面向时空交通数据修复及预测的低秩机器学习模型

陈新宇 (https://xinychen.github.io)
伍元凯 (https://kaimaoge.github.io)
孙立君 (https://lijunsun.github.io)

发布时间：2023 年 2 月
更新时间：2023 年 2 月

目录

第一章 前言 **5**

1.1 基本代数结构 5

1.2 向量与矩阵 5

1.2.1 向量 5

1.2.2 矩阵 6

1.2.3 矩阵向量化 6

1.3 高阶张量 6

1.3.1 高阶张量结构 6

1.3.2 高阶张量矩阵化 8

1.3.3 高阶张量向量化 8

1.4 时空交通数据类型 8

第二章 矩阵分解基础 **9**

第三章 张量分解基础 **11**

3.1 Kronecker 积定义 11

3.1.1 基本定义 11

3.1.2 Khatri-Rao 积 12

3.2 Kronecker 积基本性质 13

3.2.1 结合律与分配律 13

3.2.2 矩阵相乘 14

3.2.3 求逆矩阵 15

3.2.4 向量化 16

3.3 Kronecker 积特殊性质 17

3.3.1 矩阵的迹 17

3.3.2 矩阵的 Frobenius 范数 18

3.3.3 矩阵的行列式 19

3.3.4 矩阵的秩 19

3.4 向量外积 20

3.4.1 定义 20

3.4.2 性质 21

3.5 CP 张量分解 22

3.5.1 CP 分解形式 22

第四章 时序矩阵分解 **23**

第五章 贝叶斯张量分解 **25**

第六章 低秩张量填充 **27**

第七章 低秩拉普拉斯卷积模型 **29**

第八章 基于延迟嵌套的张量分解 **31**

第一章　前言

1.1　基本代数结构

长期以来，线性代数一直作为机器学习中最为重要的数学工具之一，被人们广泛用于开发各类机器学习算法。线性代数本质上是以向量与矩阵为基本代数结构，本文要讨论的张量分解等模型则主要以张量为基本代数结构。在过去的数十年间，借助线性代数这一基本数学工具，机器学习中涌现出了很多经典的代数模型，这其中不乏矩阵分解、主成分分析，而张量分解在某种程度上可看作是矩阵分解的一种衍生物。

近年来，张量分解在机器学习的众多问题中得到了很好的应用 [Kolda and Bader, 2009, Sidiropoulos et al., 2017]，但关于张量的一些计算与我们所熟悉的线性代数却大相径庭，同时，张量计算相比以矩阵计算为主导的线性代数更为抽象，这使得很多与张量分解相关的内容看起来晦涩难懂。实际上，向量与矩阵都是张量的特例，可以被定义为低阶张量。一般而言，向量是一阶张量，英文表述为 first-order tensor；矩阵是二阶张量，英文表述为 second-order tensor；三阶或者更高阶数的张量被称为高阶张量，英文表述为 higher-order tensor。在各类文献中，通常提到的张量都是特指高阶张量，当然，这在本文的叙述中也不例外。需要注意的是，在各类程序语言中，人们更愿意将张量称为多维数组。

在一个矩阵中，某一元素的位置可以说是“第 i 行、第 j 列”，即要描述某一元素的位置需用到行和列索引构成的组合 (i,j)。类似地，在一个三阶张量中，描述某一元素的位置需用到三个索引构成的组合，例如 (i,j,k)。在处理稀疏矩阵或稀疏张量时，用索引来标记元素的位置会节省下一些不必要的存储开支。

1.2　向量与矩阵

1.2.1　向量

向量包括行向量与列向量。在写法上，为避免混淆，向量在没有特别申明的情况下是指列向量，给定任意向量 $\boldsymbol{x}\in\mathbb{R}^n$ 表示大小为 n 的向量，写作

$$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^\top \tag{1.1}$$

或

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} \tag{1.2}$$

其中，符号 $\cdot^\top$ 表示转置 (transpose)。

1.2.2 矩阵

一般而言，给定任意矩阵 $\boldsymbol{X} \in \mathbb{R}^{m\times n}$，矩阵的行数为 m、列数为 n，其第 (i,j) 个元素（即矩阵的第 i 行、第 j 列元素）为

$$x_{i,j} = \boldsymbol{X}_{i,j} \tag{1.3}$$

其中，$i=1,2,\ldots,m$ 与 $j=1,2,\ldots,n$。

单位矩阵一般记作 $\boldsymbol{I}_n$，大小为 $n\times n$，其对角线上的元素均为 1、其他位置上的元素均为 0。

1.2.3 矩阵向量化

给定任意矩阵 $\boldsymbol{X} \in \mathbb{R}^{m\times n}$，若矩阵的列向量为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n \in \mathbb{R}^m$，即

$$\boldsymbol{X} = \begin{bmatrix} | & | & & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \\ | & | & & | \end{bmatrix} \tag{1.4}$$

则可对矩阵按列进行向量化，得到的向量为

$$\operatorname{vec}(\boldsymbol{X}) = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix} \in \mathbb{R}^{mn} \tag{1.5}$$

其中，符号 $\operatorname{vec}(\cdot)$ 表示向量化操作。

与矩阵向量化相反，也可定义向量的矩阵化规则。

1.3 高阶张量

1.3.1 高阶张量结构

高阶张量

一般而言，高阶张量可写成 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m_1\times m_2\times\cdots\times m_d}$，张量的阶数为 d，大小为 $m_1\times m_2\times\cdots\times m_d$。

三阶张量中的元素

这里以三阶张量为例，给定任意三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$，其第 (i,j,k) 个元素可写作如下形式：

$$x_{i,j,k} = \boldsymbol{\mathcal{X}}_{i,j,k} \tag{1.6}$$

其中，$i=1,2,\ldots,m$、$j=1,2,\ldots,n$ 与 $k=1,2,\ldots,t$。

图1.1直观地展现了三阶张量元素的示意图，可以看出：描述三阶张量中的某一元素需用到三个索引构成的组合，例如 (i,j,k)。

三阶张量中的纤维

给定任意三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$，其各个方向的纤维 (fiber) 都是向量，如图1.2所示，这些纤维分别为向量 $\boldsymbol{\mathcal{X}}_{:,j,k} \in \mathbb{R}^m$、$\boldsymbol{\mathcal{X}}_{i,:,k} \in \mathbb{R}^n$ 与 $\boldsymbol{\mathcal{X}}_{i,j,:} \in \mathbb{R}^t$，其中，$i=1,2,\ldots,m$、$j=1,2,\ldots,n$ 与 $k=1,2,\ldots,t$。与矩阵中的行向量、列向量类似，纤维是张量的基本组成部分。

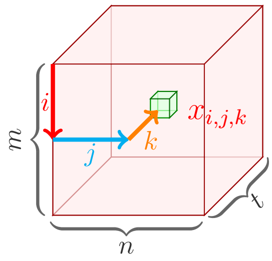
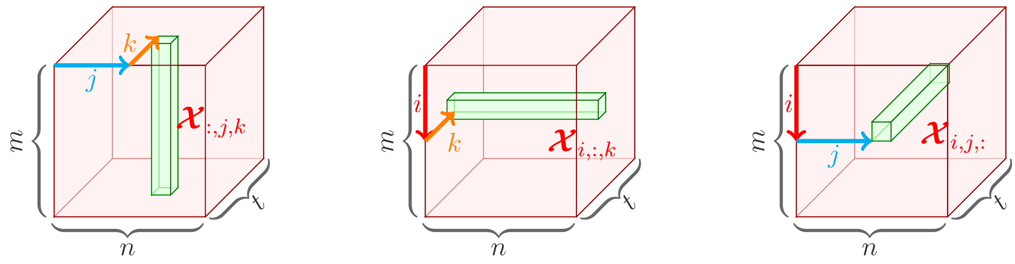


图 1.1: 三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$ 及其第 (i,j,k) 个元素 $x_{i,j,k}$



(a) 向量 $\boldsymbol{\mathcal{X}}_{:,j,k} \in \mathbb{R}^m$ (b) 向量 $\boldsymbol{\mathcal{X}}_{i,:,k} \in \mathbb{R}^n$ (c) 向量 $\boldsymbol{\mathcal{X}}_{i,j,:} \in \mathbb{R}^t$

图 1.2: 三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$ 自三个维度的纤维

三阶张量中的切片

对于任意三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$，可用三个维度的切片 (slice) 书写该张量，其中，horizontal 切片共有 m 个，分别为

$$\boldsymbol{\mathcal{X}}_{1,:,:}, \boldsymbol{\mathcal{X}}_{2,:,:}, \ldots, \boldsymbol{\mathcal{X}}_{m,:,:} \in \mathbb{R}^{n\times t} \tag{1.7}$$

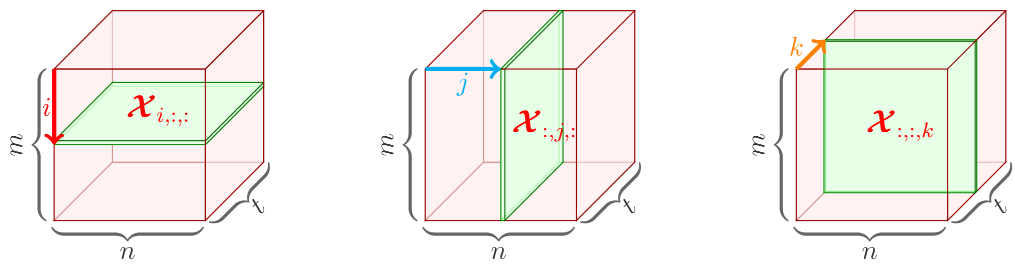
lateral 切片共有 n 个，分别为

$$\boldsymbol{\mathcal{X}}_{:,1,:}, \boldsymbol{\mathcal{X}}_{:,2,:}, \ldots, \boldsymbol{\mathcal{X}}_{:,n,:} \in \mathbb{R}^{m\times t} \tag{1.8}$$

frontal 切片共有 t 个，分别为

$$\boldsymbol{\mathcal{X}}_{:,:,1}, \boldsymbol{\mathcal{X}}_{:,:,2}, \ldots, \boldsymbol{\mathcal{X}}_{:,:,t} \in \mathbb{R}^{m\times n} \tag{1.9}$$

如图1.3所示，这些矩阵结构的切片是张量的基本组成部分。



(a) horizontal 切片 $\boldsymbol{\mathcal{X}}_{i,:,:}$ (b) lateral 切片 $\boldsymbol{\mathcal{X}}_{:,j,:}$ (c) frontal 切片 $\boldsymbol{\mathcal{X}}_{:,:,k} \in \mathbb{R}^t$

图 1.3: 三阶张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times t}$ 自三个维度的切片

例 1. *给定张量 $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{2\times 2\times 2}$，若其 frontal 切片为*

$$\boldsymbol{\mathcal{X}}_{:,:,1} = \begin{bmatrix} x_{111} & x_{121} \\ x_{211} & x_{221} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \boldsymbol{\mathcal{X}}_{:,:,2} = \begin{bmatrix} x_{112} & x_{122} \\ x_{212} & x_{222} \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \tag{1.10}$$

试写出张量 $\boldsymbol{\mathcal{X}}$ 的 *lateral* 切片与 *horizontal* 切片。

解. 张量 $\boldsymbol{\mathcal{X}}$ 的 *lateral* 切片为

$$\boldsymbol{\mathcal{X}}_{:,1,:}=\begin{bmatrix}x_{111} & x_{112}\\ x_{211} & x_{212}\end{bmatrix}=\begin{bmatrix}1 & 5\\ 3 & 7\end{bmatrix}\quad \boldsymbol{\mathcal{X}}_{:,2,:}=\begin{bmatrix}x_{121} & x_{122}\\ x_{221} & x_{222}\end{bmatrix}=\begin{bmatrix}2 & 6\\ 4 & 8\end{bmatrix} \tag{1.11}$$

张量 $\boldsymbol{\mathcal{X}}$ 的 *horizontal* 切片为

$$\boldsymbol{\mathcal{X}}_{1,:,:}=\begin{bmatrix}x_{111} & x_{112}\\ x_{121} & x_{122}\end{bmatrix}=\begin{bmatrix}1 & 5\\ 2 & 6\end{bmatrix}\quad \boldsymbol{\mathcal{X}}_{2,:,:}=\begin{bmatrix}x_{211} & x_{212}\\ x_{221} & x_{222}\end{bmatrix}=\begin{bmatrix}3 & 7\\ 4 & 8\end{bmatrix} \tag{1.12}$$

1.3.2 高阶张量矩阵化

1.3.3 高阶张量向量化

给定任意张量 $\boldsymbol{\mathcal{X}}\in\mathbb{R}^{m_1\times m_2\times\cdots\times m_d}$, 阶数为 d, 若其以第一个维度展开得到的矩阵为 $\boldsymbol{X}_{(1)}$, 则张量向量化可写作如下形式:

$$\operatorname{vec}(\boldsymbol{\mathcal{X}})=\operatorname{vec}(\boldsymbol{X}_{(1)}) \tag{1.13}$$

在张量 $\boldsymbol{\mathcal{X}}$ 中，第 $(i_1,i_2,\ldots,i_d)$ 个元素通过张量向量化之后，该元素在向量中的位置为

$$\Big(\sum_{k=1}^{d-1}m_k\Big)\cdot i_d+\Big(\sum_{k=1}^{d-2}m_k\Big)\cdot i_{d-1}+\cdots+m_1\cdot i_2+i_1 \tag{1.14}$$

例 2. 给定张量 $\boldsymbol{\mathcal{X}}\in\mathbb{R}^{2\times2\times2}$，若其 *frontal* 切片为

$$\boldsymbol{\mathcal{X}}_{:,:,1}=\begin{bmatrix}x_{111} & x_{121}\\ x_{211} & x_{221}\end{bmatrix}=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\quad \boldsymbol{\mathcal{X}}_{:,:,2}=\begin{bmatrix}x_{112} & x_{122}\\ x_{212} & x_{222}\end{bmatrix}=\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix} \tag{1.15}$$

试写出张量向量化的结果 $\operatorname{vec}(\boldsymbol{\mathcal{X}})$。

解. 根据张量向量化规则，有

$$\begin{aligned}\operatorname{vec}(\boldsymbol{\mathcal{X}}) &=\operatorname{vec}(\boldsymbol{X}_{(1)})\\ &=\operatorname{vec}([\boldsymbol{\mathcal{X}}_{:,:,1}\quad \boldsymbol{\mathcal{X}}_{:,:,2}])\\ &=(1,3,2,4,5,7,6,8)^\top\end{aligned} \tag{1.16}$$

1.4 时空交通数据类型

第二章　矩阵分解基础

第三章　张量分解基础

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

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

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。

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。

例 3. *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 4. 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。

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。

例 5. 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 6 (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**??**）:

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

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 7. 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

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

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 8. 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 9. 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 10 (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 11. 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:

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

由此，也可得到

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

例 12. 试证明公式(3.32)。

解.

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

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 13. 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

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

恒成立。

解. 根据 *Kronecker* 积性质，有

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

例 14. 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 15. *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

$$\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{C} \tag{3.39}$$

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：

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

因此，*Sylvester* 方程的解析解[1]为

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

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 16. 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 $\operatorname{tr}(\boldsymbol{X}\otimes\boldsymbol{Y})$。

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

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}) \tag{3.46}$$

及

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

例 17. 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

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

[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 $\operatorname{vec}(\boldsymbol{X})=(\boldsymbol{A}\oplus\boldsymbol{B}^\top)^{-1}\operatorname{vec}(\boldsymbol{C})$。

解. 根据矩阵的迹与 *Kronecker* 积性质，有

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

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 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 $\|\boldsymbol{X}\|_F$、$\|\boldsymbol{Y}\|_F$ 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 19. 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 20. 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 $\operatorname{rank}(\cdot)$ 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

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

恒成立。

例 21. 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 向量外积

3.4.1 定义

在线性代数中，两向量之间的外积可得到一个矩阵。对于任意向量 $\boldsymbol{a}\in\mathbb{R}^m$ 与 $\boldsymbol{b}\in\mathbb{R}^n$，则两者之间的外积为

$$\boldsymbol{c}=\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}=\boldsymbol{a}\boldsymbol{b}^\top\in\mathbb{R}^{m\times n}\tag{3.61}$$

其中，符号 $\otimes_{\text{outer}}$ 表示向量外积。

依此类推，对任意 d 个向量 $\boldsymbol{x}^{(k)}\in\mathbb{R}^{n_k},k=1,2,\ldots,d$，其外积 (outer product) 可定义为

$$\boldsymbol{\mathcal{Y}}=\boldsymbol{x}^{(1)}\otimes_{\text{outer}}\boldsymbol{x}^{(2)}\otimes_{\text{outer}}\cdots\otimes_{\text{outer}}\boldsymbol{x}^{(d)}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}\tag{3.62}$$

其中，符号 $\otimes_{\text{outer}}$ 表示向量外积。在张量 $\boldsymbol{\mathcal{Y}}$ 中，任意第 $(i_1,i_2,\ldots,i_d)$ 个元素为

$$y_{i_1,i_2,\ldots,i_d}=\prod_{k=1}^{d}x_{i_k}^{(k)}\tag{3.63}$$

其中，$i_k=1,2,\ldots,n_k$，$k=1,2,\ldots,d$。

需要注意的是，由于张量 $\boldsymbol{\mathcal{Y}}$ 是由向量外积得到的，故常被称为秩一张量 (rank-one tensor)。当 $d=3$ 时，向量外积得到的三阶张量 $\boldsymbol{\mathcal{Y}}=\boldsymbol{x}^{(1)}\otimes_{\text{outer}}\boldsymbol{x}^{(2)}\otimes_{\text{outer}}\boldsymbol{x}^{(3)}\in\mathbb{R}^{n_1\times n_2\times n_3}$ 如图3.1所示，在这里，张量 $\boldsymbol{\mathcal{Y}}$ 的任意第 (i_1,i_2,i_3) 个元素为

$$y_{i_1,i_2,i_3}=\prod_{k=1}^{3}x_{i_k}^{(k)}\tag{3.64}$$

其中，$i_k=1,2,\ldots,n_k$，$k=1,2,3$。

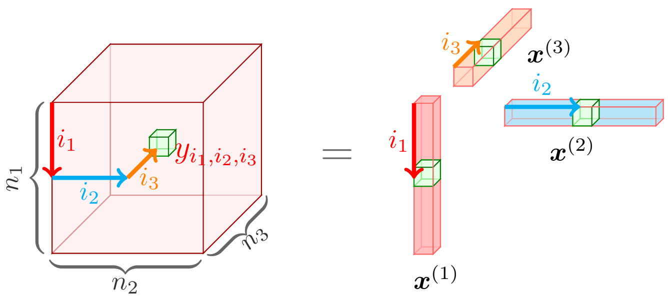


图 3.1: 向量外积得到的三阶张量

当 $d=2$ 时，向量外积为

$$\boldsymbol{Y}=\boldsymbol{x}^{(1)}\otimes_{\text{outer}}\boldsymbol{x}_{(2)}=\boldsymbol{x}^{(1)}(\boldsymbol{x}^{(2)})^\top\in\mathbb{R}^{n_1\times n_2}\tag{3.65}$$

在矩阵 $\boldsymbol{Y}$ 中，任意第 (i,j) 个元素为

$$y_{i,j}=x_i^{(1)}x_j^{(2)}\tag{3.66}$$

其中，$i=1,2,\ldots,n_1$ 与 $j=1,2,\ldots,n_2$。

例 22. 给定向量 $\boldsymbol{x}=(1,2)^\top$ 与 $\boldsymbol{y}=(3,4)^\top$，试写出 $\boldsymbol{x}\otimes_{\text{outer}}\boldsymbol{y}$。

解. 根据外积定义，有

$$\boldsymbol{x}\otimes_{\text{outer}}\boldsymbol{y}=\begin{bmatrix}3&4\\6&8\end{bmatrix}\tag{3.67}$$

例 23. 给定向量 $\boldsymbol{a}=(1,2)^\top$、$\boldsymbol{b}=(3,4,5)^\top$ 与 $\boldsymbol{c}=(6,7,8,9)^\top$，试写出 $\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}\otimes_{\text{outer}}\boldsymbol{c}$。

解. 令 $\boldsymbol{\mathcal{Y}}=\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}\otimes_{\text{outer}}\boldsymbol{c}\in\mathbb{R}^{2\times3\times4}$，根据外积定义，有

$$\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}=\begin{bmatrix}3&4&5\\6&8&10\end{bmatrix}\tag{3.68}$$

由此，可得张量 $\boldsymbol{\mathcal{Y}}$ 的 *frontal* 切片为

$$\begin{aligned}\boldsymbol{\mathcal{Y}}_{:,:,1}=(\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b})\cdot c_1=\begin{bmatrix}18&24&30\\36&48&60\end{bmatrix}\quad \boldsymbol{\mathcal{Y}}_{:,:,2}=(\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b})\cdot c_2=\begin{bmatrix}21&28&35\\42&56&70\end{bmatrix}\\ \boldsymbol{\mathcal{Y}}_{:,:,3}=(\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b})\cdot c_3=\begin{bmatrix}24&32&40\\48&64&80\end{bmatrix}\quad \boldsymbol{\mathcal{Y}}_{:,:,4}=(\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b})\cdot c_4=\begin{bmatrix}27&36&45\\54&72&90\end{bmatrix}\end{aligned}\tag{3.69}$$

3.4.2 性质

张量矩阵化

根据 Khatri-Rao 积定义与张量矩阵化规则，由向量 $\boldsymbol{x}^{(k)}\in\mathbb{R}^{n_k},k=1,2,\ldots,d$ 的外积得到的张量 $\boldsymbol{\mathcal{Y}}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$，其自第 k 维度展开得到的矩阵可写作如下形式：

$$\begin{aligned}\boldsymbol{Y}_{(k)}=&\boldsymbol{x}^{(k)}\otimes_{\text{outer}}(\boldsymbol{x}^{(d)}\odot\cdots\odot\boldsymbol{x}^{(k+1)}\odot\boldsymbol{x}^{(k-1)}\odot\cdots\odot\boldsymbol{x}^{(1)})\\ =&\boldsymbol{x}^{(k)}(\boldsymbol{x}^{(d)}\odot\cdots\odot\boldsymbol{x}^{(k+1)}\odot\boldsymbol{x}^{(k-1)}\odot\cdots\odot\boldsymbol{x}^{(1)})^\top\end{aligned}\tag{3.70}$$

其中，$\odot$ 表示 Khatri-Rao 积。

例 24. 给定向量 $\boldsymbol{a}=(1,2)^\top$、$\boldsymbol{b}=(3,4,5)^\top$ 与 $\boldsymbol{c}=(6,7,8,9)^\top$，若 $\boldsymbol{\mathcal{Y}}=\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}\otimes_{\text{outer}}\boldsymbol{c}$，试写出张量 $\boldsymbol{\mathcal{Y}}$ 的矩阵化形式 $\boldsymbol{Y}_{(1)}$、$\boldsymbol{Y}_{(2)}$ 与 $\boldsymbol{Y}_{(3)}$。

解. 根据 *Khatri-Rao* 积定义，有

$$\begin{cases}\boldsymbol{c}\odot\boldsymbol{b}=(18,24,30,21,28,35,24,32,40,27,36,45)^\top\\ \boldsymbol{c}\odot\boldsymbol{a}=(6,12,7,14,8,16,9,18)^\top\\ \boldsymbol{b}\odot\boldsymbol{a}=(3,6,4,8,5,10)^\top\end{cases}\tag{3.71}$$

从而，可得到

$$\boldsymbol{Y}_{(1)}=\boldsymbol{a}(\boldsymbol{c}\odot\boldsymbol{b})^\top=\begin{bmatrix}18&24&30&21&28&35&24&32&40&27&36&45\\36&48&60&42&56&70&48&64&80&54&72&90\end{bmatrix}\tag{3.72}$$

$$\boldsymbol{Y}_{(2)}=\boldsymbol{a}(\boldsymbol{c}\odot\boldsymbol{b})^\top=\begin{bmatrix}18&36&21&42&24&48&27&54\\24&48&28&56&32&64&36&72\\30&60&35&70&40&80&45&90\end{bmatrix}\tag{3.73}$$

$$\boldsymbol{Y}_{(3)}=\boldsymbol{c}(\boldsymbol{b}\odot\boldsymbol{a})^\top=\begin{bmatrix}18&36&24&48&30&60\\21&42&28&56&35&70\\24&48&32&64&40&80\\27&54&36&72&45&90\end{bmatrix}\tag{3.74}$$

张量向量化

根据 Khatri-Rao 积定义与张量向量化规则，由向量 $\boldsymbol{x}^{(k)}\in\mathbb{R}^{n_k}, k=1,2,\ldots,d$ 的外积得到的张量 $\boldsymbol{\mathcal{Y}}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$，其自第 k 维度展开得到的矩阵可写作如下形式：

$$\operatorname{vec}(\boldsymbol{\mathcal{Y}})=\boldsymbol{x}^{(d)}\odot\boldsymbol{x}^{(d-1)}\odot\cdots\odot\boldsymbol{x}^{(2)}\odot\boldsymbol{x}^{(1)} \tag{3.75}$$

其中，$\odot$ 表示 Khatri-Rao 积；$\operatorname{vec}(\cdot)$ 表示向量化操作。

例 25. 给定向量 $\boldsymbol{a}=(1,2)^\top$、$\boldsymbol{b}=(3,4,5)^\top$ 与 $\boldsymbol{c}=(6,7,8,9)^\top$，若 $\boldsymbol{\mathcal{Y}}=\boldsymbol{a}\otimes_{\text{outer}}\boldsymbol{b}\otimes_{\text{outer}}\boldsymbol{c}$，试写出张量 $\boldsymbol{\mathcal{Y}}$ 的向量化形式 $\operatorname{vec}(\boldsymbol{\mathcal{Y}})$。

解. 根据 *Khatri-Rao* 积定义，有

$$\boldsymbol{c}\odot\boldsymbol{b}=(18,24,30,21,28,35,24,32,40,27,36,45)^\top \tag{3.76}$$

从而，可得到

$$\begin{aligned}\operatorname{vec}(\boldsymbol{\mathcal{Y}})=\boldsymbol{c}\odot\boldsymbol{b}\odot\boldsymbol{a}=(&18,24,30,21,28,35,24,32,40,27,36,45,\\&36,48,60,42,56,70,48,64,80,54,72,90)^\top\end{aligned} \tag{3.77}$$

3.5 CP 张量分解

3.5.1 CP 分解形式

给定任意张量 $\boldsymbol{\mathcal{Y}}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$，若令其秩为 R，则 CP 分解可写作如下形式：

$$\boldsymbol{\mathcal{X}}=\sum_{r=1}^{R}\boldsymbol{u}_r^{(1)}\otimes_{\text{outer}}\boldsymbol{u}_r^{(2)}\otimes_{\text{outer}}\cdots\otimes_{\text{outer}}\boldsymbol{u}_r^{(d)} \tag{3.78}$$

其中，因子矩阵为

$$\boldsymbol{U}^{(k)}=\begin{bmatrix}| & | & & | \\ \boldsymbol{u}_1^{(k)} & \boldsymbol{u}_2^{(k)} & \cdots & \boldsymbol{u}_R^{(k)} \\ | & | & & |\end{bmatrix}\in\mathbb{R}^{n_k\times R},\ k=1,2,\ldots,d \tag{3.79}$$

第四章　时序矩阵分解

第五章　贝叶斯张量分解

第六章　低秩张量填充

第七章　低秩拉普拉斯卷积模型

第八章　基于延迟嵌套的张量分解

参考文献

Tamara G Kolda and Brett W Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.

Nicholas D Sidiropoulos, Lieven De Lathauwer, Xiao Fu, Kejun Huang, Evangelos E Papalexakis, and Christos Faloutsos. Tensor decomposition for signal processing and machine learning. *IEEE Transactions on Signal Processing*, 65(13):3551–3582, 2017.