Kronecker积专题讲义

矩阵理论中的强大工具及其应用

专题讲义

2025 年 5 月 25 日

摘要

Kronecker积是矩阵理论中一个极其重要但经常被低估的工具。它不仅为处理矩阵方程提供了强有力的代数框架，更在统计学、信号处理、控制理论等领域有着广泛应用。本讲义从基础定义出发，系统地介绍Kronecker积的性质、与向量化操作的关系、特征值理论、伪逆性质，以及在求解矩阵方程中的应用。特别关注其在矩阵最小二乘问题中的核心作用。

目录

1 引言

Kronecker积（也称为张量积或直积）是由德国数学家Leopold Kronecker在19世纪提出的一种矩阵运算。虽然其定义相对简单，但它在现代矩阵理论中扮演着核心角色，特别是在处理矩阵方程、多维数据分析、以及量子力学等领域。

Kronecker积的威力在于它能够将复杂的矩阵运算转化为向量运算，从而利用成熟的线性代数理论。这种转化不仅在理论分析中有重要意义，在数值计算中也提供了新的算法思路。

2 基本定义与直观理解

2.1 Kronecker积的定义

Kronecker积

设$A \in \mathbb{R}^{m\times n}$，$B \in \mathbb{R}^{p\times q}$，则$A$和$B$的Kronecker积$A \otimes B$定义为$mp \times nq$矩阵：

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

其中a_{ij}是矩阵A的(i,j)元素。

2.2 简单例子

例 2.1 (2×2矩阵的Kronecker积). 设$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$，$B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$，则：

$$A \otimes B = \begin{bmatrix} 1 \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} & 2 \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \\ 3 \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} & 4 \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{bmatrix}$$

2.3 几何直观

Kronecker积可以理解为一种”缩放和复制”操作：

- 矩阵A的每个元素a_{ij}都”扩展”为一个子块$a_{ij}B$
- 矩阵B在每个位置都被A的对应元素缩放
- 结果矩阵保持了两个原矩阵的结构信息

3 基本性质

3.1 代数性质

定理 3.1 (Kronecker积的基本代数性质)**.** 设A, A_1, A_2是适当维数的矩阵，B, B_1, B_2也是适当维数的矩阵，c是标量，则：

(i) **双线性性：**

$$(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B \tag{1}$$

$$A \otimes (B_1 + B_2) = A \otimes B_1 + A \otimes B_2 \tag{2}$$

(ii) **标量分配律：**

$$c(A \otimes B) = (cA) \otimes B = A \otimes (cB)$$

(iii) **结合律：**

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$$

（当矩阵乘法有定义时）

(iv) **转置性质：**

$$(A \otimes B)^T = A^T \otimes B^T$$

(v) **一般不满足交换律：**

$$A \otimes B \neq B \otimes A$$

（除非特殊情况）

性质(iii)的证明. 设$A_1 \in \mathbb{R}^{m_1 \times n_1}$，$A_2 \in \mathbb{R}^{n_1 \times p_1}$，$B_1 \in \mathbb{R}^{m_2 \times n_2}$，$B_2 \in \mathbb{R}^{n_2 \times p_2}$。

$(A_1 \otimes B_1)(A_2 \otimes B_2)$的$(i, j)$块为：

$$\sum_{k=1}^{n_1} (A_1)_{ik} B_1 \cdot (A_2)_{kj} B_2 \tag{3}$$

$$= \sum_{k=1}^{n_1} (A_1)_{ik} (A_2)_{kj} (B_1 B_2) \tag{4}$$

$$= \left(\sum_{k=1}^{n_1} (A_1)_{ik} (A_2)_{kj} \right) (B_1 B_2) \tag{5}$$

$$= (A_1 A_2)_{ij} (B_1 B_2) \tag{6}$$

这正是$(A_1 A_2) \otimes (B_1 B_2)$的$(i, j)$块。 □

4 向量化操作与Kronecker积

4.1 向量化操作

向量化操作

对于矩阵$X \in \mathbb{R}^{m\times n}$，向量化操作$\text{vec}(X)$将矩阵$X$的列依次堆叠成一个$mn \times 1$的列向量：

$$\text{vec}(X) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

其中x_j是X的第j列。

4.2 核心定理：矩阵乘积的向量化

这是Kronecker积理论中最重要的定理之一：

矩阵乘积的向量化公式

对于矩阵$A \in \mathbb{R}^{m\times p}$，$X \in \mathbb{R}^{p\times q}$，$B \in \mathbb{R}^{q\times n}$，有：

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$$

特别地：

$$\text{vec}(AX) = (I \otimes A)\text{vec}(X) \tag{7}$$
$$\text{vec}(XB) = (B^T \otimes I)\text{vec}(X) \tag{8}$$

详细证明. 我们逐步构建这个重要公式的证明。

步骤1：分析矩阵乘积的结构

设$AXB = Y$，其中$Y \in \mathbb{R}^{m\times n}$。矩阵Y的第$j$列为：

$$y_j = AXb_j$$

其中b_j是B的第j列。

步骤2：向量化矩阵Y

$$\text{vec}(Y) = \text{vec}(AXB) = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} AXb_1 \\ AXb_2 \\ \vdots \\ AXb_n \end{bmatrix}$$

步骤3：表示AXb_j

注意到AXb_j是矩阵AX右乘向量b_j的结果。我们有：

$$AXb_j = A(Xb_j) = A\sum_{k=1}^{q} x_k (b_j)_k = \sum_{k=1}^{q} (b_j)_k A x_k$$

其中x_k是X的第k列，$(b_j)_k$是b_j的第k个分量。

步骤4：重新排列

$$AXb_j = \sum_{k=1}^{q} (b_j)_k A x_k = [A, A, \dots, A] \begin{bmatrix} (b_j)_1 x_1 \\ (b_j)_2 x_2 \\ \vdots \\ (b_j)_q x_q \end{bmatrix}$$

这可以写成：

$$AXb_j = (b_j^T \otimes A)\text{vec}(X)$$

步骤5：组装最终结果

$$\text{vec}(AXB) = \begin{bmatrix} (b_1^T \otimes A)\text{vec}(X) \\ (b_2^T \otimes A)\text{vec}(X) \\ \vdots \\ (b_n^T \otimes A)\text{vec}(X) \end{bmatrix} = \begin{bmatrix} b_1^T \otimes A \\ b_2^T \otimes A \\ \vdots \\ b_n^T \otimes A \end{bmatrix} \text{vec}(X)$$

注意到：

$$\begin{bmatrix} b_1^T \otimes A \\ b_2^T \otimes A \\ \vdots \\ b_n^T \otimes A \end{bmatrix} = B^T \otimes A$$

因此：

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$$

□

4.3 公式的直观理解

向量化公式的几何意义

公式$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$的几何意义是：

- 左乘矩阵A对应Kronecker积的右因子
- 右乘矩阵B对应Kronecker积的左因子（转置后）
- 这种"交叉"结构反映了矩阵乘法的非交换性
- $B^T \otimes A$将向量空间$\mathbb{R}^{pq}$映射到$\mathbb{R}^{mn}$

5 特征值和奇异值理论

5.1 特征值的Kronecker积性质

定理 5.1 (Kronecker积的特征值). 设$A \in \mathbb{R}^{m\times m}$的特征值为$\lambda_1, \lambda_2, \ldots, \lambda_m$，$B \in \mathbb{R}^{n\times n}$的特征值为$\mu_1, \mu_2, \ldots, \mu_n$，则$A \otimes B$的特征值为：

$$\{\lambda_i \mu_j : i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n\}$$

证明. 设$Av_i = \lambda_i v_i$，$Bu_j = \mu_j u_j$，其中v_i, u_j分别是对应的特征向量。

考虑向量$v_i \otimes u_j$（定义为$(v_i \otimes u_j) = \text{vec}(u_j v_i^T)$）：

$$\begin{aligned}
(A \otimes B)(v_i \otimes u_j) &= (A \otimes B)\text{vec}(u_j v_i^T) && (9)\\
&= \text{vec}(A \cdot u_j v_i^T \cdot B^T) \quad \text{（使用向量化公式）} && (10)\\
&= \text{vec}(A u_j v_i^T B^T) && (11)\\
&= \text{vec}(\lambda_i u_j v_i^T B^T) \quad \text{（因为}Av_i = \lambda_i v_i\text{，而}u_j v_i^T\text{的每列都是}v_i\text{的倍数）} && (12)\\
&= \lambda_i \text{vec}(u_j v_i^T B^T) && (13)\\
&= \lambda_i \text{vec}(u_j (Bv_i)^T) && (14)\\
&= \lambda_i \text{vec}(u_j (\mu_j v_i)^T) && (15)\\
&= \lambda_i \mu_j \text{vec}(u_j v_i^T) && (16)\\
&= \lambda_i \mu_j (v_i \otimes u_j) && (17)
\end{aligned}$$

因此$v_i \otimes u_j$是$A \otimes B$的特征向量，对应特征值$\lambda_i \mu_j$。 □

5.2 奇异值的性质

定理 5.2 (Kronecker积的奇异值). 设$A \in \mathbb{R}^{m\times p}$的奇异值为$\sigma_1^{(A)}, \ldots, \sigma_{\min(m,p)}^{(A)}$，$B \in \mathbb{R}^{n\times q}$的奇异值为$\sigma_1^{(B)}, \ldots, \sigma_{\min(n,q)}^{(B)}$，则$A \otimes B$的奇异值为：

$$\{\sigma_i^{(A)} \sigma_j^{(B)} : \text{所有有效的 } i, j\}$$

6 伪逆的Kronecker积性质

这是回答你最初问题的核心部分。

6.1 Moore-Penrose伪逆的Kronecker积性质

Kronecker积的伪逆公式

对于矩阵$A \in \mathbb{R}^{m\times p}$和$B \in \mathbb{R}^{n\times q}$，有：

$$(A \otimes B)^+ = A^+ \otimes B^+$$

详细证明. 我们需要验证$A^+ \otimes B^+$满足Moore-Penrose伪逆的四个条件。

设$P = A^+ \otimes B^+$，我们需要证明：

1. $(A \otimes B)P(A \otimes B) = A \otimes B$
2. $P(A \otimes B)P = P$
3. $((A \otimes B)P)^T = (A \otimes B)P$
4. $(P(A \otimes B))^T = P(A \otimes B)$

验证条件1：

$$\begin{aligned}
(A \otimes B)(A^+ \otimes B^+)(A \otimes B) &= (A \otimes B)((A^+ \otimes B^+)(A \otimes B)) && (18)\\
&= (A \otimes B)(AA^+ \otimes BB^+) \quad \text{（结合律）} && (19)\\
&= (A(AA^+)) \otimes (B(BB^+)) && (20)\\
&= A \otimes B \quad \text{（因为}AA^+A = A\text{，}BB^+B = B\text{）} && (21)
\end{aligned}$$

验证条件2：

$$\begin{aligned}
(A^+ \otimes B^+)(A \otimes B)(A^+ \otimes B^+) &= (A^+A \otimes B^+B)(A^+ \otimes B^+) && (22)\\
&= (A^+AA^+) \otimes (B^+BB^+) && (23)\\
&= A^+ \otimes B^+ \quad \text{（因为}A^+AA^+ = A^+\text{，}B^+BB^+ = B^+\text{）} && (24)
\end{aligned}$$

验证条件3：

$$\begin{aligned}
((A \otimes B)(A^+ \otimes B^+))^T &= (AA^+ \otimes BB^+)^T && (25)\\
&= (AA^+)^T \otimes (BB^+)^T && (26)\\
&= AA^+ \otimes BB^+ \quad \text{（因为}AA^+\text{和}BB^+\text{都是对称的）} && (27)\\
&= (A \otimes B)(A^+ \otimes B^+) && (28)
\end{aligned}$$

验证条件4： 类似地可以验证$(P(A \otimes B))^T = P(A \otimes B)$。

因此$(A \otimes B)^+ = A^+ \otimes B^+$。 □

6.2 为什么这个性质成立？

伪逆Kronecker积性质的深层原因

这个性质成立的深层原因在于：

1. **结构保持**：Kronecker积保持了原矩阵的几何结构
2. **子空间分解**：$(A \otimes B)$的四个基本子空间可以表示为原矩阵对应子空间的Kronecker积
3. **正交性保持**：如果A和B的SVD分解中包含正交矩阵，Kronecker积仍保持正交性

4. **最优化性质：**伪逆的最小范数性质在Kronecker积下得到保持

7 矩阵方程的求解

7.1 一般线性矩阵方程

考虑矩阵方程：

$$AXB = C$$

其中$A \in \mathbb{R}^{m\times p}$，$B \in \mathbb{R}^{q\times n}$，$C \in \mathbb{R}^{m\times n}$已知，$X \in \mathbb{R}^{p\times q}$待求。

定理 7.1 (矩阵方程的Kronecker积解法)**.** 矩阵方程$AXB = C$等价于线性方程组：

$$(B^T \otimes A)\text{vec}(X) = \text{vec}(C)$$

因此解为：

$$\text{vec}(X) = (B^T \otimes A)^+\text{vec}(C) = (B^{+T} \otimes A^+)\text{vec}(C)$$

即：

$$X = \text{unvec}((B^{+T} \otimes A^+)\text{vec}(C)) = A^+C(B^T)^+ = A^+CB^{+T}$$

7.2 在最小二乘中的应用

这正是你在原始问题中遇到的情况！

矩阵最小二乘问题的Kronecker积解法

考虑矩阵最小二乘问题：

$$\min_X \|AXB - C\|_F^2$$

步骤1：向量化

$$\min_{\text{vec}(X)} \|\text{vec}(AXB) - \text{vec}(C)\|_2^2$$

步骤2：应用向量化公式

$$\min_{\text{vec}(X)} \|(B^T \otimes A)\text{vec}(X) - \text{vec}(C)\|_2^2$$

步骤3：标准最小二乘解

$$\text{vec}(X^*) = ((B^T \otimes A)^T(B^T \otimes A))^{-1}(B^T \otimes A)^T\text{vec}(C)$$

步骤4：应用Kronecker积性质

$$(B^T \otimes A)^T = (B^T)^T \otimes A^T = B \otimes A^T \tag{29}$$

$$(B^T \otimes A)^T(B^T \otimes A) = (B \otimes A^T)(B^T \otimes A) = BB^T \otimes A^TA \tag{30}$$

步骤5：应用伪逆性质

$$\text{vec}(X^*) = (B^T \otimes A)^+\text{vec}(C) = (B^{+T} \otimes A^+)\text{vec}(C)$$

步骤6：反向量化

$$X^* = A^+ C B^{+T}$$

8 高级性质和应用

8.1 块矩阵的Kronecker积

定理 8.1 (块矩阵的Kronecker积). 设$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，$B$为任意矩阵，则：

$$A \otimes B = \begin{bmatrix} A_{11} \otimes B & A_{12} \otimes B \\ A_{21} \otimes B & A_{22} \otimes B \end{bmatrix}$$

8.2 条件数的性质

定理 8.2 (Kronecker积的条件数).

$$\kappa(A \otimes B) = \kappa(A) \cdot \kappa(B)$$

其中$\kappa(\cdot)$表示条件数。

这个性质在数值计算中很重要，因为它告诉我们Kronecker积可能会显著放大数值误差。

9 数值计算考虑

9.1 存储和计算复杂性

注 9.1 (计算复杂性).

- 存储$A \otimes B$需要$O(mnpq)$空间（$A \in \mathbb{R}^{m\times n}$，$B \in \mathbb{R}^{p\times q}$）
- 直接计算$(A \otimes B)x$需要$O(mnpq)$时间
- 利用结构可以将复杂度降低到$O(mn + pq)$

因此，实际计算中应避免显式构造Kronecker积矩阵。

9.2 高效算法

避免显式构造的计算方法

计算$(B^T \otimes A)\text{vec}(X)$时：

方法1：直接方法

$$y = (B^T \otimes A)\text{vec}(X) = \text{vec}(AXB)$$

计算AXB然后向量化，复杂度$O(mpq + pqn)$。

方法2：逐步计算

1. 计算$Y_1 = AX$，复杂度$O(mpq)$
2. 计算$Y_2 = Y_1 B$，复杂度$O(mqn)$

3. $y = \text{vec}(Y_2)$，复杂度$O(mn)$

总复杂度：$O(mpq + mqn)$，远小于$O(mnpq)$。

10 在统计学中的应用

10.1 多元线性模型

在多元线性回归中，我们有：

$$Y = XB + E$$

其中$Y \in \mathbb{R}^{n\times p}$是响应矩阵，$X \in \mathbb{R}^{n\times q}$是设计矩阵，$B \in \mathbb{R}^{q\times p}$是系数矩阵。
使用Kronecker积，可以将这个矩阵回归问题转化为向量回归：

$$\text{vec}(Y) = (I_p \otimes X)\text{vec}(B) + \text{vec}(E)$$

10.2 协方差结构建模

在具有Kronecker积协方差结构的模型中：

$$\text{Cov}(\text{vec}(Y)) = \Sigma_2 \otimes \Sigma_1$$

这种结构在空间-时间数据分析中很常见，其中Σ_1和Σ_2分别描述不同维度的相关性。

11 总结与深入理解

11.1 Kronecker积的本质

Kronecker积的数学本质

Kronecker积是：

1. **张量积的矩阵实现**：它实际上是两个线性变换的张量积
2. **直积空间的线性算子**：$A \otimes B : \mathbb{R}^n \otimes \mathbb{R}^q \to \mathbb{R}^m \otimes \mathbb{R}^p$
3. **双线性映射的线性化**：将双线性问题转化为线性问题
4. **结构保持的运算**：保持原矩阵的代数和几何性质

11.2 为什么Kronecker积如此重要？

- **统一框架**：为处理矩阵方程提供统一的代数框架
- **降维工具**：将高维矩阵问题转化为向量问题
- **结构利用**：充分利用问题的结构特性
- **理论桥梁**：连接线性代数、多重线性代数和张量分析

11.3 学习建议

1. **从定义开始**：熟练掌握基本定义和简单计算
2. **理解向量化**：深入理解向量化公式及其证明
3. **掌握性质**：特别是伪逆、特征值等重要性质
4. **实践应用**：在具体问题中应用这些理论
5. **数值意识**：了解计算复杂性和数值稳定性问题

12 习题与进一步阅读

12.1 理论练习

1. 证明$(A \otimes B)(C \otimes D) = AC \otimes BD$当矩阵乘法有定义时。
2. 设A和B都是正定矩阵，证明$A \otimes B$也是正定的。
3. 证明$\det(A \otimes B) = (\det A)^n (\det B)^m$，其中$A \in \mathbb{R}^{m\times m}$，$B \in \mathbb{R}^{n\times n}$。
4. 求解矩阵方程$AXB + CXD = E$，其中所有矩阵维数适配。

12.2 数值实验

1. 比较直接构造$A \otimes B$与利用结构计算$(A \otimes B)x$的效率。
2. 研究Kronecker积矩阵的条件数如何影响线性方程组的求解精度。
3. 实现矩阵最小二乘问题的Kronecker积算法，并与传统方法比较。

通过这个系统的学习，你应该能够深入理解Kronecker积在矩阵理论中的核心地位，特别是它在解决矩阵方程和最小二乘问题中的重要作用。