

高代选期末复习

水至月生

2026 年 6 月 1 日

前言

高代选这一门课在我眼里比较混沌邪恶，每年内容变动极大，内容和线代夹杂在一起略显混乱，而且每年题目变化大，还没办法往年题拟合。

好消息是考试是比较简单的（真的简单吗），盛洁老师给出的样卷是 7 计算 +3 证明主要是计算题占大头，但是证明题比较邪恶，很考验注意力，但也不能完全放弃，最好还是能拿下两道。

本份讲义将按照讲课的顺序过一遍可能会考到的基础知识点，之后给出一些基础的技巧和证明题的方法。碍于篇幅限制，本讲义主要是知识点和考点的罗列，涉及到一些理解的内容就略过了。

建议复习的话直接阅读各种分解和计算，之后掌握后再去看基础知识点和其它的一些小技巧。

需要的话可以上小班辅导系统找我约小班辅导/来答疑坊线下找我

如果对你有帮助的话请说谢谢水至月生。

一. 基础知识

核心内容在计算部分和技巧部分，这一部分的内容主要作用不是复习基础知识，而是说明我这没列的就基本可以不看了)

1.1 线性空间和线性映射

关于线性空间，线性映射基础的内容和定义这里碍于篇幅限制，暂且略去。这里面的重点在于线性映射的矩阵表示，极大线性无关组。如果你能顺利解决下面这两道例题，就无需再复习线性空间，线性映射的基础概念了。

例 1. 设 $V = \mathbb{P}_2(\mathbb{R})$ 为所有次数小于等于 2 的实系数多项式组成的向量空间. 定义 V 上的内积为:

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx.$$

考虑算子 $T: V \rightarrow V$, $T(p) = (x^2 - 1)p'' + 2xp'$.

求 T 在基 $\{1, x, x^2\}$ 下的矩阵;

例 2. 设 $V = M_{2 \times 2}(\mathbb{C})$ 为所有 2×2 复矩阵组成的向量空间. 对于任意 $A, B \in V$, 定义 V 上的内积为:

$$\langle A, B \rangle = \frac{1}{2} \operatorname{tr}(A\bar{B}^T).$$

其中 $\operatorname{tr}(M)$ 为 M 的对角线元素之和.

设 V 的子空间 $U = \{A \in M_{2 \times 2}(\mathbb{C}) \mid \operatorname{tr}(A) = 0\}$, 求 U 的一组基;

1.2 内积

这一部分内容也相当简单, 也就直接略过了。

1.3 Hermite 二次型

这一块内容大概不考, 看看就行了

1.4 不变子空间

不变子空间和直和的基础定义这里就不赘述了, 这里强调一下比较重要的定理。

1.4.1. 性质: 若 $W \subseteq V$ 是 T 的不变子空间, 则存在 T 的某一特征向量属于 W 。

1.4.2. 定理: 设 V 是 n 维复空间, $T: V \rightarrow V$ 是线性变换, A 是 T 在一组基下的矩阵。

(1) 设 V_1 是 T 的不变子空间, 则存在 V 的子空间 V_2 使得

$$V = V_1 \oplus V_2.$$

(2) 设 v_1, \dots, v_r 为 V_1 的一组基, w_1, \dots, w_{n-r} 为 V_2 的一组基, 则 T 在基 $v_1, \dots, v_r, w_1, \dots, w_{n-r}$ 下的矩阵形如

$$B = \begin{pmatrix} B_{11} & B_{12} \\ 0_{(n-r) \times r} & B_{22} \end{pmatrix},$$

其中 B_{11} 是 r 阶方阵, B_{22} 是 $n-r$ 阶方阵, 且 B 与 A 相似。

(3) 若 V_2 也是 T 的不变子空间, 则 $B_{12} = 0$ 。

和不变子空间直接相关的是 Schur 分解

1.5 伴随

这一部分简单回顾一下基础概念

1.5.1. 定理: 设 V 是 n 维内积空间, φ 是 V 上的线性变换, 则存在 V 上唯一的线性变换 φ^* , 使得对任意的 $x, y \in V$, 都有

$$(\varphi(x), y) = (x, \varphi^*(y)).$$

上述 φ^* 称为线性变换 φ 的伴随.

1.5.2. 伴随的表示矩阵: 在标准正交基下, 原来的变换为 A 的话, 则其伴随即 (共轭) 转置

1.6 正交变换和酉变换

这一部分要注意一下正交变换的一些性质

1.6.1. 定义: 设 V 是 n 维内积空间, φ 是 V 上的线性变换, 若 φ 保持内积, 即对任意的 $x, y \in V$, $(\varphi(x), \varphi(y)) = (x, y)$, 则当 V 是欧氏空间时, 称 φ 是 V 上的正交变换; 当 V 是酉空间时, 称 φ 是 V 上的酉变换.

1.6.2. 正交矩阵和酉矩阵: 若 n 阶实矩阵 P 适合 $P'P = PP' = I_n$, 则称为正交矩阵; 若 n 阶复矩阵 U 适合 $\bar{U}'U = U\bar{U}' = I_n$, 则称为酉矩阵.

显然, 一组标准正交基下, 正交变换的表示矩阵是正交矩阵

1.6.3. 性质:

- (1) 正交矩阵的每一行 (列) 向量的长度都是 1
- (2) 正交矩阵的行与行之间相互正交, 列与列之间相互正交
- (3) 正交变化保内积

1.7 正规算子

这一部分应该只会考一个计算谱分解, 但下面的基础定义还是要了解

1.7.1. 自伴随算子: 设 φ 是 n 维内积空间 V 上的线性变换, 若 $\varphi = \varphi^*$, 则称 φ 是 V 上的自伴随算子.

当 V 是欧氏空间时, φ 是自伴随算子的充要条件是 φ 在某一组 (任一组) 标准正交基下的表示矩阵是对称矩阵;

当 V 是酉空间时, φ 是自伴随算子的充要条件是 φ 在某一组 (任一组) 标准正交基下的表示矩阵是 Hermite 矩阵.

注当 V 是欧氏空间时, 自伴随算子又称为对称变换; 当 V 是酉空间时, 自伴随算子又称为 Hermite 变换.

1.7.2. 定理: 自伴随算子可以正交对角化

: 这里的正交对角化即谱分解

1.7.3. 定理: 实对称矩阵和 Hermite 矩阵的特征值都是实数.

1.7.4. 正规算子与正规矩阵: 设 φ 是内积空间 V 上的线性变换, 若 $\varphi\varphi^* = \varphi^*\varphi$, 则称 φ 是正规算子.

若 n 阶复矩阵 A 适合 $A\bar{A}' = \bar{A}'A$, 则称为复正规矩阵; 若 n 阶实矩阵 A 适合 $AA' = A'A$, 则称为实正规矩阵.

1.7.5. 定理: 任一复正规矩阵均酉相似于复对角矩阵.

1.7.6. 定理: 任一 n 阶酉矩阵均酉相似于下列形状的对角矩阵:

$$\text{diag}\{c_1, c_2, \dots, c_n\},$$

其中 c_i 为模长等于 1 的复数.

1.8 Jordan 标准型

这一部分知识比较重要, 同样的我们略过不考点

1.8.1. 特征值与特征向量

设 $A \in \mathbb{P}^{n \times n}$, $\lambda \in \mathbb{P}$, 非零列向量 $\xi \in \mathbb{P}^n$, 若满足

$$A\xi = \lambda\xi,$$

则称 λ 为矩阵 A 的特征值, ξ 为 A 对应于特征值 λ 的特征向量。

1.8.2. 特征多项式

称行列式

$$f(\lambda) = |\lambda I - A|$$

为矩阵 A 的特征多项式, 特征多项式的根即为 A 的全体特征值。

1.8.3. 代数重数

设 λ_0 是矩阵 A 的一个特征值, 若 λ_0 作为特征多项式 $f(\lambda)$ 的根是 k 重根, 则称 k 为特征值 λ_0 的代数重数。

1.8.4. 几何重数

对特征值 λ_0 , 齐次线性方程组

$$(\lambda_0 I - A)\mathbf{x} = \mathbf{0}$$

的解空间 (特征子空间) 的维数, 称为 λ_0 的几何重数。

根据定义我们自然可以得出几何重数等于

$$\dim \text{Ker}(\lambda_0 I - A) = \dim \text{Im}(\lambda_0 I - A) = n - r(\lambda_0 I - A)$$

1.8.5. 极小多项式

设 $A \in \mathbb{P}^{n \times n}$, 在所有满足 $m(A) = 0$ 的非零多项式 $m(x) \in \mathbb{P}[x]$ 中, 次数最低且首项系数为 1 的多项式, 称为矩阵 A 的极小多项式。

极小多项式的根也一定包含所有特征值

1.8.6. Jordan 块

形如

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}_{r \times r}$$

的 r 阶方阵称为对应于特征值 λ 的 r 阶 Jordan 块。

7. Jordan 标准型

由若干个 Jordan 块对角拼接而成的分块对角矩阵

$$J = \text{diag}(J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_s}(\lambda_s))$$

称为 Jordan 标准型。复数域上任意方阵 A 都相似于唯一的 Jordan 标准型 (不计 Jordan 块排列次序)。

二. 各种分解和计算

高代选会考很多分解，我并不认为挨个讲分解是一种什么好的学习路径，但是都是期末复习了，没招了

2.1 Suchr 分解

一句话解决：期末不考这个，要是考了直接算 Jordan 分解就好，后面的内容看不看吧。

Suchr 分解即将矩阵 A 分解为

$$A = PUP^{-1}$$

其中 U 是一个上三角阵

这里简要说明一下其原理，即运用到 1.4 中不变子空间的性质，先找到 A 的一个特征向量 e_1 ，之后考虑 e_1 的补空间，然后找到 A 在其上的限制的一个特征向量 e_2 ，之后重复上述操作，直到找到 e_n ，将这组向量作为一组基即可得到 Suchr 分解。

计算过程也即上面原理，太麻烦了，真考出来算 Jordan 就好。

2.2 满秩分解 (CR 分解)

对于一个 $m \times n$ 的矩阵 A ，设其秩为 r ，其满秩分解即将其分解为一个行满秩矩阵和列满秩矩阵的乘积，

$$A_{m \times n} = C_{m \times r} R_{r \times n}$$

其的一个典型算法即 CR 分解：

2.2.1. 定义： 设 $A_{m \times n}$ ， $\dim C(A) = r$ 。构造 $C_{m \times r}$ 如下：依次取 A 的列，若当前列与已选列线性无关则放入 C ，否则舍弃。最终 C 的 r 列构成 $C(A)$ 的一组基，记为 $C = (\gamma_1, \dots, \gamma_r)$ 。再构造 $R_{r \times n}$ ： A 的每一列可由 C 的列唯一线性表示，即存在 b_{ij} 使得

$$\beta_j = \sum_{i=1}^r b_{ij} \gamma_i, \quad j = 1, \dots, n,$$

令 $R = (b_{ij})_{r \times n}$ ，则有 $A = CR$ 。

用人话说，即将 A 看作若干个列向量，之后取其中的一组基，得到 C ，之后直接算出 R 。

上面没看明白没关系，可以直接看这个例子：

例子：

$$A = \begin{pmatrix} 1 & 3 & 8 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix}.$$

先取第一列，第二列和第一列线性无关，取第二列，第三列和前两列线性相关，不取，综上 C 就是前两列平在一起。

之后计算 R ， A 的第一列就是 C 的第一列，所以 R 的第一列就是 $(1, 0)$ ， A 的第二列就是 C 的第二列，所以 R 的第二列就是 $(0, 1)$ ， A 的第三列是 C 的第一列的 2 倍加上第二列的 2 倍，所以 R 的第三列就是 $(2, 2)$ ，

2.3 秩 1 分解

秩 1 分解即将矩阵 A 拆为 $r(A)$ 个秩 1 矩阵的和, 由于秩 1 矩阵可以表示为一个列向量和一个行向量的乘积, 所以也即将矩阵 A 拆为 $r(A)$ 个列向量和行向量的乘积的和

秩 1 分解可由满秩分解直接得到: 设 $A = CR$, 其中 $C = [u_1, u_2, \dots, u_r] \in \mathbb{R}^{m \times r}$, $R = [v_1, v_2, \dots, v_r]^T \in \mathbb{R}^{r \times n}$, 则:

$$A = CR = u_1 v_1^T + u_2 v_2^T + \dots + u_r v_r^T$$

2.4 LU 分解

2.4.1. 定义: 将方阵 A 分解为

$$A = LU$$

其中 L 是单位下三角矩阵 (对角线全为 1), U 是上三角矩阵

其计算相对简单, 对于矩阵 A 进行行倍加变化 (不要交换), 得到一个上三角矩阵即 U , 之后就可以直接求出 L , 或者根据行倍加变化时的乘子直接看出。

也可以强行列方程求解。

2.5 LDL 分解和 cholesky 分解

2.5.1. 定义 (LDL 分解) 设 A 为 n 阶实对称矩阵, 若存在 n 阶单位下三角矩阵 L 与 n 阶对角矩阵 D , 满足

$$A = LDL^T$$

则称此式为矩阵 A 的 LDL^T 分解。

其中: L 为主对角线全为 1 的单位下三角矩阵; D 为实对角矩阵; L^T 是 L 的转置, 为单位上三角矩阵。

这个矩阵的计算和 LU 分解是基本一样的, 我们可以直接算 LU 分解, 之后把 U 拆为 DL^T 即可

2.5.2 定义 (cholesky 分解) 设 A 为实对称正定矩阵, 如果 $A = GG^T$, 其中 G 是下三角矩阵, 则称分解式 $A = GG^T$ 为实对称矩阵 A 的楚列斯基 (Cholesky) 分解

cholesky 分解基本等同于 LDL^T , 我们只需要将 D 拆为 CC^T , 之后计算 LC 即可得到

2.6 QR 分解 (UR 分解)

2.6.1. 定义: 对任意 n 阶可逆实矩阵 A , 存在一个 n 阶正交矩阵 Q 及一个主对角元素为正数的 n 阶上三角阵 R , 使 $A = QR$, 称为可逆实矩阵 A 的 QR 分解, 且该分解唯一。

说人话即分解为一个正交矩阵和一个上三角阵的乘积, 其计算方法一般靠的是 Gram-Schmidt 正交化方法, gram 正交化方法可以这样简化记忆, 每一次减去在之前向量组上的投影

2.6.2. 定理: n 维欧氏空间 \mathbb{R}^n 中, 任意 $s \leq n$ 个线性无关的向量 $\alpha_1, \alpha_2, \dots, \alpha_s$, 均可转化成

一个正交向量组 $\beta_1, \beta_2, \dots, \beta_s$, 其中

$$\begin{aligned}\beta_1 &= \alpha_1, \\ \beta_2 &= \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1, \\ \beta_3 &= \alpha_3 - \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)}\beta_2, \\ &\vdots \\ \beta_s &= \alpha_s - \frac{(\alpha_s, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_s, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \dots - \frac{(\alpha_s, \beta_{s-1})}{(\beta_{s-1}, \beta_{s-1})}\beta_{s-1}.\end{aligned}$$

而且 $L(\alpha_1, \alpha_2, \dots, \alpha_i) = L(\beta_1, \beta_2, \dots, \beta_i)$, $i = 1, 2, \dots, s$, 进而通过把 $\beta_1, \beta_2, \dots, \beta_s$ 单位化可得标准正交向量组 $\gamma_1, \gamma_2, \dots, \gamma_s$.

QR 分解具体的计算过程即对 A 的列向量算出其按照 Gram 正交化方法转化得来的正交向量组, 并将其单位化, 从而直接得到 Q , 之后直接计算 R

2.7 谱分解

2.7.1. 定义: 设 A 为 n 阶实对称矩阵, 其特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 对应的标准正交特征向量为 q_1, q_2, \dots, q_n , 则 A 可分解为

$$A = \sum_{i=1}^n \lambda_i q_i q_i^T = Q \Lambda Q^T$$

其中每个 $q_i q_i^T$ 是秩 1 正交投影矩阵, Q 为正交矩阵, Λ 为对角阵。该分解称为矩阵 A 的谱分解。

换言之即正规算子的正交对角化。

计算的时候直接算特征值特征向量就好, 唯一需要注意的是要把特征向量处理成正交且单位的

2.8 Jordan 分解

Jordan 分解的形式在前一部分已经提及, 这里就不重复了, 这里说明 Jordan 分解的计算:

- (1) 计算特征值
- (2) 计算 $n - r(\lambda - A)$, 极小多项式等分析 A 的 Jordan 标准型
- (3) 计算特征向量和广义特征向量

下面给出一个例子:

2.8.1. 例: 设复数域上的四维线性空间 V 上的线性变换 φ 在一组基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix},$$

求 V 的一组基, 使 φ 在这组基下的表示矩阵为 Jordan 标准型, 并求出从原来的基到新基的过渡矩阵。

解: 计算其特征多项式为 $|\lambda I - A| = (\lambda - 1)^4$

再计算其极小多项式为 $m(A) = (x - 1)^2$

以及计算得 $n - r(I - A) = 2$

因此, A 的特征值只有 1, 只有两个 Jordan 块, 每个块的大小为 2, 于是 A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}.$$

设矩阵 P 是从 $\{e_1, e_2, e_3, e_4\}$ 到新基的过渡矩阵, 则

$$P^{-1}AP = J,$$

此即

$$AP = PJ.$$

设 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 其中 α_i 是四维列向量, 代入得

$$(A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix},$$

化成方程组为

$$\begin{cases} (A - I)\alpha_1 = 0, \\ (A - I)\alpha_2 = \alpha_1, \\ (A - I)\alpha_3 = 0, \\ (A - I)\alpha_4 = \alpha_3. \end{cases}$$

由于 α_1, α_3 都是 A 的属于特征值 1 的特征向量, 故 α_2, α_4 称为属于特征值 1 的广义特征向量. 我们可取方程组 $(A - I)x = 0$ 的两个线性无关的解分别作为 α_1, α_3 (注意不能取线性相关的两个解, 因为 P 是非异阵), 然后再分别求出 α_2, α_4 (注意诸 α_i 的解可能不唯一, 只需取比较简单的一组解即可). 经计算可得

$$\alpha_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -5 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

于是

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix}.$$

因此新基为

$$\{e_1 - 2e_2 + e_3 - 5e_4, e_2, e_3 - e_4, e_4\}.$$

2.9 矩阵函数

按照我们学习的定义，矩阵能直接进行的计算只有加减乘，所以对于各种函数，我们矩阵函数的定义方式一般就是直接用多项式函数（级数）。

既然是多项式，就要算幂次，这里简化计算的方式就是用 Jordan 标准型，这也就是矩阵函数的第一种计算，展开为多项式函数，之后用 Jordan 标准型计算。

在上一种方式的基础上我们还能简化一些，应用下面的命题：

2.9.1. 命题

设多项式 $f(x) = a_m x^m + \cdots + a_1 x + a_0$ ，定义

$$f(J_i) = a_m J_i^m + \cdots + a_1 J_i + a_0 I,$$

其中 J_i 是对应于特征值 λ_i 的 r 阶 Jordan 块，则

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{1}{1!} f'(\lambda_i) & \frac{1}{2!} f''(\lambda_i) & \cdots & \frac{1}{(r-1)!} f^{(r-1)}(\lambda_i) \\ 0 & f(\lambda_i) & \frac{1}{1!} f'(\lambda_i) & \cdots & \frac{1}{(r-2)!} f^{(r-2)}(\lambda_i) \\ 0 & 0 & f(\lambda_i) & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \frac{1}{1!} f'(\lambda_i) \\ 0 & 0 & 0 & 0 & f(\lambda_i) \end{pmatrix}_{r \times r}.$$

此矩阵的第 k 条上对角线上的元素均为 $\frac{1}{(k-1)!} f^{(k-1)}(\lambda_i)$ ，其中 $1 \leq k \leq r$ 。

应用该命题我们可以迅速得到 jordan 块的矩阵函数，极大简化计算，这也是我们最常用的计算 $f(A)$ 的方式：

- (1) 计算 A 的 Jordan 分解 $A = PJP^{-1}$
- (2) 对于每个 Jordan 块计算 $f(J_i)$
- (3) 计算 $A = Pf(J)P^{-1}$

此外还有两个需要注意的点：

- (1) 矩阵没有交换性，所有矩阵函数也没有
- (2) 对于一些矩阵系数的常微分方程，我们可以类比标量的解法，最后将问题转化到矩阵函数的计算，最典型的是这个，最好直接记住：

$$\frac{dX(t)}{dt} = AX(t) \text{ 的一般解是 } X(t) = e^{At}\vec{c} = Pe^{Jt}P^{-1}\vec{c} = Pe^{Jt}\vec{c}$$

2.10 奇异值分解

2.10.1. 奇异值： 设 V, U 分别为 n, m 维欧氏空间（酉空间）， $\varphi: V \rightarrow U$ 是线性映射，若存在非负实数 σ 以及非零向量 $v \in V, u \in U$ ，使得

$$\varphi(v) = \sigma u, \quad \varphi^*(u) = \sigma v,$$

则称 σ 是 φ 的奇异值, v, u 分别称为 φ 关于 σ 的右奇异向量与左奇异向量.

2.10.2. 奇异值分解: 设 A 为 $m \times n$ 实矩阵 (复矩阵), 则存在 m 阶正交矩阵 (酉矩阵) P , n 阶正交矩阵 (酉矩阵) Q , 使得

$$A = P \begin{pmatrix} S & O \\ O & O \end{pmatrix} Q,$$

其中 $S = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_r\}$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ 是 A 的非零奇异值.

2.10.3. 定理:

$$\lambda_{\min} \left(\frac{A + A^T}{2} \right) \leq \frac{x^T A x}{x^T x} \leq \lambda_{\max} \left(\frac{A + A^T}{2} \right), \quad \forall x \neq \mathbf{0}$$

如果 A 是对称矩阵则可化简为

$$\lambda_{\min}(A) \leq \frac{x^T A x}{x^T x} \leq \lambda_{\max}(A), \quad \forall x \neq \mathbf{0}$$

2.10.10. 计算流程:

- (1) 计算 $A^T A$
- (2) 计算 $A^T A$ 的特征值和特征向量, 特征向量正交单位处理化后即 Q , 特征值开方后即 S
- (3) 直接计算 P

2.10.4. 例: 求下列矩阵的奇异值分解:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix};$$

解:

经计算可得

$$A' A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.$$

可求得其正交相似标准型 B 和过渡矩阵 Q 分别为:

$$B = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix},$$

即 $Q' A' A Q = B$. 设 Q 的 2 个列向量依次为 α_1, α_2 , 令

$$\sigma_1 = \sqrt{5}, \quad \beta_1 = \frac{1}{\sqrt{5}} A \alpha_1 = \left(\frac{2}{\sqrt{10}}, \frac{1}{\sqrt{10}}, \frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}} \right)',$$

$$\sigma_2 = 1, \quad \beta_2 = A \alpha_2 = \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right)'.$$

添加单位向量 $\beta_3 = \left(-\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right)'$ 与 $\beta_4 = \left(-\frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}, \frac{2}{\sqrt{10}}, -\frac{1}{\sqrt{10}}\right)'$, 使 $\beta_1, \beta_2, \beta_3, \beta_4$ 成为 \mathbb{R}^4 的一组标准正交基. 令 $P = (\beta_1, \beta_2, \beta_3, \beta_4)$, 则 P 为正交矩阵, 于是 A 的奇异值分解为

$$A = \begin{pmatrix} 2 & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} \\ \frac{\sqrt{10}}{1} & 1 & 0 & \frac{\sqrt{10}}{2} \\ \frac{\sqrt{10}}{1} & \frac{\sqrt{2}}{1} & 0 & \frac{\sqrt{10}}{2} \\ \frac{\sqrt{10}}{1} & -\frac{1}{\sqrt{2}} & 0 & \frac{\sqrt{10}}{2} \\ \frac{2}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{\sqrt{2}}{1} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

2.11 最小二乘法

这个直接记住结论就好: 对于 $Ax = b$, 若其不可解, 则最接近的解即为

$$x = (A^T A)^{-1} A^T b$$

(首先上面的式子不一定是直接可算的, 因为 $A^T A$ 可能不满秩.)

学习的各种方法知识在不同的情况下求这个, 这里主要介绍伪逆.

2.11.1. 伪逆: 设 $A \in \mathbb{R}^{m \times n}$, 若矩阵 $A^+ \in \mathbb{R}^{n \times m}$ 满足

1. $AA^+A = A$,
2. $A^+AA^+ = A^+$,
3. $(AA^+)^T = AA^+$,
4. $(A^+A)^T = A^+A$,

则称 A^+ 为 A 的伪逆. 对任意矩阵 A , 伪逆唯一存在.

在 $A^T A$ 满秩时, 伪逆即 $(A^T A)^{-1} A^T$

2.11.2. SVD 求伪逆

设 $A = U\Sigma V^T$ 为奇异值分解, 记

$$\Sigma^+ = \begin{pmatrix} \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}\right) & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix},$$

则

$$A^+ = V\Sigma^+U^T.$$

我们的最小解也就可以直接表示为 $x = A^+b$

有时候题目不是让我们直接算 $Ax = b$, 而是给我们一些点, 让我们用一个若干次曲线拟合, 这时候我们的 x 就是这个曲线的系数组成的向量, A 中每一行的数据对应一组 $1, x, x^2, \dots, b$ 就是 y

举个例子，比如 $(1, 1), (2, 3), (4, 6), (5, 11)$ 用二次曲线拟合 $y = a_0 + a_1x + a_2x^2$ ，那对应的就是

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 6 \\ 11 \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}, \quad A\boldsymbol{\alpha} = \mathbf{b}$$

三. 一些技巧

3.1 Jordan 标准型

证明题肯定是有了一道和 Jordan 标准型相关的。

Jordan 标准型的题目一般我们从两个部分下手：

首先我们关注各种量和 Jordan 标准型之间的联系

3.1.1. 命题

(1) 特征多项式 $|\lambda I - A|$ 给出了特征值以及特征值的代数重数，这给出了 Jordan 标准型的对角线元素

(2) 极小多项式 $m(A)$ 给出了所有的特征值，对于 $m(A)$ 中每个单项式 $(x - \lambda_1)^{k_1}$ ，指数项 k_1 给出了特征值 Jordan 块的最大大小

推论：如果极小多项式无重根，即每个指数都是 1，对于 Jordan 标准型大小都是 1，即矩阵可对角化

(3) 几何重数为 $\dim \text{Ker}(\lambda_0 I - A) = n - \dim \text{Im}(\lambda_0 I - A) = n - r(\lambda_0 I - A)$ ，给出了 Jordan 块的数量

题目显然不会直接告诉我们几何重数，一般的形式比如： A 满足 $A^2 = A$ (A 的极小多项式是 $x^2 - x$ 的因子)，或者 $r(A)$ (从而可以得到特征值 0 的几何重数)

利用这些关系以及一些默认的信息我们可以推导出关于 Jordan 块的一些信息。从而解决题目。

默认的信息一般就是指，矩阵大小为 n ，所有很多值的和不会超过 n ，比如几何重数

如果这方面没法入手，我们就直接考虑将题目中的矩阵变为 Jordan 标准型，从而考虑是否能给出一些突破口

3.2 计算特征值

在上面各种分解的计算中，大部分都要求计算特征值，但这一步可能没那么好算，这里提供一些技巧。

3.2.1. 瞪眼 + 蒙直接观察，看 $\lambda I - A$ 是否是满秩的 (行列线性无关)

一般我们先计算 $\text{tr}(A)$ ，考试的时候不会算太抽象的，一般都是整数，得到 $\text{tr}(A)$ 之后就基本可以猜一猜，比如 3 阶矩阵 $\text{tr}(A) = 2$ 就可能是特征值 1, 1, 0. 猜完之后就瞪眼验证。

这个方法真的还比较好用) 但是如果瞪眼超过 3-5 分钟没搞出来就别瞪了

3.2.2. 适当的行列变化

在计算行列式前适当的进行初等行列变化