Chapter 3
Multilinear Mappings Among Vector Spaces

In this chapter we consider how to use STP to describe and analyze milti-linear mappings among vector spaces. First, we consider the cross product on $\mathbb{R}^3$. Using this simple example, the structure matrix of multi-linear mappings is introduced. Then the Lie algebra and the linear mappings of matrices are investigated, as two typical multi-linear mappings. Finally, two applications are considered: One is the general linear group and its algebra, and the other one is matrix equations, including the Hautus and Sylvester Equations.

3.1 Cross Product on $\mathbb{R}^3$

In Chapter 2 we have investigated multi-linear functions. By introducing the structure matrices, it was shown that the semi-tensor product is a proper tool to describe and analyze multi-linear functions. A multi-linear mapping can also be expressed via structure matrix and STP.

Definition 3.1. Let V_i, $i = 0, 1, \cdots, k$ be n_i-dimensional vector spaces with $\{e^i_1, \cdots, e^i_{n_i}\}$ as the fixed bases of V_i, and $\phi : V_1 \ltimes \cdots \ltimes V_k \to V_0$ be a multi-linear mapping. Denote

$$\phi(e^1_{i_1}, \cdots, e^k_{i_k}) = \sum_{i_0=1}^{n_0} \mu^{i_0}_{i_1,\cdots,i_k} e^0_{i_0}, \quad i_j = 1, \cdots, n_j,\ j = 1, \cdots, k.$$

Then the matrix

$$M_\phi = \begin{bmatrix} \mu^1_{11\cdots1} & \cdots & \mu^1_{11\cdots n_k} & \cdots & \mu^1_{n_1 n_2 \cdots n_{k-1} 1} & \cdots & \mu^1_{n_1 1 n_2 \cdots n_k} \\ \mu^2_{11\cdots1} & \cdots & \mu^2_{11\cdots n_k} & \cdots & \mu^2_{n_1 n_2 \cdots n_{k-1} 1} & \cdots & \mu^2_{n_1 1 n_2 \cdots n_k} \\ \vdots \\ \mu^{n_0}_{11\cdots1} & \cdots & \mu^{n_0}_{11\cdots n_k} & \cdots & \mu^{n_0}_{n_1 n_2 \cdots n_{k-1} 1} & \cdots & \mu^{n_0}_{n_1 1 n_2 \cdots n_k} \end{bmatrix} \tag{3.1}$$

is called the structure matrix of ϕ.

Let $X_i = (x^i_1, \cdots, x^i_{n_i})^T \in V_i$, $i = 1, \cdots, k$. (Precisely, $X_i = \sum_{j=1}^{n_i} x^i_j e^i_j$ and we use its coefficients as a column vector expression of X_i, when the basis of V_i is fixed.) Then the following result is obvious.

Proposition 3.1. *Let $X_i \in V_i$, $i = 1, \cdots, k$. Then*

$$\phi(X_1, \cdots, X_k) = M_\phi \ltimes_{i=1}^k X_i. \tag{3.2}$$

This expression is very convenient in investigating multi-linear mappings. In this section we mainly consider the cross product on $\mathbb{R}^3$.

Consider the cross product $\times_c$ on $\mathbb{R}^3$ again. It is easy to verify that $\times_c : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is a bilinear mapping. Fix the canonical basis of $\mathbb{R}^3$ as $\{e_1 = \delta_3^1, e_2 = \delta_3^2, e_3 = \delta_3^3\}$. Denote

$$e_i \times e_j = \mu^1_{ij} e_1 + \mu^2_{ij} e_2 + \mu^3_{ij} e_3, \quad i, j = 1, 2, 3.$$

It is easy to calculate that

$$\begin{array}{llllll}
\mu^1_{11} = 0, & \mu^2_{11} = 0, & \mu^3_{11} = 0, & \mu^1_{12} = 0, & \mu^2_{12} = 0, & \mu^3_{12} = 1, \\
\mu^1_{13} = 0, & \mu^2_{13} = -1, & \mu^3_{13} = 0, & \mu^1_{21} = 0, & \mu^2_{21} = 0, & \mu^3_{21} = -1, \\
\mu^1_{22} = 0, & \mu^2_{22} = 0, & \mu^3_{22} = 0, & \mu^1_{23} = 1, & \mu^2_{23} = 0, & \mu^3_{23} = 0, \\
\mu^1_{31} = 0, & \mu^2_{31} = 1, & \mu^3_{31} = 0, & \mu^1_{32} = -1, & \mu^2_{32} = 0, & \mu^3_{32} = 0, \\
\mu^1_{33} = 0, & \mu^2_{33} = 0, & \mu^3_{33} = 0. & & &
\end{array}$$

Then the structure matrix of $\times_c$ is

$$M_c = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.3}$$

Assume $X = (x_1, x_2, x_3)^T$ and $Y = (y_1, y_2, y_3)^T$. Then

$$\begin{aligned}
X \times_c Y &= M_c \ltimes X \ltimes Y \\
&= M_c \ltimes X \ltimes Y \\
&= M_c [x_1y_1\ x_1y_2\ x_1y_3\ x_2y_1\ x_2y_2\ x_2y_3\ x_3y_1\ x_3y_2\ x_3y_3]^T \\
&= [x_2y_3 - x_3y_2, -x_1y_3 + x_3y_1, x_1y_2 - x_2y_1]^T.
\end{aligned} \tag{3.4}$$

Using structure matrix to calculate the cross product of two vectors is not an efficient way. But the analytic expression of cross product, as

$$X \times_c Y = M_c XY,$$

could be very convenient in theoretical analysis. A simple example is for multi-cross product. We have that

$$Y_1 \times_c Y_2 \times_c \cdots \times_c Y_k = M_c^k \ltimes_{i=1}^k Y_i. \tag{3.5}$$

Particularly, we have

$$\underbrace{Y \times_c Y \times_c \cdots \times_c Y}_{k} = M_c^k Y^k. \tag{3.6}$$

Following example may be more convincing.

Example 3.1. In mechanics it is easy to see that the angular momentum of a rigid body about its mass center is

$$H = \int r \times_c (\omega \times_c r) \mathrm{d}m, \tag{3.7}$$

where $r = (x, y, z)$ is the position arrow, starting from the mass center; $\omega = (\omega_x, \omega_y, \omega_z)^T$ is the angular speed. We want to prove the following equation for angular momentum (3.8), which appears so often to many books and papers.

$$\begin{bmatrix} H_x \\ H_y \\ H_z \end{bmatrix} = \begin{bmatrix} I_x & -I_{xy} & -I_{zx} \\ -I_{xy} & I_y & -I_{yz} \\ I_{zx} & -I_{yz} & I_z \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}, \tag{3.8}$$

where

$$\begin{array}{lll} I_x = \int (y^2 + z^2)\mathrm{d}m, & I_y = \int (z^2 + x^2)\mathrm{d}m, & I_z = \int (x^2 + y^2)\mathrm{d}m, \\ I_{xy} = \int xy\mathrm{d}m, & I_{yz} = \int yz\mathrm{d}m, & I_{zx} = \int zx\mathrm{d}m. \end{array}$$

Let M be the moment of the force, acting on the rigid body. We first prove that the dynamic equation of a rotating solid body is

$$\frac{\mathrm{d}H}{\mathrm{d}t} = M. \tag{3.9}$$

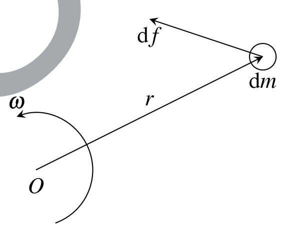

Fig. 3.1 Rotation

Consider a mass dm, with O as its rotating center, r as the position vector (from O to dm). df is the force acting on it. (Refer to Fig. 3.1) From Newton's second law:

$$\mathrm{d}f = a\mathrm{d}m = \frac{\mathrm{d}v}{\mathrm{d}t}\mathrm{d}m = \frac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r)\mathrm{d}m.$$

Now consider the moment of force on it, which is

$$\mathrm{d}M = r \times_c \mathrm{d}f = r \times_c \frac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r)\mathrm{d}m.$$

Integrating it over the solid body, we have

$$M = \int r \times_c \frac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r)\mathrm{d}m. \tag{3.10}$$

We claim that

$$r \times_c \frac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r) = \frac{\mathrm{d}}{\mathrm{d}t}\left[r \times_c (\omega \times_c r)\right]. \tag{3.11}$$

$$\begin{aligned} RHS &= \tfrac{\mathrm{d}}{\mathrm{d}t}(r) \times_c (\omega \times_c r) + r \times_c \tfrac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r) \\ &= (\omega \times_c r) \times_c (\omega \times_c r) + r \times_c \tfrac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r) \\ &= 0 + r \times_c \tfrac{\mathrm{d}}{\mathrm{d}t}(\omega \times_c r) = LHS. \end{aligned}$$

Applying this to (3.10), we have

$$\begin{aligned} M &= \int \tfrac{\mathrm{d}}{\mathrm{d}t}[r \times_c (\omega \times_c r)]\mathrm{d}m \\ &= \tfrac{\mathrm{d}}{\mathrm{d}t} \int r \times_c (\omega \times_c r)\mathrm{d}m \\ &= \tfrac{\mathrm{d}}{\mathrm{d}t} H. \end{aligned}$$

Next, we prove the angular momentum equation (3.8). According to (3.3), we have

$$\begin{aligned} H &= \int r \times_c (\omega \times_c r)\mathrm{d}m \\ &= \int M_c r M_c \omega r \mathrm{d}m \\ &= \int M_c (I_3 \otimes M_c) r \omega r \mathrm{d}m \\ &= \int M_c (I_3 \otimes M_c) W_{[3,9]} r^2 \omega \mathrm{d}m \\ &= \int M_c (I_3 \otimes M_c) W_{[3,9]} r^2 \mathrm{d}m \omega. \\ &:= \int \Psi r^2 \mathrm{d}m \omega, \end{aligned}$$

where

$$\begin{aligned} \Psi &= M_c (I_3 \otimes M_c) W_{[3,9]} \\ &= \begin{bmatrix} 0&0&0&0&-1&0&0&0&-1&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&1&0&0&0&0&0&0&0&-1&0&0&0&0&0&0&0&-1&0&0&0&0&0&0&0&1&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&-1&0&0&0&-1&0&0&0&0 \end{bmatrix}. \end{aligned}$$

Then we have

$$\Psi r^2 = \begin{bmatrix} y^2+z^2 & xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{bmatrix}.$$

(3.8) follows immediately.

3.2 General Linear Algebra

This section considers the structure matrices of general linear algebra and its sub-algebras. General linear algebra is an important Lie algebra. We first introduce Lie algebra.

Definition 3.2. 1. Assume V is a vector space over $\mathbb{R}$ and there is a mapping $*: V \times V \to V$ satisfies:

$$(\alpha X + \beta Y) * Z = \alpha (X * Z) + \beta (Y * Z), \quad \alpha, \beta \in \mathbb{R};\ X, Y, Z \in V. \tag{3.12}$$

Then $(V, *)$ is called an algebra.

2. Let $(V, *)$ be an algebra. It is called a Lie algebra if it satisfies

(i) (Skew-symmetry)

$$X * Y = -Y * X, \quad X, Y \in V; \tag{3.13}$$

(ii) (Jacobi Identity)

$$(X * Y) * Z + (Y * Z) * X + Z * X * Y = 0, \quad X, Y, Z \in V. \tag{3.14}$$

3. Let $(V, *)$ be a Lie algebra and $W \subset V$ be a subspace. W is called a Lie sub-algebra of $(V, *)$, if

$$X * Y \in W, \quad \forall X, Y \in W. \tag{3.15}$$

Consider $\mathscr{M}_n$. It is obviously a vector space over $\mathbb{R}$. We then can define a product over $\mathscr{M}_n$, called a Lie bracket, as

$$[A, B] := AB - BA, \quad A, B \in \mathscr{M}_n. \tag{3.16}$$

We leave to the reader the verifying $(\mathscr{M}_n, [\cdot, \cdot])$ is a Lie algebra. This algebra is called the nth order general linear algebra, and denoted by $gl(n, \mathbb{R})$. It is very important because any finite dimensional Lie algebra is a sub-algebra of a general linear algebra [9].

Let $\{\Delta_{i,j} \mid i, j = 1, \cdots, n\}$ be a basis of $gl(n, \mathbb{R})$, where $\Delta_{ij} = (d^{i,j}_{p,q})$ and

$$d^{i,j}_{p,q} = \begin{cases} 1, & p = i, \text{ and } q = j \\ 0, & \text{otherwise.} \end{cases}$$

Then it is clear that

$$[\Delta_{i,j}, \Delta_{\alpha,\beta}] = \begin{cases} \Delta_{i,\beta}, & j = \alpha, \text{ and } i \neq \beta \\ -\Delta_{\alpha,j}, & \beta = i, \text{ and } j \neq \alpha \\ \Delta_{i,\beta} - \Delta_{\alpha,j}, & j = \alpha \neq i = \beta \\ 0, & \text{otherwise.} \end{cases} \tag{3.17}$$

Now assume the basis is arranged into a row in the order of $\mathrm{id}(i,j;n,n)$, then the structure matrix, denoted by M_{L^n}, can easily be constructed using (3.13). Then we have

$$V_r([A,B]) = M_{L^n} V_r(A) V_r(B), \quad A,B \in gl(n,\mathbb{R}). \tag{3.18}$$

We give a numerical example for this.

Example 3.2. Consider $gl(2,\mathbb{R})$. The basis is:

$$\Delta_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \Delta_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \Delta_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Using (3.13), it is easy to calculate M_{L^2} as

$$M_{L^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.19}$$

Proposition 3.2. *Consider $gl(n,\mathbb{R})$.*

1. *Define*

$$sl(n,\mathbb{R}) := \{A \in gl(n,\mathbb{R}) \mid \mathrm{tr}(A) = 0\}. \tag{3.20}$$

$sl(n,\mathbb{R})$ is a sub-algebra of $gl(n,\mathbb{R})$, which is called the special linear algebra.

2. *Define*

$$o(n,\mathbb{R}) := \left\{A \in gl(n,\mathbb{R}) \mid A^T = -A\right\}. \tag{3.21}$$

$o(n,\mathbb{R})$ is a sub-algebra of $gl(n,\mathbb{R})$, which is called the orthogonal algebra.

3. *Define*

$$sp(2n,\mathbb{R}) := \left\{A \in gl(2n,\mathbb{R}) \mid A^T J + JA = 0\right\}, \tag{3.22}$$

where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

$sp(n,\mathbb{R})$ is a sub-algebra of $gl(2n,\mathbb{R})$, which is called the symplectic algebra.

We leave the proof of Proposition 3.2 to the reader.

Example 3.3. Consider $sl(2,\mathbb{R})$. Choose a basis as

$$e_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

From Table 3.1 the structure matrix of $sl(2,\mathbb{R})$ is obtained easily as

Table 3.1 Lie bracket on $sl(2,\mathbb{R})$

	e_1	e_2	e_3
e_1	0	$-2e_1$	e_2
e_2	$2e_1$	0	$-2e_3$
e_3	$-e_2$	$2e_3$	0

$$M_{sl} = \begin{bmatrix} 0 & -2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 \end{bmatrix}. \tag{3.23}$$

Definition 3.3. Let $(V,*)$ be a Lie algebra and $X \in V$. Then $\mathrm{ad}_X : V \to V$ defined by

$$\mathrm{ad}_X(Y) := X * Y, \quad Y \in V, \tag{3.24}$$

is called the adjoint representation of X.

Since $*$ is a bilinear mapping, ad_X is a linear mapping. Assume V is an n-dimensional vector space and fix a basis of V as $\{e_1,\cdots,e_n\}$. Then the corresponding structure matrix is uniquely expressed as M_V. Then we have

$$\mathrm{ad}_X(Y) = M_V XY, \quad \forall Y \in V.$$

Then we have

Proposition 3.3. *Under a fixed basis the matrix expression of the adjoint representation is*

$$M_{\mathrm{ad}_X} = M_V X. \tag{3.25}$$

Example 3.4. Let

$$X = \begin{bmatrix} -2 & 0 \\ 1 & 2 \end{bmatrix} \in sl(2,\mathbb{R}).$$

Using the basis in Example 3.3, $X = -2e_2 + e_3 \sim (0\ -2\ 1)^T$. Then the matrix expression of ad_X is

$$M_{\mathrm{ad}_X} = M_{sl}X = \begin{bmatrix} -4 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 4 \end{bmatrix}.$$

3.3 Mappings over Matrices

The row stacking and column stacking forms of a matrix have been investigated in Chapter 1. These expressions are sometimes very convenient. In this section we consider the matrix expression of linear and polynomial mappings between matrices, which are expressed in their row or column stacking forms.

We denote by $L(V,W)$ the set of linear mappings from vector space V to vector space W. Particularly, $L(\mathcal{M}_{p\times q},\mathcal{M}_{m\times n})$ is the set of linear mappings from $\mathcal{M}_{p\times q}$ to $\mathcal{M}_{m\times n}$. We start from some examples.

Example 3.5 (Lyapunov mapping). Given a square matrix $A\in\mathcal{M}_n$. Consider the following mapping $L_A:\mathcal{M}_n\to\mathcal{M}_n$, defined as

$$L_A(X)=AX+XA^{\mathrm{T}},\quad X\in\mathcal{M}_n. \tag{3.26}$$

A square matrix A is said to be Hurwitz , if all the eigenvalues of A have negative real parts. The following result is well known [10]: A is a Hurwitz matrix, if and only if, for any negative definite matrix $Q<0$, $L_A(X)=Q$ has unique solution, which is positive definite.

As a linear mapping on $\mathcal{M}_n$, L_A has a matrix expression as [8]

$$M^c_{L_A}=A\otimes I+I\otimes A. \tag{3.27}$$

The precise meaning of this matrix expression is that as both the argument matrix and the image matrix are expressed into column stacking form, we have

$$V_c(L_A(X))=M^c_{L_A}V_c(X). \tag{3.28}$$

Note that we use superscript c to indicate that the matrix expression is available for matrices under column stacking forms. When the row stacking forms are considered, we should have

$$V_r(L_A(X))=M^r_{L_A}V_r(X).$$

Since

$$L^T_A(X)=X^TA^T+AX^T=L_A(X^T),$$

we have

$$V_r(L_A(X))=V_c(L^T_A(X))=V_c(L_A(X^T))=M^c_{L_A}V_c(X^T)=M^c_{L_A}V_r(X).$$

It follows that

$$M^r_{L_A}=M^c_{L_A}. \tag{3.29}$$

Example 3.6 (Symplectic mapping). Recall that a $2n\times 2n$ matrix $X\in sp(2n,\mathbb{R})$, iff, X satisfies (3.18). In general, we can replace J by an arbitrary matrix $N\in gl(n,\mathbb{R})$ and define a mapping [4]

$$L_N(X)=NX+X^{\mathrm{T}}N,\quad X\in\mathcal{M}_n. \tag{3.30}$$

It was proved that [4]

$$\mathcal{G}_N=\{X\in gl(n,\mathbb{R})\,|\,L_N(X)=0\}$$

is a sub-algebra of $gl(n,\mathbb{R})$.

It is not difficult to show that the matrix expression of L_N is

$$M^c_{L_N} = I_n \otimes N + (N^T \otimes I_n) W_{[n]}. \tag{3.31}$$

In fact, $\mathscr{G}_N$ is the Lie algebra of Lie group [4]

$$G_N = \{Z \in GL(n,R) \,|\, Z^{\mathrm{T}} N Z = N\}.$$

The following property of L_N is interesting.

Proposition 3.4 ([4]). *For any $N \in \mathscr{M}_n$, $L_N(X) = 0$ has at least a solution $X \neq 0$. In other words, 0 is an eigenvalue of L_N.*

Let $\rho : \mathscr{M}_{p\times q} \to \mathscr{M}_{m\times n}$. We always have two matrix expressions M^c_ρ and M^r_ρ corresponding to column stacking form and row stacking form respectively. The following proposition shows that these two expressions are easily convertible.

Proposition 3.5. *Let $\rho \in L(\mathscr{M}_{p\times q}, \mathscr{M}_{m\times n})$. Then*

$$\begin{cases} M^r_\rho = W_{[n,m]} M^c_\rho W_{[p,q]}, \\ M^c_\rho = W_{[m,n]} M^r_\rho W_{[q,p]}. \end{cases} \tag{3.32}$$

Particularly, if $\rho \in L(\mathscr{M}_n, \mathscr{M}_n)$, then (3.32) becomes

$$\begin{cases} M^r_\rho = W_{[n]} M^c_\rho W_{[n]}, \\ M^c_\rho = W_{[n]} M^r_\rho W_{[n]}. \end{cases} \tag{3.33}$$

Proof. Consider the first equality of (3.32): Using the formulas of Proposition 1.13, we have

$$\begin{aligned} V_r(\rho(X)) &= W_{[n,m]} V_c(\rho(X)) \\ &= W_{[n,m]} M^c_\rho V_c(X) \\ &= W_{[n,m]} M^c_\rho W_{[p,q]} V_r(X). \end{aligned}$$

Since $X \in \mathscr{M}_{p\times q}$ is arbitrary, the first equality of (3.32) follows.

Left multiplying both sides of the first equality of (3.32) by $W_{[m,n]}$ and right multiplying both sides by $W_{[q,p]}$ and using equality (1.53), we have the second equality of (3.32). □

For convenience, we take M^c_ρ as a default matrix expression of the linear mapping ρ on matrices.

Let $Z \in \mathscr{M}_{n\times p}$. We consider a general linear mapping, formed by matrix products. Precisely, let $\rho : \mathscr{M}_{n\times p} \to \mathscr{M}_{m\times q}$, defined as

$$Z \mapsto AZB + CZ^{\mathrm{T}} D, \tag{3.34}$$

where $A \in \mathscr{M}_{m\times n}$, $B \in \mathscr{M}_{p\times q}$, $C \in \mathscr{M}_{m\times p}$, $D \in \mathscr{M}_{n\times q}$.

It is not difficult to verify that several useful linear matrix mappings, such as Lyapunov mapping, symplectic mapping etc., are particular forms of ρ, defined by (3.34). When we find the matrix expression of ρ, we have the matrix expressions of all such linear matrix mappings. In fact, we have the following:

Proposition 3.6. *The matrix expression of ρ defined by (3.34), is*

$$M^c = (B^{\mathrm{T}} \otimes A) + (D^{\mathrm{T}} \otimes C)W_{[p,n]}. \tag{3.35}$$

Proof. We first prove matrix expressions of 4 fundamental linear matrix mappings. In the following we assume the concerned matrices have the dimensions as in (3.34).

Table 3.2 Expression under column stacking forms

ρ:	M^c_ρ
$Z \mapsto AZ$	$I_p \otimes A$
$Z \mapsto ZB$	$B^{\mathrm{T}} \otimes I_n$
$Z \mapsto CZ^{\mathrm{T}}$	$(I_n \otimes C)W_{[p,n]}$
$Z \mapsto Z^{\mathrm{T}}D$	$(D^{\mathrm{T}} \otimes I_p)W_{[p,n]}$

First two mappings are well known. Using Lemma 2.3, their matrix expressions are easily verifiable. For the third one, we have

$$\begin{aligned} V_c(CZ^{\mathrm{T}}) &= (I_n \otimes C)V_c(Z^{\mathrm{T}}) = (I_n \otimes C)V_r(Z) \\ &= (I_n \otimes C)W_{[p,n]}W_{[n,p]}V_r(Z) = (I_n \otimes C)W_{[p,n]}V_c(Z). \end{aligned}$$

To get the last one, we can prove that

$$\begin{aligned} V_c(Z^{\mathrm{T}}D) &= (D^{\mathrm{T}} \otimes I_p)V_c(Z^{\mathrm{T}}) = (D^{\mathrm{T}} \otimes I_p)V_r(Z) \\ &= (D^{\mathrm{T}} \otimes I_p)W_{[p,n]}W_{[n,p]}V_r(Z) = (D^{\mathrm{T}} \otimes I_p)W_{[p,n]}V_c(Z). \end{aligned}$$

The mapping ρ of (3.34) can be considered as a linear combination of two compounded mappings, and use the matrix expressions of the four fundamental mappings can get the matrix expressions of the two compounded mappings of ρ. Then adjusting the dimensions of the matrices in the mapping (3.34), we can have

$$M^c = (B^{\mathrm{T}} \otimes I_m)(I_p \otimes A) + (D^{\mathrm{T}} \otimes I_m)(I_n \otimes C)W_{[p,n]}.$$

A simplification leads to (3.35). □

In the following we give a numerical example to illustrate the formula (3.35).

Example 3.7. Assume $A, C \in \mathscr{M}_{3\times 2}$, $B, D \in \mathscr{M}_{2\times 4}$, and

$$A=\begin{bmatrix}1&-1\\2&1\\0&1\end{bmatrix},\quad C=\begin{bmatrix}1&2\\0&1\\-1&1\end{bmatrix},$$
$$B=\begin{bmatrix}1&-1&0&1\\2&1&1&0\end{bmatrix},\quad D=\begin{bmatrix}1&1&1&-1\\1&0&2&1\end{bmatrix}.$$

1. Assume $\rho : M_2 \to M_{3\times 2}$, defined as $Z \mapsto AZ$. Then

$$M_\rho = I_2 \otimes A = \begin{bmatrix}1&-1&0&0\\2&1&0&0\\0&1&0&0\\0&0&1&-1\\0&0&2&1\\0&0&0&1\end{bmatrix}.$$

   Hence

$$V_c(AZ) = M_\rho V_c(Z) = \begin{bmatrix}1&-1&0&0\\2&1&0&0\\0&1&0&0\\0&0&1&-1\\0&0&2&1\\0&0&0&1\end{bmatrix}\begin{bmatrix}z_{11}\\z_{21}\\z_{12}\\z_{22}\end{bmatrix} = \begin{bmatrix}z_{11}-z_{21}\\2z_{11}+z_{21}\\z_{21}\\z_{12}-z_{22}\\2z_{12}+z_{22}\\z_{22}\end{bmatrix}. \tag{3.36}$$

   A direct computation shows that

$$AZ = \begin{bmatrix}1&-1\\2&1\\0&1\end{bmatrix}\begin{bmatrix}z_{11}&z_{12}\\z_{21}&z_{22}\end{bmatrix} = \begin{bmatrix}z_{11}-z_{21}&z_{12}-z_{22}\\2z_{11}+z_{21}&2z_{12}+z_{22}\\z_{21}&z_{22}\end{bmatrix},$$

   which verifies (3.36).
2. Assume $\rho : M_2 \to M_{2\times 4}$, defined as $Z \mapsto ZB$. Then

$$M_\rho = B^{\mathrm{T}} \otimes I_2 = \begin{bmatrix}1&0&2&0\\0&1&0&2\\-1&0&1&0\\0&-1&0&1\\0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix}.$$

   Hence

$$
\begin{aligned}
V_c(ZB) &= M_\rho V_c(Z) \\
&= \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11}+2z_{12} \\ z_{21}+2z_{22} \\ -z_{11}+z_{12} \\ -z_{21}+z_{22} \\ 2z_{12}+z_{22} \\ z_{12} \\ z_{22} \\ z_{11} \\ z_{21} \end{bmatrix}.
\end{aligned} \tag{3.37}
$$

A direct computation shows that

$$
\begin{aligned}
ZB &= \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} z_{11}+2z_{12} & -z_{11}+z_{12} & z_{12} & z_{11} \\ z_{21}+2z_{22} & -z_{21}+z_{22} & z_{22} & z_{21} \end{bmatrix},
\end{aligned}
$$

which verifies (3.37).

3. Assume $\rho : M_2 \to M_{3\times 2}$, defined as $Z \mapsto CZ^{\mathrm{T}}$. Note that

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

that is,

$$
M_\rho = (I_2 \otimes C) W_{[2,2]} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix}.
$$

Hence

$$
\begin{aligned}
V_c(CZ^{\mathrm{T}}) &= M_\rho V_c(Z) \\
&= \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11}+2z_{12} \\ z_{12} \\ -z_{11}+z_{12} \\ z_{21}+2z_{22} \\ z_{22} \\ -z_{21}+z_{22} \end{bmatrix}.
\end{aligned} \tag{3.38}
$$

A direct computation shows that

$$CZ^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \end{bmatrix} = \begin{bmatrix} z_{11}+2z_{12} & z_{21}+2z_{22} \\ z_{12} & z_{22} \\ -z_{11}+z_{12} & -z_{21}+z_{22} \end{bmatrix},$$

which verifies (3.38).

4. Assume $\rho : M_2 \to M_{2\times 4}$, defined as $Z \mapsto Z^{\mathrm{T}} D$. Then

$$M_\rho = (D^{\mathrm{T}} \otimes I_2) W_{[2,2]} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Hence

$$\begin{aligned} V_c(Z^{\mathrm{T}} D) &= M_\rho V_c(Z) \\ &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11}+z_{21} \\ z_{12}+z_{22} \\ z_{11} \\ z_{12} \\ z_{11}+2z_{21} \\ z_{12}+2z_{22} \\ -z_{11}+z_{21} \\ -z_{12}+z_{22} \end{bmatrix}. \end{aligned} \tag{3.39}$$

A direct computation shows that

$$\begin{aligned} Z^{\mathrm{T}} D &= \begin{bmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} z_{11}+z_{21} & z_{11} & z_{11}+2z_{21} & -z_{11}+z_{21} \\ z_{12}+z_{22} & z_{12} & z_{12}+2z_{22} & -z_{12}+z_{22} \end{bmatrix}, \end{aligned}$$

which verifies (3.39).

5. Assume $\rho : M_2 \to M_{3\times 4}$, defined as $Z \mapsto AZB + CZ^{\mathrm{T}} D$. Using (3.35), we have

$$M_\rho = (B^{\mathrm{T}} \otimes A) + (D^{\mathrm{T}} \otimes C) W_{[2,2]} = \begin{bmatrix} 2 & 0 & 4 & 0 \\ 2 & 1 & 5 & 3 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 3 & -1 \\ -2 & -1 & 3 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 2 & 3 & 3 \\ 0 & 0 & 3 & 3 \\ -1 & -2 & 1 & 3 \\ 0 & 0 & -2 & 2 \\ 2 & 1 & -1 & 1 \\ 1 & 0 & -1 & 1 \end{bmatrix}.$$

Hence

$$\begin{aligned} & V_c(AZB + CZ^{\mathrm{T}} D) \\ &= M_\rho V_c(Z) \\ &= \begin{bmatrix} 2 & 0 & 4 & 0 \\ 2 & 1 & 5 & 3 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 3 & -1 \\ -2 & -1 & 3 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 2 & 3 & 3 \\ 0 & 0 & 3 & 3 \\ -1 & -2 & 1 & 3 \\ 0 & 0 & -2 & 2 \\ 2 & 1 & -1 & 1 \\ 1 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} 2z_{11} + 4z_{12} \\ 2z_{11} + z_{21} + 5z_{12} + 3z_{22} \\ -z_{11} + z_{12} + 3z_{22} \\ z_{21} + 3z_{12} - z_{22} \\ -2z_{11} - z_{21} + 3z_{12} + z_{22} \\ -z_{11} - z_{21} + z_{12} + z_{22} \\ z_{11} + 2z_{21} + 3z_{12} + 3z_{22} \\ 3z_{12} + 3z_{22} \\ -z_{11} - 2z_{21} + z_{12} + 3z_{22} \\ -2z_{12} + 2z_{22} \\ 2z_{11} + z_{21} - z_{12} + z_{22} \\ z_{11} - z_{12} + z_{22} \end{bmatrix}. \end{aligned} \tag{3.40}$$

A direct computation shows that

$$AZB + CZ^{\mathrm{T}} D = \begin{bmatrix} 2z_{11} + 4z_{12} & z_{21} + 3z_{12} - z_{22} \\ 2z_{11} + z_{21} + 5z_{12} + 3z_{22} & -2z_{11} - z_{21} + 3z_{12} + z_{22} \\ -z_{11} + z_{12} + 3z_{22} & -z_{11} - z_{21} + z_{12} + z_{22} \end{bmatrix}$$

$$\begin{bmatrix} z_{11} + 2z_{21} + 3z_{12} + 3z_{22} & -2z_{12} + 2z_{22} \\ 3z_{12} + 3z_{22} & 2z_{11} + z_{21} - z_{12} + z_{22} \\ -z_{11} - 2z_{21} + z_{12} + 3z_{22} & z_{11} - z_{12} + z_{22} \end{bmatrix}.$$

This verifies (3.40).

Mapping ρ defined by (3.34) consists of two types of terms. In fact, it can be used for any mapping with finite such terms. A particular case is when all the matrices involved are square matrices of same dimensions. In the following we give another numerical example.

Example 3.8. Given a set of matrices as:

$$A=\begin{bmatrix}1&1\\-1&1\end{bmatrix},\quad B=\begin{bmatrix}0&-1\\1&2\end{bmatrix},$$

$$C=\begin{bmatrix}1&0\\1&-1\end{bmatrix},\quad D=\begin{bmatrix}1&2\\-2&1\end{bmatrix}.$$

Consider a mapping

$$L(Z)=AZ+ZB+CZ^{\mathrm{T}}D+AZB. \tag{3.41}$$

Using (3.35) and Table 3.2, the matrix expression of L can be obtained as

$$\begin{aligned}
M_L^c &= I_2\otimes A+B^{\mathrm{T}}\otimes I_2+(D^{\mathrm{T}}\otimes C)W_{[2]}+B^{\mathrm{T}}\otimes A\\
&=\begin{bmatrix}1&1&0&0\\-1&1&0&0\\0&0&1&1\\0&0&-1&1\end{bmatrix}+\begin{bmatrix}0&0&1&0\\0&0&0&1\\-1&0&2&0\\0&-1&0&2\end{bmatrix}\\
&\quad+\begin{bmatrix}1&0&-2&0\\1&-1&-2&2\\2&0&1&0\\2&-2&0&-1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}+\begin{bmatrix}0&0&1&1\\0&0&-1&1\\-1&-1&2&2\\1&-1&-2&2\end{bmatrix}\\
&=\begin{bmatrix}2&-1&2&1\\0&-1&-2&4\\0&0&5&3\\3&-2&5&4\end{bmatrix}.
\end{aligned}$$

For convenience in use, we listed the matrix expressions of some useful matrix mappings in Table 3.3.

Table 3.3 Matrix Expression of Some Mappings

mapping	notation	ρ_A	$M^c_{\rho_A}$
Lyapunov mapping	L_A	$Z\mapsto AZ+ZA^{\mathrm{T}}$	$I\otimes A+A\otimes I$
Generalized Lyapunov mapping	L_{AB}	$Z\mapsto AZ+ZB$	$I\otimes A+B^{\mathrm{T}}\otimes I$
symplectic mapping	S_A	$Z\mapsto AZ+Z^{\mathrm{T}}A$	$I\otimes A+(A^{\mathrm{T}}\otimes I)W$
adjoint mapping	ad_A	$Z\mapsto AZ-ZA$	$I\otimes A-A^{\mathrm{T}}\otimes I$
conjugate mapping	C_j^A	$Z\mapsto AZA^{-1}$	$A^{-T}\otimes A$
cogradient mapping	C_g^A	$Z\mapsto AZA^{\mathrm{T}}$	$A\otimes A$

Using the above matrix expressions, some useful formulas can be deduced.

Proposition 3.7. *Let* $A\in M_{m\times n}$, $B\in M_{p\times q}$. *Then*

$$(I_p\otimes A)W_{[n,p]}=W_{[m,p]}(A\otimes I_p), \tag{3.42}$$

$$W_{[m,p]}(A\otimes B)W_{[q,n]}=(B\otimes A). \tag{3.43}$$

Proof. Assume $Z \in M_{p\times n}$, and consider the matrix expression of the mapping $Z \mapsto AZ^{\mathrm{T}}$, which can be obtained through the following two ways:

1. $Z \mapsto Z^{\mathrm{T}} \mapsto AZ^{\mathrm{T}}$: Note that where we use swap matrix $W_{[n,p]}$ first, and then use $I_p \otimes A$. It follows that the matrix expression of the mapping is $(I_p \otimes A)W_{[n,p]}$.
2. $Z \mapsto ZA^{\mathrm{T}} \mapsto (ZA^{\mathrm{T}})^{\mathrm{T}} = AZ^{\mathrm{T}}$: Now we first use $A \otimes I_p$, and then use $W_{[m,p]}$. Hence, the same mapping can be expressed as $W_{[m,p]}(A \otimes I_p)$.

Since the matrix expression of a linear mapping is unique, (3.42) follows.

As for (3.43), let $Z \in \mathscr{M}_{q\times n}$, and consider the matrix expression of $Z \mapsto AZ^{\mathrm{T}}B^{\mathrm{T}}$, which can also be obtained through two ways:

1. $Z \mapsto ZA^{\mathrm{T}} \mapsto BZA^{\mathrm{T}} \mapsto (BZA^{\mathrm{T}})^{\mathrm{T}}$: It is realized by $W_{[m,p]}(I_m \otimes B)(A \otimes I_q)$, which equals to $W_{[m,p]}(A \otimes B)$.
2. $Z \mapsto Z^{\mathrm{T}} \mapsto AZ^{\mathrm{T}} \mapsto AZ^{\mathrm{T}}B^{\mathrm{T}}$: It is realized alternatively by $(B \otimes I_m)(I_q \otimes A)W_{[n,q]}$, which equals to $(B \otimes A)W_{[n,q]}$.

Hence we have

$$W_{[m,p]}(A \otimes B) = (B \otimes A)W_{[n,q]}.$$

Right multiplying the above equation by $W_{[q,n]}$ yields (3.43). □

In fact, it is not difficult to see that (3.42) can be deduced from (3.43). Both of them will be used in the sequel.

The following proposition shows the matrix expression of a general linear matrix mapping under row stacking forms. We give them a direct proof. In fact, they can also be obtained from the expression under column stacking forms, using Proposition 3.5.

Proposition 3.8. *The matrix expression of the mapping ρ defined by (3.34) under row stacking form is*

$$M^r = (A \otimes B^{\mathrm{T}}) + (C \otimes D^{\mathrm{T}})W_{[n,p]}. \tag{3.44}$$

Proof. We first also give the matrix expressions of four fundamental linear matrix mappings, which themselves are also useful.

Table 3.4 Matrix expressions under row stacking form

ρ:	M^r_ρ
$Z \mapsto AZ$	$A \otimes I_p$
$Z \mapsto ZB$	$I_n \otimes B^{\mathrm{T}}$
$Z \mapsto CZ^{\mathrm{T}}$	$(C \otimes I_n)W_{[n,p]}$
$Z \mapsto Z^{\mathrm{T}}D$	$(I_p \otimes D^{\mathrm{T}})W_{[n,p]}$

We prove them one by one. Using Lemma 2.3, we have

$$V_r(AZ) = AV_r(Z) = (A \otimes I_p)V_r(Z),$$

which proves the first equality. As for the second one, we have

$$V_r(ZB) = V_c(B^{\mathrm{T}}Z^{\mathrm{T}}) = (I_n \otimes B^{\mathrm{T}})V_c(Z^{\mathrm{T}}) = (I_n \otimes B^{\mathrm{T}})V_r(Z).$$

Then for the third one, we have

$$\begin{aligned} V_r(CZ^{\mathrm{T}}) &= V_c(ZC^{\mathrm{T}}) = (C \otimes I_n)V_c(Z) \\ &= (C \otimes I_n)W_{[n,p]}W_{[p,n]}V_c(Z) \\ &= (C \otimes I_n)W_{[n,p]}V_r(Z). \end{aligned}$$

Finally, we have

$$\begin{aligned} V_r(Z^{\mathrm{T}}D) &= V_c(D^{\mathrm{T}}Z) = (I_p \otimes D^{\mathrm{T}})V_c(Z) \\ &= (I_p \otimes D^{\mathrm{T}})W_{[n,p]}W_{[p,n]}V_c(Z) \\ &= (I_p \otimes D^{\mathrm{T}})W_{[n,p]}V_r(Z). \end{aligned}$$

This is the proof for the last equality. Using Table 3.4, we have, for the compounded mappings, that

$$\begin{aligned} & V_r(AZB + CZ^{\mathrm{T}}D) \\ &= [(I_m \otimes B^{\mathrm{T}})(A \otimes I_p) + (I_m \otimes D^{\mathrm{T}})(C \otimes I_n)W_{[n,p]}]V_r(Z) \\ &= [A \otimes B^{\mathrm{T}} + (C \otimes D^{\mathrm{T}})W_{[n,p]}]V_r(Z). \end{aligned} \tag{3.45}$$

(3.44) follows. □

3.4 Conversion of Matrix Polynomials

From Chapter 1 one sees that in addition to the standard form, a matrix A can also be expressed by its row stacking form $V_r(A)$ or column stacking form $V_c(A)$. Sometimes we need to convert a matrix from one form to another. These conversions are particularly important as the matrix is a matrix expression of a linear mapping between vector spaces.

Let $A \in \mathscr{M}_{m\times n}$. We define two mappings

$$\pi_s^r, \pi_s^c : \mathscr{M}_{m\times n} \to \mathscr{M}_{m\times ns^2}$$

as follows:

$$\pi_s^r(A) = A[I_n \otimes V_r^{\mathrm{T}}(I_s)], \tag{3.46}$$

and

$$\pi_s^c(A) = A[I_n \otimes \mathrm{Row}_1(I_s)\ \ I_n \otimes \mathrm{Row}_2(I_s)\ \ \cdots\ \ I_n \otimes \mathrm{Row}_s(I_s)], \tag{3.47}$$

Then we have

Proposition 3.9.

$$\pi_s^c(A) = \pi_s^r(A) W_{[s,n]}, \quad \forall A \in M_{m\times n}. \tag{3.48}$$

Proof. A straightforward computation shows that

$$\begin{aligned}
\pi_s^r(A) = \Big[& \underbrace{\underbrace{\mathrm{Col}_1(A)\, 0 \cdots 0}_{s}\ \underbrace{0\ \mathrm{Col}_1(A)\, 0 \cdots 0}_{s} \cdots \underbrace{0 \cdots 0\ \mathrm{Col}_1(A)}_{s}}_{s} \\
& \vdots \\
& \underbrace{\underbrace{\mathrm{Col}_n(A)\, 0 \cdots 0}_{s}\ \underbrace{0\ \mathrm{Col}_n(A)\, 0 \cdots 0}_{s} \cdots \underbrace{0 \cdots 0\ \mathrm{Col}_n(A)}_{s}}_{s} \Big],
\end{aligned} \tag{3.49}$$

and

$$\begin{aligned}
\pi_s^c(A) = \Big[& \underbrace{\underbrace{\mathrm{Col}_1(A)\, 0 \cdots 0}_{s}\ \underbrace{\mathrm{Col}_2(A)\, 0 \cdots 0}_{s} \cdots \underbrace{\mathrm{Col}_n(A)\, 0 \cdots 0}_{s}}_{n} \\
& \vdots \\
& \underbrace{\underbrace{0 \cdots \mathrm{Col}_1(A)}_{s}\ \underbrace{0 \cdots \mathrm{Col}_2(A)}_{s} \cdots \underbrace{0 \cdots 0\ \mathrm{Col}_n(A)}_{s}}_{n} \Big].
\end{aligned} \tag{3.50}$$

Denote by

$$H_{ij} = [\underbrace{0 \cdots 0}_{j-1} \mathrm{Col}_i(A) \underbrace{0 \cdots 0}_{s-j}], \quad i = 1, \cdots, n;\ j = 1, \cdots, s.$$

Then we have the following alternative expressions of $\pi_s^r(A)$ and $\pi_s^r(A)$ as follows.

$$\begin{aligned}
\pi_s^r(A) &= \Big[H_{11}\, H_{12} \cdots H_{1s} \cdots H_{n1}\, H_{n2} \cdots H_{ns} \Big]; \\
\pi_s^c(A) &= \Big[H_{11}\, H_{21} \cdots H_{n1} \cdots H_{1s}\, H_{2s} \cdots H_{ns} \Big].
\end{aligned}$$

That is, in (3.49) $\{H_{ij}\}$ are arranged in the order of $\mathrm{id}(i,j;n,s)$, while in (3.50) $\{H_{ij}\}$ are arranged in the order of $\mathrm{id}(j,i;s,n)$. Then (3.48) follows from Proposition 2.9 immediately. □

Next, we convert the product of a constant matrix with an unknown matrix as a product of a coefficient matrix with unknown vectors, which is conventional in linear algebra.

Proposition 3.10. *Assume* $A \in \mathscr{M}_{m\times n}$ *and* $X \in \mathscr{M}_{n\times s}$*, then*

$$AX = \pi_s^r(A)V_r(X), \quad or \quad AX = \pi_s^r(A)W_{[s,n]}V_c(X). \tag{3.51}$$

Alternatively,

$$AX = \pi_s^c(A)V_c(X), \quad or \quad AX = \pi_s^c(A)W_{[n,s]}V_r(X). \tag{3.52}$$

Proof. Using (3.49), a straightforward computation shows that

$$\pi_s^r(A)V_r(X) = \left[\sum_{k=1}^n \mathrm{Col}_k(A)x_{1k}, \sum_{k=1}^n \mathrm{Col}_k(A)x_{2k}, \cdots, \sum_{k=1}^n \mathrm{Col}_k(A)x_{nk}\right] = AX.$$

This is the first equality in (3.51).

Using Proposition 3.5 to the first equality of (3.51), the second equality is obtained. The proof of (3.52) is similar. □

Combining the above formulas with some previous results about the linear mappings of matrices, some useful formulas can be obtained.

Proposition 3.11. *Assume $X \in \mathscr{M}_{m\times n}$ and $A \in \mathscr{M}_{n\times s}$, then*

$$XA = (I_m \otimes V_r^{\mathrm{T}}(I_s))W_{s,m}A^{\mathrm{T}}V_c(X). \tag{3.53}$$

Proof. Using (3.52) and Table 3.2, we have

$$\begin{aligned} XA &= I_m(XA) = \pi_s^c(I_m)V_c(XA) \\ &= \pi_s^c(I_m)(A^{\mathrm{T}} \otimes I_m)V_c(X). \end{aligned} \tag{3.54}$$

Using (3.46) and (3.48), we have

$$\pi_s^c(I_m) = \left(I_m \otimes V_r^{\mathrm{T}}(I_s)\right) \otimes W_{[s,m]}.$$

Using Proposition 2.4, we have

$$(A^{\mathrm{T}} \otimes I_m)V_c(X) = A^{\mathrm{T}}V_c(X).$$

Plugging them into (3.54) yields (3.53). □

The following example shows how to convert a higher-order matrix product form of a unknown matrix into the standard form as coefficient matrix times power of unknowns.

Example 3.9. Let $A,B,C,Z \in \mathscr{M}_n$. Consider a mapping $p(Z) : Z \mapsto AZBZC$. We intend to express $p(x)$ into a "quadratic form" of Z.

Using (3.35) and (3.53), we have

$$\begin{aligned} V_c(p(Z)) &= (C^{\mathrm{T}} \otimes A)V_c(ZBZ) \\ &= (C^{\mathrm{T}} \otimes A) \ltimes (BZ)^{\mathrm{T}} \ltimes V_c(Z). \end{aligned} \tag{3.55}$$

Applying (3.43) to $(BZ)^{\mathrm{T}}$ yields

$$
\begin{aligned}
(BZ)^{\mathrm{T}} = Z^{\mathrm{T}}B^{\mathrm{T}} &= \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) W_{[n]} \ltimes B \ltimes V_c(Z^{\mathrm{T}}) \\
&= \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) W_{[n]} (B \otimes I_n) W_{[n]} V_r(Z^{\mathrm{T}}) \\
&= \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) (I_n \otimes B) V_c(Z).
\end{aligned}
$$

Finally, we have

$$
V_c(p(Z)) = (C^{\mathrm{T}} \otimes A) \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) (I_n \otimes B) V_c^2(Z). \tag{3.56}
$$

The polynomial expression of a matrix product with higher order of unknown matrix, as obtained in the previous example, is very convenient in some investigations. In the following we give an example about control system.

Example 3.10. Consider a linear control system

$$
\dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m. \tag{3.57}
$$

When a state feedback

$$
u = Fx + v
$$

is used for solving the decoupling problem, we have to exam the decoupling matrix [6]. The key issue for calculating the decoupling matrix is to calculate

$$
(A + BF)^k. \tag{3.58}
$$

$(A+BF)^k$ can be considered as a matrix product form of kth order of F. We intend to express it as a polynomial of F. First of all, Using Proposition 3.10, we express $A+BF$ as

$$
A + BF = A + Ef,
$$

where $f = V_r(F)$, and $E = \pi_m^r(B)$, which can be constructed by (3.46). Hence we have

$$
\begin{aligned}
(A + E \ltimes f)^2 &= (A + E \ltimes f)(A + E \ltimes f) \\
&= A^2 + [A \ltimes E + E \ltimes (I_n \otimes A)] \ltimes f + E \ltimes (I_n \otimes E) \ltimes f^2.
\end{aligned}
$$

It is clear that we can have the following polynomial form

$$
(A + BF)^k = A^k + P_1^k \ltimes f + P_2^k \ltimes f^2 + \cdots + P_k^k \ltimes f^k. \tag{3.59}
$$

Hence we have

$$
\begin{aligned}
(A + BF)^{k+1} &= (A + E \ltimes f)(A^k + P_1^k \ltimes f + P_2^k \ltimes f^2 + \cdots + P_k^k \ltimes f^k) \\
&= A^{k+1} + [A \ltimes P_1^k + E \ltimes (I_n \otimes A^k)] \ltimes f \\
&\quad + [A \ltimes P_2^k + E \ltimes (I_n \otimes P_1^k)] \ltimes f^2 \\
&\quad + \cdots + [A \ltimes P_k^k + E \ltimes (I_n \otimes P_k^k)] \ltimes f^{k+1}.
\end{aligned}
$$

For notational consistence, we define $P_0^k := A^k$. Then for the coefficients in (3.59) we have a recursive formula as

$$\begin{cases} P_1^1 = E, \\ P_i^k = A \ltimes P_i^{k-1} + E \ltimes (I_n \otimes P_{i-1}^{k-1}), \quad i = 1, 2, \cdots, k;\ k = 2, 3, \cdots. \end{cases} \tag{3.60}$$

Note that $P_i^k \in \mathcal{M}_{n \times n^{i+1}}$, which are easily computable. In fact, all the entries of the matrices in (3.58) are also functions of $\{f_{ij}\}$. The problem is, it is very difficult to calculate them. But the formulas (3.59) and (3.60) provide a convenient way to calculate the functions.

Finally, we consider how to convert a matrix into its stacking form and vise versa. The problem seems stupid because everybody knows how to do this. But what we are interested is the converting formulas, which are very useful in theoretical expressions and/or deductions.

Proposition 3.12. *Assume $A \in \mathcal{M}_{m \times n}$, then*

$$V_r(A) = A \ltimes V_r(I_n), \tag{3.61}$$

$$V_c(A) = W_{[m,n]} \ltimes A \ltimes V_c(I_n). \tag{3.62}$$

Proof. A straightforward computation yields (3.61). Using (1.53) to (3.61) yields (3.62). □

Conversely, we can reconstruct A from its row or column stacking form.

Proposition 3.13. *Assume $A \in \mathcal{M}_{m \times n}$, then*

$$A = [I_m \otimes V_r^{\mathrm{T}}(I_n)] \ltimes V_r(A) = [I_m \otimes V_r^{\mathrm{T}}(I_n)] \ltimes W_{[n,m]} \ltimes V_c(A). \tag{3.63}$$

Proof. Replacing A and X in (3.51) by I_m and A respectively, and then using (3.46) yield the first equality of (3.63). Applying (1.53) to the first equality yields the second equality. □

We give a numerical example to depict this.

Example 3.11. Given a matrix $A \in \mathcal{M}_{3 \times 2}$ as

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Now $m = 3$, $n = 2$. Using Lemma 2.4, we have

$$V_r(I_n) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

$$A \ltimes V_r(I_n) = \begin{bmatrix} \begin{bmatrix} a_{11} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{12} \end{bmatrix} \\ \begin{bmatrix} a_{21} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{22} \end{bmatrix} \\ \begin{bmatrix} a_{31} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{32} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{31} \end{bmatrix} = V_r(A).$$

Using (3.62), we have (We refer to Example 1.10 for the numerical forms of $W_{[3,2]}$ and $W_{[2,3]}$.)

$$W_{[3,2]} \ltimes A = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ a_{21} & 0 & a_{22} & 0 \\ a_{31} & 0 & a_{32} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ 0 & a_{21} & 0 & a_{22} \\ 0 & a_{31} & 0 & a_{32} \end{bmatrix}.$$

Then

$$W_{[3,2]} \ltimes A \ltimes V_r(I_n) = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ a_{21} & 0 & a_{22} & 0 \\ a_{31} & 0 & a_{32} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ 0 & a_{21} & 0 & a_{22} \\ 0 & a_{31} & 0 & a_{32} \end{bmatrix} \ltimes \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = V_c(A).$$

Next, we verify (3.63). For the first part we have

$$\begin{aligned} [I_3 \otimes V_r^{\mathrm{T}}(I_2)] \ltimes V_r(A) &= [I_3 \otimes (1\,0\,0\,1)] \ltimes V_r(A) \\ &= \begin{bmatrix} 1&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&1 \end{bmatrix} \ltimes \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{31} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = A. \end{aligned}$$

To verify the second part of (3.63), we have

$$
\begin{aligned}
&[I_3 \otimes (1\,0\,0\,1)] \ltimes W_{[2,3]} \\
&= \begin{bmatrix} 1&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&1 \end{bmatrix} \ltimes \begin{bmatrix} 1&0&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&1&0&0&0 \\ 0&0&0&0&0&1 \end{bmatrix} \\
&= \begin{bmatrix} 1&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&1 \end{bmatrix}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&[I_3 \otimes (1\,0\,0\,1)] \ltimes W_{[2,3]} \ltimes V_c(A) \\
&= \begin{bmatrix} 1&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&1 \end{bmatrix} \ltimes \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = A.
\end{aligned}
$$

3.5 Two Applications

3.5.1 General Linear Group and Its Algebra

We refer to [1] for differentiable manifold. (A reader who is not familiar with the basic concepts in differential geometry may skip this subsection.)

Definition 3.4. A set G with an operator $*: G \times G \to G$ is called a Lie group, if

1. $(G, *)$ is a group;
2. G is an analytic manifold;
3. Both the product $*$ and the inverse $g \mapsto g^{-1}$ are analytic.

Consider the set

$$G := \{A \in \mathscr{M}_n \mid \det(A) \neq 0\}.$$

G can be naturally imbedded into $\mathbb{R}^{n^2}$. Hence as an open subset of $\mathbb{R}^{n^2}$, G is an n^2-dimensional analytic manifold. More over $(G, \cdot)$ is a group, where $\cdot$ is the matrix product over $\mathscr{M}_n$. Eventually, it is easy to check that $(G, \cdot)$, with the differential structure inherited from $\mathbb{R}^{n^2}$, is a Lie group, which is called the general linear group and denoted by $GL(n, \mathbb{R})$.

Let M be a k-dimensional analytic manifold, $f(x)$, $x \in M$ is called a vector field if at each $x \in M$ it is a k-dimensional vector. The set of (analytic) vector fields on M is denoted by $V^{\omega}(M)$, which is a vector space. The Lie bracket on $V^{\omega}(M)$ is defined as (within each coordinate chart)

$$[f(x),g(x)] = \frac{\partial g(x)}{\partial x} f(x) - \frac{\partial f(x)}{\partial x} g(x), \quad f(x),g(x) \in C^{\omega}(M), \tag{3.64}$$

where $\frac{\partial g(x)}{\partial x}$ is the Jacobian matrix of $g(x)$. That is,

$$\frac{\partial g(x)}{\partial x} = J_{g(x)} := \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \cdots & \frac{\partial g_1(x)}{\partial x_k} \\ \vdots & & \\ \frac{\partial g_k(x)}{\partial x_1} & \cdots & \frac{\partial g_k(x)}{\partial x_k} \end{bmatrix}.$$

Then $V^{\omega}(M)$ with this Lie bracket becomes a Lie algebra.

Let $\psi : M \to M$ be an analytic diffeomorphism (bijective mapping). Then ψ can reduce a mapping $\psi_* : V^{\omega}(M) \to V^{\omega}(M)$ as

$$\psi_*(f(x)) := J_{\psi} f(\phi^{-1}(x)), \quad f(x) \in V^{\omega}(M). \tag{3.65}$$

Now assume G is a Lie group and $X \in G$, then we define a mapping, called the left-displacement of x, defined as

$$\phi_X^L : g \to xg, \quad g \in G. \tag{3.66}$$

Then it is easy to see that $\phi_X^L : G \to G$ is an analytic diffeomorphism.

A vector field $f(x) \in V^{\omega}(G)$ is called a left-invariant vector field, if

$$(\phi_X^L)_*(f(P)) = f(\phi_X^L(P)) = f(XP), \quad \forall X, P.$$

It is easy to prove that the set of left-invariant vector fields forms a Lie-subalgebra of $V^{\omega}(G)$. This Lie algebra is called the Lie algebra of the Lie group G.

In this subsection we intend to prove the following important fact:

Theorem 3.1. *The Lie algebra of the general linear group $GL(n,\mathbb{R})$ is the general linear algebra $gl(n,\mathbb{R})$.*

Proof. For notational compactness, denote by $G := GL(n,\mathbb{R})$ and its Lie algebra $V(G) := V^{\omega}(GL(n,\mathbb{R}))$. Let the coordinates of $X \in G$ be $V_c(X)$. Then a vector field $F(X) \in V(G)$ is also considered as an $n \times n$ matrix. In conventional way, the vector fields should be expressed as $V_c(F(X))$.

Now assume $F(X) \in V(G)$ is a left invariant vector field, and $F(I_n) = A$. Fix an $X \in G$, we define the left displacement as

$$\phi_X^L : P \mapsto XP, \quad \forall P \in GL(n,\mathbb{R}).$$

Since $X \in \mathbb{R}^{n \times n}$, according to Table 3.2 we know that the matrix expression of ϕ_X^L under column stacking form is $I_n \otimes X$. Hence

$$V_c\left(\phi_X^L(P)\right) = (I_n \otimes X)p,$$

where $p = V_c(P)$. Denote by $a = V_c(A)$. Now consider the left displacement as a diffeomorphism on $GL(n,\mathbb{R})$, it can drive the tangent vector A at $I \in G$ to the point $X \in G$, precisely, it is

$$\frac{\partial \phi_X^L(P)}{\partial p} a = (I_n \otimes X)a.$$

Now since $F(X)$ is left invariant, when A is driven to X the vector is exactly $F(X)$. Denoting $x = V_c(X)$, we, therefore, have

$$V_c(F(X)) = J_{\phi_X^L} a = \frac{\partial (I_n \otimes X)x}{\partial x} a = (I_n \otimes X)a = V_c(XA).$$

Again, the last equality comes from Table 3.2. We conclude that the left invariant vector field $F(X)$, when expressed as a matrix, is

$$F(X) = XA. \tag{3.67}$$

Now let $F(X)$ and $W(X)$ be two left invariant vector fields with $F(I_n) = A$ and $W(I_n) = B$. Using (3.67), we have that the matrix expressions of $F(X)$ and $G(X)$ are $F(X) = XA$ and $W(X) = XB$ respectively. Back to conventional vector form, we have

$$F(x) = (A^{\mathrm{T}} \otimes I_n)x, \quad W(x) = (B^{\mathrm{T}} \otimes I_n)x.$$

Using formula (11c), we have

$$\begin{aligned}
[F(x), W(x)] &= (B^{\mathrm{T}} \otimes I_n)(A^{\mathrm{T}} \otimes I_n)x - (A^{\mathrm{T}} \otimes I_n)(B^{\mathrm{T}} \otimes I_n)x \\
&= ((B^{\mathrm{T}} \otimes I_n)(A^{\mathrm{T}} \otimes I_n) - (A^{\mathrm{T}} \otimes I_n)(B^{\mathrm{T}} \otimes I_n))x \\
&= (B^{\mathrm{T}}A^{\mathrm{T}} \otimes I_n - A^{\mathrm{T}}B^{\mathrm{T}} \otimes I_n)x = ((AB - BA) \otimes I_n)x.
\end{aligned}$$

Checking Table 3.2 again, we know that the matrix expression of $U(X) := [F(x), W(x)]$ is $(AB - BA)X$, which is a left invariant vector field with $U(I_n) = AB - BA$.

We conclude that the generalized linear algebra $gl(n,\mathbb{R})$, as a Lie algebra, is the Lie algebra of the generalized linear group $GL(n,\mathbb{R})$, as a Lie group. □

Definition 3.5. Let G be a Lie group. $H < G$ is a subgroup. H is called a Lie subgroup of G, if H is a regular sub-manifold of G.

Checking a regular sub-manifold is in general a difficult job. The following result is very convenient in use.

Theorem 3.2 ([1]). *Let G be a Lie group. $H < G$ is a subgroup. If H is closed under the topology of G, H is a Lie subgroup.*

Proposition 3.14. *1. Define*

$$O(n,\mathbb{R}) = \left\{ A \in GL(n,\mathbb{R}) \mid A^T = A^{-1} \right\}.$$

Then $O(n,\mathbb{R})$ is a Lie sub-group of $GL(n,\mathbb{R})$, called the orthogonal group.

2. *Define*

$$SO(n,\mathbb{R}) = \{A \in O(n,\mathbb{R}) \mid \det(A) = 1\}.$$

Then $SO(n,\mathbb{R})$ is a Lie sub-group of $O(n,\mathbb{R})$, called the special orthogonal group.

3. *Define*

$$SP(2n,\mathbb{R}) = \left\{ A \;\middle|\; \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} A + A^T \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = 0 \right\}.$$

Then $SP(2n,\mathbb{R})$ is a Lie sub-group of $GL(2n,\mathbb{R})$, called the symplectic group.

4. *$o(n,\mathbb{R})$ is the Lie algebra of $O(n,\mathbb{R})$.*
5. *$o(n,\mathbb{R})$ is also the Lie algebra of $SO(n,\mathbb{R})$.*
6. *$sp(2n,\mathbb{R})$ is the Lie algebra of $SP(2n,\mathbb{R})$.*

We leave the proof of Proposition 3.14 to the reader.

3.5.2 Hautus and Sylvester Equations

As another application, we consider Hautus equation. Let $A_i \in \mathcal{M}_{n\times m}$, $q_i(t)$, $i = 1,\cdots,k$ be some polynomials, $S \in \mathcal{M}_{p\times p}$, $R \in \mathcal{M}_{n\times p}$, $X \in \mathcal{M}_{m\times p}$. The following equation about unknown matrix X is called a Hautus equation:

$$A_1 X q_1(S) + \cdots + A_k X q_k(S) = R. \tag{3.68}$$

To show the importance of Hautus equation, we consider some of its special cases.

Let $A \in \mathcal{M}_{n\times n}$, $S \in \mathcal{M}_{p\times p}$, and $R \in \mathcal{M}_{n\times p}$. The following equation is called the Sylvester equation, which is very useful in control theory.

$$AX - XS = R. \tag{3.69}$$

It is easy to see that (3.69) is obtained from (3.68) with $A_1 = A$, $A_2 = I$, $q_1(t) = 1$, $q_2(t) = -t$.

Assume A and S are square matrices, and B, P, C, Q are matrices with proper dimensions. The following equation is called the regulation equation, which was proposed when investigating regulation problem of control systems.

$$\begin{aligned} \Pi S &= A\Pi + B\Gamma + P, \\ 0 &= C\Pi + Q. \end{aligned} \tag{3.70}$$

(3.70) can be converted into a Hautus equation:

$$A_1 X - A_2 X S = R,$$

where

$$A_1 = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}; A_2 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}; R = \begin{bmatrix} -P \\ -Q \end{bmatrix}; X = \begin{bmatrix} \Pi \\ \Gamma \end{bmatrix}.$$

Theorem 3.3. *The Hautus equation has solution for each R, if and only if, the n rows of matrix*

$$A(\lambda) = A_1 q_1(\lambda) + \cdots + A_k q_k(\lambda) \tag{3.71}$$

are linearly independent for each eigenvalue of S. Moreover, if $n = m$, the solution is unique.

Proof. Assume $T \in O(p, \mathbb{R})$, and let

$$\tilde{X} = XT, \quad \tilde{S} = T^{-1}ST, \quad \tilde{R} = R.$$

Then the equation (3.68) can be converted as

$$A_1 \tilde{X} q_1(\tilde{S}) + \cdots + A_k \tilde{X} q_k(\tilde{S}) = \tilde{R}. \tag{3.72}$$

It is clear that the existence of solution X with respect to each R is equivalent to the existence of $\tilde{X}$ with respect to $\tilde{R}$. Moreover, $X = \tilde{X}T^{-1}$. Hence, without loss of generality, we can assume S is in its Jordan canonical form.

Using Proposition 3.6, (3.68) can be converted to

$$\left[q_1(S^{\mathrm{T}}) \otimes A_1 + \cdots + q_k(S^{\mathrm{T}}) \otimes A_k\right] x = r, \tag{3.73}$$

where $x = V_c(X)$, $r = V_c(R)$. Since S has Jordan canonical form

$$S = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix},$$

hence (3.73) can be expressed as $Ex = r$, where

$$E = \begin{bmatrix} Q(\lambda_1) & 0 & \cdots & 0 \\ * & Q(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ * & \cdots & * & Q(\lambda_p) \end{bmatrix},$$

where

$$Q(t) = q_1(t)A_1 + \cdots + q_k(t)A_k.$$

The conclusion follows. □

Corollary 3.1. *Sylvester equation (3.69) has solution with respect to each R, if and only if, A and S have no common eigenvalue. Moreover, in this case the solution is unique.*

Proof. For Sylvester equation, we have

$$A(\lambda) = A - I\lambda.$$

Then it is clear that for each $\lambda \in \sigma(S)$, $A(\lambda)$ is nonsingular, if and only if, A and S has no common eigenvalue. □

Remark 3.1. 1. Lyapunov mapping is closely related to the stability of linear systems. There are many discussions. For instance, Its many properties have been discussed in [2]. Its applications to switched linear systems have been discussed in [3]. The following is an interesting problem: Denote by S the set of $n \times n$ symmetric matrices and K the set of $n \times n$ skew symmetric matrices. Both of them are subspaces of $\mathscr{M}_{n\times n}$. In addition, it is easy to prove that they are invariant subspaces under the Lyapunov L_A. Hence, we can restrict L_A on these two invariant subspaces respectively, and denote them as L_A^S, L_A^K. It was proved in [2] that

$$\|L_A\| = \max\{\|L_A^S\|, \|L_A^K\|\}.$$

Based on this we propose a conjecture

$$\|L_A\| = \|L_A^S\|. \tag{3.74}$$

It is still an open conjecture.

2. Hautus equation and Sylvester equation have many applications in control theory, particularly, in the investigation of output regulation problem. More discussion about Hautus equation and Sylvester equation can be found in [7]. We refer the reader, who is interested in output regulation, to [5].

Exercise 3

1. Prove $(\mathbb{R}^3, \times_c)$ is a Lie algebra.
2. Prove $(\mathscr{M}_n, [\cdot,\cdot])$ is a Lie algebra.
3. Prove Proposition 3.2.
4. Let V be a finite dimensional vector space over $\mathbb{R}$. Denote by $End(V)$ the set of linear mappings $\phi : V \to V$. (In general this set is called the endomorphisms of V.) Let $\mathscr{G}$ be a Lie algebra. A mapping $\pi : \mathscr{G} \to End(V)$ is called a representation of $\mathscr{G}$ in V if π satisfies[9]

   (i) π is linear,
   (ii) $\pi([X,Y]) = \pi(X)\pi(Y) - \pi(Y)\pi(X), \quad X, Y \in \mathscr{G}.$

   Prove that the adjoint representation ad_X is a representation.
5. Consider $o(3,\mathbb{R})$.

   (i) Find a basis of $o(3,\mathbb{R})$.
   (ii) Find the structure matrix of $o(3,\mathbb{R})$ with respect to the basis you obtained.
   (iii) Let

$$X = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \in o(3,\mathbb{R}).$$

Find the matrix expression of ad_X with respect to the basis.

6. Consider $sp(4,\mathbb{R})$.

   (i) Find a basis of $sp(4,\mathbb{R})$.
   (ii) Find the structure matrix of $sp(4,\mathbb{R})$ with respect to the basis you obtained.
   (iii) Let

$$X = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

   Check whether $X \in sp(4,\mathbb{R})$? If "yes", find the matrix expression of ad_X with respect to the basis.

7. Assume $(V,[\cdot,\cdot])$ is a Lie algebra and $e = \{e_1,\cdots,e_n)$ is a basis of V. Moreover, $X \in V$ and the adjoint representation ad_X has matrix expression as $M_e \in \mathscr{M}$ with respect to basis e. Let $\alpha = \{\alpha_1,\cdots,\alpha_n\}$ be another basis of V and

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = T \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}.$$

   Find the matrix expression of ad_X with respect to the basis α.
8. Prove that

$$\mathscr{G}_N := \{X \in gl(n,\mathbb{R}) \mid L_N(X) = 0\},$$

   where L_N is defined in (3.30), is a Lie sub-algebra of $gl(n,\mathbb{R})$.
9. Prove Proposition 3.4.
10. Let $A \in \mathscr{M}_{m\times n}$, $X \in \mathscr{M}_{n\times q}$, $Y \in \mathscr{M}_{p\times m}$. Prove the following two formulas:

$$V_c(AX) = W_{[m,q]}AW_{[q,n]}V_c(X). \tag{3.75}$$

$$V_r(YA) = W_{[n,p]}A^{\mathrm{T}}W_{[p,m]}V_r(Y). \tag{3.76}$$

11. a. Consider $\mathscr{M}_n$ and define a product as

$$\langle A,B\rangle = BA - AB.$$

    Prove that $(\mathscr{M}_n,\langle\cdot,\cdot\rangle)$ is a Lie algebra.
    b. Assume the Lie algebra of a Lie group is generalized by right invariant vector fields. Then prove that the Lie algebra of $GL(n,\mathbb{R})$ is $(\mathscr{M}_n,\langle\cdot,\cdot\rangle)$.
    c. Consider $V(\mathbb{R}^n)$, the set of smooth vector fields on $\mathbb{R}^n$. Define the Lie brocker as in . Prove that $V(\mathbb{R}^n)$ with Lie brocket is a Lie algebra.

d. Proof of Proposition 3.14. (Ignore this if you are not familiar with Differential Geometry.)

12. Consider the Riccati equation

$$A^{\mathrm{T}}X + XA + Q + XRX = 0,$$

where $A, Q, R \in \mathscr{M}_n$. Express it into a polynomial form as

$$C_0x^2 + C_1x + C_2 = 0,$$

where $x = V_c(X)$.

13. Given a matrix equation

$$AXBX^{\mathrm{T}}CX = XDX^{\mathrm{T}},$$

where $A, B, C, D, X \in \mathscr{M}_n$. Express it into a polynomial form as

$$Fx^3 - Gx^2 = (Fx - G)x^2 = 0,$$

where (i) $x = V_r(X)$, (ii) $x = V_c(X)$.

References

1. Boothby, W.: An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd edn. Academic Press, Inc., Orlando (1986)
2. Cheng, D.: On Lyapunov mapping and its applications. Commu. on Inform. Sys. **1**(3), 255–272 (2001)
3. Cheng, D., Guo, L., Huan, J.: On quadratic Lyapunov functions. IEEE Trans. Aut. Contr. **48**(5), 885–890 (2003)
4. Cheng, D., Xue, W., Huang, J.: On general Hamiltonian systems. In: Proc. ICARCV'98, pp. 185–189. Singapore (1998)
5. Huang, J.: Nonlinear Output Regulation, Theory and Applications. Siam, Philadelphia (2004)
6. Isidori, A.: Nonlinear Control Systems, 3rd edn. Springer (1995)
7. Knobloch, H., Isidori, A., Flockerzi, D.: Topics in Control Theory. Birkhäuser Verlag, Basel (1993)
8. Ooba, T., Funahashi, Y.: Two conditions concerning common quadratic Lyapunov functions for linear systems. IEEE Trans. Aut. Contr. **42**(5), 719–721 (1997)
9. Varadarajan, V.: Lie Groups, Lie Algebras, and Their Representations. Springer-Verlag, New York (1984)
10. Willems, J.: Stability Theory of Dynamical Systems. John Wiley & Sons Inc. (1970)