第九章　　Kronecker积与线性矩阵方程

- 本章讨论含有未知矩阵的线性矩阵代数方程。
  - 9.1 矩阵的Kronecker积
  - 9.2 矩阵的拉直与线性矩阵方程
  - 9.3 线性方程$AXB = C$与矩阵最佳逼近问题
  - 9.4* 矩阵方程$AX = B$的Hermite解与矩阵最佳逼近问题
  - 9.5* 矩阵方程$AX + XB = C$和$X - AXB = C$

9.1 矩阵的Kronecker积

- **定义9.1.1** 设$A=(a_{ij})\in C^{m\times n}$,$B=(b_{ij})\in C^{p\times q}$，则称如下分块矩阵

$$A\otimes B=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}\in C^{mp\times nq}$$

为A与B的Kronecker积（**直积，张量积**），简记为$A\otimes B=(a_{ij}B)$.

- 显然$A\otimes B$与$B\otimes A$是同阶矩阵，但一般说来，$A\otimes B\neq B\otimes A$，即矩阵的Kronecker积不满足交换律。

- 对单位矩阵，有

$$I_m\otimes I_n=I_n\otimes I_m=I_{mn}$$

- **定理9.1.1** 矩阵的Kronecker积具有下列基本性质:

(1) 对任意复数k，$(kA)\otimes B = A\otimes(kB) = k(A\otimes B)$;

(2) $(A\otimes B)\otimes C = A\otimes(B\otimes C)$;

(3) $A\otimes(B+C) = A\otimes B + A\otimes C$,$(B+C)\otimes A = B\otimes A + C\otimes A$;

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

(5) $(A\otimes B)^H = A^H\otimes B^H$;

(6) 若$A\in C^{n\times m}$，$B\in C^{p\times s}$，$C\in C^{m\times n}D\in C^{s\times q}$，则$(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$;

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

(8) 如果A和B都是对角矩阵、上（下）三角矩阵、实对称矩阵、Hermite矩阵、正交矩阵、酉矩阵，则$A\otimes B$也分别是这种类型的矩阵。

- **证明**
  - (1)~(5)由定义9.1.1即可证明.
  - 下面证明(6)

$$\begin{aligned}(A\otimes B)(C\otimes D) &= \begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{k1}B & \cdots & a_{km}B \end{pmatrix}\begin{pmatrix} c_{11}D & \cdots & c_{1n}D \\ \vdots & & \vdots \\ c_{m1}D & \cdots & c_{mm}D \end{pmatrix} \\ &= \left(\sum_{l=1}^{m} a_{il}c_{lj}BD\right) = \left(\left(\sum_{l=1}^{m} a_{il}c_{lj}\right)BD\right) = (AC)\otimes(BD)\end{aligned}$$

下面证明(7)

$$\begin{aligned}(A\otimes B)(A^+\otimes B^+)(A\otimes B) &= (AA^+)\otimes(BB^+)(A\otimes B)\\ &= (AA^+A)\otimes(BB^+B)=A\otimes B,\end{aligned}$$

同理可证$(A^+\otimes B^+)(A\otimes B)(A^+\otimes B^+)=A^+\otimes B^+$

$$\begin{aligned}[(A\otimes B)(A^+\otimes B^+)]^H &= [(AA^+)\otimes(BB^+)]^H=(AA^+)^H\otimes(BB^+)^H\\ &= (AA^+)\otimes(BB^+)=(A\otimes B)(A^+\otimes B^+)\end{aligned}$$

同理可证$[(A^+\otimes B^+)(A\otimes B)]^H=(A^+\otimes B^+)(A\otimes B)$

这说明$(A\otimes B)^+=A^+\otimes B^+$

- **定理9.1.2** 设$f(x,y) = \sum_{i,j=0}^{K} c_{ij}x^i y^j$是变量$x$,$y$的复变量二元多项式，对$A\in C^{m\times m}$，$B\in C^{n\times n}$，定义$mn$阶矩阵$f(A,B)=\sum_{i,j=0}^{K} c_{ij}A^i\otimes B^j$，其中$A^0 = I_m$，$B^0 = I_n$。如果$A$和$B$的特征值分别为$\lambda_1$，$\lambda_2$，$\cdots$，$\lambda_m$和$\mu_1$，$\mu_2$，$\cdots$，$\mu_n$，则$f(A,B)$的特征值为$f(\lambda_i,\mu_j)(i=1,2,\cdots,m,j=1,2,\cdots,n)$.

- **证明** 由Schur定理知，存在酉矩阵P,Q使得

$$P^HAP=\begin{pmatrix}\lambda_1 & & & *\\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\equiv A_1, Q^HBQ=\begin{pmatrix}\mu_1 & & & *\\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n\end{pmatrix}\equiv B_1$$

其中A_1,B_1均为上三角矩阵，由定理9.1.1（8）知，$P\otimes Q$是酉矩阵，并且$A_1^i\otimes B_1^i$是上三角矩阵，则

$$
\begin{aligned}
(P\otimes Q)^{-1}f(A,B)(P\otimes Q) &= (P\otimes Q)^{H}f(A,B)(P\otimes Q)\\
&= \sum_{i,j=0}^{K} c_{ij}(P\otimes Q)^{H}(A^{i}\otimes B^{j})(P\otimes Q)\\
&= \sum_{i,j=0}^{K} c_{ij}(P^{H}\otimes Q^{H})(A^{i}\otimes B^{j})(P\otimes Q)\\
&= \sum_{i,j=0}^{K} c_{ij}(P^{H}A^{i}P)\otimes(Q^{H}B^{j}Q) = \sum_{i,j=0}^{K} c_{i,j}A_1^{i}\otimes B_1^{j} = f(A_1,B_1)
\end{aligned}
$$

也是上三角矩阵。因为

$$
A_1^{i}\otimes B_1^{j} = \begin{pmatrix} \lambda_1^{i}B_1^{j} & & * \\ & \ddots & \\ & & \lambda_m^{i}B_1^{j} \end{pmatrix}, \lambda_l^{i}B_1^{j} = \begin{pmatrix} \lambda_l^{i}\mu_1^{j} & & * \\ & \ddots & \\ & & \lambda_l^{i}\mu_n^{j} \end{pmatrix}
$$

则$f(A_1,B_1)$的对角元，即$f(A,B)$的特征值为$f(\lambda_i,\mu_j)(i=1,2,\cdots,m,j=1,2,\cdots,n)$.

由定理9.1.2可得如下结论

- **定理9.1.3** 设A, B分别为$m\times m$和$n\times n$矩阵，并且其特征值分别为$\lambda_1, \lambda_2, \cdots, \lambda_m$和$\mu_1, \mu_2, \cdots, \mu_n$，则

(1) $A\otimes B$的mn个特征值为$\lambda_i\mu_j(i=1,2,\cdots,m; j=1,2,\cdots,n)$；

(2) $A\otimes I_n + I_m\otimes B$的$mn$个特征值为$\lambda_i+\mu_j(i=1,2,\cdots,m; j=1,2,\cdots,n)$；

(3) $\det(A\otimes B) = (\det(A))^n(\det(B))^m$；

(4) $tr(A\otimes B) = tr(A)tr(B)$；

(5) 若A, B均为非奇异矩阵，则$A\otimes B$是非奇异矩阵，并且

$$(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$$

- **证明**　由定理9.1.2，即得（1）和（2），并且

$$
\begin{aligned}
|A \otimes B| &= \prod_{i=1}^{m}\left(\prod_{j=1}^{n}\lambda_i\mu_j\right) = \prod_{i=1}^{m}\left(\lambda_i^n\prod_{j=1}^{n}\mu_j\right) \\
&= \left(\prod_{i=1}^{m}\lambda_i^n\right)\left(\prod_{j=1}^{n}\mu_j\right)^m = |A|^n|B|^m
\end{aligned}
$$

$$
tr(A \otimes B) = \sum_{i=1}^{m}\sum_{j=1}^{n}\lambda_i\mu_j = \left(\sum_{i=1}^{m}\lambda_i\right)\left(\sum_{j=1}^{n}\mu_j\right) = tr(A)tr(B)
$$

- **定理9.1.4** 设$A \in C^{m\times n}$，$B \in C^{p\times q}$，则

$$\operatorname{rank}(A \otimes B) = \operatorname{rank}(A)\operatorname{rank}(B)$$

- **证明** 由定理2.1.10可知，存在非奇异矩阵$P \in C^{m\times m}$，$Q \in C^{n\times n}$，$S \in C^{p\times p}$，$T \in C^{q\times q}$使得

$$PAQ = \begin{pmatrix} I_{r_A} & 0 \\ 0 & 0 \end{pmatrix} = A_1, SBT = \begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix} = B_1,$$

其中$r_A = \operatorname{rank}(A), r_B = \operatorname{rank}(B)$。由定理9.1.1有

$$A \otimes B = \left(P^{-1}A_1Q^{-1}\right) \otimes \left(S^{-1}B_1T^{-1}\right) = \left(P^{-1} \otimes S^{-1}\right)(A_1 \otimes B_1)\left(Q^{-1} \otimes T^{-1}\right)$$

由定理9.1.3知，$P^{-1} \otimes S^{-1}$，$Q^{-1} \otimes T^{-1}$均为非奇异矩阵，则

$$\operatorname{rank}(A \otimes B) = \operatorname{rank}(A_1 \otimes B_1)$$

而$\operatorname{rank}(A_1 \otimes B_1) = r_A r_B = \operatorname{rank}(A)\operatorname{rank}(B)$，于是$\operatorname{rank}(A \otimes B) = \operatorname{rank}(A)\operatorname{rank}(B)$

- **定理9.1.5** 设A,B分别为$m\times m$和$n\times n$矩阵，则存在一个mn阶排列矩阵P使得

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

- 对Kronecker积也有幂的概念，记

$$A^{[k]}=\underbrace{A\otimes A\otimes\cdots\otimes A}_{k}$$

- 关于Kronecker积的幂，有如下结果：
- **定理9.1.6** 设$A\in C^{m\times n}$，$B\in C^{p\times q}$，则

$$(AB)^{[k]}=A^{[k]}B^{[k]}$$

- **证明** 对k作数学归纳法。当$k=1$时，结论显然成立。设对$k-1$结论成立，则由定理9.1.1有

$$\begin{aligned}(AB)^{[k]} &= (AB)\otimes(AB)^{[k-1]}=(AB)\otimes(A^{[k-1]}B^{[k-1]})\\ &= (A\otimes A^{[k-1]})(B\otimes B^{[k-1]})=A^{[k]}B^{[k]}\end{aligned}$$

9.2 矩阵的拉直与线性矩阵方程

9.2.1 矩阵的拉直

- 定义9.2.1 设$A = (a_{ij}) \in C^{m\times n}$，记$a_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} \quad (i = 1, 2, \cdots, n)$。令

$$\mathrm{vec}(A) = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$$

则称vec(A)为矩阵A的**列拉直(列展开)**。

- 类似的，可以考虑矩阵的行拉直（行展开）。

- **定理9.2.1** 设$A \in C^{m\times n}$，$B \in C^{n\times p}$，$C \in C^{p\times q}$，则

$$\mathrm{vec}(ABC) = (C^T \otimes A)\mathrm{vec}(B)$$

- **证明** 记$B = (b_1, \cdots, b_p), b_i \in C^n (i = 1, \cdots, p), C = (c_1, \cdots, c_q), c_j \in C^p (j = 1, \cdots, q)$，则

$$\mathrm{vec}(ABC) = \mathrm{vec}(ABc_1, ABc_2, \cdots, ABc_q) = \begin{pmatrix} ABc_1 \\ ABc_2 \\ \vdots \\ ABc_q \end{pmatrix}$$

而

$$ABc_i = c_{1i}Ab_1 + c_{2i}Ab_2 + \cdots + c_{pi}Ab_p = (c_{1i}A, c_{2i}A, \cdots, c_{pi}A)\mathrm{vec}(B)$$

故

$$\mathrm{vec}(ABC) = \begin{pmatrix} c_{11}A & c_{21}A & \cdots & c_{p1}A \\ c_{12}A & c_{22}A & \cdots & c_{p2}A \\ \vdots & \vdots & & \vdots \\ c_{1q}A & c_{2q}A & \cdots & c_{pq}A \end{pmatrix} \mathrm{vec}(B) = (C^T \otimes A)\mathrm{vec}(B)$$

由定理9.2.1直接得如下结论

- 推论9.2.1　设$A \in C^{m\times m}$，$B \in C^{n\times n}$，$X \in C^{m\times n}$，则

(1)　$\mathrm{vec}(AX) = (I_n \otimes A)\mathrm{vec}(X)$;

(2)　$\mathrm{vec}(XB) = (B^T \otimes I_m)\mathrm{vec}(X)$;

(3)　$\mathrm{vec}(AX + XB) = (I_n \otimes A + B^T \otimes I_m)\mathrm{vec}(X)$

9.2.2 线性矩阵方程

- 一般的线性矩阵方程可以表示为

$$A_1XB_1 + A_2XB_2 + \cdots + A_pXB_p = C \tag{1}$$

其中$A_i \in C^{m\times m}, B_i \in C^{n\times n}(i = 1, 2, \cdots, p), C \in C^{m\times n}$是已知矩阵，而$X \in C^{m\times n}$是未知矩阵。利用矩阵的kronecker积和拉直，可以给出上述线性方程的可解性及其算法。

- **定理9.2.2** 矩阵$X \in C^{m\times n}$是上述矩阵方程的解的充分必要条件为$x = \mathrm{vec}(X)$是如下线性方程组的解

$$Gx = \mathrm{vec}(C) \tag{2}$$

其中$G = \sum_{i=1}^{p} B_i^T \otimes A_i$.

- **证明** 对矩阵方程(1)两边拉直，并利用定理9.2.1，有

$$\begin{aligned}\mathrm{vec}(C) &= \mathrm{vec}\left(\sum_{i=1}^{p} A_iXB_i\right) = \sum_{i=1}^{p} \mathrm{vec}(A_iXB_i) \\ &= \sum_{i=1}^{p} (B_i^T \otimes A_i)\mathrm{vec}(X) = G\mathrm{vec}(X)\end{aligned}$$

因此矩阵方程(1)的解与线性方程组(2)的解相同，故定理得证。

- 由定理9.2.2和线性方程组的可解性理论直接得如下推论.
- **推论9.2.2** 矩阵方程(1)有解的充分必要条件是rank$[G, \mathrm{vec}(C)] = \mathrm{rank}(G)$；矩阵方程组(1)有唯一解的充分必要条件是$G$非奇异。
- 如果矩阵方程(1)有解，对线性方程组(2)应用定理8.2.5，可得解$x = \mathrm{vec}(X)$的通式。如果矩阵方程(1)无解，则可考虑下列最小二乘问题的解

$$\|A_1XB_1 + A_2XB_2 + \cdots + A_pXB_p - C\|_F = \min$$

上式可以写成

$$\|G\mathrm{vex}(X) - \mathrm{vec}(C)\|_2 = \min$$

对此最小二乘问题应用定理8.5.5，可得最小二乘解$x = \mathrm{vec}(X)$的通式。
- 同样可以考虑(1)的极小最小二乘解。

9.3 矩阵方程$AXB=C$与矩阵最佳逼近问题

9.3.1 矩阵方程$AXB=C$

- 考虑矩阵方程
$$AXB=C$$
其中A,B,C分别是$m\times n$，$p\times q$，$m\times q$已知矩阵，X为$n\times p$矩阵。
- 定理8.2.4给出了矩阵方程$AXB=C$有解的条件及有解时通解的表达式。
- 利用Moore-Penrose广义逆，可以给出上述矩阵方程有解的另一个条件。
- 利用定理9.2.1将上述矩阵方程两端列展开，可得线性方程组
$$(B^T\otimes A)\text{vec}(X)=\text{vec}(C)$$
- **定理9.3.1** 设$A\in C^{m\times n}$，$B\in C^{p\times q}$，$C\in C^{m\times q}$，矩阵方程$AXB=C$有解的充分必要条件是
$$AA^+CCB+B=C$$
并且在有解的情况下，其通解为
$$X=A^+CB^++Y-A^+AYBB^+$$

- 证明　由定理9.2.2知矩阵方程$AXB=C$有解的充分必要条件是线性方程组$(B^T\otimes A)\text{vec}(X)=\text{vec}(C)$ 有解,而由定理8.5.4可知,$(B^T\otimes A)\text{vec}(X)=\text{vec}(C)$有解的充分必要条件是

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

由定理9.1.1和定理9.2.1，上式化为

$$\begin{aligned}\text{vec}(C) &= (B^T\otimes A)(B^T\otimes A)^+\text{vec}(C)=(B^T\otimes A)((B^+)^T\otimes A^+)\text{vec}(C)\\ &= [(B^+B)^T\otimes(AA^+)]\text{vec}(C)=\text{vec}(AA^+CB^+B)\end{aligned}$$

故线性方程$AXB=C$有解的充分必要条件是$AA^+CCB+B=C$成立。
若线性方程组$(B^T\otimes A)\text{vec}(X)=\text{vec}(C)$有解，由定理8.5.4知其通解为

$$\begin{aligned}\text{vec}(X) &= (B^T\otimes A)^+\text{vec}(C)+[I-(B^T\otimes A)^+(B^T\otimes A)]\text{vec}(Y)\\ &= ((B^+)^T\otimes A^+)\text{vec}(C)+[I-((B^+)^T\otimes A^+)(B^T\otimes A)]\text{vec}(Y)\\ &= \text{vec}[A^+CB^+]+\text{vec}(Y)-[((B^+)^TB^T)\otimes(A^+A)]\text{vec}(Y)\\ &= \text{vec}[A^+CB^+]+\text{vec}(Y)-\text{vec}[A^+AYBB^+]\end{aligned}$$

因此矩阵方程$AXB=C$的通解为

$$X=A^+CB^++Y-A^+AYBB^+$$

- 如果矩阵方程$AXB = C$无解，则可以考虑如下最小二乘问题的解

$$\|AXB - C\|_F = \min$$

上式可以改写为

$$\|(B^T \otimes A)\mathrm{vec}(X) - \mathrm{vec}(C))\|_2 = \min$$

对上述最小二乘问题应用定理8.5.5，并由

$$\begin{aligned}
\mathrm{vec}(X) &= (B^T \otimes A)^+\mathrm{vec}(C) + [I - (B^T \otimes A)^+(B^T \otimes A)]\mathrm{vec}(Y) \\
&= ((B^+)^T \otimes A^+)\mathrm{vec}(C) + [I - ((B^+)^T \otimes A^+)(B^T \otimes A)]\mathrm{vec}(Y) \\
&= \mathrm{vec}[A^+CB^+] + \mathrm{vec}(Y) - [((B^+)^T B^T) \otimes (A^+A)]\mathrm{vec}(Y) \\
&= \mathrm{vec}[A^+CB^+] + \mathrm{vec}(Y) - \mathrm{vec}[A^+AYBB^+]
\end{aligned}$$

可得最小二乘问题$\|AXB - C\|_F = \min$的通解为

$$X = A^+CB^+ + Y - A^+AYBB^+$$

其中$Y \in C^{n\times p}$是任意矩阵。

- 对最小二乘问题$\|(B^T \otimes A)\mathrm{vec}(X) - \mathrm{vec}(C))\|_2 = \min$应用定理8.5.6，则得矩阵方程$AXB = C$的极小最小二乘解

$$X = A^+CB^+$$

9.3.2 带约束的矩阵最佳逼近问题

- 在所有满足矩阵方程$AXB = C$或最小二乘问题$\|AXB - C\|_F = \min$的矩阵中，选取一个与对X的初步估计$\tilde{X} = (\tilde{x}_{ij}) \in C^{n\times p}$最接近的矩阵$\hat{X}$，作为矩阵$X$的最佳估计。这类问题称为**带约束的矩阵最佳逼近问题**。

- 记
$$S_X = \{X \in C^{n\times p} \mid \|AXB - C\|_F = \min\}$$
则带约束的矩阵最佳逼近问题可表述为：对$\tilde{X} = \tilde{x}_{ij} \in C^{n\times p}$，在$S_X$上求$\tilde{X}$的最佳逼近$\hat{X}$，即
$$\|\tilde{X} - \hat{X}\|_F = \inf_{X\in S_X} \|\tilde{X} - X\|_F \tag{3}$$

- **定理9.3.2** 设$A \in C^{m\times n}$，$B \in C^{p\times q}$，$C \in C^{m\times q}$，则$\tilde{X} = (\tilde{x}_{ij} \in C^{n\times p})$在$S_X$上存在唯一的最佳逼近$\hat{X}$，并且
$$\hat{X} = A^+CB^+ + \tilde{X} - A^+A\tilde{X}BB^+ \tag{4}$$

- 如果$\tilde{X} = 0$，则得矩阵方程$AXB = C$的极小二乘解$X = A^+CB^+$.

Thank you!