Outline:

1, Matrix-M (3种方式, mit).

2. ~~矩阵 vector~~

3. ~~分块矩阵 (数量, 加减乘除)~~

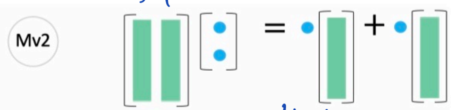
4, 左乘 (行) 、右乘 (列).

5. 特殊矩阵

Matrix-vector product. (2 ways)

output: m×1 的vector

The row vectors of A are multiplied by a vector $\boldsymbol{x}$ and become the three dot-product elements of $A\boldsymbol{x}$.



The product $A\boldsymbol{x}$ is a linear combination of the column vectors of A.

右乘向量
↓
列变换

$$A\boldsymbol{x} = \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}(x_1+2x_2)\\(3x_1+4x_2)\\(5x_1+6x_2)\end{bmatrix}$$

$$A\boldsymbol{x} = \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = x_1\begin{bmatrix}1\\3\\5\end{bmatrix} + x_2\begin{bmatrix}2\\4\\6\end{bmatrix}$$

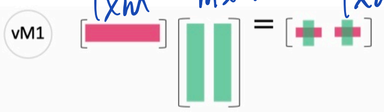
Figure 3: Matrix times Vector - (Mv1), (Mv2)

行列: dot product

列行: linear combination.

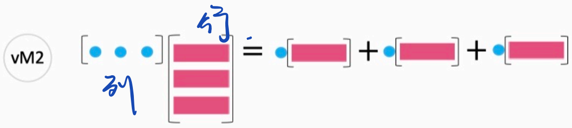
Vector-matrix product.

左乘向量
↓
行变换



$$\boldsymbol{y}A = [y_1 \quad y_2 \quad y_3]\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} = [(y_1+3y_2+5y_3) \quad (2y_1+4y_2+6y_3)]$$

A row vector $\boldsymbol{y}$ is multiplied by the two column vectors of A and become the two dot-product elements of $\boldsymbol{y}A$.



$$\boldsymbol{y}A = [y_1 \quad y_2 \quad y_3]\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} = y_1[1 \quad 2] + y_2[3 \quad 4] + y_3[5 \quad 6]$$

The product $\boldsymbol{y}A$ is a linear combination of the row vectors of A.

Figure 4: Vector times Matrix - (vM1), (vM2)

vector-vector product.
(2 ways)

outer product.

秩1矩.

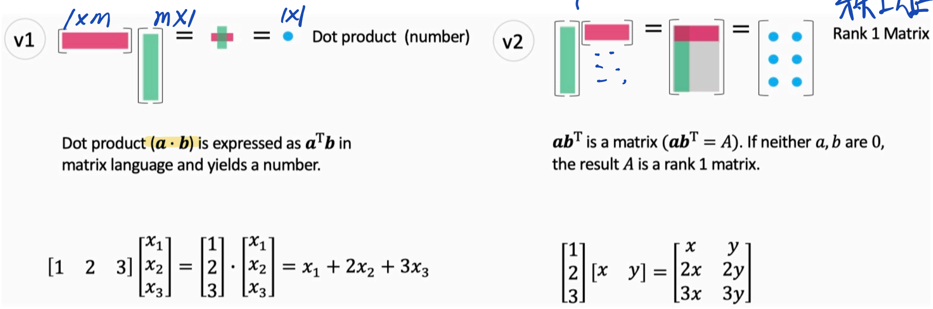


Figure 2: Vector times Vector - (v1), (v2)

(v1) is a elementary operation of two vectors, but (v2) multiplies the column to the row and produce a rank 1 matrix. Knowing this outer product (v2) is the key for the later sections.

例1

☆【例1.10】 已知 $\boldsymbol{\alpha}=(1,2,3)$，$\boldsymbol{\beta}=(1,\frac{1}{2},\frac{1}{3})$，设 $\boldsymbol{A}=\boldsymbol{\alpha}^T\boldsymbol{\beta}$，其中 $\boldsymbol{\alpha}^T$ 是 $\boldsymbol{\alpha}$ 的转置，则 $\boldsymbol{A}^n=$

$$A^n=\underbrace{(\alpha^T\beta)(\alpha^T\beta)\cdots(\alpha^T\beta)}_{n}$$

$$=\alpha^T\underbrace{(\beta\alpha^T)(\beta\alpha^T)\cdots(\beta\alpha^T)}_{n-1}\beta$$

$$=\begin{bmatrix}1\\2\\3\end{bmatrix}3^{n-1}(1,\tfrac{1}{2},\tfrac{1}{3})$$

$$[1,\tfrac{1}{2},\tfrac{1}{3}]$$

$$=3^{n-1}\begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}1 & \frac{1}{2} & \frac{1}{3}\\ 2 & 1 & \frac{2}{3}\\ 3 & \frac{3}{2} & 1\end{bmatrix}$$

α^T β

α^T β

✍14 Characterize the matrices AA^T and A^TA when A is of dimension $n \times 1$.

Solution: When $A \in \mathbb{R}^{n\times 1}$ then AA^T is a rank-one $n \times n$ matrix and A^TA is a scalar: the inner product of the column A with itself.

记住这2个样子, $\begin{cases} AA^T: \text{秩1阵, 3种应用} \\ A^TA: \text{scalar.} \end{cases}$

$$xy^T = \begin{pmatrix} x_1y_1 & & & \\ & x_2y_2 & & \\ & & \ddots & \\ & & & x_ny_n \end{pmatrix} \quad \text{秩1矩.}$$

$$x^Ty \quad (\quad) \begin{bmatrix} \\ \\ \end{bmatrix} \quad \text{scalar.}$$

x^Ty 恰好是 xy^T 主对角线上元素的和.

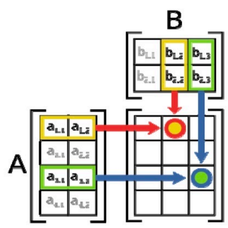
分块矩阵的运算.

1, 转置.

$$A=\begin{pmatrix} C & C^T \\ -C^T & C \end{pmatrix}$$

$$A^T=\begin{pmatrix} C^T & (-C^T)^T \\ (C^T)^T & C^T \end{pmatrix}=\begin{pmatrix} C^T & -C \\ C & C^T \end{pmatrix}$$

2, 乘法.



3, 加法和数乘 与矩阵的加法和数乘一样

2. Let $\mathcal{E}$ the $n \times n$ matrix that has ones on its anti-diagonal and zeros elsewhere ($a_{ij} = 1$ if $i + j = n + 1$, otherwise 0). (a) : What is the entry (i, j) of $\mathcal{E}A\mathcal{E}$? (b) : A matrix A is said to be *persymmetric* if $\mathcal{E}A\mathcal{E} = A^T$. Is a Toeplitz matrix persymmetric? [Hint: a persymmetric A is one that is symmetric about its antidiagonal.] (c) : Find an example of a small matrix that is persymmetric but not Toeplitz. (d) : Consider a matrix A of the following form

$$A = \begin{pmatrix} C & C^T \\ -C^T & C \end{pmatrix}.$$

Assuming that C is persymmetric, show that A is persymmetric. Assuming that C is Toeplitz is A Toeplitz? [prove if true, find a counter example if false.]

特殊矩阵.

酉矩阵.

Unitary.

Kronecker products of matrices

在下边作了一个补充（网上资料）.

第三章　Kronecker 积与 Kronecker 分解

Kronecker 积是张量计算中非常重要的一种运算规则，不同于常见的矩阵运算规则，给定任意两个矩阵，两者之间进行 Kronecker 积得到的是一个分块矩阵。Kronecker 分解是一种以 Kronecker 积为基础的分解形式，又被称为 Kronecker 积分解、Kronecker 积逼近 (Kronecker product approximation)、最近 Kronecker 积 (nearest Kronecker product) 等，它是矩阵计算与张量计算中十分重要的逼近问题。本章首先介绍 Kronecker 积的定义与性质，然后引出 Kronecker 分解的一般形式、优化问题、求解过程等，最后给出以 Kronecker 分解为基础的模型参数压缩问题。

3.1　Kronecker 积定义

3.1.1　基本定义

Kronecker 积是以德国数学家 Leopold Kronecker 的名字命名的运算规则，已广泛应用于各类矩阵计算以及张量计算算法中。从定义出发，给定任意矩阵 $\boldsymbol{X} \in \mathbb{R}^{m\times n}$ 与 $\boldsymbol{Y} \in \mathbb{R}^{p\times q}$，则两者之间的 Kronecker 积为

定义

$$\overset{m\times n}{\boldsymbol{X}} \otimes \overset{p\times q}{\boldsymbol{Y}} = \begin{bmatrix} x_{11}\boldsymbol{Y} & x_{12}\boldsymbol{Y} & \cdots & x_{1n}\boldsymbol{Y} \\ x_{21}\boldsymbol{Y} & x_{22}\boldsymbol{Y} & \cdots & x_{2n}\boldsymbol{Y} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1}\boldsymbol{Y} & x_{m2}\boldsymbol{Y} & \cdots & x_{mn}\boldsymbol{Y} \end{bmatrix} \in \mathbb{R}^{(mp)\times(nq)} \tag{3.1}$$

其中，符号 $\otimes$ 表示 Kronecker 积。这里的 Kronecker 积得到的矩阵大小为 $(mp)\times(nq)$，在写法上符合线性代数中对分块矩阵 (block matrix) 的定义，其中，分块矩阵的子矩阵是由矩阵 $\boldsymbol{X}$ 的每个元素与矩阵 $\boldsymbol{Y}$ 相乘得到。

kronecker 积存在先后次序 → 不满足交换律.

矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 之间的 Kronecker 积存在前后顺序，根据 Kronecker 积的定义，可得到矩阵 $\boldsymbol{Y}$ 与 $\boldsymbol{X}$ 之间的 Kronecker 积为

$$\boldsymbol{Y} \otimes \boldsymbol{X} = \begin{bmatrix} y_{11}\boldsymbol{X} & y_{12}\boldsymbol{X} & \cdots & y_{1q}\boldsymbol{X} \\ y_{21}\boldsymbol{X} & y_{22}\boldsymbol{X} & \cdots & y_{2q}\boldsymbol{X} \\ \vdots & \vdots & \ddots & \vdots \\ y_{p1}\boldsymbol{X} & y_{p2}\boldsymbol{X} & \cdots & y_{pq}\boldsymbol{X} \end{bmatrix} \in \mathbb{R}^{(mp)\times(nq)} \tag{3.2}$$

尽管矩阵 $\boldsymbol{X}\otimes\boldsymbol{Y}$ 与矩阵 $\boldsymbol{Y}\otimes\boldsymbol{X}$ 大小一致，但两者并不相等，因此，Kronecker 积不存在交换律。

矩阵的kronecker积.

例 10. *给定矩阵* $\boldsymbol{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *与* $\boldsymbol{Y} = \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}$，*试写出两者之间的 Kronecker 积* $\boldsymbol{X}\otimes\boldsymbol{Y}$ *与* $\boldsymbol{Y}\otimes\boldsymbol{X}$。

解. 根据 *Kronecker* 积定义，有

$$
\boldsymbol{X}\otimes\boldsymbol{Y}=\begin{bmatrix}1\times\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix} & 2\times\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}\\ 3\times\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix} & 4\times\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}\end{bmatrix}=\begin{bmatrix}5&6&7&10&12&14\\8&9&10&16&18&20\\15&18&21&20&24&28\\24&27&30&32&36&40\end{bmatrix}\tag{3.3}
$$

$$
\boldsymbol{Y}\otimes\boldsymbol{X}=\begin{bmatrix}5\times\begin{bmatrix}1&2\\3&4\end{bmatrix} & 6\times\begin{bmatrix}1&2\\3&4\end{bmatrix} & 7\times\begin{bmatrix}1&2\\3&4\end{bmatrix}\\ 8\times\begin{bmatrix}1&2\\3&4\end{bmatrix} & 9\times\begin{bmatrix}1&2\\3&4\end{bmatrix} & 10\times\begin{bmatrix}1&2\\3&4\end{bmatrix}\end{bmatrix}=\begin{bmatrix}5&10&6&12&7&14\\15&20&18&24&21&28\\8&16&9&18&10&20\\24&32&27&36&30&40\end{bmatrix}\tag{3.4}
$$

例 11. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}$，试问等式 $(\boldsymbol{X}\otimes\boldsymbol{Y})^\top=\boldsymbol{X}^\top\otimes\boldsymbol{Y}^\top$ 是否成立。

性质.

解. 根据 *Kronecker* 积定义，有

$$
\boldsymbol{X}^\top\otimes\boldsymbol{Y}^\top=\begin{bmatrix}1\times\begin{bmatrix}5&8\\6&9\\7&10\end{bmatrix} & 3\times\begin{bmatrix}5&8\\6&9\\7&10\end{bmatrix}\\ 2\times\begin{bmatrix}5&8\\6&9\\7&10\end{bmatrix} & 4\times\begin{bmatrix}5&8\\6&9\\7&10\end{bmatrix}\end{bmatrix}=\begin{bmatrix}5&8&15&24\\6&9&18&27\\7&10&21&30\\10&16&20&32\\12&18&24&36\\14&20&28&40\end{bmatrix}\tag{3.5}
$$

在这里，等式 $(\boldsymbol{X}\otimes\boldsymbol{Y})^\top=\boldsymbol{X}^\top\otimes\boldsymbol{Y}^\top$ 是成立的。

向量的kronecker积

例 12. 给定向量 $\boldsymbol{x}=(1,2)^\top$ 与 $\boldsymbol{y}=(3,4)^\top$，试写出 $\boldsymbol{x}\otimes\boldsymbol{y}$ 与 $\boldsymbol{x}\otimes\boldsymbol{y}^\top$。

解. 根据 *Kronecker* 积定义，有

$$
\boldsymbol{x}\otimes\boldsymbol{y}=\begin{bmatrix}1\times\begin{bmatrix}3\\4\end{bmatrix}\\ 2\times\begin{bmatrix}3\\4\end{bmatrix}\end{bmatrix}=\begin{bmatrix}3\\4\\6\\8\end{bmatrix}\tag{3.6}
$$

$$
\boldsymbol{x}\otimes\boldsymbol{y}^\top=\begin{bmatrix}1\times\begin{bmatrix}3&4\end{bmatrix}\\ 2\times\begin{bmatrix}3&4\end{bmatrix}\end{bmatrix}=\begin{bmatrix}3&4\\6&8\end{bmatrix}\tag{3.7}
$$

在这里，$\boldsymbol{x}\otimes\boldsymbol{y}^\top=\boldsymbol{x}\boldsymbol{y}^\top$，即向量外积。

例 13 (向量自回归). 对于多元时间序列，向量自回归可写作如下形式（参见例6）:

$$
\boldsymbol{X}\boldsymbol{\Psi}_0^\top=\sum_{k=1}^{d}\boldsymbol{A}_k\boldsymbol{X}\boldsymbol{\Psi}_k^\top+\boldsymbol{E}\tag{3.8}
$$

若令

$$
\begin{aligned}\boldsymbol{A}&=\begin{bmatrix}\boldsymbol{A}_1&\boldsymbol{A}_2&\cdots&\boldsymbol{A}_d\end{bmatrix}\in\mathbb{R}^{N\times(dN)}\\ \boldsymbol{\Psi}&=\begin{bmatrix}\boldsymbol{\Psi}_1&\boldsymbol{\Psi}_2&\cdots&\boldsymbol{\Psi}_d\end{bmatrix}\in\mathbb{R}^{(T-d)\times(dT)}\end{aligned}\tag{3.9}
$$

则向量自回归可进一步写作如下形式:

$$
\boldsymbol{X}\boldsymbol{\Psi}_0^\top=\boldsymbol{A}(\boldsymbol{I}_d\otimes\boldsymbol{X})\boldsymbol{\Psi}^\top+\boldsymbol{E}\tag{3.10}
$$

3.1.2 Khatri-Rao 积

以 Kronecker 积为基础，可定义另一种十分重要的运算规则，即 Khatri-Rao 积。给定任意矩阵

$$\boldsymbol{X}=\begin{bmatrix} | & | & & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_d \\ | & | & & | \end{bmatrix}\in\mathbb{R}^{m\times d}\quad \boldsymbol{Y}=\begin{bmatrix} | & | & & | \\ \boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_d \\ | & | & & | \end{bmatrix}\in\mathbb{R}^{n\times d} \tag{3.11}$$

若两个矩阵列数相同，则两者之间的 Khatri-Rao 积为

$$\boldsymbol{X}\odot\boldsymbol{Y}=\begin{bmatrix} | & | & & | \\ \boldsymbol{x}_1\otimes\boldsymbol{y}_1 & \boldsymbol{x}_2\otimes\boldsymbol{y}_2 & \cdots & \boldsymbol{x}_d\otimes\boldsymbol{y}_d \\ | & | & & | \end{bmatrix}\in\mathbb{R}^{(mn)\times d} \tag{3.12}$$

其中，列向量是由 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的列向量进行 Kronecker 积运算得到的。

例 14. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}5&6\\7&8\\9&10\end{bmatrix}$，试写出 $\boldsymbol{X}\odot\boldsymbol{Y}$。

解. 根据 *Khatri-Rao* 积定义，有

$$\boldsymbol{X}\odot\boldsymbol{Y}=\begin{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}\otimes\begin{bmatrix}5\\7\\9\end{bmatrix} & \begin{bmatrix}2\\4\end{bmatrix}\otimes\begin{bmatrix}6\\8\\10\end{bmatrix}\end{bmatrix}=\begin{bmatrix}5&12\\7&16\\9&20\\15&24\\21&32\\27&40\end{bmatrix} \tag{3.13}$$

3.2 Kronecker 积基本性质

3.2.1 结合律与分配律

在小学数学中，我们学习了加减乘除的运算规则。以乘法为例，不妨重温一下烙印在我们脑海中的基本概念：

- 乘法结合律：$x\times y\times z=x\times(y\times z)$
- 乘法分配律：$x\times z+y\times z=(x+y)\times z$

由于 Kronecker 积本质上也是元素间相乘，所以同样存在结合律与分配律。对于任意矩阵 $\boldsymbol{X}$、$\boldsymbol{Y}$ 与 $\boldsymbol{Z}$，结合律可归纳为

$$\boldsymbol{X}\otimes\boldsymbol{Y}\otimes\boldsymbol{Z}=\boldsymbol{X}\otimes(\boldsymbol{Y}\otimes\boldsymbol{Z}) \tag{3.14}$$

分配律可归纳为

$$\boldsymbol{X}\otimes\boldsymbol{Z}+\boldsymbol{Y}\otimes\boldsymbol{Z}=(\boldsymbol{X}+\boldsymbol{Y})\otimes\boldsymbol{Z} \tag{3.15}$$

例 15. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$、$\boldsymbol{Y}=\begin{bmatrix}5&6\\7&8\end{bmatrix}$ 与 $\boldsymbol{Z}=\begin{bmatrix}1&1\\1&1\end{bmatrix}$，试写出 $\boldsymbol{X}\otimes\boldsymbol{Y}\otimes\boldsymbol{Z}$ 与 $\boldsymbol{X}\otimes(\boldsymbol{Y}\otimes\boldsymbol{Z})$。

解. 根据 *Kronecker* 积定义，有

$$\boldsymbol{X} \otimes \boldsymbol{Y} = \begin{bmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{bmatrix} \tag{3.16}$$

$$\boldsymbol{Y} \otimes \boldsymbol{Z} = \begin{bmatrix} 5 & 5 & 6 & 6 \\ 5 & 5 & 6 & 6 \\ 7 & 7 & 8 & 8 \\ 7 & 7 & 8 & 8 \end{bmatrix} \tag{3.17}$$

从而，可得到

$$\boldsymbol{X} \otimes \boldsymbol{Y} \otimes \boldsymbol{Z} = \begin{bmatrix} 5 & 5 & 6 & 6 & 10 & 10 & 12 & 12 \\ 5 & 5 & 6 & 6 & 10 & 10 & 12 & 12 \\ 7 & 7 & 8 & 8 & 14 & 14 & 16 & 16 \\ 7 & 7 & 8 & 8 & 14 & 14 & 16 & 16 \\ 15 & 15 & 18 & 18 & 20 & 20 & 24 & 24 \\ 15 & 15 & 18 & 18 & 20 & 20 & 24 & 24 \\ 21 & 21 & 24 & 24 & 28 & 28 & 32 & 32 \\ 21 & 21 & 24 & 24 & 28 & 28 & 32 & 32 \end{bmatrix} = \boldsymbol{X} \otimes (\boldsymbol{Y} \otimes \boldsymbol{Z}) \tag{3.18}$$

例 16. 给定 $\boldsymbol{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$、$\boldsymbol{Y} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ 与 $\boldsymbol{Z} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$，试写出 $\boldsymbol{X} \otimes \boldsymbol{Z} + \boldsymbol{Y} \otimes \boldsymbol{Z}$ 与 $(\boldsymbol{X} + \boldsymbol{Y}) \otimes \boldsymbol{Z}$。

解. 根据 *Kronecker* 积定义，有

$$\boldsymbol{X} \otimes \boldsymbol{Z} + \boldsymbol{Y} \otimes \boldsymbol{Z} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \\ 3 & 3 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 5 & 6 & 6 \\ 5 & 5 & 6 & 6 \\ 7 & 7 & 8 & 8 \\ 7 & 7 & 8 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 & 8 \\ 6 & 6 & 8 & 8 \\ 10 & 10 & 12 & 12 \\ 10 & 10 & 12 & 12 \end{bmatrix} \tag{3.19}$$

$$(\boldsymbol{X} + \boldsymbol{Y}) \otimes \boldsymbol{Z} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 8 & 8 \\ 6 & 6 & 8 & 8 \\ 10 & 10 & 12 & 12 \\ 10 & 10 & 12 & 12 \end{bmatrix} \tag{3.20}$$

3.2.2 矩阵相乘

对于任意矩阵 $\boldsymbol{X} \in \mathbb{R}^{m \times n}$、$\boldsymbol{Y} \in \mathbb{R}^{s \times t}$、$\boldsymbol{U} \in \mathbb{R}^{n \times p}$ 与 $\boldsymbol{V} \in \mathbb{R}^{t \times q}$，则矩阵 $\boldsymbol{X} \otimes \boldsymbol{Y} \in \mathbb{R}^{(ms) \times (nt)}$ 的列数 nt 与矩阵 $\boldsymbol{U} \otimes \boldsymbol{V} \in \mathbb{R}^{(nt) \times (pq)}$ 的行数 nt 一致，可进行矩阵相乘，两者相乘得到的矩

阵满足:

$$
\begin{aligned}
(\boldsymbol{X}\otimes\boldsymbol{Y})(\boldsymbol{U}\otimes\boldsymbol{V}) &= \begin{bmatrix} x_{11}\boldsymbol{Y} & \cdots & x_{1n}\boldsymbol{Y} \\ \vdots & \ddots & \vdots \\ x_{m1}\boldsymbol{Y} & \cdots & x_{mn}\boldsymbol{Y} \end{bmatrix} \begin{bmatrix} u_{11}\boldsymbol{V} & \cdots & u_{1p}\boldsymbol{V} \\ \vdots & \ddots & \vdots \\ u_{n1}\boldsymbol{V} & \cdots & u_{np}\boldsymbol{V} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^{n} x_{1k}u_{k1}\boldsymbol{Y}\boldsymbol{V} & \cdots & \sum_{k=1}^{n} x_{1k}u_{kp}\boldsymbol{Y}\boldsymbol{V} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} x_{mk}u_{k1}\boldsymbol{Y}\boldsymbol{V} & \cdots & \sum_{k=1}^{n} x_{mk}u_{kp}\boldsymbol{Y}\boldsymbol{V} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^{n} x_{1k}u_{k1} & \cdots & \sum_{k=1}^{n} x_{1k}u_{kp} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} x_{mk}u_{k1} & \cdots & \sum_{k=1}^{n} x_{mk}u_{kp} \end{bmatrix} \otimes (\boldsymbol{Y}\boldsymbol{V}) \\
&= (\boldsymbol{X}\boldsymbol{U})\otimes(\boldsymbol{Y}\boldsymbol{V}) \in \mathbb{R}^{(ms)\times(pq)}
\end{aligned} \tag{3.21}
$$

例 17 (矩阵的奇异值分解). 给定任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{p\times q}$, 若奇异值分解分别为

$$
\boldsymbol{X}=\boldsymbol{W}\boldsymbol{S}\boldsymbol{Q}^\top \quad \boldsymbol{Y}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top \tag{3.22}
$$

试证明矩阵 $\boldsymbol{X}\otimes\boldsymbol{Y}$ 的奇异值分解可由矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的奇异值分解计算得到, 即

$$
\boldsymbol{X}\otimes\boldsymbol{Y}=(\boldsymbol{W}\otimes\boldsymbol{U})(\boldsymbol{S}\otimes\boldsymbol{D})(\boldsymbol{Q}\otimes\boldsymbol{V})^\top \tag{3.23}
$$

解. 根据 *Kronecker* 积性质, 有

$$
\begin{aligned}
\boldsymbol{X}\otimes\boldsymbol{Y} &= (\boldsymbol{W}\boldsymbol{S}\boldsymbol{Q}^\top)\otimes(\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top) \\
&= (\boldsymbol{W}\otimes\boldsymbol{U})((\boldsymbol{S}\boldsymbol{Q}^\top)\otimes(\boldsymbol{D}\boldsymbol{V}^\top)) \\
&= (\boldsymbol{W}\otimes\boldsymbol{U})(\boldsymbol{S}\otimes\boldsymbol{D})(\boldsymbol{Q}^\top\otimes\boldsymbol{V}^\top) \\
&= (\boldsymbol{W}\otimes\boldsymbol{U})(\boldsymbol{S}\otimes\boldsymbol{D})(\boldsymbol{Q}\otimes\boldsymbol{V})^\top
\end{aligned} \tag{3.24}
$$

3.2.3 求逆矩阵

对于任意可逆矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times m}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{n\times n}$, 由于

$$
(\boldsymbol{X}\otimes\boldsymbol{Y})\left(\boldsymbol{X}^{-1}\otimes\boldsymbol{Y}^{-1}\right)=\left(\boldsymbol{X}\boldsymbol{X}^{-1}\right)\otimes\left(\boldsymbol{Y}\boldsymbol{Y}^{-1}\right)=\boldsymbol{I}_m\otimes\boldsymbol{I}_n=\boldsymbol{I}_{mn} \tag{3.25}
$$

故有

$$
(\boldsymbol{X}\otimes\boldsymbol{Y})^{-1}=\boldsymbol{X}^{-1}\otimes\boldsymbol{Y}^{-1} \tag{3.26}
$$

恒成立。这意味着: 若计算 $\boldsymbol{X}\otimes\boldsymbol{Y}$ 的逆矩阵, 可先对 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 分别求逆矩阵, 再对得到的逆矩阵进行 Kronecker 积运算。

例 18. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix}$, 试写出 $(\boldsymbol{X}\otimes\boldsymbol{Y})^{-1}$ 与 $\boldsymbol{X}^{-1}\otimes\boldsymbol{Y}^{-1}$。

解. 根据 *Kronecker* 积定义, 有

$$
\boldsymbol{X}\otimes\boldsymbol{Y}=\begin{bmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{bmatrix} \tag{3.27}
$$

对该矩阵求逆矩阵，得到

$$(\boldsymbol{X}\otimes\boldsymbol{Y})^{-1}=\begin{bmatrix}8&-6&-4&3\\-7&5&3.5&-2.5\\-6&4.5&2&-1.5\\5.25&-3.75&-1.75&1.25\end{bmatrix}\tag{3.28}$$

对矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 分别求逆矩阵：

$$\boldsymbol{X}^{-1}=\begin{bmatrix}-2&1\\1.5&-0.5\end{bmatrix}\quad\boldsymbol{Y}^{-1}=\begin{bmatrix}-4&3\\3.5&-2.5\end{bmatrix}\tag{3.29}$$

再对得到的逆矩阵进行 *Kronecker* 积运算，有

$$\boldsymbol{X}^{-1}\otimes\boldsymbol{Y}^{-1}=\begin{bmatrix}8&-6&-4&3\\-7&5&3.5&-2.5\\-6&4.5&2&-1.5\\5.25&-3.75&-1.75&1.25\end{bmatrix}\tag{3.30}$$

对于任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times m}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{p\times q}$，由上述 Kronecker 积性质同样可得到如下性质：

$$(\boldsymbol{X}\otimes\boldsymbol{Y})^{\dagger}=\boldsymbol{X}^{\dagger}\otimes\boldsymbol{Y}^{\dagger}\tag{3.31}$$

其中，$\cdot^{\dagger}$ 表示伪逆 (Moore-Penrose pseudoinverse)。

3.2.4 向量化

对于任意矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$、$\boldsymbol{X}\in\mathbb{R}^{n\times p}$ 与 $\boldsymbol{B}\in\mathbb{R}^{p\times q}$，三者相乘满足：

$$\operatorname{vec}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B})=(\boldsymbol{B}^{\top}\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{X})\tag{3.32}$$

由此，也可得到

$$\begin{cases}\operatorname{vec}(\boldsymbol{A}\boldsymbol{X})=(\boldsymbol{I}_p\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{X})\\\operatorname{vec}(\boldsymbol{X}\boldsymbol{B})=(\boldsymbol{B}^{\top}\otimes\boldsymbol{I}_n)\operatorname{vec}(\boldsymbol{X})\end{cases}\tag{3.33}$$

例 19. 试证明公式(3.32)。

解.

$$\begin{aligned}\operatorname{vec}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B})=&\boldsymbol{A}\boldsymbol{x}_1b_{11}+\boldsymbol{A}\boldsymbol{x}_2b_{21}+\cdots+\boldsymbol{A}\boldsymbol{x}_pb_{p1}\\&+\boldsymbol{A}\boldsymbol{x}_1b_{12}+\boldsymbol{A}\boldsymbol{x}_2b_{22}+\cdots+\boldsymbol{A}\boldsymbol{x}_pb_{p2}\\&+\cdots+\boldsymbol{A}\boldsymbol{x}_1b_{1q}+\boldsymbol{A}\boldsymbol{x}_2b_{2q}+\cdots+\boldsymbol{A}\boldsymbol{x}_pb_{pq}\\=&\begin{bmatrix}\boldsymbol{A}b_{11}&\boldsymbol{A}b_{21}&\cdots&\boldsymbol{A}b_{p1}\\\boldsymbol{A}b_{12}&\boldsymbol{A}b_{22}&\cdots&\boldsymbol{A}b_{p2}\\\vdots&\vdots&\ddots&\vdots\\\boldsymbol{A}b_{1q}&\boldsymbol{A}b_{2q}&\cdots&\boldsymbol{A}b_{pq}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1\\\boldsymbol{x}_2\\\vdots\\\boldsymbol{x}_p\end{bmatrix}\\=&(\boldsymbol{B}^{\top}\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{X})\end{aligned}\tag{3.34}$$

其中，$\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_p\in\mathbb{R}^n$ 表示矩阵 $\boldsymbol{X}$ 的列向量。

例 20. 对于任意向量 $\boldsymbol{x}\in\mathbb{R}^n$、$\boldsymbol{z}\in\mathbb{R}^p$ 与矩阵 $\boldsymbol{Y}\in\mathbb{R}^{p\times q}$，试证明

$$\left(\boldsymbol{x}^{\top}\otimes\boldsymbol{Y}\right)^{\top}\boldsymbol{z}=\left((\boldsymbol{x}\boldsymbol{z}^{\top})\otimes\boldsymbol{I}_q\right)\operatorname{vec}\left(\boldsymbol{Y}^{\top}\right)\tag{3.35}$$

恒成立。

解. 根据 *Kronecker* 积性质，有

$$
\begin{aligned}
\left(\boldsymbol{x}^\top \otimes \boldsymbol{Y}\right)^\top \boldsymbol{z} &= \left(\boldsymbol{x} \otimes \boldsymbol{Y}^\top\right) \boldsymbol{z} \\
&= \operatorname{vec}\left(\boldsymbol{Y}^\top \boldsymbol{z}\boldsymbol{x}^\top\right) \\
&= \operatorname{vec}\left(\boldsymbol{I}_q \boldsymbol{Y}^\top \left(\boldsymbol{z}\boldsymbol{x}^\top\right)\right) \\
&= \left(\left(\boldsymbol{x}\boldsymbol{z}^\top\right) \otimes \boldsymbol{I}_q\right) \operatorname{vec}\left(\boldsymbol{Y}^\top\right)
\end{aligned} \tag{3.36}
$$

例 21. 对于任意矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$、$\boldsymbol{x} \in \mathbb{R}^n$ 与 $\boldsymbol{B} \in \mathbb{R}^{n\times n}$，试证明三者相乘满足：

$$
\operatorname{vec}(\boldsymbol{A} \operatorname{diag}(\boldsymbol{x})\boldsymbol{B}) = (\boldsymbol{B}^\top \odot \boldsymbol{A})\boldsymbol{x} \tag{3.37}
$$

解. 根据 *Kronecker* 积与 *Khatri-Rao* 积性质，有

$$
\begin{aligned}
\operatorname{vec}(\boldsymbol{A} \operatorname{diag}(\boldsymbol{x})\boldsymbol{B}) =& (\boldsymbol{B}^\top \otimes \boldsymbol{A}) \operatorname{vec}(\operatorname{diag}(\boldsymbol{x})) \\
=& (\boldsymbol{B}^\top \odot \boldsymbol{A})\boldsymbol{x}
\end{aligned} \tag{3.38}
$$

例 22. *Sylvester* 方程是一种著名的矩阵方程，由英国数学家 *James Joseph Sylvester* 于 *1884* 年提出。时至今日，*Sylvester* 方程已在控制理论中具有极为广泛的应用。具体而言，已知矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times m}$、$\boldsymbol{B} \in \mathbb{R}^{n\times n}$ 与 $\boldsymbol{C} \in \mathbb{R}^{m\times n}$，则 *Sylvester* 方程的一般形式为

$$
\boldsymbol{A}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{B} = \boldsymbol{C} \tag{3.39}
$$

其中，$\boldsymbol{X} \in \mathbb{R}^{m\times n}$ 为待定参数。试根据 *Kronecker* 积性质写出 *Sylvester* 方程的解析解。

解. 首先将 *Sylvester* 方程写成

$$
\boldsymbol{A}\boldsymbol{X}\boldsymbol{I}_n + \boldsymbol{I}_m\boldsymbol{X}\boldsymbol{B} = \boldsymbol{C} \tag{3.40}
$$

根据 *Kronecker* 积性质，*Sylvester* 方程可写成如下形式：

$$
\left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^\top \otimes \boldsymbol{I}_m\right) \operatorname{vec}(\boldsymbol{X}) = \operatorname{vec}(\boldsymbol{C}) \tag{3.41}
$$

因此，*Sylvester* 方程的解析解[1]为

$$
\operatorname{vec}(\boldsymbol{X}) = \left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^\top \otimes \boldsymbol{I}_m\right)^{-1} \operatorname{vec}(\boldsymbol{C}) \tag{3.42}
$$

尽管该解析解形式简洁，但复杂度却很高。在实际问题中，往往需要借助更为高效的数值计算方法（如 *Bartels-Stewart* 算法）对 *Sylvester* 方程进行求解。

3.3 Kronecker 积特殊性质

3.3.1 矩阵的迹

在线性代数中，矩阵的迹 (trace) 表示方阵对角线元素之和，数学符号为 $\operatorname{tr}(\cdot)$。对于任意矩阵 $\boldsymbol{X} \in \mathbb{R}^{m\times m}$ 与 $\boldsymbol{Y} \in \mathbb{R}^{n\times n}$，矩阵 $\boldsymbol{X} \otimes \boldsymbol{Y}$ 的迹等于矩阵 $\boldsymbol{X}$ 的迹乘以矩阵 $\boldsymbol{Y}$ 的迹，即

$$
\operatorname{tr}(\boldsymbol{X} \otimes \boldsymbol{Y}) = \operatorname{tr}(\boldsymbol{X}) \cdot \operatorname{tr}(\boldsymbol{Y}) \tag{3.43}
$$

恒成立。

例 23. 给定矩阵 $\boldsymbol{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ 与 $\boldsymbol{Y} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$，试写出 $\operatorname{tr}(\boldsymbol{X})$、$\operatorname{tr}(\boldsymbol{Y})$ 与 $\operatorname{tr}(\boldsymbol{X} \otimes \boldsymbol{Y})$。

[1] 有时候，可定义 Kronecker 和 (Kronecker sum，数学符号通常为 $\oplus$) 令 $\boldsymbol{A} \oplus \boldsymbol{B}^\top = \boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^\top \otimes \boldsymbol{I}_m$，将该解析解简记为 $\operatorname{vec}(\boldsymbol{X}) = \left(\boldsymbol{A} \oplus \boldsymbol{B}^\top\right)^{-1} \operatorname{vec}(\boldsymbol{C})$。

解. 根据定义，矩阵 $\boldsymbol{X}$ 的迹与矩阵 $\boldsymbol{Y}$ 的迹分别为

$$\operatorname{tr}(\boldsymbol{X})=1+4=5\quad \operatorname{tr}(\boldsymbol{Y})=5+8=13 \tag{3.44}$$

由于

$$\boldsymbol{X}\otimes\boldsymbol{Y}=\begin{bmatrix}5&6&10&12\\7&8&14&16\\15&18&20&24\\21&24&28&32\end{bmatrix} \tag{3.45}$$

故 $\operatorname{tr}(\boldsymbol{X}\otimes\boldsymbol{Y})=5+8+20+32=65$。

在矩阵计算中，矩阵的迹有两条重要性质，给定任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{n\times m}$，满足

$$\operatorname{tr}(\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BA}) \tag{3.46}$$

及

$$\operatorname{tr}(\boldsymbol{AB})=\operatorname{vec}(\boldsymbol{A}^\top)^\top\operatorname{vec}(\boldsymbol{B}) \tag{3.47}$$

例 24. 给定矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$、$\boldsymbol{B}\in\mathbb{R}^{n\times p}$、$\boldsymbol{C}\in\mathbb{R}^{p\times q}$ 与 $\boldsymbol{D}\in\mathbb{R}^{q\times m}$，试证明

$$\operatorname{tr}(\boldsymbol{ABCD})=\operatorname{vec}(\boldsymbol{B})^\top(\boldsymbol{C}\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{D}^\top) \tag{3.48}$$

解. 根据矩阵的迹与 *Kronecker* 积性质，有

$$\begin{aligned}\operatorname{tr}(\boldsymbol{ABCD})&=\operatorname{tr}(\boldsymbol{D}(\boldsymbol{ABC}))\\&=\operatorname{vec}(\boldsymbol{D}^\top)^\top\operatorname{vec}(\boldsymbol{ABC})\\&=\operatorname{vec}(\boldsymbol{D}^\top)^\top(\boldsymbol{C}^\top\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{B})\\&=\operatorname{vec}(\boldsymbol{B})^\top(\boldsymbol{C}\otimes\boldsymbol{A})\operatorname{vec}(\boldsymbol{D}^\top)\end{aligned} \tag{3.49}$$

3.3.2 矩阵的 Frobenius 范数

从定义出发，矩阵的 Frobenius 范数表示矩阵元素的平方和开根号，一般用 $\|\cdot\|_F$ 表示。对于任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$，其 Frobenius 范数为

$$\|\boldsymbol{X}\|_F=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}x_{ij}^2} \tag{3.50}$$

据此定义，给定任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{p\times q}$，有

$$\|\boldsymbol{X}\otimes\boldsymbol{Y}\|_F=\|\boldsymbol{X}\|_F\cdot\|\boldsymbol{Y}\|_F \tag{3.51}$$

恒成立。

例 25. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}5&6\\7&8\end{bmatrix}$，试写出 $\|\boldsymbol{X}\|_F$、$\|\boldsymbol{Y}\|_F$ 与 $\|\boldsymbol{X}\otimes\boldsymbol{Y}\|_F$。

解. 根据定义，矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的 *Frobenius* 范数分别为

$$\|\boldsymbol{X}\|_F=\sqrt{1^2+2^2+3^2+4^2}=\sqrt{30}\quad\|\boldsymbol{Y}\|_F=\sqrt{5^2+6^2+7^2+8^2}=\sqrt{174} \tag{3.52}$$

由于

$$\boldsymbol{X}\otimes\boldsymbol{Y}=\begin{bmatrix}5&6&10&12\\7&8&14&16\\15&18&20&24\\21&24&28&32\end{bmatrix} \tag{3.53}$$

故 $\|\boldsymbol{X}\otimes\boldsymbol{Y}\|_F=\sqrt{5220}$。

Frobenius 范数这一概念不适用于向量，对于任意向量 $\boldsymbol{x}\in\mathbb{R}^m$，其元素的平方和开根号是 ℓ_2 范数，即

$$\|\boldsymbol{x}\|_2=\sqrt{\sum_{i=1}^m x_i^2} \tag{3.54}$$

例 26. 给定向量 $\boldsymbol{x}=(1,2)^\top$ 与 $\boldsymbol{y}=(3,4)^\top$，试写出 $\|\boldsymbol{x}\|_2$、$\|\boldsymbol{y}\|_2$ 与 $\|\boldsymbol{x}\otimes\boldsymbol{y}\|_2$。

解. 根据定义，向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的 ℓ_2 范数分别为

$$\|\boldsymbol{x}\|_2=\sqrt{1^2+2^2}=\sqrt{5}\quad \|\boldsymbol{y}\|_2=\sqrt{3^2+4^2}=5 \tag{3.55}$$

由于 $\boldsymbol{x}\otimes\boldsymbol{y}=(3,4,6,8)^\top$，故 $\|\boldsymbol{x}\otimes\boldsymbol{y}\|_2=\sqrt{3^2+4^2+6^2+8^2}=5\sqrt{5}$。

3.3.3 矩阵的行列式

矩阵的行列式 (determinant) 是线性代数中非常重要的一个概念，贯穿线性代数的几乎所有内容，一般使用符号 $\det(\cdot)$ 表示。若给定矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times m}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{n\times n}$，则

$$\det(\boldsymbol{X}\otimes\boldsymbol{Y})=\det(\boldsymbol{X})^n\cdot\det(\boldsymbol{Y})^m \tag{3.56}$$

恒成立。

例 27. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}1&3&2\\4&1&3\\2&5&2\end{bmatrix}$，试写出矩阵的行列式 $\det(\boldsymbol{X})$、$\det(\boldsymbol{Y})$ 与 $\det(\boldsymbol{X}\otimes\boldsymbol{Y})$。

解. 矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的行列式分别为

$$\det(\boldsymbol{X})=\begin{vmatrix}1&2\\3&4\end{vmatrix}=-2\quad \det(\boldsymbol{Y})=\begin{vmatrix}1&3&2\\4&1&3\\2&5&2\end{vmatrix}=17 \tag{3.57}$$

故 $\det(\boldsymbol{X})^3\cdot\det(\boldsymbol{Y})^2=-2312$。

矩阵 $\boldsymbol{X}\otimes\boldsymbol{Y}$ 的行列式为

$$\det(\boldsymbol{X}\otimes\boldsymbol{Y})=\begin{vmatrix}1&3&2&2&6&4\\4&1&3&8&2&6\\2&5&2&4&10&4\\3&9&6&4&12&8\\12&3&9&16&4&12\\6&15&6&8&20&8\end{vmatrix}=-2312 \tag{3.58}$$

3.3.4 矩阵的秩

矩阵的秩 (rank) 是线性代数中非常重要的一个概念，在信号处理、图像处理等领域中应用广泛，一般使用符号 $\mathrm{rank}(\cdot)$ 表示。若给定矩阵 $\boldsymbol{X}\in\mathbb{R}^{m\times n}$ 与 $\boldsymbol{Y}\in\mathbb{R}^{p\times q}$，则

$$\mathrm{rank}(\boldsymbol{X}\otimes\boldsymbol{Y})=\mathrm{rank}(\boldsymbol{X})\cdot\mathrm{rank}(\boldsymbol{Y}) \tag{3.59}$$

恒成立。

例 28. 给定矩阵 $\boldsymbol{X}=\begin{bmatrix}1&2\\2&4\end{bmatrix}$ 与 $\boldsymbol{Y}=\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}$，试写出 $\operatorname{rank}(\boldsymbol{X})$、$\operatorname{rank}(\boldsymbol{Y})$ 与 $\operatorname{rank}(\boldsymbol{X}\otimes\boldsymbol{Y})$。

解. 在这里，$\operatorname{rank}(\boldsymbol{X})=1$，$\operatorname{rank}(\boldsymbol{Y})=2$。

由于

$$\boldsymbol{X}\otimes\boldsymbol{Y}=\begin{bmatrix}5&6&7&10&12&14\\8&9&10&16&18&20\\10&12&14&20&24&28\\16&18&20&32&36&40\end{bmatrix}\tag{3.60}$$

故 $\operatorname{rank}(\boldsymbol{X}\otimes\boldsymbol{Y})=2$。

3.4　朴素 Kronecker 分解

3.4.1　定义

一般而言，给定任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{(mp)\times(nq)}$，若 $\boldsymbol{A}\in\mathbb{R}^{m\times n},\boldsymbol{B}\in\mathbb{R}^{p\times q}$ 为朴素 Kronecker 分解中的待定参数，则可将分解过程描述为如下优化问题：

$$\min_{\boldsymbol{A},\boldsymbol{B}}\ \|\boldsymbol{X}-\boldsymbol{A}\otimes\boldsymbol{B}\|_F^2\tag{3.61}$$

其中，我们建模的目标是寻找最佳的矩阵 $\boldsymbol{A},\boldsymbol{B}$ 使得损失函数最小化。

为便于理解该优化问题，不妨用一组小矩阵一窥究竟，令 $m=3,n=p=q=2$，则此时的目标函数为

$$\|\boldsymbol{X}-\boldsymbol{A}\otimes\boldsymbol{B}\|_F^2=\left\|\left[\begin{array}{cc|cc}x_{11}&x_{12}&x_{13}&x_{14}\\x_{21}&x_{22}&x_{23}&x_{24}\\\hline x_{31}&x_{32}&x_{33}&x_{34}\\x_{41}&x_{42}&x_{43}&x_{44}\\\hline x_{51}&x_{52}&x_{53}&x_{54}\\x_{61}&x_{62}&x_{63}&x_{64}\end{array}\right]-\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\\a_{31}&a_{32}\end{bmatrix}\otimes\begin{bmatrix}b_{11}&b_{12}\\b_{21}&b_{22}\end{bmatrix}\right\|_F^2\tag{3.62}$$

3.4.2　引入 permute 概念

在这里，我们引入 permute 概念是为了对矩阵的维度按照特定规则进行调整，这一做法最早是由 Van Loan 和 Pitsianis 于 1993 年提出的 [Van Loan and Pitsianis, 1993]。在公式(3.62)中，首先使用分块矩阵表示矩阵 $\boldsymbol{X}\in\mathbb{R}^{6\times4}$：

$$\left[\begin{array}{cc|cc}x_{11}&x_{12}&x_{13}&x_{14}\\x_{21}&x_{22}&x_{23}&x_{24}\\\hline x_{31}&x_{32}&x_{33}&x_{34}\\x_{41}&x_{42}&x_{43}&x_{44}\\\hline x_{51}&x_{52}&x_{53}&x_{54}\\x_{61}&x_{62}&x_{63}&x_{64}\end{array}\right]=\begin{bmatrix}\boldsymbol{X}_{11}&\boldsymbol{X}_{12}\\\boldsymbol{X}_{21}&\boldsymbol{X}_{22}\\\boldsymbol{X}_{31}&\boldsymbol{X}_{32}\end{bmatrix}\tag{3.63}$$

其中，分块矩阵 $\boldsymbol{X}$ 拥有 3×2 个分块，即子矩阵，每个子矩阵的大小为 2×2，这些子矩阵分别写作如下形式：

$$\begin{aligned}\boldsymbol{X}_{11}&=\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}&\boldsymbol{X}_{12}&=\begin{bmatrix}x_{13}&x_{14}\\x_{23}&x_{24}\end{bmatrix}\\\boldsymbol{X}_{21}&=\begin{bmatrix}x_{31}&x_{32}\\x_{41}&x_{42}\end{bmatrix}&\boldsymbol{X}_{22}&=\begin{bmatrix}x_{33}&x_{34}\\x_{43}&x_{44}\end{bmatrix}\\\boldsymbol{X}_{31}&=\begin{bmatrix}x_{51}&x_{52}\\x_{61}&x_{62}\end{bmatrix}&\boldsymbol{X}_{32}&=\begin{bmatrix}x_{53}&x_{54}\\x_{63}&x_{64}\end{bmatrix}\end{aligned}\tag{3.64}$$

有了这些子矩阵之后，需要对这些子矩阵进行向量化，得到的向量依次为

$$\operatorname{vec}(\boldsymbol{X}_{11})=\begin{bmatrix}x_{11}\\x_{21}\\x_{12}\\x_{22}\end{bmatrix}\quad\operatorname{vec}(\boldsymbol{X}_{21})=\begin{bmatrix}x_{31}\\x_{41}\\x_{32}\\x_{42}\end{bmatrix}\quad\cdots\quad\operatorname{vec}(\boldsymbol{X}_{32})=\begin{bmatrix}x_{53}\\x_{63}\\x_{54}\\x_{64}\end{bmatrix}\tag{3.65}$$

最后，使用这些向量构造如下矩阵：

$$\tilde{\boldsymbol{X}}=\begin{bmatrix}\operatorname{vec}(\boldsymbol{X}_{11})^\top\\\operatorname{vec}(\boldsymbol{X}_{21})^\top\\\operatorname{vec}(\boldsymbol{X}_{31})^\top\\\operatorname{vec}(\boldsymbol{X}_{12})^\top\\\operatorname{vec}(\boldsymbol{X}_{22})^\top\\\operatorname{vec}(\boldsymbol{X}_{32})^\top\end{bmatrix}\in\mathbb{R}^{6\times 4}\tag{3.66}$$

在这里，将矩阵 $\boldsymbol{X}$ 构造成矩阵 $\tilde{\boldsymbol{X}}$ 的过程通常被称为 permute。

由于

$$\begin{aligned}\operatorname{vec}(\boldsymbol{X}_{11})&=a_{11}\cdot\operatorname{vec}(\boldsymbol{B})\\\operatorname{vec}(\boldsymbol{X}_{21})&=a_{21}\cdot\operatorname{vec}(\boldsymbol{B})\\&\vdots\\\operatorname{vec}(\boldsymbol{X}_{32})&=a_{32}\cdot\operatorname{vec}(\boldsymbol{B})\end{aligned}\tag{3.67}$$

此时，Kronecker 分解的优化问题可写作如下形式：

$$\underset{\boldsymbol{A},\boldsymbol{B}}{\arg\min}\ \|\boldsymbol{X}-\boldsymbol{A}\otimes\boldsymbol{B}\|_F^2=\underset{\boldsymbol{A},\boldsymbol{B}}{\arg\min}\,\|\tilde{\boldsymbol{X}}-\operatorname{vec}(\boldsymbol{A})\operatorname{vec}(\boldsymbol{B})^\top\|_F^2\tag{3.68}$$

实际上，向量化之后的待定参数 $\operatorname{vec}(\boldsymbol{A})$ 和 $\operatorname{vec}(\boldsymbol{B})$ 构成了一个标准的矩阵分解问题。

3.4.3 求解过程

对于公式(3.61)中 Kronecker 分解的优化问题，可根据 Eckhart-Young 定理对如下优化问题进行求解：

$$\min_{\boldsymbol{A},\boldsymbol{B}}\ \|\tilde{\boldsymbol{X}}-\operatorname{vec}(\boldsymbol{A})\operatorname{vec}(\boldsymbol{B})^\top\|_F^2\tag{3.69}$$

若 $\tilde{\boldsymbol{X}}$ 的奇异值分解为 $\tilde{\boldsymbol{X}}=\sum_{r=1}^{\min\{mn,pq\}}\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^\top$，其中，奇异值为 $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_{\min\{mn,pq\}}$，则矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的最优解为

$$\begin{cases}\operatorname{vec}(\hat{\boldsymbol{A}})=\sqrt{\sigma_1}\cdot\boldsymbol{u}_1\\\operatorname{vec}(\hat{\boldsymbol{B}})=\sqrt{\sigma_1}\cdot\boldsymbol{v}_1\end{cases}\tag{3.70}$$

这里的最优解恰好是秩为 1 的逼近问题。

例 29. 给定矩阵 $\boldsymbol{A}=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ 与 $\boldsymbol{B}=\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}$，试写出两者之间的 *Kronecker* 积 $\boldsymbol{X}=\boldsymbol{A}\otimes\boldsymbol{B}$，并求 *Kronecker* 分解 $\hat{\boldsymbol{A}},\hat{\boldsymbol{B}}=\arg\min\limits_{\boldsymbol{A},\boldsymbol{B}}\|\boldsymbol{X}-\boldsymbol{A}\otimes\boldsymbol{B}\|_F^2$。

解. 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间的 *Kronecker* 积为

$$\boldsymbol{X}=\boldsymbol{A}\otimes\boldsymbol{B}=\begin{bmatrix}5&6&7&10&12&14\\8&9&10&16&18&20\\15&18&21&20&24&28\\24&27&30&32&36&40\end{bmatrix}\tag{3.71}$$

令分块矩阵 $\boldsymbol{X}$ 由如下 *4* 个子矩阵构成：

$$\begin{aligned}\boldsymbol{X}_{11}&=\begin{bmatrix}5&6&7\\8&9&10\end{bmatrix}\quad\boldsymbol{X}_{12}=\begin{bmatrix}10&12&14\\16&18&20\end{bmatrix}\\\boldsymbol{X}_{21}&=\begin{bmatrix}15&18&21\\24&27&30\end{bmatrix}\quad\boldsymbol{X}_{22}=\begin{bmatrix}20&24&28\\32&36&40\end{bmatrix}\end{aligned}\tag{3.72}$$

对这些子矩阵分别进行向量化：

$$\operatorname{vec}(\boldsymbol{X}_{11})=\begin{bmatrix}5\\8\\6\\9\\7\\10\end{bmatrix}\quad\operatorname{vec}(\boldsymbol{X}_{21})=\begin{bmatrix}15\\24\\18\\27\\21\\30\end{bmatrix}\quad\operatorname{vec}(\boldsymbol{X}_{12})=\begin{bmatrix}10\\16\\12\\18\\14\\20\end{bmatrix}\quad\operatorname{vec}(\boldsymbol{X}_{22})=\begin{bmatrix}20\\32\\24\\36\\28\\40\end{bmatrix}\tag{3.73}$$

有了这些向量之后，构造如下矩阵：

$$\tilde{\boldsymbol{X}}=\begin{bmatrix}\operatorname{vec}(\boldsymbol{X}_{11})^\top\\\operatorname{vec}(\boldsymbol{X}_{21})^\top\\\operatorname{vec}(\boldsymbol{X}_{12})^\top\\\operatorname{vec}(\boldsymbol{X}_{22})^\top\end{bmatrix}=\begin{bmatrix}5&8&6&9&7&10\\15&24&18&27&21&30\\10&16&12&18&14&20\\20&32&24&36&28&40\end{bmatrix}\tag{3.74}$$

由此，*Kronecker* 分解的优化问题等价于

$$\hat{\boldsymbol{A}},\hat{\boldsymbol{B}}=\arg\min_{\boldsymbol{A},\boldsymbol{B}}\|\tilde{\boldsymbol{X}}-\operatorname{vec}(\boldsymbol{A})\operatorname{vec}(\boldsymbol{B})^\top\|_F^2\tag{3.75}$$

对矩阵 $\tilde{\boldsymbol{X}}$ 进行奇异值分解，则矩阵 $\hat{\boldsymbol{A}}$ 与 $\hat{\boldsymbol{B}}$ 分别为

$$\begin{aligned}\hat{\boldsymbol{A}}&=\begin{bmatrix}-1.85471325&-3.7094265\\-5.56413975&-7.418853\end{bmatrix}\\\hat{\boldsymbol{B}}&=\begin{bmatrix}-2.69583452&-3.23500142&-3.77416832\\-4.31333523&-4.85250213&-5.39166904\end{bmatrix}\end{aligned}\tag{3.76}$$

在这里，矩阵 $\hat{\boldsymbol{A}}$ 与 $\hat{\boldsymbol{B}}$ 的所有元素均为负数，可将这些元素全部写成相反数。

3.5 广义 Kronecker 分解

形式上说，给定任意矩阵 $\boldsymbol{X}\in\mathbb{R}^{(mp)\times(nq)}$，若 $\boldsymbol{A}_r\in\mathbb{R}^{m\times n},\boldsymbol{B}_r\in\mathbb{R}^{p\times q},r=1,2,\ldots,R$ 为广义 Kronecker 分解中的待定参数，则可将分解过程描述为如下逼近问题 [Hameed et al., 2022]:

$$\min_{\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R}\left\|\boldsymbol{X}-\sum_{r=1}^R\boldsymbol{A}_r\otimes\boldsymbol{B}_r\right\|_F^2 \tag{3.77}$$

其中，我们的建模目标是寻找最佳的矩阵 $\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R$ 使得损失函数最小化。在这里，参数数量为 $R(mn+pq)$。

与朴素 Kronecker 分解类似，可先将广义 Kronecker 分解的逼近问题写作如下形式：

$$\arg\min_{\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R}\left\|\boldsymbol{X}-\sum_{r=1}^R\boldsymbol{A}_r\otimes\boldsymbol{B}_r\right\|_F^2=\arg\min_{\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R}\left\|\tilde{\boldsymbol{X}}-\sum_{r=1}^R\operatorname{vec}(\boldsymbol{A}_r)\operatorname{vec}(\boldsymbol{B}_r)^\top\right\|_F^2 \tag{3.78}$$

其中，矩阵 $\tilde{\boldsymbol{X}}$ 是由矩阵 $\boldsymbol{X}$ 进行 permute 构造得到。

根据 Eckhart-Young 定理对上述优化问题进行求解，若矩阵 $\tilde{\boldsymbol{X}}$ 的奇异值分解为 $\tilde{\boldsymbol{X}}=\sum_{r=1}^{\min\{mn,pq\}}\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^\top$，其中，奇异值为 $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_{\min\{mn,pq\}}$，则矩阵 $\boldsymbol{A}_r$ 和 $\boldsymbol{B}_r$ 的最优解为

$$\begin{cases}\operatorname{vec}(\hat{\boldsymbol{A}}_r)=\sqrt{\sigma_r}\boldsymbol{u}_r\\ \operatorname{vec}(\hat{\boldsymbol{B}}_r)=\sqrt{\sigma_r}\boldsymbol{v}_r\end{cases} \tag{3.79}$$

例 30. *给定大小为 512 × 512 的灰度图像，如图 3.1所示，令 $\boldsymbol{A}_r\in\mathbb{R}^{16\times 32},\boldsymbol{B}_r\in\mathbb{R}^{32\times 16},r=1,2,\ldots,R$，给定 $R=5,10,50,100$，试采用广义 Kronecker 分解重构灰度图像，即*

$$\hat{\boldsymbol{X}}=\sum_{r=1}^R\boldsymbol{A}_r\otimes\boldsymbol{B}_r \tag{3.80}$$

图 3.1: 大熊猫灰度图像

解. 对于广义 *Kronecker* 分解，根据公式(3.78)与(3.79)，给定 $R=5,10,50,100$，能得到图 *7.3*所示的重构结果。当 $R=5$ 时，图像压缩率（参数数量除以观测值数量）为 $\frac{5\times 2}{512}$；当 $R=100$ 时，图像压缩率为 $\frac{100\times 2}{512}$；随着秩 R 增大，图片重构效果逐渐变好。

3.6 模型参数压缩问题

Kronecker 分解的一个重要用途是压缩模型参数。以多元线性回归 (multivariate linear regression) 为例，给定输入、输出数据为 $\mathcal{D}=\{(\boldsymbol{x}_1,\boldsymbol{y}_1),\cdots,(\boldsymbol{x}_N,\boldsymbol{y}_N)\}\in\mathbb{R}^{nq}\times\mathbb{R}^{mp}$，则多元线性回归的优化问题为

$$\min_{\boldsymbol{W}}\ \frac{1}{2}\sum_{n=1}^N\|\boldsymbol{y}_n-\boldsymbol{W}\boldsymbol{x}_n\|_2^2 \tag{3.81}$$

(a) $R=5$ (b) $R=10$

(c) $R=50$ (d) $R=100$

图 3.2: 基于广义 Kronecker 分解的灰度图像重构

令

$$\boldsymbol{X}=\begin{bmatrix} | & & | \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_N \\ | & & | \end{bmatrix}\in\mathbb{R}^{nq\times N} \tag{3.82}$$

$$\boldsymbol{Y}=\begin{bmatrix} | & & | \\ \boldsymbol{y}_1 & \cdots & \boldsymbol{y}_N \\ | & & | \end{bmatrix}\in\mathbb{R}^{mp\times N}, \tag{3.83}$$

则此时多元线性回归的等价优化问题为

$$\min_{\boldsymbol{W}}\ \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{W}\boldsymbol{X}\|_F^2 \tag{3.84}$$

不妨假设这里的系数矩阵 $\boldsymbol{W}\in\mathbb{R}^{(mp)\times(nq)}$ 存在一个广义 Kronecker 分解，且由 R 个成分构成，则基于广义 Kronecker 分解的多元线性回归可写作如下形式：

$$\min_{\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R}\ \frac{1}{2}\left\|\boldsymbol{Y}-\sum_{r=1}^R(\boldsymbol{A}_r\otimes\boldsymbol{B}_r)\boldsymbol{X}\right\|_F^2 \tag{3.85}$$

将优化问题改写为如下形式即可得到一个标准的广义 Kronecker 分解：

$$\min_{\{\boldsymbol{A}_r,\boldsymbol{B}_r\}_{r=1}^R}\ \frac{1}{2}\left\|\boldsymbol{Y}\boldsymbol{X}^{\dagger}-\sum_{r=1}^R(\boldsymbol{A}_r\otimes\boldsymbol{B}_r)\right\|_F^2 \tag{3.86}$$

从而可根据广义 Kronecker 分解的求解方法对该多元线性回归问题进行求解。

例 31 (矩阵自回归模型[2]). 对于多维时间序列 *(multidimensional time series)*, 若任意时刻 t 对应的观测数据为矩阵 $\boldsymbol{X}_t \in \mathbb{R}^{M\times N}$, 则矩阵自回归的表达式为

$$\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1}\boldsymbol{B}^\top + \boldsymbol{E}_t, t = 2, 3, \ldots, T \tag{3.87}$$

其中, $\boldsymbol{A} \in \mathbb{R}^{M\times M}$ 与 $\boldsymbol{B} \in \mathbb{R}^{N\times N}$ 为自回归过程的系数矩阵 *(coefficient matrix)*; 矩阵 $\boldsymbol{E}_t \in \mathbb{R}^{M\times N}$ 为自回归过程的残差矩阵 *(residual matrix)*。若令 $\boldsymbol{x}_t = \mathrm{vec}(\boldsymbol{X}_t)$ 与 $\boldsymbol{\epsilon}_t = \mathrm{vec}(\boldsymbol{E}_t)$, 试写出与矩阵自回归等价的向量自回归表达式。

解. 根据 *Kronecker* 积性质, 矩阵自回归等价于如下向量自回归:

$$\begin{aligned}\mathrm{vec}(\boldsymbol{X}_t) &= \mathrm{vec}(\boldsymbol{A}\boldsymbol{X}_{t-1}\boldsymbol{B}^\top) + \mathrm{vec}(\boldsymbol{E}_t)\\ &= (\boldsymbol{B}\otimes\boldsymbol{A})\,\mathrm{vec}(\boldsymbol{X}_{t-1}) + \mathrm{vec}(\boldsymbol{E}_t)\\ \Longrightarrow \quad \boldsymbol{x}_t &= (\boldsymbol{B}\otimes\boldsymbol{A})\boldsymbol{x}_{t-1} + \boldsymbol{\epsilon}_t\end{aligned} \tag{3.88}$$

在这里, 矩阵自回归的待定参数数量为 $M^2 + N^2$, 若对观测数据进行向量化且不对系数矩阵进行 *Kronecker* 分解, 则向量自回归的待定参数数量为 $(MN)^2$, 容易引发过参数化 *(over-parameterization)* 问题。

参考资料

1. Kathrin Schacke (2013). On the Kronecker Product. `https://www.math.uwaterloo.ca/~hwolkowi/henry/reports/kronthesisschaecke04.pdf`

2. Are Hjørungnes (2011). Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications. Cambridge University Press.

[2] `http://www.stat.rutgers.edu/home/rongchen/publications/20JoE_Matrix_AR.pdf`