What is STP of Matrices in Hypermatrix Pearspective
从超矩阵的视角看矩阵半张量积

Daizhan Cheng

Research Center of STP Theory and Applications
LiaoCheng University
LiaoCheng, Shandong, China

2023 山东科学大讲堂
LiaoCheng, Shandong
December 27, 2023

Outline of Presentation

I. Hypermatrix(超矩阵)

☞ Hypermatrix: Multi-indexed Data

Definition 1.1[1]

(i) A set of order d data

$$A := \{a_{i_1,i_2,\cdots,i_d} \mid i_s \in [1, n_s],\ s \in [1, d]\} \in \mathbb{F}^{n_1\times\cdots\times n_d} \quad (1)$$

is called an order d hypermatrix (briefly, d-hypermatrix) with dimensions $n_1 \times n_2 \times \cdots \times n_d$. The set of d-hypermatrix with dimension $n_1 \times n_2 \times \cdots \times n_d$ is denoted by $\mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$, where $a_{i_1,i_2,\cdots,i_d} \in \mathbb{F}$ and $\mathbb{F}$ can be $\mathbb{R}$, $\mathbb{C}$, or other fields.

Definition 1.1(cont'd)

(ii) $A \in \mathbb{F}^{\overbrace{n \times n \times \cdots \times n}^{d}}$ is called a d-hypercubic (超立方阵).

(iii) $A \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_d}$ with $n_1 = n_d$ is called a d-hypersquare (超矩形阵).

[1] Lek-Heng Lim, Tensors and Hypermatrices, in L. Hogben (Ed.) *Handbook of Linear Algebra* (2nd ed.), Chapter 15, Chapman and Hall/CRC.https://doi.org/10.1201/b16113, 2013.

☞ Special Cases

(i) $d = 1$ (Vector):

$$A = \{x_i \mid i \in [1, n]\} \in \mathbb{F}^n.$$

Express A into vector form:

$$A = (x_1, \cdots, x_n)$$

or

$$A = (x_1, \cdots, x_n)^T$$

(ii) $d = 2$ (Matrix):

$$A = \{x_{i,j} \mid i \in [1, m],\ j \in [1, n]\} \in \mathbb{F}^{m \times n}.$$

Express A into matrix form:

$$A = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & & \cdots & \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix};$$

or

$$A^T;$$

or

$$V_r(A) = (x_{1,1}, x_{1,2}, \cdots, x_{1,n}, \cdots, x_{m,n})^T;$$

$$V_c(A) = (x_{1,1}, x_{2,1}, \cdots, x_{m,1}, \cdots, x_{m,n})^T.$$

(iii) $d = 3$ (Cubic Matrix):

$$A = \{d_{i,j,k} \mid i \in [1,p],\ j \in [1,m],\ k \in [1,n]\} \in \mathbb{F}^{m \times n}.$$

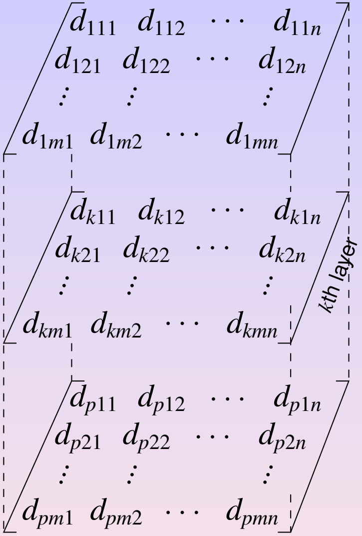

Figure 1: 一个立体阵

☞ Tensor

Definition 1.2

Let V be an n-dimensional vector space.

$$d = \{d_1, \cdots, d_n\}$$

a basis of V.

$$e = \{e_1, \cdots, e_n\}$$

a basis of V^*, which is dual to d. That is,

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

A multi-linear mapping $t : \underbrace{V \times \cdots \times V}_{r} \times \underbrace{V^* \times \cdots \times V^*}_{s} \to \mathbb{R}$ is called a tensor of covariant order r and contra-variant order s.

Definition 1.2(cont'd)

$$\mu^{i_1,\cdots,i_r}_{j_1,\cdots,j_s} := t(d_{i_1},\cdots,d_{i_r};e_{j_1},\cdots,e_{j_s}),$$

$$i_\alpha, j_\beta \in [1,n], \alpha \in [1,r], \beta \in [1,s].$$

$$D_t := \left\{\mu^{i_1,\cdots,i_r}_{j_1,\cdots,j_s} \mid i_\alpha, j_\beta \in [1,n], \alpha \in [1,r], \beta \in [1,s]\right\}$$

is the set of structure constants.

$$\text{order}(D_t) = r + s.$$

II. STP Approach to Hypermatrix

☞ Application of STP

Definition 2.1

Let $A \in \mathcal{M}_{m\times n}$ $B \in \mathcal{M}_{p\times q}$, $t = \operatorname{lcm}(n,p)$. The semi-tensor product (STP) of A and B is

$$A \ltimes B := \left(A \otimes I_{t/n}\right)\left(B \otimes I_{t/p}\right). \quad (2)$$

Example 2.2

Consider multi-linear mappings.

(i) Let $\pi_1 : \mathbb{R}^n \to \mathbb{R}$. Say,

$$\pi_1(\delta_n^i) = a_i, \quad i \in [1,n].$$

Example 2.2(cont'd)

Set

$$V_A = (a_1, \cdots, a_n).$$

Then

$$\pi(x) = V_A x.$$

(ii) Let $\pi_2 : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$. Say,

$$\pi_2(\delta_m^i, \delta_n^j) = a_{i,j}, \quad i \in [1, m], j \in [1, n].$$

Set

$$M_A = (a_{i,j}) \in \mathcal{M}_{m \times n}.$$

Let $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Then

$$\pi_2(x, y) = x^T M_A y.$$

Example 2.2(cont'd)

(iii) Let $\pi_3 : \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$. Say,

$$\pi_3(\delta_p^k, \delta_m^i, \delta_n^j) = d_{k,i,j}, \quad k \in [1,p], i \in [1,m], j \in [1,n].$$

Set

$$\begin{aligned} V_A &= [d_{1,1,1}, \cdots, d_{1,1,n}, \cdots, d_{1,m,1}, \cdots, d_{1,m,n}, \\ &\quad \cdots, d_{p,1,1}, \cdots, d_{p,m,n}] \in \mathbb{R}^{pmn}. \end{aligned}$$

Let $x \in \mathbb{R}^p$, $y \in \mathbb{R}^m$, $z \in \mathbb{R}^n$. Then

$$\pi_3(x,y,z) = V_A \ltimes x \ltimes y \ltimes z.$$

Example 2.3

Let $V = \mathbb{R}^n$. Consider a tensor $t \in \mathcal{T}_s^r$ with

$$\mu_{j_1,\cdots,j_s}^{i_1,\cdots,i_r} := t(\delta_n^{i_1},\cdots,\delta_n^{i_r};(\delta_n^{j_1})^T,\cdots,(\delta_n^{j_s})^T),$$

$$i_\alpha, j_\beta \in [1,n], \alpha \in [1,r], \beta \in [1,s].$$

Construct the structure matrix of t as

$$M_t = \begin{bmatrix} \mu_{1,1,\cdots,1}^{1,1,\cdots,1} & \mu_{1,1,\cdots,1}^{1,1,\cdots,2} & \cdots & \mu_{1,1,\cdots,1}^{1,1,\cdots,n} \\ \mu_{1,1,\cdots,2}^{1,1,\cdots,1} & \mu_{1,1,\cdots,2}^{1,1,\cdots,2} & \cdots & \mu_{1,1,\cdots,2}^{1,1,\cdots,n} \\ & & \cdots & \\ \mu_{n,n,\cdots,1}^{1,1,\cdots,1} & \mu_{n,n,\cdots,2}^{1,2,\cdots,1} & \cdots & \mu_{n,n,\cdots,n}^{n,n,\cdots,n} \end{bmatrix}$$

Let $x_i \in \mathbb{R}^n$, $i \in [1,r]$, $\omega_j \in (\mathbb{R}^n)^*$, $j \in [1,s]$. Then

$$t(x_1,\cdots,x_r;\omega_1,\cdots,\omega_s) = \omega_s \ltimes \cdots \ltimes \omega_1 \ltimes M_t \ltimes x_1 \ltimes \cdots \ltimes x_r.$$

☞ Summary

(i) **Classical Matrix Theory is used for Matrices and Vectors.**

(ii) **STP Theory can be used for Hypermatrices.**

(iii) **The multi-linear mapping over Hypermatrices can be realized by STP as:**

Hypermatrix $\rightarrow$ Matrix M_H

$$\Downarrow$$

STP： $M_H \ltimes M_H \rightarrow M_H$

Figure 2: Using STP For Hypermatrices

III. Matrix Expression of Hypermatrix

☞ Set Point of View for Hypermatrix

A hypermatrix consists of two ingredients:

(i) a set of data

$$D_A := \{a_{i_1,\cdots,i_d} \mid i_s \in [1, n_s],\ s \in [1, d]\}; \tag{3}$$

(ii) an ordered set of indexes.

$$\mathbf{r} := \{r_1, r_2, \cdots, r_d\}.$$

Definition 3.1

Given a hypermatrix $A = [a_{r_1,r_2,\cdots,r_d}]$. For each partition

$$\begin{aligned} \mathbf{r} &= \{r_1, r_2, \cdots, r_d\} = (r_{i_1}, r_{i_2}, \cdots, r_{i_p}) \\ &\quad \cup \{r_{j_1}, r_{j_2}, \cdots, r_{j_q}\} := \mathbf{r}_1 \cup \mathbf{r}_2, \end{aligned} \tag{4}$$

there is a matrix expression of A, denoted by

Definition 3.1(cont'd)

$$M_A^{\mathbf{r}_1 \times \mathbf{r}_2} := M_A^{\mathbf{r}_1} \in \mathbb{F}^{s \times t}, \tag{5}$$

where $s = \prod_{k=1}^{p} n_{i_k}$, $t = \prod_{k=1}^{q} n_{j_k}$. Moreover, the elements in $M_A^{\mathbf{r}_1 \times \mathbf{r}_2}$ are $\{a_{r_1, r_2, \cdots, r_d}\}$, which are arranged by $ID(\mathbf{r}_1; n_{r_{i_1}}, n_{r_{i_2}}, \cdots, n_{r_{i_p}})$ for rows, and by $ID(\mathbf{r}_2; n_{r_{j_1}}, n_{r_{j_2}}, \cdots, n_{r_{j_q}})$ for columns.

Example 3.2

Given $A = [a_{i_1,i_2,i_3}] \in \mathbb{F}^{2\times3\times2}$. Then

(i)

$$M_A^{\emptyset} = [a_{111}, a_{112}, a_{121}, a_{122}, a_{131}, a_{132}, a_{211}, a_{212}, a_{221}, a_{222}, a_{231}, a_{232}].$$

(ii)

$$M_A^{(1)} = \begin{bmatrix} a_{111} & a_{112} & a_{121} & a_{122} & a_{131} & a_{132} \\ a_{211} & a_{212} & a_{221} & a_{222} & a_{231} & a_{232} \end{bmatrix};$$

$$M_A^{(2)} = \begin{bmatrix} a_{111} & a_{112} & a_{211} & a_{212} \\ a_{121} & a_{122} & a_{221} & a_{222} \\ a_{131} & a_{132} & a_{231} & a_{232} \end{bmatrix}; \quad \text{etc.}$$

Example 3.2(cont'd)

(iii)

$$M_A^{(1,2)} = \begin{bmatrix} a_{111} & a_{112} \\ a_{121} & a_{122} \\ a_{131} & a_{132} \\ a_{211} & a_{212} \\ a_{221} & a_{222} \\ a_{231} & a_{232} \end{bmatrix}; \quad \text{etc.}$$

$$M_A^{(1,3)} = \begin{bmatrix} a_{111} & a_{121} & a_{131} \\ a_{112} & a_{122} & a_{132} \\ a_{211} & a_{221} & a_{231} \\ a_{212} & a_{222} & a_{232} \end{bmatrix}; \quad \text{etc.}$$

(iv)

$$M_A^{(1,2,3)} = \left(M_A^{\emptyset}\right)^{\mathrm{T}}.$$

Definition 3.3

(i)

$$V_A := M_A^{\emptyset \times \mathbf{r}}$$

is called the (row) vector expression of hypermatrix A.

(ii)

$$M_A := M_A^{\{1\} \times \mathbf{r} \setminus \{1\}}$$

is called the matrix-1 expression of hypermatrix A.

☞ Hypervectors

Definition 3.4

Let $x_i \in \mathbb{F}^{n_i}$, $i \in [1, d]$. Then

$$x := \ltimes_{i=1}^{d} x_i \tag{6}$$

is called a hypervector of order d.

The set of hypervectors is denoted by

$$\mathbb{F}^{n_1 \ltimes \cdots \ltimes n_d} := \{x \mid x \text{ is obtaied by (6)}\}.$$

Note that the components of x can be expressed as

$$D_x := \left\{x_{i_1,\cdots,i_d} = x_1^{i_1} x_2^{i_2} \cdots x_d^{i_d} \mid i_j \in [1, n_j], j \in [1, d]\right\},$$

where $x_r^{i_r}$ is the i_r component of x_r.

It is clear that D_x (or briefly hypervector x) is a hypermatrix.

Proposition 3.5

$$\mathbb{F}^{n_1 \ltimes \cdots \ltimes n_d} \subset \mathbb{F}^{n_1 \times \cdots \times n_d} \qquad (7)$$

is a subset of hypermatrices.

Remark 3.6

Since the set of hypervectors $\mathbb{F}^{n_1 \ltimes \cdots \ltimes n_d}$ contains a set of basis of the set of hypermatrices $\mathbb{F}^{n_1 \times \cdots \times n_d}$, any multi-linear mapping over $\mathbb{F}^{n_1 \times \cdots \times n_d}$ is uniquely determined by its restriction on the set of hypervectors $\mathbb{F}^{n_1 \ltimes \cdots \ltimes n_d}$.

Definition 3.7

Assume $V \in (\mathbb{F}^n)^s$ is an s dimensional vector subspace of the n dimensional vector space on $\mathbb{F}$.

(i) A hypervector $x = \ltimes_{i=1}^{t} x_i$ with $x_i \in V$ is said to be a hypervector over V, denoted by $x \in V^t$.

(ii) If $x = \ltimes_{i=1}^{t} x_i \in V^t$ and

$$\mathrm{rank}[x_1, \cdots, x_t] = \dim(V) \quad (= s),$$

x is said to be of full rank.

(iii) If $x = \ltimes_{i=1}^{t} x_i \in V^t$ is of full rank and $\{x_{i_1}, \cdots, x_{i_s}\} \subset \{x_1, \cdots, x_t\}$ is the first set of basis of V searching from left, it is called the first basis subset.

IV. σ-transpose of Hypermatrices

☞ Permutation Group

Definition 4.1

The symmetric group of order n, denoted by $\mathbf{S}_n$, is the set of permutations of n objects. The product over $\mathbf{S}_n$ is the compounded permutations.

Example 4.2

Consider $\mathbf{S}_3$.

(i) $\sigma \in \mathbf{S}_3$ can be expressed by

$$\sigma : [1, 2, 3] \to [2, 1, 3]; \quad \text{or } \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}; \quad \text{or } (1, 2)$$

.

Example 4.2(cont'd)

(ii) $\mu : [1,2,3] \to [3,1,2] \in \mathbf{S}_3$. Then

$$\mu \circ \sigma = [1,2,3] \xrightarrow{\sigma} [2,1,3] \xrightarrow{\mu} [1,3,2].$$

That is,

$$\mu \circ \sigma = (2,3).$$

(iii)

$$\sigma^{-1} = (2,1); \quad \mu^{-1} = (1,2,3).$$

Remark 4.3

Each $\sigma \in \mathbf{S}_n$ can be expressed as a product of swaps， say,

$$(1,2,3) = (1,2)(2,3).$$

If σ can be expressed as a product of even swaps, then $sign(\sigma) = 1$; otherwise, $sign(\sigma) = -1$. The expression is not unique, but the odevity (odd and even) remain unchanged.

Definition 4.4

(i) Consider a d-hypermatrix $A = [a_{j_1,j_2,\cdots,j_d}] \in \mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$, and assume $\sigma \in \mathbf{S}_d$. The

$$A^\sigma := \left[a_{j_{\sigma(1)}\cdots j_{\sigma(d)}}\right] \in \mathbb{F}^{n_{\sigma(1)}\times\cdots\times n_{\sigma(d)}}. \tag{8}$$

(ii) If a d-hypercubic $A \in \mathbb{F}^{\overbrace{n\times\cdots\times n}^{d}}$ satisfies

$$A^\sigma = A, \quad \forall\sigma \in \mathbf{S}_d,$$

then A is said to be a symmetric d-heypercubic.

(iii) A d-hypercubic $A \in \mathbb{F}^{\overbrace{n\times\cdots\times n}^{d}}$ is said to be skew-symmetric if

$$A^\sigma = sign(\sigma)A, \quad \forall\sigma \in \mathbf{S}_d.$$

Proposition 4.5

A $d = 2$ hypercubic $A \in \mathbb{F}^{n \times n}$ is (skew-)symmetric, if and only if, M_A is (skew-)symmetric.

Proposition 4.6

Let $A \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ and $\mathbf{r} \subset \mathbf{d} =< d >$. Then

$$\left[M_A^{\mathbf{r} \times (\mathbf{d} \backslash \mathbf{r})} \right]^T = M_A^{(\mathbf{d} \backslash \mathbf{r}) \times \mathbf{r}}. \tag{9}$$

Permutation Matrix

Algorithm 4.7

Let $n = \prod_{i=1}^d n_i$, $n_i \geq 2$, $\sigma \in \mathbf{S}_n$. A logical matrix $W^{\sigma}_{[n_1,n_2,\cdots,n_d]} \in \mathcal{L}_{n\times n}$, called a σ-permutation matrix, is constructed as follows:

- Step 1: Define

$$D = D^{\sigma}_{[n_1,n_2,\cdots,n_d]} := \left\{ \delta^{j_1}_{n_{\sigma(1)}} \delta^{j_2}_{n_{\sigma(2)}} \cdots \delta^{j_d}_{n_{\sigma(d)}} \;\middle|\; j_i \in [1, n_{\sigma(i)}], i = 1, 2, \cdots, d \right\}.$$

Exampl 4.8

Assume $d = 3$; $n_1 = 2$, $n_2 = 3$, $n_3 = 5$. $\sigma = (1,2,3)$. We have

$$\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1.$$

$$n_{\sigma_1} = 3,\ n_{\sigma_2} = 5,\ n_{\sigma_3} = 2.$$

$$D^{\sigma}_{[2,3,5]} = \left[\delta_3^{j_1} \delta_5^{j_2} \delta_2^{j_3} \mid j_1 \in [1,3], j_2 \in [1,5], j_3 \in [1,2] \right].$$

Algorithm 4.7(cont'd)

- Step 2: Arrange $\{\sigma(i) \mid i \in [1,d]\}$ into an increasing sequence as

$$1 = \sigma(i_1) < \sigma(i_2) < \cdots < \sigma(i_d) = d.$$

That is,

$$i_j = \sigma^{-1}(j), \quad j \in [1,d].$$

Exampl 4.8 (cont'd)

$$1 = \sigma(3) < \sigma(1) < \sigma(2) = 3.$$

Hence,

$$i_1 = 3,\ i_2 = 2,\ i_3 = 2.$$

Algorithm 4.7(cont'd)

- Step 3:
  Set an index order as

$$
\begin{aligned}
ID_\sigma &:= ID\left(j_{i_1}, j_{i_2}, \cdots, j_{i_d}; k_{\sigma(i_1)}, k_{\sigma(i_2)}, \cdots, k_{\sigma(i_d)}\right) \\
&= ID\left(j_{\sigma^{-1}(1)}, j_{\sigma^{-1}(2)}, \cdots, j_{\sigma^{-1}(d)}; k_1, k_2, \cdots, k_d\right).
\end{aligned}
$$

Exampl 4.8 (cont'd)

$$ID_\sigma = ID\left(j_3, j_1, j_2 \mid k_1, k_2, k_3\right)$$

Algorithm 4.7(cont'd)

- Step 4:

$$W^{\sigma}_{[n_1,n_2,\cdots,n_d]} := \left[d^{j_1}_{n_{\sigma(1)}} d^{j_2}_{n_{\sigma(2)}} \cdots d^{j_d}_{n_{\sigma(d)}} \right| \text{arranged by the order of } ID_{\sigma} \Big]. \qquad (10)$$

Exampl 4.8 (cont'd)

$$\begin{aligned} W_{\sigma} &= \left\{ \delta_3^{j_1}\delta_5^{j_2}\delta_2^{j_3} \mid j_1 \in [1,3], j_2 \in [1,5], j_3 \in [1,2] \right\} \\ &= \{ \delta_3^1\delta_5^1\delta_2^1, \delta_3^1\delta_5^2\delta_2^1, \cdots, \delta_3^1\delta_5^5\delta_2^1, \\ &\quad \delta_3^1\delta_5^1\delta_2^2, \delta_3^1\delta_5^2\delta_2^2, \cdots, \delta_3^1\delta_5^5\delta_2^2, \\ &\quad \delta_3^2\delta_5^1\delta_2^1, \delta_3^2\delta_5^2\delta_2^1, \cdots, \delta_3^2\delta_5^5\delta_2^1, \\ &\quad \delta_3^2\delta_5^1\delta_2^2, \delta_3^2\delta_5^2\delta_2^2, \cdots, \delta_3^2\delta_5^5\delta_2^2, \\ &\quad \delta_3^3\delta_5^1\delta_2^1, \delta_3^3\delta_5^2\delta_2^1, \cdots, \delta_3^3\delta_5^5\delta_2^1, \\ &\quad \delta_3^3\delta_5^1\delta_2^2, \delta_3^3\delta_5^2\delta_2^2, \cdots, \delta_3^3\delta_5^5\delta_2^2 \} \\ &= \delta_{30}[1,3,5,7,9,11,13,15,17,19,21,23,25,27, \\ &29,2,4,6,8,10,12,14,16,18,20,22,24,26,28,30]. \end{aligned}$$

Example 4.8

Consider $d = 3$, $n_1 = 2$, $n_2 = 3$, and $n_3 = 5$. We construct $W^\sigma := W^\sigma_{[2,3,5]}$.

(1) $\sigma_1 = id$ (i.e., $[1,2,3] \to [1,2,3]$): We have

$$W^{\sigma_1} = I_{30}.$$

(2) $\sigma_2 = (2,3)$ (i.e., $[1,2,3] \to [1,3,2]$): Then

$$D = \{\delta_2^{j_1}\delta_5^{j_2}\delta_3^{j_3} \mid j_1 \in [1,2], j_2 \in [1,5], j_3 \in [1,3]\}$$

$$\begin{aligned} W^{\sigma_2} = \big[& \delta_2^1\delta_5^1\delta_3^1, \delta_2^1\delta_5^2\delta_3^1, \delta_2^1\delta_5^3\delta_3^1, \delta_2^1\delta_5^4\delta_3^1, \delta_2^1\delta_5^5\delta_3^1, \\ & \delta_2^1\delta_5^1\delta_3^2, \delta_2^1\delta_5^2\delta_3^2, \delta_2^1\delta_5^3\delta_3^2, \delta_2^1\delta_5^4\delta_3^2, \delta_2^1\delta_5^5\delta_3^2, \\ & \delta_2^1\delta_5^1\delta_3^3, \delta_2^1\delta_5^2\delta_3^3, \delta_2^1\delta_5^3\delta_3^3, \delta_2^1\delta_5^4\delta_3^3, \delta_2^1\delta_5^5\delta_3^3, \\ & \delta_2^2\delta_5^1\delta_3^1, \delta_2^2\delta_5^2\delta_3^1, \delta_2^2\delta_5^3\delta_3^1, \delta_2^2\delta_5^4\delta_3^1, \delta_2^2\delta_5^5\delta_3^1, \\ & \delta_2^2\delta_5^1\delta_3^2, \delta_2^2\delta_5^2\delta_3^2, \delta_2^2\delta_5^3\delta_3^2, \delta_2^2\delta_5^4\delta_3^2, \delta_2^2\delta_5^5\delta_3^2, \\ & \delta_2^2\delta_5^1\delta_3^3, \delta_2^2\delta_5^2\delta_3^3, \delta_2^2\delta_5^3\delta_3^3, \delta_2^2\delta_5^4\delta_3^3, \delta_2^2\delta_5^5\delta_3^3 \big] \end{aligned}$$

Example 4.8(cont'd)

$$W^{\sigma_2} = \delta_{30}[1, 4, 7, 10, 13, 2, 5, 8, 11, 14, 3, 6, 9, 12, 15, 16, 19, 22, 25, 28, 17, 20, 23, 26, 29, 18, 21, 24, 27, 30].$$

(3) $\sigma_3 = (1, 2)$ (i.e., $[1, 2, 3] \to [2, 1, 3]$):
Similarly, we have

$$D = \{\delta_3^{j_1}\delta_2^{j_2}\delta_5^{j_3} \mid j_1 \in [1, 3], j_2 \in [1, 2], j_3 \in [1, 5]\}$$

$$\begin{aligned} W^{\sigma_3} &= \left[\delta_3^1\delta_2^1\delta_5^1, \delta_3^1\delta_2^1\delta_5^2, \cdots, \delta_3^2\delta_2^1\delta_5^1, \right. \\ &\quad \left. \cdots, d_3^2\delta_2^1\delta_5^5, \cdots, \delta_3^3\delta_2^2\delta_5^5\right] \\ &= \delta_{30}[1, 2, 3, 4, 5, 11, 12, 13, 14, 15, 21, 22, 23, 24, \\ &\quad 25, 6, 7, 8, 9, 10, 16, 17, 18, 19, 20, 26, 27, 28, 29, 30]. \end{aligned}$$

Example 4.8(cont'd)

(4) $\sigma_4 = (1,2,3)$ (i.e., $[1,2,3] \to [2,3,1]$): We have

$$W^{\sigma_4} = \left[\delta_3^1\delta_5^1\delta_2^1, \cdots, \delta_3^3\delta_5^5\delta_2^2\right]$$
$$= \delta_{30}[1,3,5,7,9,11,13,15,17,19,21,23,25,27,$$
$$29,2,4,6,8,10,12,14,16,18,20,22,24,26,28,30].$$

(5) $\sigma_5 = (1,3,2)$ (i.e., $[1,2,3] \to [3,1,2]$): Then

$$W^{\sigma_5} = \left[\delta_5^1\delta_2^1\delta_3^1, \cdots, \delta_5^5\delta_2^2\delta_3^3\right]$$
$$= \delta_{30}[1,7,13,19,25,2,8,14,20,26,3,9,15,21,27$$
$$4,10,16,22,28,5,11,17,23,29,6,12,18,24,30].$$

(6) $\sigma_6 = (1,3)$ (i.e., $[1,2,3] \to [3,2,1]$): Then

$$W^{\sigma_6} = \left[\delta_5^1\delta_3^1\delta_2^1, \cdots, \delta_5^5\delta_3^3\delta_2^2\right]$$
$$= \delta_{30}[1,7,13,19,25,3,9,15,21,27,5,11,17,23,$$
$$29,2,8,14,20,26,4,10,16,22,28,6,12,18,24,30].$$

Proposition 4.9

(i)

$$\left[W^{\sigma}_{[n_1,\cdots,n_d]}\right]^T = \left[W^{\sigma}_{[n_1,\cdots,n_d]}\right]^{-1} = W^{\sigma^{-1}}_{[n_1,\cdots,n_d]}. \tag{11}$$

(ii) Let $\sigma, \mu \in \mathbf{S}_d$. Then

$$W^{\sigma}_{[n_1,n_2,\cdots,n_d]} W^{\mu}_{[n_1,n_2,\cdots,n_d]} = W^{\sigma\circ\mu}_{[n_1,n_2,\cdots,n_d]}. \tag{12}$$

Proposition 4.10

Assume $x_i \in \mathbb{F}^{n_i}$, $i \in <d>$, $\sigma \in \mathbf{S}_d$. Then

$$\ltimes_{i=1}^{d} x_{\sigma(i)} = W^{\sigma}_{[n_1,n_2,\cdots,n_d]} \ltimes_{i=1}^{d} x_i. \tag{13}$$

Corollary 3.11

Let $A \in \mathbb{F}^{n_1\times\cdots\times n_d}$ be a hypermatrix of order d. Then

$$V_{A^{\sigma}} = V_A \left[W^{\sigma}_{[n_1,\cdots,n_d]}\right]^T = V_A W^{\sigma^{-1}}_{[n_1,\cdots,n_d]}. \tag{14}$$

☞ Conversion of Matrix Expressions

Definition 4.12

(i) Let $A = [a_{i,j}] \in \mathbb{F}^{m\times n}$ be a matrix. Then

$$\mathbf{V}_r(A) := [a_{1,1}, a_{1,2}, \cdots, a_{1,n}, a_{2,1}, \cdots, a_{m,n}] \tag{15}$$

is called the row stacking form of A;

$$\mathbf{V}_c(A) := [a_{1,1}, a_{2,1}, \cdots, a_{m,1}, a_{1,2}, \cdots, a_{m,n}] \tag{16}$$

is called the column stacking form of A.

Definition 4.12(cont'd)

(ii) Let $x \in \mathbb{F}^n$ and $s|n$. Say, $n = st$. Then

$$\mathbf{V}_r^s(x) := \begin{bmatrix} x_1 & x_2 & \cdots & x_s \\ x_{s+1} & x_{s+2} & \cdots & x_{2s} \\ \vdots & & & \\ x_{(t-1)s+1} & x_{(t-1)s+2} & \cdots & x_{ts} \end{bmatrix} \tag{17}$$

$$\mathbf{V}_c^s(x) := \begin{bmatrix} x_1 & x_{s+1} & \cdots & x_{(t-1)s+1} \\ x_2 & x_{s+2} & \cdots & x_{(t-1)s+2} \\ \vdots & & & \\ x_s & x_{2s} & \cdots & x_{ts} \end{bmatrix} \tag{18}$$

Definition 4.12(cont'd)

(iii) Let $A \in \mathbb{F}^{m\times n}$ and $s|(mn)$. Then

$$\mathbf{V}_r^s(A) := \mathbf{V}_r^s\left(\mathbf{V}_r(A)\right) \tag{19}$$

is called the s-row stacking form.

$$\mathbf{V}_c^s(A) := \mathbf{V}_c^s\left(\mathbf{V}_c(A)\right) \tag{20}$$

is called the s-column stacking form.

Proposition 4.13

Let $A \in \mathbb{F}^{m\times n}$, $X \in \mathbb{F}^{n\times q}$, and $Y \in \mathbb{F}^{p\times m}$. Then

$$\mathbf{V}_r(AX) = A \ltimes \mathbf{V}_r(X), \tag{21}$$

$$\mathbf{V}_c(YA) = A^T \ltimes \mathbf{V}_c(Y). \tag{22}$$

Denote by

$$\delta_n^I := \mathbf{V}_r(I_n) = \mathbf{V}_c(I_n) = [(\delta_n^1)^T, (\delta_n^2)^T, \cdots, (\delta_n^n)^T]^T.$$

Proposition 4.14

Let $A \in \mathbb{F}^{m\times n}$. Then

$$\mathbf{V}_r(A) = A \ltimes \delta_n^I. \tag{23}$$

$$\mathbf{V}_c(A) = A^T \ltimes \delta_m^I. \tag{24}$$

Conversely,

$$A = \mathbf{V}_r^n(\mathbf{V}_r(A)) = \mathbf{V}_c^m(\mathbf{V}_c(A)). \tag{25}$$

Proposition 4.15

Given $A = [a_{i_1,\cdots,i_d}] \in \mathbb{F}^{n_1\times\cdots\times n_d}$, $\mathbf{i_r} = (i_1,\cdots,i_r) \subset \mathbf{d} =< d >$, and

$$\sigma_{\mathbf{i_r}} : \mathbf{d} \to (\mathbf{i_r}, \mathbf{d}\backslash\mathbf{i_r}), \quad n_{\mathbf{i_r}} = \prod_{s=1}^{r} n_{i_s}, \quad n_{\mathbf{d}\backslash\mathbf{i_r}} = \prod_{i_j\in\mathbf{d}\backslash\mathbf{i_r}} n_{i_j}.$$

Then

(i) (Vector Form to Matrix Form:)

$$M_A^{\mathbf{i_r}\times(\mathbf{d}\backslash\mathbf{i_r})} = \mathbf{V}_r^{n_{\mathbf{d}\backslash\mathbf{i_r}}}\left(V_A W_{[n_1,\cdots,n_d]}^{\sigma_{\mathbf{i_r}}^{-1}}\right). \tag{26}$$

(ii) (Matrix Form to Vector Form:)

$$V_A = \left(M_A^{\mathbf{i_r}\times(\mathbf{d}\backslash\mathbf{i_r})} \ltimes \delta^I_{n_{\mathbf{d}\backslash\mathbf{i_r}}}\right)^T W_{[n_1,\cdots,n_d]}^{\sigma_{\mathbf{i_r}}}. \tag{27}$$

Corollary 4.16

Let $\mathbf{i_r}$, $\sigma_{\mathbf{i_r}}$ be as in Proposition 4.15, and $\mathbf{j_s} = (j_1, \cdots, j_s)$ and $\sigma_{\mathbf{j_s}} : \mathbf{d} \to (\mathbf{j_s}, \mathbf{d}\backslash\mathbf{j_s})$. Then

$$\begin{aligned} M_A^{\mathbf{j_s}\times(\mathbf{d}\backslash\mathbf{j_s})} \quad = \quad & \mathbf{V}_r^{n_{\mathbf{d}\backslash\mathbf{j_s}}} \left[\left(M_A^{\mathbf{i_r}\times(\mathbf{d}\backslash\mathbf{i_r})} \ltimes \delta^I_{\mathbf{d}\backslash\mathbf{i_r}} \right)^T \right. \\ & \left. \times W^{\sigma_{\mathbf{i_r}}}_{[n_1,\cdots,n_d]} W^{\sigma_{\mathbf{j_s}}^{-1}}_{[n_1,\cdots,n_d]} \right]. \end{aligned} \tag{28}$$

V. STP of Hypermatrices

Definition 5.1

Let $A \in \mathbb{F}^{m_1 \times s \times n_1}$ and $B \in \mathbb{F}^{m_2 \times s \times n_2}$. The M-1 expressions of A and B are

$$M_A = [A_1, A_2, \cdots, A_s],$$
$$M_B = [B_1, B_2, \cdots, B_s],$$

where $A_i \in \mathcal{M}_{m_1 \times n_1}$, $B_i \in \mathcal{M}_{m_2 \times n_2}$, $i \in [1, s]$. The STP of A and B, denoted by $C = A \ltimes B$ is defined by

$$M_C := [A_1 \ltimes B_1, A_2 \ltimes B_2, \cdots, A_s \ltimes B_s]. \tag{29}$$

Denote by

$$\mathbb{F}^{\infty \times s \times \infty} := \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \mathbb{F}^{m \times s \times n}.$$

Then

$$\ltimes : \mathbb{F}^{\infty \times s \times \infty} \times \mathbb{F}^{\infty \times s \times \infty} \to \mathbb{F}^{\infty \times s \times \infty}.$$

☞ DK-STP of Matrices

Definition 5.2

Let $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{p\times q}$, $t = \operatorname{lcm}(n,p)$. The DK-STP of A and B, denoted by $A \times B \in \mathcal{M}_{m\times q}$, is defined as follows.

$$A \times B := \left(A \otimes \mathbf{1}_{t/n}^{T}\right)\left(B \otimes \mathbf{1}_{t/p}\right). \tag{30}$$

Remark 5.3

(i) When $n = p$,

$$A \times B = AB.$$

(ii) If $A, B \in \mathcal{M}_{m\times n}$, then $A \times B \in \mathcal{M}_{m\times n}$.

D. Cheng, From DK-STP to Non-square General Linear Algebra and General Linear Group, (preprint: http:arxiv.org/abs/2305.19794v2), 2023.

Remark 5.3(cont'd)

(iii) It is MM-, MV-, and VV- STP.

(iv) $(\mathcal{M}_{m\times n}, +, \ltimes)$ is a ring.

Proposition 5.4

Let $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{p\times q}$, $t = \operatorname{lcm}(n, p)$.

$$\begin{aligned} A \ltimes B &= A\left(I_n \otimes \mathbf{1}^T_{t/n}\right)\left(I_p \otimes \mathbf{1}_{t/p}\right)B \\ &:= A\Psi_{n\times p}B, \end{aligned} \tag{31}$$

where

$$\Psi_{n\times p} = \left(I_n \otimes \mathbf{1}^T_{t/n}\right)\left(I_p \otimes \mathbf{1}_{t/p}\right) \in \mathcal{M}_{n\times p}$$

is called the bridge matrix of dimension $n \times p$.

Definition 5.5

Assume $A \in \mathcal{M}_{m\times n}$. Consider $A : \mathbb{R}^\infty \to \mathbb{R}^\infty$ by $x \mapsto A \ltimes x$.
Then $\mathbb{R}^m \subset \mathbb{R}^\infty$ is an invariant subspace.
Denote by Π_A the restriction of $A|_{\mathbb{R}^m} = \Pi_A$. That is

$$A \ltimes x = \Pi_A x, \quad \forall x \in \mathbb{R}^m. \tag{32}$$

Proposition 5.6

$$\Pi_A = A \ltimes I_m = A\Psi_{n\times m}. \tag{33}$$

☞ Generalized Cayley-Hamilton Theorem

Definition 5.7

(i)

$$A^{<k>} := \underbrace{A \barwedge \cdots \barwedge A}_{k}. \tag{34}$$

(ii) Let $A \in \mathcal{M}_{m\times n}$ and $A|_{\mathbb{R}^m} = \Pi_A$. The characteristic polynomial of Π_A is called the characteristic polynomial of A.

Theorem 5.8

Let $A \in \mathcal{M}_{m\times n}$ and $A|_{\mathbb{R}^m} = \Pi_A$. Denote by $p(x) = x^m + p_{m-1}x^{m-1} + \cdots + p_0$ the characteristic polynomial of $\Pi(A)$. Then

$$A^{<m+1>} + p_{r-1}A^{<m>} + \cdots + p_0 A = 0. \tag{35}$$

Definition 5.9

Consider $\mathcal{M}_{m\times n}$, a Lie bracket over $\mathcal{M}_{m\times n}$, defined by using $\ltimes$, is

$$[A,B]_{\ltimes} := A\ltimes B - B\ltimes A, \quad A,B\in\mathcal{M}_{m\times n}. \tag{36}$$

Proposition 5.10

(i) $\mathcal{M}_{m\times n}$ with Lie bracket defined by (36) is a Lie algebra, denoted by $\mathrm{gl}(m\times n,\mathbb{F})$.

(ii) There exists the corresponding Lie group, denoted by $\mathrm{GL}(m\times n,\mathbb{F})$, which has $\mathrm{gl}(m\times n,\mathbb{F})$ as its Lie algebra.

☞ DK-STP of Hypermatrices

Definition 5.11

Let $A, B \in \mathbb{F}^{\infty \times s \times \infty}$ with M-1 expressions of A and B as

$$M_A = [A_1, A_2, \cdots, A_s],$$
$$M_B = [B_1, B_2, \cdots, B_s].$$

The DK-STP of A and B, denoted by $C = A \ltimes B$, is defined by

$$M_C := [A_1 \ltimes B_1, A_2 \ltimes B_2, \cdots, A_s \ltimes B_s]. \tag{37}$$

Definition 5.2

Let $A \in \mathbb{F}^{m\times s\times n}$ with M-1 expressions of A as $M_A = [A_1, A_2, \cdots, A_s]$.

(i)

$$A^{<k>} := \underbrace{A \ltimes \cdots \ltimes A}_{k}.$$

(ii) Let $p_i(x)$ be the characteristic function of A_i, $i \in [1, s]$. $p(x) := \prod_{i=1}^{s} p_i(x)$ is the characteristic function of A.

☞ Generalized Cayley-Hamilton Theorem for Hypermatrices

Theorem 5.12

Let $A \in \mathbb{F}^{m\times s\times n}$ with its characteristic function $p(x) = x^\mu + p_{\mu-1}x^{\mu-1} + \cdots + p_0$. Then

$$A^{<\mu+1>} + p_{\mu-1}A^{<\mu>} + \cdots + p_0 A = 0. \tag{38}$$

Definition 5.13

Consider $\mathbb{F}^{m\times s\times n}$, a Lie bracket over $\mathbb{F}^{m\times s\times n}$, defined by using $\bar{\times}$, is

$$[A,B]_{\bar{\times}} := A \bar{\times} B - B \bar{\times} A, \quad A,B \in \mathbb{F}^{m\times s\times n}. \tag{39}$$

Proposition 5.14

(i) $\mathbb{F}^{m\times s\times n}$ with Lie bracket defined by (39) is a Lie algebra, denoted by $\mathrm{gl}(m\times s\times n,\mathbb{F})$.

(ii) There exists the corresponding Lie group, denoted by $\mathrm{GL}(m\times s\times n,\mathbb{F})$, which has $\mathrm{gl}(m\times s\times n,\mathbb{F})$ as its Lie algebra.

References

- [1] D. Cheng, X. Zhang, Z. L. Ji，Semi-tensor product of htpermatrices with application to compound hypermatrices. *Proc. 42nd CCC*, 239-245, 2023.
- [2] D. Cheng, M. Meng, X. Zhang, Z. Ji, Contracted product of hypermatrices via STP of matrices, *Contr. Theory & Tech.*, Vol. 21, 265-280, 2023.
- [3] D. Cheng, M. Meng, X. Zhang, Z. Ji, Compound product of of hypermatrices, *J. Syst. Sci. Complex*, to appear.
- [4] D. Cheng, Z. Ji, On universal eigenvalues and eigenvectors of hypermatrices, *arXiv:2310.14266*, 2023.

VI. Conclusion

(i) Classical Matrix Theory is a dimension-restricted matrix theory.
STP Theory is a dimension-free matrix theory.

(ii) Classical Matrix Theory is used for matrices and vectors.
STP Theory is used for hypermatrices.

Hypermatrix is a wide field for STP to demonstrate her ability!

Thank you for your attention!

Question?