psi_n^1 = I_n$. Then these $\Psi_n^k$ are the matrices of the mappings $\psi_n^k$.

**Proposition 15.3.** *A k-th degree polynomial*

$$P(x) = P_0 + P_1 x + P_2 x^2 + \cdots + P_k x^k, \quad x \in \mathbb{R}^n$$

*has its symmetric expression as*

$$P(x) = \tilde{P}_0 + \tilde{P}_1 x + \tilde{P}_2 x^2 + \cdots + \tilde{P}_k x^k,$$

*where*

$$\tilde{P}_i = P_i \Psi_n^i, \quad i = 0, 1, \cdots, k. \tag{15.7}$$

*Moreover,* $P(x) \equiv 0$*, if and only if*

$$P_i \Psi_n^i = 0, \quad i = 0, 1, \cdots, k.$$

*Example 15.5.* 1. Let $n = 3$ and $k = 2$. Then $\Psi_3^2$ is

$$\Psi_3^2 = \begin{array}{c} \begin{array}{ccccccccc} 11 & 12 & 13 & 21 & 22 & 23 & 31 & 32 & 33 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ (11) \\ (12) \\ (13) \\ (21) \\ (22) \\ (23) \\ (31) \\ (32) \\ (33) \end{array}.$$

2. Let $n=2$ and $k=3$. Then $\Psi_2^3$ is

$$\Psi_3^2 = \begin{array}{cc} \begin{array}{cccccccc} 111 & 112 & 121 & 122 & 211 & 212 & 221 & 222 \end{array} & \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{c} (111) \\ (112) \\ (121) \\ (122) \\ (211) \\ (212) \\ (221) \\ (222) \end{array} \end{array}.$$

3. Consider the polynomial (15.2). It is easy to check that both (15.3) and (15.4) are not symmetric. Choosing any one and using its coefficient matrix $F$, and using formula (15.7), we can get the symmetric coefficient matrix of (15.2) as

$$f(x) = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} & -1 \end{bmatrix} \ltimes x^3.$$

Let $P[x]$, $x \in \mathbb{R}^n$ be the vector space of real polynomials. Then its is a direct sum of the spaces of $k$-th homogeneous polynomials with $k = 0, 1, \cdots$. That is,

$$P[x] = B_n^0 \oplus B_n^1 \oplus B_n^2 \oplus \cdots. \tag{15.8}$$

**Proposition 15.4.** *Let $P(x) \in B_n^p$, $Q(x) \in B_n^q$ with $P(x) = M_P x^p$ $Q(x) = M_Q x^q$. Then their product $P(x)Q(x) \in B_n^{p+q}$ has its coefficient matrix as*

$$M_{PQ} = M_P \ltimes M_Q. \tag{15.9}$$

*Proof.*

$$\begin{aligned} M_P \ltimes M_Q \ltimes x^{p+q} &= M_P \ltimes M_Q \ltimes x^q \ltimes x^p = M_P \ltimes (M_Q \ltimes x^q) \ltimes x^p \\ &= (M_Q \ltimes x^q) \ltimes (M_P \ltimes x^p) = M_{PQ} \ltimes x^{p+q}. \end{aligned}$$

□

The following is an immediate consequence of Proposition 15.4.

**Proposition 15.5.** *Let $M_P \in \mathbb{R}^{n^p}$ and $M_Q \in \mathbb{R}^{n^q}$ be two row vectors. Then*

$$M_P \ltimes M_Q \ltimes x^{p+q} = M_Q \ltimes M_P \ltimes x^{p+q}. \tag{15.10}$$

Since $x^k$ is not a basis of $B_n^k$, (15.10) does not imply that $M_P \ltimes M_Q = M_Q \ltimes M_P$.

According to Proposition 15.4, all the algebraic formulas in Elementary Algebra can be extended to the corresponding multi-variable case. For instance, since $(a+b)^2 = a^2 + 2ab + b^2$, it is also true that

$$(F_1 \ltimes x + F_2 \ltimes x^2)^2 = F_1^2 \ltimes x^2 + 2F_1 \ltimes F_2 x^3 + F_2^2 \ltimes x^4.$$

Equality (15.10) has the following generalization:

**Proposition 15.6.** *Let $F \in R^{n^{\alpha}}$ be a row vector and $t > \alpha$. Then for any row vector $G \in R^{n^{t-\alpha}}$*

$$(G \ltimes F) \ltimes x^t = F \ltimes x^{\alpha} \ltimes G \ltimes x^{t-\alpha}. \tag{15.11}$$

We give an example to demonstrate the advantage of this expression.

*Example 15.6.* Given a polynomial

$$P(x) = A_0 + A_1 x + A_2 x^2 + \cdots + A_t x^t, \quad x \in \mathbb{R}^n.$$

We want to check whether it has a linear factor $Q(x) = C - Dx$, where $C \neq 0$.

For simplicity, we set $C = 1$. Assume we have

$$\begin{aligned}&(1 - Dx)(B_0 + B_1 x + \cdots + B_{t-1} x^{t-1})\\ &\quad = A_0 + A_1 x + A_2 x^2 + \cdots + A_t x^t.\end{aligned} \tag{15.12}$$

Comparing the coefficients on both sides of (15.12) yields

$$\begin{cases} B_0 = A_0, \\ B_k = A_k + DB_{k-1}, \quad k = 1, \cdots, t-1. \end{cases} \tag{15.13}$$

To satisfy (15.12), if and only if the inductively defined $B_{t-1}$ satisfies

$$-DB_{t-1} x^t = A_t x^t. \tag{15.14}$$

Calculating $B_{t-1}$ from (15.13) and plugging it into (15.14), we have

$$(A_t + DA_{t-1} + D^2 A_{t-2} + \cdots + D^t A_0) x^t = 0. \tag{15.15}$$

Using Example 15.6 and Proposition 15.3, we have the following proposition.

**Proposition 15.7.** *The polynomial $P(x) = A_0 + A_1 x + A_2 x^2 + \cdots + A_t x^t$ has a linear factor $1 - Dx$, if and only if*

$$(A_t + DA_{t-1} + D^2 A_{t-2} + \cdots + D^t A_0) \Psi_n^t = 0. \tag{15.16}$$

Since the expression of a polynomial is not unique, in the above factorization we have to make some additional calculations. Note that for a zero term, it could be expressed into the sum of some non-zero terms. But it does not affect the final result, because in the whole calculation process there are only additions and productions. Then the zero terms remain zero at the final result. Hence the above factorization is independent the choice of equivalent coefficients. Here we say the coefficients $F_1^t$ $F_2^t$ of $x^t$ are equivalent, if $\psi_n^t(F_1^t - F_2^t) = 0$, that is, $(F_1^t - F_2^t)\Psi_n^t = 0$, because now we have $F_1^t x^t = F_2^t x^t$.

Note that this result is similar to the case of single variable case. The only difference is the "scalar time" in single-variable case is replaced by semi-tensor product for multi-variable case.

The following is a numerical example.

*Example 15.7.* Given a polynomial

$$\begin{aligned} P(x) &= 1-(3x+2y-z)+2x^2-3xy+z^2+6x^2y-2x^2z \\ &\quad +7xy^2-9xyz+xz^2-2y^2z+y^3+z^3, \\ Q(x) &:= 1-D(x)=1-(x+y-z), \end{aligned}$$

where $x=(x,y,z)^{\mathrm{T}} \in \mathbb{R}^3$. Then

$$D=(1\ \ 1\ \ -1),$$

$$A_1=(-3\ \ -2\ \ 1), \quad A_2=(2\ \ -3\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1),$$

$$A_3=(0\,6\,-2\,0\,7\,-9\,0\,0\,1\,0\,0\,0\,0\,1\,-2\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1).$$

Assume there is an $R(x)$ such that $Q(x)R(x)=P(x)$. Then $R(x)=1+B_1x+B_2x^2$. According to (15.13), we can calculate $B_1$ and $B_2$ as follows:

$$B_1=A_1+D=(-3\ \ -2\ \ 1)+(1\ \ 1\ \ -1)=(-2\ \ -1\ \ 0),$$

$$B_2=A_2+DB_1=(0\ \ -5\ \ 2\ \ -1\ \ -1\ \ 1\ \ 0\ \ 0\ \ 1).$$

Hence

$$\begin{aligned} DB_2 = (&0\,0\,0\,-5\,-5\,5\,2\,2\,-2\,-1\,-1\,1\,-1\,-1\,1 \\ &1\,1\,-1\,0\,0\,0\,0\,0\,0\,1\,1\,-1). \end{aligned}$$

It is clear that $(A_3+DB_2)\Psi_3^3=0$. Then we have

$$\begin{aligned} R(x) &= 1+B_1x+B_2x^2=1-(2x+y) \\ &\quad +(-6xy-y^2+2xz+yz+z^2). \end{aligned}$$

Next, we consider a polynomial mapping. An $m$-dimensional and $k$-th order polynomial mapping can be expressed as

$$P(x)=A_0+A_1x+A_2x^2+\cdots+A_kx^k; \quad x \in \mathbb{R}^n,\ P(x) \in \mathbb{R}^m,$$

where $A_j \in M_{m \times n^j}$, $j=0,\cdots,k$.

In the follow we consider the semi-tensor product of two polynomial mappings.

**Proposition 15.8.** *Let* $P(x)=A_0+A_1x+A_2x^2+\cdots+A_px^p$, $Q(x)=B_0+B_1x+B_2x^2+\cdots+B_qx^q$. *Then*

$$P(x) \ltimes Q(x) = \sum_{i=0}^{p+q} \sum_{j=0}^{i} A_j \otimes B_{i-j} x^i. \tag{15.17}$$

*Proof.* Recall (2.19) in Example 2.4, for two vectors $x$ and $y$ we have

$$(Ax) \ltimes (By) = (A \otimes B)(x \ltimes y).$$

Using this formula and the distributive law of the semi-tensor product, (15.17) follows. □

The following example shows that semi-tensor product is a powerful tool in dealing with the polynomial matrices.

*Example 15.8.* Let $A(x)$ be a square matrix with its entries as $t$-th degree polynomials, where $x \in \mathbb{R}^n$. Let $B(x) = I - Dx$ be a linear form, where $D \in M_{n\times n^2}$. The question is: when there exists a polynomial matrix $C(x)$, such that $A(x) = B(x)C(x)$? We are looking a similar result as in Proposition 15.7.

First, we show that $A(x)$ can be expressed as

$$A(x) = A_0 + A_1 x + \cdots + A_t x^t. \tag{15.18}$$

To this end, we first express each entries of $A$ into the sum of homogeneous matrix forms as before. That is, $A(x)$ can be expressed as

$$A(x) = A_0 + A^1 + \cdots + A^t,$$

where

$$A^k = \begin{bmatrix} a_{11}x^k & \cdots & a_{1n}x^k \\ \vdots & \ddots & \vdots \\ a_{n1}x^k & \cdots & a_{nn}x^k \end{bmatrix}, \quad a_{ij} \in M_{1\times n^k}, \quad k = 1,2,\cdots,t.$$

It is easy to verify that

$$A^k = E_k(I_n \otimes x^k) = E_k \rtimes x^k, \tag{15.19}$$

where

$$E_k = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M_{n\times n^{k+1}}.$$

Using (4.21) of Proposition 4.3 to (15.19), we have

$$E_k \rtimes x^k = E_k \ltimes W_{[n^k,n]} \ltimes x^k \ltimes W_{[n,1]}.$$

Note that

$$W_{[n,1]} = W_{[1,n]} = I_n. \tag{15.20}$$

Using (15.20) and setting

$$A_k = E_k \ltimes W_{[n^k,n]}, \quad k = 1,\cdots,t$$

yield (15.18).

Now similar to polynomial case, we assume there exists $C(x) = C_0 + C_1x + \cdots + C_{t-1}x^{t-1}$, such that $A(x) = B(x)C(x)$. Comparing the coefficients of both sides of

$$(I - Dx)(C_0 + C_1x + \cdots + C_{t-1}x^{t-1}) = A_0 + A_1x + \cdots + A_tx^t,$$

we have

$$\begin{cases} C_0 = A_0, \\ C_kx^k = A_kx^k + DxC_{k-1}x^{k-1}, \quad k = 1, \cdots, t. \end{cases} \tag{15.21}$$

Note that

$$xC_{k-1} = (I_n \otimes C_{k-1}) \ltimes x.$$

Plugging it into (15.21) yields that

$$\begin{cases} C_0 = A_0, \\ C_k = A_k + D \rtimes C_{k-1}, \quad k = 1, \cdots, t-1. \end{cases} \tag{15.22}$$

An iterative calculation shows that

$$C_{t-1} = \sum_{i=1}^{t} D^{\rtimes(i-1)} \rtimes A_{t-i},$$

where $D^{\rtimes i} = \underbrace{D \rtimes D \rtimes \cdots \rtimes D}_{i}$. Finally, matching the coefficients of $x^t$, we need

$$Dx \ltimes C_{t-1}x^{t-1} = (D \rtimes C_{t-1})x^t + A_tx^t = 0.$$

That is,

$$\left[\sum_{i=0}^{t} D^{\rtimes i} \rtimes A_{t-i}\right] x^t = 0, \tag{15.23}$$

which assures that $B(x)$ is a left factor of $A(x)$.

Next, we consider when there is a matrix $C(x)$, such that $A(x) = C(x)B(x)$? A similar argument shows that the condition is

$$\left[\sum_{i=0}^{t} A_{t-i} \rtimes D^{\rtimes i}\right] x^t = 0. \tag{15.24}$$

Summarizing the results obtained in the above example, we have the following conclusion.

**Theorem 15.1.** *Assume $A(x)$ is a square matrix with its entries as $t$-th degree polynomials with $x \in \mathbb{R}^n$. Let $B(x) = I - Dx$ be a linear form. Then $B(x)$ is a left factor* *of* $A(x)$*, if and only if*

$$\left[\sum_{i=0}^{t} D^{\rtimes i} \rtimes A_{t-i}\right] \Psi_n^t = 0. \tag{15.25}$$

$B(x)$ *is a right factor of* $A(x)$*, if and only if*

$$\left[\sum_{i=0}^{t} A_{t-i} \rtimes D^{\rtimes i}\right] \Psi_n^t = 0. \tag{15.26}$$

*Proof.* In fact (15.25) and (15.26) are from (15.23) and (15.24) respectively. The only thing which is different from the polynomial case is how to multiply the symmetrization matrix $\Psi_n^t$. What do we need now is to multiply it to each entries of the matrix. That is, it must be multiplied to $(x_{i,1}, x_{i,n+1}, x_{i,2n+1}, \cdots)$ . We, hence, need

$$\left[\sum_{i=0}^{t} D^{\rtimes i} \rtimes A_{t-i}\right] (\Psi_n^t \otimes I_n) = 0,$$

etc. According to Proposition 2.4, the above equation is equivalent to (15.25). □

## 15.2 Differential Form of Functional Matrices

**Definition 15.4.** 1. Let $f(x) : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. Its differential is defined as

$$Df(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n}\right); \tag{15.27}$$

and its gradient is defined as

$$\nabla f(x) = (Df(x))^{\mathrm{T}}. \tag{15.28}$$

2. Let $M(x) \in M_{p\times q}$ be a functional matrix with its entries as functions of $x \in \mathbb{R}^n$. The differential of $M(x)$ , denoted by $DM(x) \in \mathscr{M}_{p\times nq}$, is obtained by replacing each entries $m_{i,j}$ of $M(x)$ by their differentials $\left(\frac{\partial d m_{i,j}}{\partial x_1}, \cdots, \frac{\partial m_{i,j}}{\partial x_n}\right)$. That is,

$$DM(x) = \begin{bmatrix} \frac{\partial m_{11}(x)}{\partial x_1} & \cdots & \frac{\partial m_{11}(x)}{\partial x_n} & \cdots & \frac{\partial m_{1n}(x)}{\partial x_1} & \cdots & \frac{\partial m_{1n}(x)}{\partial x_n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \frac{\partial m_{n1}(x)}{\partial x_1} & \cdots & \frac{\partial m_{n1}(x)}{\partial x_n} & \cdots & \frac{\partial m_{nn}(x)}{\partial x_1} & \cdots & \frac{\partial m_{nn}(x)}{\partial x_n} \end{bmatrix}; \tag{15.29}$$

and its gradient , denoted by $\nabla M(x) \in \mathscr{M}_{np\times q}$, is obtained by replacing each entries $m_{i,j}$ of $M(x)$ by their gradients $\left(\frac{\partial dm_{i,j}}{\partial x_1}, \cdots, \frac{\partial m_{i,j}}{\partial x_n}\right)^{\mathrm{T}}$.

The higher order differentials or gradients can be defined inductively as

$$
\begin{aligned}
D^{k+1}M &= D(D^kM) \in M_{p\times n_{k+1}q}, \quad k \geq 1, \\
\nabla^{k+1}M &= \nabla(\nabla^kM) \in M_{pn_{k+1}\times q}, \quad k \geq 1.
\end{aligned}
$$

To show the usefulness of these definitions we give some basic discussions about pseudo Hamiltonian control system [5].

*Example 15.9.* 1. In $\mathbb{R}^n$ an $n\times n$ functional matrix $M(x)$ is given. It can be used to define a pseudo Poisson bracket : $C^\infty(\mathbb{R}^n)\times C^\infty(\mathbb{R}^n)\to C^\infty(\mathbb{R}^n)$, which has $M(x)$ as its structure matrix, as

$$
\{f(x), g(x)\} = Df(x)M(x)\nabla g(x), \quad f(x), g(x) \in C^\infty(\mathbb{R}^n). \tag{15.30}
$$

$\mathbb{R}^n$ with the structure matrix $M(x)$ is called a pseudo Poisson manifold.

2. In a pseudo Poisson manifold let $f(x)$ be a smooth function with $x\in\mathbb{R}^n$. A pseudo Hamilton vector field , with $f(x)$ as its pseudo Hamiltonian function is defined as

$$
X_F(x) = M(x)\nabla F(x). \tag{15.31}
$$

3. The following dynamic control system is called a pseudo Hamiltonian system:

$$
\dot{x} = X_f(x), \quad x \in \mathbb{R}^n. \tag{15.32}
$$

$$
\tag{15.33}
$$

Consider a control system as

$$
\begin{cases}
\dot{x} = X_f(x) + \sum\limits_{i=1}^m X_{g_i}(x)u_i(x), \quad x\in\mathbb{R}^n \\
y = h(x) = (\{g_1, f\}, \{g_2, f\}, \cdots, \{g_m, f\})^T.
\end{cases} \tag{15.34}
$$

If we can find controls $u_i(x)$, $i = 1, \cdots, m$, such that the closed-loop system is a pseudo Hamilton system with the original structure matrix $M(x)$ unchanged, then it is called a pseudo Hamilton realization.
It is clear that what do we need is

$$
\sum_{i=1}^m X_{g_i}(x)u_i(x) = M(x)\nabla\phi(x). \tag{15.35}
$$

4. Assume $dg_i$, $i = 1, \cdots, m$ are linearly independent. Then we can find $\{z_j(x),$ $j = 1,\ \cdots,\ n-m\}$, such that $\{dg_i,\ i = 1, \cdots, m,\ dz_j(x),\ j = 1, \cdots, n-m\}$

are linearly independent. By Implicit Function Theory, we can locally express $\phi(x) = \phi(g_i, z_j)$.
In addition, assume $M(x)$ is nonsingular, then

$$\nabla\phi(x) = \sum_{i=1}^{m} \nabla G_i \frac{\partial \phi(G_i, z_j)}{\partial G_i} + \sum_{j=1}^{n-m} \nabla z_j \frac{\partial \phi(G_i, z_j)}{\partial z_j} = \sum_{i=1}^{m} \nabla G_i u_i.$$

Hence

$$\begin{cases} \dfrac{\partial \phi(G_i, z_j)}{\partial z_j} = 0, & j = 1, \cdots, n-m, \\ u_i = \dfrac{\partial \phi(G_i, z_j)}{\partial G_i}, & i = 1, \cdots, m. \end{cases}$$

We conclude that assume $M(x)$ is nonsingular, then the system (15.34) has a local pseudo Hamilton realization, if and only if there exists a smooth function $\phi = \phi(g_1, \cdots, g_m)$, such that

$$u_i = \frac{\partial \phi(g_1, \cdots, g_m)}{\partial G_i}, \quad i = 1, \cdots, m. \tag{15.36}$$

5. Assume the system (15.34) is output detectable [4], that is the output satisfies $\lim_{t\to\infty} h(x(t)) = 0$, $\lim_{t\to\infty} x(t) = 0$. Then we can assume there is a Lyapunov function $L = h^{\mathrm{T}}(x)P(x)h(x)$, where $P(x) > 0$ is positive definite. Then for the closed-loop system we have

$$\dot{L} = D(h^{\mathrm{T}}(x)P(x)h(x))M(x)(\nabla F(x) + \nabla\phi(G(x))).$$

Hence if we can find $P(x) > 0$ and $\phi(g_1, \cdots, g_m)$, such that $\dot{L} < 0$, then the system is stabilized at the origin by the controls (15.36).

Above example shows that the differential of functional matrix is very useful. In the following we consider the differential of product of functional matrices. First, we consider $D(A(x) \ltimes B(x))$, where $A(x) \succ B(x)$.

**Proposition 15.9.** *Assume $A(x) \succ_t B(x)$, where $x \in \mathbb{R}^n$. Then*

$$D(A(x) \ltimes B(x)) = DA(x) \ltimes B(x) + A(x) \ltimes DB(x) \ltimes (I_s \otimes W_{[t,n]}), \tag{15.37}$$

*where $s$ is the column number of $B(x)$.*

*Proof.* Without loss of generality, we assume

$$A(x) = (a_{11}, \cdots, a_{1t}, \cdots, a_{q1}, \cdots, a_{qt})$$

is a row vector. First, we assume $s = 1$, then $B(x) = (b_1, \cdots, b_q)^{\mathrm{T}}$ is a column vector. It follows that

$$A \ltimes B = \left( \sum_{k=1}^{q} a_{k1} b_k, \cdots, \sum_{k=1}^{q} a_{kt} b_k \right).$$

$$
\begin{aligned}
&D(A \ltimes B) = \\
&\Big(\sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_1} b_k + \sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_n} b_k + \sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_n}, \cdots, \\
&\sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_1} b_k + \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_n} b_k + \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_n}\Big) \\
&= \left(\sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_1} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_n} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_1} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_n} b_k\right) \\
&\quad + \left(\sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_n}, \cdots, \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_n}\right) \\
&:= I + II.
\end{aligned}
\tag{15.38}
$$

A straightforward computation shows that

$$
(DA) \ltimes B = \left(\sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_1} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{k1}}{\partial x_n} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_1} b_k, \cdots, \sum_{k=1}^{q} \frac{\partial a_{kt}}{\partial x_n} b_k\right),
$$

which is the first part ($I$) of (15.38).

Similarly, we can calculate that

$$
A \ltimes DB = \left(\sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_1}, \cdots, \sum_{k=1}^{q} a_{k1} \frac{\partial b_k}{\partial x_n}, \cdots, \sum_{k=1}^{q} a_{kt} \frac{\partial b_k}{\partial x_n}\right).
$$

Now both the $A \ltimes DB$ and the second part ($II$) of (15.38) consist of the elements of $\sum_{k=1}^{q} a_{ki} \frac{\partial b_k}{\partial x_j}$. But in the prior the elements are arranged by the order of $Id(j,i;n,t)$, and in the later they are arranged by the orede of $Id(i,j;t,n)$. We, therefore, have

$$
(II)^{\mathrm{T}} = W_{[n,t]} (A \ltimes DB)^{\mathrm{T}}.
$$

It follows that

$$
II = (A \ltimes DB) W_{[t,n]}.
$$

As for the general case, i.e., $s > 1$, denote the $i$-th row of $B$ by $B_i$, then

$$
\begin{aligned}
&D(A \ltimes B) \\
&= (D(A \ltimes B_1), \cdots, D(A \ltimes B_s)) \\
&= (DA \ltimes B_1 + (A \ltimes DB_1) W_{[n,t]}, \cdots, DA \ltimes B_s + (A \ltimes DB_s) W_{[n,t]}) \\
&= (DA \ltimes B_1, \cdots, DA \ltimes B_s) + ((A \ltimes DB_1) W_{[n,t]}, \cdots, (A \ltimes DB_s) W_{[n,t]}) \\
&= (DA \ltimes B) + ((A \ltimes DB_1), \cdots, (A \ltimes DB_s))(I_s \ltimes W_{[n,t]}) \\
&= (DA \ltimes B) + (A \ltimes DB)(I_s \ltimes W_{[n,t]}).
\end{aligned}
$$

Hence (15.37) holds. □

We use the following example to depict the application of the above formula.

*Example 15.10.* Assume

$$A = \begin{bmatrix} x_1^2 & x_1x_2 & x_1x_2 & x_2^2 \\ x_2^2 & x_1x_2 & x_1x_2 & x_1^2 \end{bmatrix}, \quad B = \begin{bmatrix} \sin(x_1+x_2) & \cos(x_1+x_2) \\ -\cos(x_1+x_2) & \sin(x_1+x_2) \end{bmatrix}.$$

Then

$$DA = \begin{bmatrix} 2x_1 & 0 & x_2 & x_1 & x_2 & x_1 & 0 & 2x_2 \\ 0 & 2x_2 & x_2 & x_1 & x_2 & x_1 & 2x_1 & 0 \end{bmatrix},$$

$$DB = \begin{bmatrix} \cos(x_1+x_2) & \cos(x_1+x_2) & -\sin(x_1+x_2) & -\sin(x_1+x_2) \\ \sin(x_1+x_2) & \sin(x_1+x_2) & \cos(x_1+x_2) & \cos(x_1+x_2) \end{bmatrix}.$$

For notational brevity, we denote $S := \sin(x_1+x_2)$, $C := \cos(x_1+x_2)$. Then

$$\begin{aligned} &D(A(x) \ltimes B(x)) \\ &= \begin{bmatrix} 2x_1 & 0 & x_2 & x_1 & x_2 & x_1 & 0 & 2x_2 \\ 0 & 2x_2 & x_2 & x_1 & x_2 & x_1 & 2x_1 & 0 \end{bmatrix} \ltimes \begin{bmatrix} S & C \\ -C & S \end{bmatrix} \\ &\quad + \begin{bmatrix} x_1^2 & x_1x_2 & x_1x_2 & x_2^2 \\ x_2^2 & x_1x_2 & x_1x_2 & x_1^2 \end{bmatrix} \ltimes \begin{bmatrix} C & C & -S & -S \\ S & S & C & C \end{bmatrix} \ltimes \left(I_2 \otimes W_{[2]}\right) := [a_{ij}], \end{aligned}$$

hence, we have

$$\begin{array}{ll} a_{11} = 2x_1S - x_2C + x_1^2C + x_1x_2S & a_{12} = -x_1C + x_1^2C + x_1x_2S \\ a_{13} = x_2S + x_1x_2C + x_1^2S & a_{14} = x_1C + 2x_2S + x_1x_2C + x_2^2S \\ a_{15} = 2x_1C + x_2S - x_1^2S + x_1x_2C & a_{16} = x_1S - x_1^2S + x_1x_2C \\ a_{17} = x_2C - x_1x_2S + x_2^2C & a_{18} = x_1C + 2x_2S - x_1x_2S + x_2^2C \\ a_{21} = -x_2C + x_2^2C + x_1x_2S & a_{22} = 2x_2S - x_1C + x_2^2C + x_1x_2S \\ a_{23} = x_2S - 2x_1C + x_1x_2C + x_1^2S & a_{24} = x_1S + x_1x_2C + x_1^2S \\ a_{25} = x_1S - x_2^2S + x_1x_2C & a_{26} = 2x_2C + x_1S - x_2^2S + x_1x_2C \\ a_{27} = x_2C + 2x_1S - x_1x_2S + x_1^2C & a_{28} = x_1C - x_1x_2S + x_1^2C. \end{array}$$

To verify the above result a direct calculation yields

$$\begin{aligned} &A(x) \ltimes B(x) \\ &= \begin{bmatrix} x_1^2S - x_1x_2C & x_1x_2S - x_2^2C & x_1^2C + x_1x_2S & x_1x_2C + x_2^2S \\ x_1^2S - x_1x_2C & x_1x_2S - x_1^2C & x_2^2C + x_1x_2S & x_1x_2C + x_2^2S \end{bmatrix}. \end{aligned}$$

Differentiating it, we have the same result as in the above.

Applying Proposition 15.9 to conventional matrix product yields the following result.

**Corollary 15.1.** *For conventional product of functional matrices we have*

$$D(A(x)B(x)) = DA(x) \ltimes B(x) + A(x)DB(x). \tag{15.39}$$

*Proof.* Note that $W_{[1,k]} = W_{[k,1]} = I_k$, $k > 0$. Then it is clear that (15.39) is a special case of (15.37). □

Next, we consider the case of $A(x) \prec_t B(x)$, $t > 1$. We intend to use the formula $A(x) \ltimes B(x) = (A(x) \otimes I_t)B(x)$.

First, we give two lemmas, which themselves are useful.

**Lemma 15.1.** *Let $A$ be a block row matrix, precisely,*

$$A = (A_{11}, \cdots, A_{1n}, \cdots, A_{k1}, \cdots, A_{kn}),$$

*where the blocks are labeled by $(i, j)$ and arranged in the order of $Id(i, j; k, n)$. Moreover, we assume each blocks have the same dimension. Then*

$$AW_{[n,k]} = (A_{11}, \cdots, A_{k1}, \cdots, A_{1n}, \cdots, A_{kn}). \tag{15.40}$$

*Similarly, assume*

$$B = \mathrm{Col}(B_{11}, \cdots, B_{1n}, \cdots, B_{k1}, \cdots, B_{kn}),$$

*and its blocks also have the same dimension. Then*

$$W_{[k,n]}B = \mathrm{Col}(B_{11}, \cdots, B_{k1}, \cdots, B_{1n}, \cdots, B_{kn}). \tag{15.41}$$

*Proof.* Assume the column number of $A_{ij}$ is $r$. Then it is easy to see that

$$AW_{[n,k]} = A \ltimes W_{[n,k]} = A(W_{[n,k]} \otimes I_r).$$

Where the $W_{[n,k]}$ has changed the arrange order of the blocks in $A$. Similarly, we can prove (15.41). □

**Lemma 15.2.** *Assume $A(x) \in M_{p\times q}$, $x \in \mathbb{R}^n$. Then*

$$D(A \otimes I_k) = (DA \otimes I_k)(I_q \otimes W_{[k,n]}). \tag{15.42}$$

*Proof.* First we split the columns of both $D(A \otimes I_k)$ and $(DA \otimes I_k)$ into $p \times q$ equal blocks, and denote them as $D(A \otimes I_k) = E = \{E_{ij}\}$ and $(DA \otimes I_k) = F = \{F_{ij}\}$ respectively. Then

$$E_{ij} = D(a_{ij}I_k), \quad F_{ij} = da_{ij} \otimes I_k.$$

Assume the columns of $F_{ij}$ are labeled by $(p, q)$ and arranged in the order of $Id(p, q; n, k)$. Then it is easy to verify that $E_{ij}$ consists of the same set of columns, they are indexed by $(p, q)$ but in the order of $Id(q, p; k, n)$. According to the Lemma 15.1 we have $E_{ij} = F_{ij}W_{[k,n]}$. Now set

$$W = \mathrm{diag}(\underbrace{W_{[k,n]}, \cdots, W_{[k,n]}}_{q}),$$

Then it is clear that $E = FW$, which leads to (15.42). □

**Proposition 15.10.** *Let $A(x) \in M_{p\times q}$, and $A(x) \prec_t B(x)$, $x \in \mathbb{R}^n$. Then*

$$D(A(x)\ltimes B(x)) = DA(x)\ltimes (I_q\otimes W_{[t,n]})\ltimes B(x) + A\ltimes DB(x). \tag{15.43}$$

*Proof.* Since $A(x)\ltimes B(x) = (A(x)\otimes I_t)B(x)$, using (15.39) and (15.42), we have

$$\begin{aligned} D(A(x)\ltimes B(x)) &= D[(A(x)\otimes I_t)B(x)] \\ &= (DA\otimes I_t)(I_q\otimes W_{[t,n]})\ltimes B(x) + A(x)\ltimes DB(x). \end{aligned}$$

Equation (15.43) follows. □

Following example shows how to use semi-tensor product and the differential of functional matrix to calculate the Lie derivatives of vector fields.

*Example 15.11.* Let $f(x), g(x)\in V(M)$ be two vector fields on $M$.

1. The Lie derivative of $g(x)$ with respect to $f(x)$ can locally expressed (in a coordinate chart) as [4]

$$\mathrm{ad}_f(g) = [f,g] = Dgf - Dfg. \tag{15.44}$$

2. Consider the second order Lie derivative. Using (15.37), we have

$$\begin{aligned} \mathrm{ad}_f^2(g) &= D(Dgf - Dfg)f - Df(Dgf - Dfg) \\ &= D^2g\ltimes f^2 + DgDff - D^2f\ltimes g\ltimes f - 2DfDgf + (Df)^2g. \end{aligned}$$

In the following we consider the higher order differential of products. We need some preparations. First, we define an $i+1$-th index as

$$S(i,k) = \left\{ d = (d_1,\cdots,d_{i+1})\in \mathbb{Z}_+^{i+1} \;\middle|\; \sum_{j=1}^{i+1} d_j = k \right\}.$$

If $d = (d_1,\cdots,d_{i+1})\in S(i,k)$, We define the differential $D_d(B)$ as

$$D_dB = D[(\cdots D(D(B\otimes I_{n^{d_1}})\otimes I_{n^{d_2}})\cdots\otimes I_{n^{d_i}})\otimes I_{n^{d_{i+1}}}]. \tag{15.45}$$

Using it, we define

$$D^{S(i,k)}B = \sum_{d\in S(i,k)} D_dB. \tag{15.46}$$

With these notations we have

**Corollary 15.2.** *For conventional matrix product we have*

$$D^k(A(x)B(x)) = \sum_{j=0}^{k} D^jA(x)D^{S(k-j,j)}(B(x)). \tag{15.47}$$

*Proof.* We prove it by mathematical induction. When $k=1$, it reduces to (15.39). Assume (15.47) holds for $k$. For $k+1$ we consider the terms of $D^{j+1}AD^{S(k-j,j+1)}B$, which are produced from $D^{k+1}(A(x)B(x))$. They come from two groups of terms of $D^k(A(x)B(x))$, namely, $D^jAD^{S(k-j,j)}B$ and $D^{j+1}AD^{S(k-j-1,j+1)}B$. Differentiating the $A(x)$ of the terms in first group and the $B(x)$ of the terms in the second group, yields the required terms. Then it suffices to prove that

$$D^{S(k-j,j)}B \otimes I_n + D[D^{S(k-j-1,j+1)}B] = D^{S(k-j,j+1)}B. \tag{15.48}$$

Note that for

$$d = (d_1, \cdots, d_{k-j}, d_{k-j+1}+1), \quad d_1 + \cdots + d_{k-j+1} = j,$$

the first group of the left hand side contains all $D_d$, meanwhile, for

$$d = (d_1, \cdots, d_{k-j}, 0), \quad d_1 + \cdots + d_{k-j} = j+1,$$

the second group of the left hand side also contains all $D_d$. The non-overlapped part of these two groups of $d$ forms the set of $S(k-j, j+1)$. Hence, one sees easily that (15.47) follows. □

We give an example for the calculation.

*Example 15.12.* Assume $k = 3$. We calculate the differential

$$\begin{cases} D^{S(3,0)}B = D^3B, \\ D^{S(2,1)}B = D^2(B \otimes I_n) + D(DB \otimes I_n) + D^2(B) \otimes I_n, \\ D^{S(1,2)}B = D(B \otimes I_{n^2}) + D(B \otimes I_n) \otimes I_n + DB \otimes I_{n^2}, \\ D^{S(0,3)}B = D^3B \otimes I_{n^3}. \end{cases}$$

Hence,

$$\begin{aligned} D^3(AB) = {} & AD^3B \\ & + DA \ltimes (D^2(B \otimes I_n) + D(DB \otimes I_n) + D^2(B) \otimes I_n) \\ & + D^2A \ltimes (D(B \otimes I_{n^2}) + D(B \otimes I_n) \otimes I_n + DB \otimes I_{n^2}) \\ & + D^3A(B \otimes I_{n^3}). \end{aligned}$$

Next, we consider the gradient.

**Lemma 15.3.** *1. Let $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$ be a column vector. Then*

$$x^{\mathrm{T}} = V_r^{\mathrm{T}}(I_n) \rtimes x. \tag{15.49}$$

*2. Let $x(x_1, \cdots, x_n) \in \mathbb{R}^n$ be a row vector. Then*

$$x^{\mathrm{T}} = x \ltimes V_r(I_n). \tag{15.50}$$

A straightforward computation can verify this lemma, and we leave it to the reader. Using this lemma, we have the following formulas, which convert the differential to gradient and vise versa.

**Proposition 15.11.** *Assume $A(x) \in M_{p\times q}$, where $x \in \mathbb{R}^n$. Then*

$$DA(x) = (I_p \otimes V_r^{\mathrm{T}}(I_n)) \rtimes \nabla A(x). \tag{15.51}$$

*Conversely, we have*

$$\nabla A(x) = DA(x) \ltimes (I_q \otimes V_r(I_n)). \tag{15.52}$$

*Proof.* Splitting both $DA(x)$ and $\nabla A(x)$ into $p \times q$ equal blocks in a natural way. Then what we have to do is to transpose each block, which is an $n$-dimensional row vector, into a column vector, or vise versa. Using (15.49) and (15.50) and the block multiplication, both formulas can be easily verified. □

Before the end of this section we give the formula for the gradient of product of matrices.

**Proposition 15.12.** *Given two functional matrices $A(x)$ and $B(x)$, where $x \in \mathbb{R}^n$.*

1. *If $A(x) \prec B(x)$, then*

$$\nabla(A(x)B(x)) = (\nabla A(x))B(x) + A(x)(\nabla B(x)). \tag{15.53}$$

2. *if $A(x) \succ_t B(x)$, precisely, let $A(x) \in M_{m\times tp}$, $B(x) \in M_{p\times q}$. Then*

$$\nabla(A(x)B(x)) = (\nabla A(x))B(x) + A(x)DB(x)[I_{qt} \otimes V_r(I_n)]. \tag{15.54}$$

*Proof.* Using the block multiplication rule of semi-tensor product, we can verify (15.53) by straightforward computation. Consider (15.54): the first term can be obtained by block multiplication; as for the second term, without loss the generality, we can assume $A(x) = A = \text{const.}$, Since $A \succ B(x)$, then

$$\nabla(AB(x)) = D(AB(x)) \ltimes (I_{qt} \otimes V_r(I_n)) = ADB(x)[I_{qt} \otimes V_r(I_n)].$$

□

## 15.3 Conversion of Generators

As we discussed before that $x^k$ is a generator of $B_n^k$, but not a basis. It follows that the $F$ in (15.1) is not unique. Hence, we need a basis. A basis of $B_n^k$, called the natural basis and denoted by $N_n^k$, is defined as

$$N_n^k := \left\{ x_1^{d_1} \cdots x_n^{d_n} \;\middle|\; \sum_{j=1}^n d_j = k \right\}.$$

For convenience $N_n^k$ is also used for the matrix consists of its elements arranging in the alphabetic order. That is,

$$x_1^{d_1} \cdots x_n^{d_n} \prec x_1^{t_1} \cdots x_n^{t_n},$$

if and only if there is an index $i \le n$ such that $d_k = t_k$, $k < i$ and $d_i > t_i$.

Briefly denote $Id(I;n^k) := Id(i_1,\cdots,i_k;\underbrace{n,\cdots,n}_{k})$. It is easy to see that the

$$x_{i_1} \cdots x_{i_k} \in x^k$$

is arranged in the order of $Id(I;n^k)$.

Define an index subset of $Id(I;n^k)$ as

$$Is(I;n^k) = \{(i_1,\cdots,i_k) \in Id(I;n^k) \,|\, i_1 \le i_2 \le \cdots \le i_k\},$$

which is called the symmetric index set. Note that for the set of elements in a generator, symmetric index can pick out different elements and ignore repeated elements. Precisely, if we pick out only the elements which have symmetric index, we pick out all the different elements without repeats. For instance, $x_1x_1x_2$, $x_1x_2x_1$, and $x_2x_1x_1$ are components (or elements) of $x_2^3$, in other words, $(112)$ $(121)$ $(211) \in Id(I;2^3)$. But only $(112) \in Is(I;2^3)$.

Recall Definition 15.2, and let $I = (i_1,\cdots,i_k) \in Id(I;n^k)$. Then the index frequency of $I$, denoted by $C_I \in \mathbb{Z}_+^n$, is defined as $C_I(k)$ is the times for $k$ appearing into $I$. For instance, assume $I = (11424) \in Id(I;6^5)$, then we have $C_I = (210200)$.

Now we define a mapping $\xi : Id(I;n^k) \to \mathbb{Z}_+^n$ by $I \mapsto C_I$. It follows from the definition that

**Proposition 15.13.** *1. Denote the region of $\xi$ as*

$$R(n^k) = \left\{ (c_1,\cdots,c_n) \in \mathbb{Z}_+^n \,\middle|\, \sum_{j=1}^n c_j = k \right\}. \tag{15.55}$$

*2. If $\xi$ is restricted on $Is(I;n^k)$, then $\xi|_{Is(I;n^k)} : Is(I;n^k) \to R(n^k)$ is an bijective.*

*3. Define the order at $R(n^k)$, denoted by $\prec$ as*

$$(c_1,\cdots,c_n) \prec (d_1,\cdots,d_n),$$

*if there is an $s \le n$ such that $c_1 = d_1$, $\cdots$, $c_{s-1} = d_{s-1}$ and $c_s > d_s$.*

*The order at $Id(I;n^k)$ is defined in the same way (only the above $n$ is now replaced by $k$). Under these orders it is easy to verify that $\xi : Is(I;n^k) \to R(n^k)$ is an order-reserve mapping.*

Note that $R(n^k)$ is also a set of indexes, which are used to label the natural basis $N_n^k$. For instance, assume $I = (11335) \in Id(I;6^5)$, then it is used to label an element in $B_n^k = B_6^5$, which is $x_1x_1x_3x_3x_5$. If we use the symmetric index set $Is(I;6^5)$ to label the elements in $B_6^5$, the indexes and the elements hare one-one correspondence. But it is still not convenient because the monomial forms are not very natural. According to the Proposition 15.13, we can use $C_I$ to label the elements in $N_n^k$. For instance, for $I=(11335)$ we have $C_I(I)=(2,0,2,0,1,0)$, which labels the element $x_1^2x_2^0x_3^2x_4^0x_5^2x_6^0=x_1^2x_3^2x^5$. Using this set of indexes, we denote the basis of $N_n^K$ as

$$\left\{ x^{C_I}=\prod_{j=1}^n x_j^{c_j} \,\middle|\, I\in Is(I;n^k) \right\}.$$

**Proposition 15.14.** *The cardinality (size) of $Is(I;n^k)$ is*

$$|Is(I;n^k)|=\frac{(n+k-1)!}{k!(n-1)!}, \quad k\ge 0,\ n\ge 1. \tag{15.56}$$

*Proof.* Since $\xi: Is(I;n^k)\to R(n^k)$ is a bijective, we need only to consider the size of

$$R(n^k)=\left\{ (c_1,\cdots,c_n)\in\mathbb{Z}_+^n \,\middle|\, \sum_{j=1}^n c_j=k \right\}.$$

Let $c_1$ runs from 0 to $k$, it obvious that

$$|R(n^k)|=\sum_{j=0}^k \left|R((n-1)^j)\right|.$$

Using equality

$$\binom{n-1}{0}+\binom{n}{1}+\cdots+\binom{n+k-1}{k}=\binom{n+k}{k}, \tag{15.57}$$

then the conclusion can be proved by mathematical induction. □

It is interesting that the cardinality of $Is(I;n^k)$ can be checked from the following Table 15.1. In this table all the elements in the first row and first column are 1. All other elements are the sum of their upper element and left element.

**Table 15.1** Cardinality of $Is(I;n^k)$

| $k\backslash \dim n$ | 1 | 2 | 3 | 4 | 5 | 6 | $\cdots$ |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | |
| 1 | 1 | 2 | 3 | 4 | 5 | | |
| 2 | 1 | 3 | 6 | 10 | | | |
| 3 | 1 | 4 | 10 | | | | |
| 4 | 1 | 5 | | | | | |
| 5 | 1 | | | | | | |
| $\vdots$ | | | | | | | |

Next, we consider how to convert the coefficient of a polynomial with resp generator $B_n^k$ to the coefficient with respect to natural basis $T_N(n,k)$ and vise versa.

Denote by $s = |Is(I;n^k)|$, $t = n^k$, then $T_B(n,k) \in M_{s\times t}$, $T_N(n,k) \in M_{t\times s}$ are constructed as follows.

1. Constructing $T_B(n,k)$:
   *Step 1*. Label the columns of an $s \times t$ matrix by $Id(J;n^k)$, and the rows by $Is(I;n^k)$.
   *Step 2* . Assign its entries as follows:
   Assume the row index is $I = (i_1, i_2, \cdots, i_k) \in Is(I;n^k)$, where $i_1 \le i_2 \le \cdots \le i_k$. For any element in this column, whose row multi-index is $J \in P_I$, that is, $J = (i_{\sigma(1)}, i_{\sigma(2)}, \cdots, i_{\sigma(k)})$, for certain $\sigma \in S_k$, precisely, the column multi-index is a permutation of its row multi-index, then set the entry to be
$$\alpha_I = \frac{C_I!}{k!} = \frac{\prod_{j=1}^n C_I(j)!}{k!},$$
   otherwise, set it to be 0.
   The matrix $T_B(n,k)$ is constructed. In other words, the element $\beta_{I,J}$ of $T_B(n,k)$ is
$$\beta_{I,J} = \begin{cases} \alpha_I, & J \in P_I, \\ 0, & \text{otherwise.} \end{cases}$$
2. Constructing $T_N(n,k)$:
   *Step 1* . Let $T$ be a $t \times s$ matrix. Label its rows by $Id(J;n^k)$, and its columns by $Is(I;n^k)$.
   *Step 2* . Assign its entries as follows:
   Assume the column multi-index is $J = (j_1, j_2, \cdots, j_k) \in Is(J;n^k)$. For each element in this column, if its row index $I \in P_J$, set it to be 1, otherwise, set it to be 0.
   The matrix $T_N(n,k)$ is constructed. In other words, the element $n_{I,J}$ of $T_N(n,k)$ is
$$n_{I,J} = \begin{cases} 1, & I \in P_J, \\ 0, & \text{otherwise.} \end{cases}$$

We simply denote the basis $N_n^k$ by $x_{(k)}$. For instance, when $n = 3$, $x_{(2)}$ is expressed as $x_{(2)} = (x_1^2,\ x_1x_2,\ x_1x_3,\ x_2^2,\ x_2x_3,\ x_3^2)^{\mathrm{T}}$.

The following proposition is an immediate consequence of the construction.

**Proposition 15.15.** *1. $x_{(k)}$ and $x^k$ satisfy the following relation:*
$$\begin{cases} x_{(k)} = T_B(n,k)x^k, \\ x^k = T_N(n,k)x_{(k)}. \end{cases} \tag{15.58}$$

*2. Assume $p(x) \in B_n^k$ is a k-th degree homogeneous polynomial, and $p(x) = Fx^k = Sx_{(k)}$, then*

$$S = F T_N(n,k). \tag{15.59}$$

*Moreover, $ST_B(n,k)$ is a symmetric expression of $F$. Particularly, if $F$ is symmetric, then $F = ST_B(n,k)$.*

*Example 15.13.* Let $n = 2$ and $k = 3$.

1. The matrix $T_B(n,k)$ is

$$T_B(2,3) = \begin{array}{c} \begin{array}{cccccccc} (111) & (112) & (121) & (122) & (211) & (212) & (221) & (222) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ (111) \\ (112) \\ (122) \\ (222) \end{array}. \tag{15.60}$$

Similarly, we have

$$T_N(2,3) = \begin{array}{c} \begin{array}{cccc} (111) & (112) & (122) & (222) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ (111) \\ (112) \\ (121) \\ (122) \\ (211) \\ (212) \\ (221) \\ (222) \end{array}. \tag{15.61}$$

2. Assume $f(x) = x_1^3 + 3x_1^2 x_2 - x_1 x_2^2 + x_2^3$, it can be rewritten as $f(x) = (1\,3\,-1\,1)x_{(3)}$. Using (15.60), we have

$$f(x) = (1\;3\;-1\;1)T_B(2,3)x^3 = \left(1\;1\;1\;-\frac{1}{3}\;1\;-\frac{1}{3}\;-\frac{1}{3}\;1\right)x^3.$$

3. Assume $f(x) = (1\;2\;1\;1\;-1\;-1\;-2\;-1)x^3$, using (15.61), we have

$$f(x) = (1\;2\;1\;1\;-1\;-1\;-2\;-1)T_N(2,3)x_{(3)} = (1\;2\;-2\;-1)x_{(3)}.$$

Then we can have the symmetric expression of $f(x)$ as

$$\begin{aligned} (1\;2\;-2\;-1)x_3 &= (1\;2\;-2\;-1)T_B(2,3)x^3 \\ &= \left(1\;\tfrac{2}{3}\;\tfrac{2}{3}\;-\tfrac{2}{3}\;\tfrac{2}{3}\;-\tfrac{2}{3}\;-\tfrac{2}{3}\;-1\right)x^3. \end{aligned}$$

## 15.4 Taylor Expansion of Multi-variable Functions

The differential of semi-tensor product of matrices provides a condensed form of Taylor expansion of multi-variable functions. The following expansion is easily verifiable.

**Theorem 15.2 (Taylor Series Expansion).** *Assume $F : \mathbb{R}^m \to \mathbb{R}^n$ is an analytic mapping, then its Taylor series expansion is*

$$F(x) = F(x_0) + \sum_{k=1}^{\infty} \frac{1}{k!} D^k F(x_0) \ltimes (x - x_0)^k, \quad x \in \mathbb{R}^n. \tag{15.62}$$

There is a similar expression in [1], where terms of $(x-x_0)^k$ are understood as tensor product of vectors. But without semi-tensor product, it can hardly be used in both theoretical analysis and numerical computation. An obvious advantage of (15.62) is: it has the same expression as single variable mappings. Another advantage is that within each term there is only one product, i.e., the semi-tensor product, hence the factors are associative etc.

As an application of the Taylor series expansion and the calculation of semi-tensor form of multi-variable polynomials, we consider the Taylor series expansion of the inverse mapping of a local diffeomorphism.

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be an analytic mapping with $F(0) = 0$. Otherwise, replace $F$ by $F - F(0)$. Using Taylor series expansion, we can express it as

$$y = F_1 x + F_2 x^2 + F_3 x^3 + \cdots, \tag{15.63}$$

where

$$F_k = \frac{1}{k!} D^k F|_0, \quad k \geq 1.$$

Using (15.17), we have

$$\begin{bmatrix} y \\ y^2 \\ \vdots \\ y^k \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1k} \\ 0 & A_{22} & A_{23} & \cdots & A_{2k} \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & A_{kk} \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} + O(\|x\|^{k+1}), \tag{15.64}$$

where $A_{1i} = F_i$, $i = 1, 2, \cdots$, and the $A_{ki}$, $k > 1$ can be calculated as

$$A_{ki} = \sum_{j_1 + j_2 + \cdots + j_k = i} (F_{j_1} \otimes F_{j_2} \otimes \cdots \otimes F_{j_k}), \quad i \geq k. \tag{15.65}$$

Assume $y = F(x)$ is a local diffeomorphism, then $F_1 = A_{11} = J_F$, as the Jacobi matrix of $F$ at the origin, is invertible. Moreover, because $A_{kk} = \underbrace{F_1 \otimes \cdots \otimes F_1}_{k}$, it is also invertible. Then it follows from (15.64) that

$$
\begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} & \cdots & B_{1k} \\ 0 & B_{22} & B_{23} & \cdots & B_{2k} \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & B_{kk} \end{bmatrix} \begin{bmatrix} y \\ y^2 \\ \vdots \\ y^k \end{bmatrix} + R_{k+1}. \tag{15.66}
$$

Denote the coefficient matrix on the right hand side of (15.66) is $B^{kk}$, then $B^{kk}$ can be inductively expressed as

$$
\begin{cases} B^{11} = B_{11} = F_1^{-1}, \\ B^{t+1,t+1} = \begin{bmatrix} B^{tt} & -B^{tt} A^{t,t+1} A^{-1}_{t+1,t+1} \\ 0 & A^{-1}_{t+1,t+1} \end{bmatrix}, \quad t \geq 1, \end{cases} \tag{15.67}
$$

where

$$
A^{-1}_{t+1,t+1} = \underbrace{F_1^{-1} \otimes \cdots \otimes F_1^{-1}}_{t+1}, \quad A^{t,t+1} = \begin{bmatrix} A_{1,t+1} \\ \vdots \\ A_{t,t+1} \end{bmatrix}.
$$

**Theorem 15.3.** *Let $y = F(x) : \mathbb{R}^n \to \mathbb{R}^n$, $F(0) = 0$ be a local diffeomorphism around the origin. Then its inverse mapping $x = F^{-1}y$ has the following Taylor series expansion as*

$$
x = B_{11}y + B_{12}y^2 + \cdots + B_{1k}y^k + O(\|y\|^{k+1}),
$$

*where $B_{1k}$ is shown in (15.67).*

*Proof.* In fact, this form is a summary of the above discussion. The only thing we need to notice is: the Remaining $R_{k+1}$ of (15.66). Note that for a local diffeomorphism $y = F(x)$, $y(0) = 0$, we have $O(\|x\|^k) = O(\|y\|^k)$, which means $R_{k+1} = O(\|y\|^{k+1})$. □

In solving practical engineering problems it is convenient to express a Taylor series expression of a multi-variable mapping over the natural form, which has no redundant terms. To this end, we define two matrices as follows.

$$
\begin{cases} T^N(n,k) = \mathrm{diag}(I_n, T_N(n,2), T_N(n,3), \cdots, T_N(n,k)), \\ T^B(n,k) = \mathrm{diag}(I_n, T_B(n,2), T_B(n,3), \cdots, T_B(n,k)). \end{cases}
$$

Using some properties of the matrices $T_B(n,k)$ and $T_N(n,k)$, it is easy to see that if

$$
\begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} = B^{kk} \begin{bmatrix} Y \\ Y^2 \\ \vdots \\ Y^k \end{bmatrix},
$$

then under the natural basis it becomes

$$\begin{bmatrix} x \\ x_{(2)} \\ \vdots \\ x_{(k)} \end{bmatrix} = T^B(n,k)B^{kk}T^N(n,k) \begin{bmatrix} Y \\ Y_{(2)} \\ \vdots \\ Y_{(k)} \end{bmatrix}. \tag{15.68}$$

*Example 15.14.* Consider a mapping $y = F(x)$ as

$$\begin{cases} y_1 = \sin(x_1) + x_2 - x_2^2, \\ y_2 = \log(1 + x_1 - x_2). \end{cases} \tag{15.69}$$

Using Taylor series expansion, (15.69) can be expressed as

$$\begin{cases} y_1 = x_1 + x_2 - x_2^2 - \frac{1}{6}x_1^3 + O(\|x\|^4), \\ y_2 = x_1 - x_2 - \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{3}(x_1 - x_2)^3 + O(\|x\|^4). \end{cases} \tag{15.70}$$

A straightforward computation yields the following result.

**Table 15.2** Coefficients of The Taylor Series Expansion

| y\x | $x_1$ | $x_2$ | $x_1^2$ | $x_1x_2$ | $x_2^2$ | $x_1^3$ | $x_1^2x_2$ | $x_1x_2^2$ | $x_2^3$ | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | 1 | 1 | -1 | 0 | 0 | -1/6 | 0 | 0 | 0 | $\cdots$ |
| $y_2$ | 1 | -1 | -1/2 | 1 | -1/2 | 1/3 | -1 | 1 | -1/3 | $\cdots$ |
| $y_1^2$ | 0 | 0 | 1 | 2 | 1 | 0 | -2 | 0 | -2 | $\cdots$ |
| $y_1y_2$ | 0 | 0 | 1 | 0 | -1 | -1/2 | 1/2 | 1/2 | -1/2 | $\cdots$ |
| $y_2^2$ | 0 | 0 | 1 | -2 | 1 | -1 | 3 | -3 | 1 | $\cdots$ |
| $y_1^3$ | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 3 | 1 | $\cdots$ |
| $y_1^2y_2$ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | -1 | -1 | $\cdots$ |
| $y_1y_2^2$ | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 1 | $\cdots$ |
| $y_2^3$ | 0 | 0 | 0 | 0 | 0 | 1 | -3 | 3 | 1 | $\cdots$ |
| $\vdots$ | | | | | | | | | | |

Then we have the coefficient matrix of the inverse mapping as in 15.3.

**Table 15.3** The coefficients of the Inverse Mapping

| x\y | $y_1$ | $y_2$ | $y_1^2$ | $y_1y_2$ | $y_2^2$ | $y_1^3$ | $y_1^2y_2$ | $y_1y_2^2$ | $y_2^3$ |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0.5 | 0.5 | 0.125 | 0.25 | 0.375 | 0.0208 | 0.125 | 0.0625 | 0.2083 |
| $x_2$ | 0.5 | -0.5 | 0.125 | 0.25 | -0.125 | 0.0764 | -0.0417 | 0.2294 | -0.0139 |
| $x_1^2$ | 0 | 0 | 0.25 | 0.5 | 0.25 | 0.0833 | 0 | 0.25 | 0.1667 |
| $x_1x_2$ | 0 | 0 | 0.25 | 0 | -0.25 | 0.25 | -0.5 | 0.5 | -0.5 |
| $x_2^2$ | 0 | 0 | 0.25 | -0.5 | 0.25 | 0.0833 | 0 | -0.25 | 0.1667 |
| $x_1^3$ | 0 | 0 | 0 | 0 | 0 | 0.0833 | 0.5 | 0.25 | 0.1667 |
| $x_1^2x_2$ | 0 | 0 | 0 | 0 | 0 | 0.1667 | 0 | 0 | -0.1667 |
| $x_1x_2^2$ | 0 | 0 | 0 | 0 | 0 | 0.0833 | 0 | -0.25 | 0.1667 |
| $x_2^3$ | 0 | 0 | 0 | 0 | 0 | 0.1667 | -0.5 | 0.5 | -0.1667 |
| $\vdots$ | | | | | | | | | |

From the first block of the Table 15.3 we can find the inverse mapping of $y = F(x)$, defined in (15.69), as follows.

$$
\begin{aligned}
x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= F^{-1}(y) \\
&= \begin{bmatrix} 0.5y_1 + 0.5y_2 + 0.125y_1^2 + 0.25y_1y_2 + 0.375y_2^2 \\ +0.0208y_1^3 + 0.125y_1^2y_2 + 0.0625y_1y_2^2 + 0.2083y_2^3 \\ \\ 0.5y_1 - 0.5y_2 + 0.125y_1^2 + 0.25y_1y_2 - 0.125y_2^2 \\ +0.764y_1^3 - 0.0417y_1^2y_2 + 0.2292y_1y_2^2 - 0.0139y_2^3 \end{bmatrix} + O(\|y\|^4).
\end{aligned}
$$

## 15.5 Fundamental Formula of Differential

Consider an analytic function $h(x) \in C^{\omega}(\mathbb{R}^n)$. Using Taylor series expansion, we have

$$h(x) = h_0 + h_1 x + h_2 x^2 + \cdots,$$

where $h_k$ are $1 \times n^k$ constant matrices (precisely, row vectors). Hence, the differential of $h(x)$ can be expressed as

$$Dh(x) = h_1 + h_2 D(x^2) + h_3 D(x^3) + \cdots. \tag{15.71}$$

Similarly, consider an analytic vector field $X(x) \in V^{\omega}(M)$. Using Taylor series expansion, we have

$$X(x) = X_0 + X_1 x + X_2 x^2 + \cdots,$$

where $X_k$ are $n \times n^k$ constant matrices. Then the Jacobian matrix of $X(x)$ is expressed as

$$J_X(x) = X_1 + X_2 D(x^2) + X_3 D(x^3) + \cdots. \tag{15.72}$$

Observing the above expressions, one sees easily that in geometric calculations the differential $D(x^k)$ plays a key role. This section aims on its formula.

**Lemma 15.4.**

$$\begin{aligned} D(x^k) = {} & W_{[n^{k-1},n]}x^{k-1} + xW_{[n^{k-2},n]}x^{k-2} + \cdots \\ & + x^{k-2}W_{[n,n]}x + x^{k-1} \otimes I_n, \qquad k \geq 2. \end{aligned} \tag{15.73}$$

*Proof.* We use mathematical induction. It is trivial that

$$Dx = I_n.$$

Using (15.37), we have

$$\begin{aligned} D(x^2) &= Dx \ltimes (1 \otimes W_{[n,n]}) \ltimes x + x \ltimes I_n \\ &= I_n \ltimes W_{[n,n]} \ltimes x + (x \otimes I_n)I_n = W_{[n,n]} \ltimes x + x \otimes I_n. \end{aligned}$$

Assume (15.73) holds for $k$. Invoking (15.42), we have

$$D(x \otimes I_{n^k}) = (I_n \otimes I_{n^k})(1 \otimes W_{[n^k,n]}) = I_{n^{k+1}}W_{[n^k,n]} = W_{[n^k,n]}.$$

Hence

$$\begin{aligned} D(x^{k+1}) &= D[(x \otimes I_{n^k})x^k] = D(x \otimes I_{n^k})x^k + (x \otimes I_{n^k})D(x^k) \\ &= W_{[n^k,n]}x^k + (x \otimes I_{n^k})(W_{[n^{k-1},n]}x^{k-1} \\ &\quad + xW_{[n^{k-2},n]}x^{k-2} + \cdots + x^{k-1} \otimes I_n) \\ &= W_{[n^k,n]}x^k + xW_{[n^{k-1},n]}x^{k-1} + \cdots + x^k \otimes I_n. \end{aligned}$$

□

The following theorem provides the fundamental differential formula.

**Theorem 15.4.** *Let $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. The differential of $x^m$ satisfies the following formula.*

$$D(x^{k+1}) = \Phi_k^n x^k, \quad k \geq 0, \tag{15.74}$$

*where*

$$\Phi_k^n = \sum_{s=0}^{k} I_{n^s} \otimes W_{[n^{k-s},n]}. \tag{15.75}$$

*Proof.* Invoking Lemma 15.4 and using the following swap formula

$$x^p W_{[n^s,n]} = (I_{n^p} \otimes W_{[n^s,n]})x^p,$$

(15.74) follows immediately. □

*Remark 15.2.* Since $I_1 = 1$ is a scalar and $W[1,n] = I_n$, It is easy to see that

$$\Phi_0^n = I_n.$$

If there is no possible confusion, $\Phi_k^n$ can be denoted briefly as $\Phi_k$.

## 15.6 Lie Derivative

Lie derivatives of functions, vector fields, and forms with respect to vector fields are fundamental in Nonlinear Control Theory. To avoid concerning the differential times, all the objects are assumed to be analyticity. They are denoted respectively by $C^\omega(\mathbb{R}^n)$, $V^\omega(\mathbb{R}^n)$, and $V^{*\omega}(\mathbb{R}^n)$ respectively. Moreover, we consider only the calculating formulas on $\mathbb{R}^n$, which are also available for general manifolds but within a coordinate chart.

**Definition 15.5.** Let $F : M = \mathbb{R}^n \to N = \mathbb{R}^n$ be a diffeomorphism.

1. For a smooth function $h(x) \in C^\omega(N)$, $F$ deduces a mapping $F^* : C^\omega(N) \to C^\omega(M)$, defined as
$$F^*(h) = h \circ F \in C^\omega(M).$$
2. For a vector field $X \in V^\omega(M)$, $F$ deduces a mapping $F_* : V^\omega(M) \to V^\omega(N)$, defined as
$$F_*(X)(h) = X(h \circ F), \quad \forall h \in C^\omega(N).$$
3. For a co-vector field $\alpha \in V^{*\omega}(N)$, $F$ deduces a mapping $F^* : V^{*\omega}(N) \to V^{*\omega}(M)$, defined as
$$\langle F^*(\alpha), X\rangle = \langle \alpha, F_*(X)\rangle, \quad \forall X \in V^r(M).$$

If $F$ is a local diffeomorphism, all the above mappings are also locally defined.

Consider a vector field $X \in V^\omega(\mathbb{R}^n)$, Its integral curve with initial value $x(0) = x_0$ is denoted as $\phi_t^X(x_0)$. For convenience, we assume $X$ is complete, that is, it is defined for all $t \in \mathbb{R}$. Then for each fixed $t$, the $\phi_t^X : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism [2].

**Definition 15.6.** Let $X \in V^\omega(M)$ and $h \in C^\omega(M)$. Then the Lie derivative of $h$ with respect to $X$, denoted by $L_X(h)$, is defined by

$$L_X(h) = \lim_{t\to 0} \frac{1}{t}\left[(\phi_t^X)^* f(x) - f(x)\right]. \tag{15.76}$$

**Proposition 15.16.** *Under local coordinates (15.76) can be expressed as*

$$L_X(h) = \langle dh, X\rangle = \sum_{i=1}^n X_i \frac{\partial h}{\partial x_i}. \tag{15.77}$$

*Proof.* According to the definition, we have $(\phi_t^X)^* h(x) = h(\phi_t^X(x))$. Hence its Taylor expansion with respect to $t$ is

$$h\left(\phi_t^X(x)\right) = h(x) + t\,df \cdot X(x) + O(t^2),$$

Plugging it into (15.76) yields (15.77). □

**Definition 15.7.** Let $X,\ Y \in V(M)$. The Lie derivative of $Y$ with respect to $X$, denoted by $\mathrm{ad}_X(Y)$, is defined as

$$\mathrm{ad}_X(Y) = \lim_{t\to 0} \frac{1}{t}\left[(\phi_{-t}^X)_* Y(\phi_t^X(x)) - Y(x)\right]. \tag{15.78}$$

**Proposition 15.17.** *Under local coordinates (15.78) can be expressed by*

$$\mathrm{ad}_X(Y) = J_Y X - J_X Y = [X, Y]. \tag{15.79}$$

*Where $J_Y$ is the Jacobian matrix of $Y$. Precisely,*

$$J_Y = \begin{bmatrix} \frac{\partial Y_1}{\partial x_1} & \cdots & \frac{\partial Y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Y_n}{\partial x_1} & \cdots & \frac{\partial Y_n}{\partial x_n} \end{bmatrix}.$$

*Proof.* Using Taylor series expansion, we have

$$\phi_t^X(x) = x + (tX) + O(t^2). \tag{15.80}$$

$$Y(\phi_t^X(x)) = Y(x) + J_Y(tX) + O(t^2). \tag{15.81}$$

Using (15.80), the Jacobian matrix of $\phi_{-t}^X$ is

$$J_{\phi_{-t}^X} = I - tJ_X + O(t^2). \tag{15.82}$$

Invoking (15.80)∼(15.82), we have

$$\begin{aligned}(\phi_{-t}^X)_* Y(\phi_t^X(x)) &= (I - tJ_X + O(t^2))(Y(x) + J_Y(tX) + O(t^2)) \\ &= Y(x) + t(J_Y X - J_X Y) + O(t^2).\end{aligned}$$

Plugging it into (15.78) yields (15.79). □

**Definition 15.8.** Let $X \in V(M)$ and $\alpha \in V^{*\omega}(M)$. The Lie derivative of the co-vector field $\alpha$ with respect to $X$, denoted by $L_X(\alpha)$, is defined as

$$L_X(\alpha) = \lim_{t\to 0} \frac{1}{t}\left[(\phi_t^X)^* \alpha(e_t^X(x)) - \omega(x)\right]. \tag{15.83}$$

**Proposition 15.18.** *Under local coordinates (15.80) can be expressed as*

$$L_X(\alpha) = (J_{\alpha^{\mathrm{T}}} X)^{\mathrm{T}} + \alpha J_X. \tag{15.84}$$

*Proof.* Similar to the proof of Proposition 15.17, we first use Taylor series expansion to have

$$\begin{aligned}(\phi_t^X)^*\alpha(\phi_t^X(x)) &= (\alpha(x) + t(J_{\alpha^{\mathrm{T}}}X)^{\mathrm{T}} + O(t^2))(I + tJ_x + O(t^2)) \\ &= \alpha(x) + t(J_{\alpha^{\mathrm{T}}}X)^{\mathrm{T}} + t\alpha(x)J_X + O(t^2),\end{aligned}$$

Here the transpose comes from the following convention: in local coordinate frame the co-vector field is always expressed as a row vector. Plugging the above equation into (15.83) yields (15.84). □

The higher order Lie derivatives can be defined iteratively as follows:

$$L_X^{k+1}h = L_X^k(L_Xh), \quad k \geq 1; \tag{15.85}$$

$$\mathrm{ad}_X^{k+1}Y = \mathrm{ad}_X^k(\mathrm{ad}_XY), \quad k \geq 1; \tag{15.86}$$

$$L_X^{k+1}\alpha = L_X^k(L_X\alpha), \quad k \geq 1. \tag{15.87}$$

Next, we consider the numerical calculation of Lie derivatives. First, we express function $h \in C^{\omega}(M)$, vector fields $X, Y \in V^{\omega}(M)$, and co-vector field $\alpha \in V^{*\omega}(M)$ into their Taylor series expansions as

$$\begin{aligned} h &= h_0 + h_1x + h_2x^2 + \cdots; \\ X &= X_0 + X_1x + X_2x^2 + \cdots; \\ Y &= Y_0 + Y_1x + Y_2x^2 + \cdots; \\ \alpha^{\mathrm{T}} &= \alpha_0 + \alpha_1x + \alpha_2x^2 + \cdots. \end{aligned}$$

To give the formulas we need the following lemma, which itself is interesting. We leave its proof to the reader.

**Lemma 15.5.** *Let $X \in \mathbb{R}^n$ be a column vector. Then*

$$X^{\mathrm{T}} = V_c^{\mathrm{T}}(I_n)X; \tag{15.88}$$

$$X = X^{\mathrm{T}}V_c^{\mathrm{T}}(I_n). \tag{15.89}$$

Now we are ready to present the Taylor series expansions of Lie derivatives.

**Proposition 15.19.** *1.*

$$L_Xh = \sum_{i=0}^{\infty} c_ix^i, \tag{15.90}$$

*where*

$$c_i = \sum_{k=0}^{i} h_{k+1}\Phi_k^n(I_{n^k} \otimes X_{i-k}).$$

*2.*

$$L_X Y = \sum_{i=0}^{\infty} d_i x^i, \tag{15.91}$$

*where*

$$d_i = \sum_{k=0}^{i} \left[ Y_{k+1} \Phi_k^n (I_{n^k} \otimes X_{i-k}) - X_{k+1} \Phi_k^n (I_{n^k} \otimes Y_{i-k}) \right].$$

*3.*

$$(L_X \alpha)^{\mathrm{T}} = \sum_{i=0}^{\infty} e_i x^i, \tag{15.92}$$

*where*

$$e_i = \sum_{k=0}^{i} \left[ \alpha_{k+1} \Phi_k^n (I_{n^k} \otimes X_{i-k}) - V_c^{\mathrm{T}} (I_{n^k} \otimes X_{k+1} \Phi_k^n)^{\mathrm{T}} (I_{n^k} \otimes \alpha_{i-k}) \right].$$

*Proof.* We prove equation (15.92) only. The proof of the other two formulas is similar. Invoking equation (15.84), we can obtain

$$(L_X \alpha)^{\mathrm{T}} = \frac{\partial \alpha^{\mathrm{T}}}{\partial x} X + \left( \frac{\partial X}{\partial x} \right)^{\mathrm{T}} \alpha^{\mathrm{T}}. \tag{15.93}$$

Consider its first term, which is

$$\begin{aligned} \tfrac{\partial \alpha^{\mathrm{T}}}{\partial x} X &= \left( \sum_{i=1}^{\infty} \alpha_i \Phi_{i-1}^n x^{i-1} \right) \left( \sum_{i=0}^{\infty} X_i x^i \right) \\ &= \sum_{k=0}^{\infty} \left[ \sum_{j=0}^{k} \alpha_{j+1} \Phi_j^n (I_{n^j} \otimes X_{k-j}) \right] x^k. \end{aligned} \tag{15.94}$$

$$\begin{aligned} \left(\tfrac{\partial X}{\partial x}\right)^{\mathrm{T}} &= \sum_{i=1}^{\infty} (x^{i-1})^{\mathrm{T}} (\Phi_{i-1}^n)^{\mathrm{T}} X_i^{\mathrm{T}} \\ &= V_c^{\mathrm{T}} (I_{n^{i-1}}) x^{i-1} (\Phi_{i-1}^n)^{\mathrm{T}} X_i^{\mathrm{T}} \\ &= V_c^{\mathrm{T}} (I_{n^{i-1}}) \left[ I_{n^{i-1}} \otimes X_i \Phi_{i-1}^n \right]^{\mathrm{T}} x^{i-1}. \end{aligned}$$

Hence

$$\begin{aligned} \left(\tfrac{\partial X}{\partial x}\right)^{\mathrm{T}} \alpha^{\mathrm{T}} &= \sum_{i=1}^{\infty} \sum_{k=0}^{i} \left[ V_c^{\mathrm{T}} I_{n^k} \left( I_{n^k} \otimes X_{k+1} \Phi_k^n \right)^{\mathrm{T}} x^k \alpha_{i-k} x^{i-k} \right] \\ &= \sum_{i=1}^{\infty} \left[ \sum_{k=0}^{i} V_c^{\mathrm{T}} I_{n^k} \left( I_{n^k} \otimes X_{k+1} \Phi_k^n \right)^{\mathrm{T}} (I_{n^k} \otimes \alpha_{i-k}) \right] x^i. \end{aligned} \tag{15.95}$$

Plugging (15.94) and (15.95) into (15.93) yields (15.92). □

*Remark 15.3.* This chapter involves many concepts and notations in Differential Geometry and Control Theory. We refer to [5] and [3] for geometric concepts, and to [4] for related control concepts.

## Exercise 15

1. Calculate the matrix $\Psi_3^3$. Then use it to check that in Example 15.7 $(A_3 + DB_2)\Psi_3^3 = 0$ holds.
2. Assume $x = (x_1, \cdots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, calculate (a) $Dx^2$; (b) $D(Fx + Gx^2 + Hx^3)$; (c) $\nabla(Fx + Gx^2 + Hx^3)$.
3. Calculate $T_N(3,3)$ and $T_B(3,3)$.
4. Assume a mapping $\pi : x \to y$ is defined as

$$\begin{cases} y_1 = x_1 + x_2^2, \\ y_2 = \sin(x_1 + x_2). \end{cases}$$

Give the Taylor series expansion of the inverse mapping $\pi^{-1}$ up to cubic terms. Precisely, find $A_1$, $A_2$, $A_3$, such that

$$x = A_1 y + A_2 y^2 + A_3 y^3 + O(\|y\|^4).$$

5. Assume $h(x) \in C^{\omega}(N)$, $X, Y \in V^{\omega}(M)$, $\alpha \in V^{*\omega}(N)$. Using their Taylor series expansions to calculate the second order Lie derivatives of $L_X^2 h(x)$, $\mathrm{ad}_X^2 Y$ and $L_X^2 \alpha$.

## References


1. Abraham, R., Marsden, J.: Foundations of Mechanics, 2nd edn. Benjamin/Cummings Pub. Com. Inc. (1978)
2. Boothby, W.: An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd edn. Academic Press, Inc., Orlando (1986)
3. Cheng, D., Xi, Z., Lu, Q., Mei, S.: Geometric structure of generalized controlled Hamiltonian systems and its application. Science in China, Series E **43**(4), 365–379 (2000)
4. Isidori, A.: Nonlinear Control Systems, 3rd edn. Springer (1995)
5. van der Schaft, A.: $L_2$-Gain and Passivity Techniques in Nonlinear Control. Springer-Verlag (2000)