Mid-Term Exam Review

夜未央

2026 年 1 月 7 日

1 矩阵的运算

1.1 矩阵的乘法与矩阵的分块

矩阵的加法与数乘的定义是很自然的, 这里不再列出. 我们在本章主要着眼于矩阵的乘法.

定义 1.1 矩阵乘法 设 $\boldsymbol{A} = (a_{ij})_{s\times n}$, $\boldsymbol{B} = (b_{ij})_{n\times m}$, 令

$$\boldsymbol{C} = (c_{ij})_{m\times n}$$

其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

称 $\boldsymbol{C}$ 是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的积, 记作 $\boldsymbol{C} = \boldsymbol{AB}$.

实际上对于上述 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, $\boldsymbol{AB}$ 的 (i,j) 元是 $\boldsymbol{A}$ 的第 i 行行向量 $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{B}$ 的第 j 列列向量 $\boldsymbol{\beta}_j$ 的积. 至少笔者认为如此记忆会带来计算上的好处. 实际上需要根据定义计算的场景几乎只有简单矩阵的运算, 这纯粹是笔头功夫, 在考试时计算这类问题时一定要注意验算.

定理 1.2 基本矩阵的乘法 设 $\boldsymbol{E}_{ij}$ 是 n 级基本矩阵. 用 $\boldsymbol{E}_{ij}$ 左乘矩阵 $\boldsymbol{A}$ 的乘积矩阵是将 $\boldsymbol{A}$ 的第 j 行移到第 i 行, 其余行均为 $\boldsymbol{0}$; 用 $\boldsymbol{E}_{ij}$ 右乘 $\boldsymbol{A}$ 的乘积矩阵是将 $\boldsymbol{A}$ 的第 i 列移到第 j 列, 其余列均为 $\boldsymbol{0}$.

然后我们来考虑矩阵的分块. 前面矩阵的乘法的操作单元是矩阵的元素, 自然地可以想到把子矩阵作为最小的操作单元进行计算. 于是把矩阵乘法定义中的元素换成对应规模的子矩阵, 结果仍然成立. 这就是矩阵分块的想法.

常用的分块操作以及相关的引理如下:

1. 将矩阵运算拆分成矩阵与列向量的运算: 考虑 $s\times n$ 矩阵 $\boldsymbol{A}$ 和 $n\times m$ 矩阵 $\boldsymbol{B}$, 将 $\boldsymbol{B}$ 分块写为列向量的形式:

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_n \end{bmatrix}$$

则有

$$\boldsymbol{AB} = \boldsymbol{A}\begin{bmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{A\beta}_1 & \cdots & \boldsymbol{A\beta}_n\end{bmatrix}$$

在后面考虑特征向量的相关问题时经常用到这样的拆分方法.

2. Schur 公式: 对于矩阵 $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}$, 如果 $\boldsymbol{A}$ 可逆且 $\boldsymbol{D}$ 为方阵, 则可以通过如下行列变换将其变为分块对角矩阵:

$$\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ -\boldsymbol{CA}^{-1} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{D}-\boldsymbol{CA}^{-1}\boldsymbol{B}\end{bmatrix}$$

1.2 矩阵乘积的秩与行列式

定理 1.3 矩阵乘积的秩 对于 $s \times n$ 级矩阵 $\boldsymbol{A}$ 和 $n \times m$ 级矩阵 $\boldsymbol{B}$, 总有

$$\operatorname{rank}\boldsymbol{AB} \leqslant \min\{\operatorname{rank}\boldsymbol{A}, \operatorname{rank}\boldsymbol{B}\}$$

定理 1.4 Sylvester 秩不等式 对于 $s \times n$ 级矩阵 $\boldsymbol{A}$ 和 $n \times m$ 级矩阵 $\boldsymbol{B}$, 总有

$$\operatorname{rank}\boldsymbol{AB} \geqslant \operatorname{rank}\boldsymbol{A} + \operatorname{rank}\boldsymbol{B} - n$$

证明. 实际上只需证

$$\operatorname{rank}\boldsymbol{AB} + \operatorname{rank}\boldsymbol{I} \geqslant \operatorname{rank}\boldsymbol{A} + \operatorname{rank}\boldsymbol{B}$$

注意到 $\operatorname{rank}\boldsymbol{X} + \operatorname{rank}\boldsymbol{Y} = \operatorname{rank}\begin{bmatrix}\boldsymbol{X} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Y}\end{bmatrix}$, 所以只需证明

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{AB} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} \geqslant \operatorname{rank}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{bmatrix}$$

而

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{AB} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} = \operatorname{rank}\begin{bmatrix}\boldsymbol{AB} & \boldsymbol{A}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} = \operatorname{rank}\begin{bmatrix}\boldsymbol{0} & \boldsymbol{A}\\ -\boldsymbol{B} & \boldsymbol{I}\end{bmatrix} = \operatorname{rank}\begin{bmatrix}\boldsymbol{B} & \boldsymbol{I}\\ \boldsymbol{0} & \boldsymbol{A}\end{bmatrix} \geqslant \operatorname{rank}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{bmatrix}$$

于是命题得证. □

定理 1.5 方阵乘积的行列式 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为数域 $\mathbb{K}$ 上的 n 级方阵, 则

$$\det\boldsymbol{AB} = \det\boldsymbol{A}\det\boldsymbol{B}$$

2 特殊矩阵

定理 2.1 可逆矩阵的求法 考虑可逆的 n 级矩阵 $\boldsymbol{A}$, 其逆有两种主要的求法:

(1) 求出 $\boldsymbol{A}$ 的伴随方阵 $\boldsymbol{A}^*$, 然后根据

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \boldsymbol{A}^*$$

求出 $\boldsymbol{A}^{-1}$. 其中 $\boldsymbol{A}^*$ 的 (i,j) 元是 $\boldsymbol{A}$ 的 (j,i) 元的代数余子式. 注意这里元素的转置对应关系, 以及是代数余子式而非余子式.

(2) 对矩阵 $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{I}_n\end{bmatrix}$ 做初等行变换使得左半部分为 $\boldsymbol{I}_n$, 此时右半部分的矩阵即为 $\boldsymbol{A}^{-1}$, 即

$$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{I}_n\end{bmatrix} \xrightarrow{\text{初等行变换}} \begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{A}^{-1}\end{bmatrix}$$

定理 2.2 基本矩阵的可逆性 初等矩阵和有限个初等矩阵的乘积都可逆.

可逆矩阵的主要用途是消去和拆分.

3 矩阵的相似和相抵

定义 3.1 相抵标准形 对于任意 $n \times m$ 级矩阵 $\boldsymbol{A}$, 总存在 n 级可逆矩阵 $\boldsymbol{P}$ 和 m 级可逆矩阵 $\boldsymbol{Q}$ 使得

$$\boldsymbol{A} = \boldsymbol{P} \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}_{n\times m} \boldsymbol{Q}$$

其中 $r = \operatorname{rank} \boldsymbol{A}$, 中间的矩阵 $\boldsymbol{\Sigma}$ 称作 $\boldsymbol{A}$ 的相抵标准型.

由于相抵标准型刻画了矩阵的秩, 因此在与秩相关的问题中用处很大.

3.1 矩阵相似的特殊结论

定理 3.2 任何复方阵都相似于上三角矩阵.

证明. 考虑 n 级复方阵 $\boldsymbol{A}$. 对于其一个特征向量 $\boldsymbol{x}$ 和对应的特征值 λ, 注意到特征方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}\lambda$ 具有相似的形式. 因此考虑对阶数 n 进行归纳.

当 $n = 1$ 时, 命题显然成立.

当 $n \geqslant 2$ 时, 考虑 $\boldsymbol{A}$ 的一个特征值 λ 和对应的特征向量 $\boldsymbol{\alpha}$, 将 $\boldsymbol{\alpha}$ 扩充为 $\mathbb{K}^n$ 的一组基, 并据此构造可逆矩阵 $\boldsymbol{P}$ 使得其第一列为 $\boldsymbol{\alpha}$. 于是有

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P} \begin{bmatrix} \lambda & \boldsymbol{x} \\ \boldsymbol{0} & \boldsymbol{X} \end{bmatrix}$$

这里 $\boldsymbol{X}$ 是 $n-1$ 级复方阵. 根据归纳假设, 存在 $n-1$ 级可逆矩阵 $\boldsymbol{Q}$ 和 $n-1$ 级上三角矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{U}\boldsymbol{Q}^{-1}$$

于是

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}\end{bmatrix}\begin{bmatrix}\lambda & \boldsymbol{x}\boldsymbol{Q}\\ \boldsymbol{0} & \boldsymbol{U}\end{bmatrix}\begin{bmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}^{-1}\end{bmatrix}\boldsymbol{P}^{-1}$$

令

$$\boldsymbol{R}=\begin{bmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}^{-1}\end{bmatrix}\boldsymbol{P}^{-1}$$

则 $\boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^{-1}$ 为上三角矩阵. □

定理 3.3 Caylay-Hamilton 定理 设 n 级矩阵 $\boldsymbol{A}$ 的特征多项式为 f, 证明: $f(\boldsymbol{A})=\boldsymbol{0}$.

证明. 由于特征多项式不会因 $\mathbb{K}$ 的改变而改变, 因此不妨令 $\mathbb{K}=\mathbb{C}$. 我们已经证明 $\boldsymbol{A}$ 相似于某一上三角矩阵 $\boldsymbol{U}$, 不妨设可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{U}\boldsymbol{P}$. 于是

$$f(\boldsymbol{A})=\boldsymbol{P}^{-1}f(\boldsymbol{U})\boldsymbol{P}$$

相似变换不改变特征多项式, 因此 $\boldsymbol{U}$ 的特征多项式也为 f. 由于 $\boldsymbol{U}$ 是上三角矩阵, 因此 $f(\lambda)=\prod\limits_{i=1}^{n}(\lambda-u_{ii})$. 于是只需证明

$$\prod_{i=1}^{n}(\boldsymbol{U}-u_{ii}\boldsymbol{I})=\boldsymbol{0}$$

即可.

对 $\boldsymbol{U}$ 的阶数 n 做归纳. 当 $n=1$ 时显然成立. 当 $n\geqslant 2$ 时有

□

3.2 对角化

定理 3.4 一般矩阵的对角化 对于可对角化的 n 级矩阵 $\boldsymbol{A}$, 求出其 n 个线性无关的特征向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 分别对应于特征值 $\lambda_1,\cdots,\lambda_n$, 然后令

$$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}$$

则有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$$

定理 3.5 实对称矩阵的正交对角化 n 级实对称矩阵 $\boldsymbol{A}$ 一定可对角化. 考虑 $\boldsymbol{A}$ 的所有特征值 $\lambda_1, \cdots, \lambda_m$, 对应于特征值 λ_i 的线性无关的特征向量 $\boldsymbol{\alpha}_{i1}, \cdots, \boldsymbol{\alpha}_{ir_i}$, 将其做 Schmidt 正交化为向量 $\boldsymbol{\eta}_{i1}, \cdots, \boldsymbol{\eta}_{ir_i}$, 然后令

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{\eta}_{11} & \cdots & \boldsymbol{\eta}_{1r_1} & \cdots & \boldsymbol{\eta}_{m1} & \cdots & \boldsymbol{\eta}_{mr_m} \end{bmatrix}$$

则 $\boldsymbol{T}$ 为正交矩阵, 并且有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \operatorname{diag}\{\lambda_1 \boldsymbol{I}_{r_1}, \cdots, \lambda_m \boldsymbol{I}_{r_m}\}$$

3.3 正交方阵

定理 3.6 Schmidt 正交化 对向量组 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$, 依次定义

$$\boldsymbol{\beta}_i = \boldsymbol{\alpha}_i - \sum_{j=1}^{i-1} \frac{\langle \boldsymbol{\beta}_j, \boldsymbol{\alpha}_i \rangle}{\langle \boldsymbol{\beta}_j, \boldsymbol{\beta}_j \rangle} \boldsymbol{\beta}_j, \quad \boldsymbol{f}_i = \frac{1}{||\boldsymbol{\beta}_i||} \boldsymbol{\beta}_i$$

则 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_i$ 是与 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$ 等价的单位正交向量组.

4 二次型和矩阵的合同

定义 4.1 合同 对于 $\mathbb{K}$ 上的 n 阶矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 如果存在 $\mathbb{K}$ 上的 n 阶可逆矩阵 $\boldsymbol{C}$ 使得

$$\boldsymbol{C}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{B}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同.

在大多数情况下, 我们主要考虑对称矩阵的合同关系.

定理 4.2 二次型的标准形的求法 设二次型 $f(x_1, \cdots, x_n)$ 的矩阵为 $\boldsymbol{A}$, 对 $\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{bmatrix}$ 的上半部分做成对行列变换, 下半部分做对应的列变换, 使得上半部分变为对角矩阵 $\boldsymbol{D}$, 即

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{bmatrix} \longrightarrow \begin{bmatrix} \boldsymbol{D} \\ \boldsymbol{C} \end{bmatrix}$$

则 $\boldsymbol{C}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{D}$. 令 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$ 即可得到 $\boldsymbol{x}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{x}$ 的一个标准形 $\boldsymbol{y}^{\mathrm{t}}\boldsymbol{D}\boldsymbol{y}$.

5　线性空间

6　线性映射

定义 6.1 过渡矩阵　n 维线性空间的一组基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 向另一组基 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 的过渡矩阵 $\boldsymbol{P}$ 定义为

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix} \boldsymbol{P}$$

对 $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$ 做初等行变换即可得 $\begin{bmatrix}\boldsymbol{I} & \boldsymbol{P}\end{bmatrix}$.
如果线性映射 $\mathcal{A}$ 在 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 下的矩阵分别为 $\boldsymbol{S}$ 和 $\boldsymbol{T}$, 则有

$$\boldsymbol{T} = \boldsymbol{P}^{-1} \boldsymbol{S} \boldsymbol{P}$$