

考试内容: 矩阵论、数值计算、数理统计各五题, 考生“任选”10题, 必须每个内容至少选三题! 每题 10 分

一、矩阵论

1. 会求线性空间两组基之间的过渡矩阵, 同一向量在两组基下的坐标; 会求线性变换的矩阵, 特征值以及特征向量。

设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 及 $\beta_1, \beta_2, \dots, \beta_n$ 是线性空间 V_n 的两个基, 且有

$$\begin{cases} \beta_1 = p_{11}\alpha_1 + p_{21}\alpha_2 + \dots + p_{n1}\alpha_n \\ \beta_2 = p_{12}\alpha_1 + p_{22}\alpha_2 + \dots + p_{n2}\alpha_n \\ \dots \\ \beta_n = p_{1n}\alpha_1 + p_{2n}\alpha_2 + \dots + p_{nn}\alpha_n \end{cases}, \quad (\beta_1 \quad \beta_2 \quad \dots \quad \beta_n) = (\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n) \underbrace{\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}}_{\text{记作 } P}$$

称此公式为基变换公式, P 为由 α 到 β 的过渡矩阵。

定理1 设 V_n 中的元素 α , 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标为 $(x_1, x_2, \dots, x_n)^T$, 在基 $\beta_1, \beta_2, \dots, \beta_n$ 下的坐标为 $(x'_1, x'_2, \dots, x'_n)^T$,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}, \quad \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

例 求坐标变换公式?

$$\begin{array}{ll} \alpha_1 = x^3 + 2x^2 - x & \beta_1 = 2x^3 + x^2 + 1 \\ \alpha_2 = x^3 - x^2 + x + 1 & \beta_2 = x^2 + 2x + 2 \\ \alpha_3 = -x^3 + 2x^2 + x + 1 & \beta_3 = -2x^3 + x^2 + x + 2 \\ \alpha_4 = -x^3 - x^2 + 1 & \beta_4 = x^3 + 3x^2 + x + 2 \end{array}$$

解: 利用一组最基本的基 $x^3, x^2, x, 1$ 过渡

因为 $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (x^3, x^2, x, 1)A$, $(\beta_1, \beta_2, \beta_3, \beta_4) = (x^3, x^2, x, 1)B$,

$$\text{其中 } A = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{pmatrix},$$

得 $(\beta_1, \beta_2, \beta_3, \beta_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) A^{-1}B$.

故坐标变换公式为

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix} = B^{-1}A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

用初等变换计算 $B^{-1}A$.

$$(B \mid A) = \left(\begin{array}{cccc|ccccc} 2 & 0 & -2 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 3 & 2 & -1 & 2 & -1 \\ 0 & 2 & 1 & 1 & -1 & 1 & 1 & 0 \\ 1 & 2 & 2 & 2 & 0 & 1 & 1 & 1 \end{array} \right) \xrightarrow{\text{初等行变换}} \sim \left(\begin{array}{cccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & -1 & 1 & -1 \end{array} \right)$$

$$\left(\begin{array}{cccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 & 1 & -1 \end{array} \right) = (E \mid B^{-1}A) \quad \text{所以} \quad \begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

例:

证明 $x^3, x^3+x, x^2+1, x+1$ 是 $P[x]_3$ 的一个基,
并求多项式 x^2+2x+3 在这个基下的坐标.

证明 令

$$\begin{aligned} k_1x^3 + k_2(x^3+x) + k_3(x^2+1) + k_4(x+1) \\ = (k_1+k_2)x^3 + k_3x^2 + (k_2+k_4)x + (k_3+k_4) \\ = 0 \\ \Rightarrow \begin{cases} k_1+k_2=0, \\ k_3=0, \\ k_2+k_4=0, \\ k_3+k_4=0 \end{cases} \Rightarrow k_1=k_2=k_3=k_4=0 \end{aligned}$$

故 $x^3, x^3+x, x^2+1, x+1$ 线性无关, 是 $P[x]_3$ 的一个基. 故 x^2+2x+3 在这个基下的坐标为 $(0,0,1,2)^T$.

又令

$$\begin{aligned} a_1x^3 + a_2(x^3+x) + a_3(x^2+1) + a_4(x+1) \\ = x^2+2x+3, \\ \begin{cases} a_1+a_2=0, \\ a_3=1, \\ a_2+a_4=2, \\ a_3+a_4=3 \end{cases} \quad \text{解之可得} \quad \begin{cases} a_1=0, \\ a_2=0, \\ a_3=1, \\ a_4=2. \end{cases} \end{aligned}$$

线性变换的矩阵:

$$\sigma(a_1, a_2, \dots, a_n) = (a_1, a_2, \dots, a_n) \begin{pmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{pmatrix} = (a_1, a_2, \dots, a_n)A,$$

A 为线性变换的矩阵

定理: 线性变换 $\sigma: V \rightarrow V$, $a_1, a_2, \dots, a_n; \beta_1, \beta_2, \dots, \beta_n$ 是 V 的两组基, 过渡阵为 P .

σ 在 a_1, a_2, \dots, a_n 与 $\beta_1, \beta_2, \dots, \beta_n$ 下的矩阵分别为 A, B . 则 $B = P^{-1}AP$

例: 设线性变换 $A: R^3 \rightarrow R^3$ 把基:

$$\alpha_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ 变为 } \alpha'_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \alpha'_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha'_3 = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}$$

1、求 σ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵

2、求 $\xi = (1, 2, 3)^T$ 及 $\sigma(\xi)$ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标

3、求 $\xi = (1, 2, 3)^T$ 及 $\sigma(\xi)$ 在基 $\alpha'_1, \alpha'_2, \alpha'_3$ 下的坐标

1、求 σ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵

$$\text{已知 } \sigma(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3) A$$

$$\sigma(\alpha_1, \alpha_2, \alpha_3) = (\alpha'_1, \alpha'_2, \alpha'_3)$$

$$\Rightarrow (\alpha_1, \alpha_2, \alpha_3) A = (\alpha'_1, \alpha'_2, \alpha'_3)$$

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

2、求 $\xi = (1, 2, 3)^T$ 及 $\sigma(\xi)$ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标

$$\xi = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\sigma(\xi) = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

3、求 $\xi = (1, 2, 3)^T$ 及 $\sigma(\xi)$ 在基 $\alpha'_1, \alpha'_2, \alpha'_3$ 下的坐标

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ -1 & -1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 3 \\ 0 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \xi = (\alpha'_1, \alpha'_2, \alpha'_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \sigma(\xi) = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = (\alpha'_1, \alpha'_2, \alpha'_3) \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

特征值与特征向量

例 1 设 V 是数域 K 上的 3 维线性空间, T 是 V 上的一个线性变换, T 在 V 的一个基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵是

$$A = \begin{bmatrix} 2 & -2 & 2 \\ -2 & -1 & 4 \\ 2 & 4 & -1 \end{bmatrix}$$

求 T 的全部特征值与特征向量。

解: A 的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 2 & -2 \\ 2 & \lambda + 1 & -4 \\ -2 & -4 & \lambda + 1 \end{vmatrix} = (\lambda - 3)^2(\lambda + 6)$$

$$|\lambda I - A| = (\lambda - 3)^2(\lambda + 6)$$

所以 A 的特征值是 3 (二重) 与 -6。

对于特征值 3, 解齐次线性方程组 $(3I - A)X = 0$

$$(3I - A) = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

得到一个基础解系: $[-2 \ 1 \ 0]^T, [2 \ 0 \ 1]^T$

从而 T 的属于 3 的极大线性无关特征向量组是

$$\xi_1 = -2\alpha_1 + \alpha_2, \quad \xi_2 = 2\alpha_1 + \alpha_3$$

$$k_1 \xi_1 + k_2 \xi_2, \quad k_1, k_2 \in K$$

对于特征值 -6, 解齐次线性方程组 $(-6I - A)X = 0$

$$(-6I - A) = \begin{bmatrix} -8 & 2 & -2 \\ 2 & -5 & -4 \\ -2 & -4 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \alpha_3 = 1/2\alpha_1 + \alpha_2 \Leftrightarrow \alpha_1 + 2\alpha_2 - 2\alpha_3 = 0$$

得到一个基础解系: $[1 \ 2 \ -2]^T$

从而 T 的属于 -6 的极大线性无关特征向量组是

$$\xi_3 = \alpha_1 + 2\alpha_2 - 2\alpha_3$$

于是 T 的属于 -6 的全部特征向量

$$k \xi_3, \quad k \in K$$

这里 k 为数域 K 中任意非零数。

2. 给定一线性空间的两组向量分别生成的子空间, 求它们的和空间与交空间的基与维数。

例: 设 $a_1 = (2, 1, 3, 1)^T$, $a_2 = (-1, 1, -3, 1)^T$, $\beta_1 = (4, 5, 3, 1)^T$, $\beta_2 = (1, 5, -3, 1)^T$,

$V_1 = \text{span}\{a_1, a_2\}$, $V_2 = \text{span}\{\beta_1, \beta_2\}$ 。求 1、 $V_1 + V_2$ 的基与维数。 2、 $V_1 \cap V_2$ 的基与维数。

$$\text{解: } V_1 + V_2 = \text{span}\{a_1, a_2, \beta_1, \beta_2\}, \quad (a_1, a_2, \beta_1, \beta_2) = \begin{bmatrix} 2 & -1 & 4 & 1 \\ 1 & 1 & 5 & 5 \\ 3 & -3 & 3 & -3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

a_1, a_2, β 是极大无关组, 是 $V_1 + V_2$ 的基, 维数为 3。

设 $\alpha \in V_1 \cap V_2$ 有 $\alpha \in V_1$ 且 $\alpha \in V_2$

$$\begin{aligned} \alpha = k_1 a_1 + k_2 a_2 &= k_3 \beta_1 + k_4 \beta_2 = k_3 \beta_1 + k_4 (-a_1 + a_2 + \beta_1) \\ (k_1 + k_4) a_1 + (k_2 - k_4) a_2 - (k_3 + k_4) \beta_1 &= 0 \end{aligned}$$

$$\begin{aligned} k_1 + k_4 &= 0 \\ k_2 - k_4 &= 0 \\ k_3 + k_4 &= 0 \end{aligned} \Rightarrow \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = c \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \alpha = c(-a_1 + a_2) = c(\beta_2 - \beta_1) = c \begin{pmatrix} -3 \\ 0 \\ -6 \\ 0 \end{pmatrix} = k \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$$

$V_1 \cap V_2$ 的基为 $(1, 0, 2, 0)^T$ 维数为 1

3.会求矩阵的各种范数: $\|A\|_F$, $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$

设 $A = (a_{ij}) \in C^{m \times n}$, $x = (x_1, x_2, \dots, x_n)^T \in C^n$, 则与向量 $1 -$ 范数, $2 -$ 范数和 $\infty -$ 范数相容的矩阵算子范数分别为

(1) $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$, 称为矩阵 A 的列模和最大值范数;

(2) $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$, 称为矩阵 A 的行模和最大值范数;

(3) $\|A\|_2 = \sqrt{\lambda_{A^H A}}$ (其中 $\lambda_{A^H A}$ 为 $A^H A$ 的最大特征值), 称为矩阵 A 的谱范数.

(4) $\|A\|_F$: 矩阵的各个元素平方之和再开平方根), 称为矩阵 A 的 F 范数.

对任何矩阵 $A \in C^{m \times n}$, 都有 $\|A^H\|_1 = \|A^T\|_1 = \|A\|_\infty$, $\|A^H\|_2 = \|A^T\|_2 = \|A\|_2$

$$\|A^H A\|_2 = \|A\|_2^2, \|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$$

例: $A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 3 \\ 1 & 2 & 0 \end{bmatrix}$, 计算 $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$ 和 $\|A\|_F$

解: $\|A\|_1 = 5, \|A\|_\infty = 5, \|A\|_F = \sqrt{23}$

$$A^H A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 9 & 6 \\ 0 & 6 & 9 \end{bmatrix}, \|A\|_2 = \sqrt{15}$$

4. 会求矩阵的奇异值分解及其伪逆矩阵 A^+

① 求奇异值

引理 1：对于任何一个矩阵 A 都有

$$\text{Rank}(AA^H) = \text{Rank}(A^H A) = \text{Rank}(A)$$

设 $A \in C_r^{m \times n}$, λ_i 是 AA^H 的特征值, μ_i 是 $A^H A$ 的特征值, 它们都是实数。如果记

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_m = 0$$

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > \mu_{r+1} = \mu_{r+2} = \cdots = \mu_n = 0$$

特征值 λ_i 与 μ_i 之间有如下关系：

定理：设 $A \in C_r^{m \times n}$, 那么 $\lambda_i = \mu_i > 0$, $i = 1, 2, \dots, r$ 。

同时, 称 $\alpha_i = \sqrt{\lambda_i} = \sqrt{\mu_i} > 0$, $i = 1, 2, \dots, r$ 为矩阵 A 的正奇异值, 简称奇异值。

例：求下列矩阵的奇异值

$$(1) \quad A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (2) \quad A = \begin{bmatrix} 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix}$$

解：(1)

由于 $AA^H = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, 显然 AA^H 的特征值为 5, 0, 0, 所以 A 的奇异值为 $\sqrt{5}, 0, 0$

(2)

由于 $AA^H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, 显然 AA^H 的特征值为 2, 4, 所以 A 的奇异值为 $\sqrt{2}, 2$

例：求下列矩阵的奇异值分解表达式

$$(1) \quad A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (2) \quad A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$$

解：(1) 容易计算 $A^H A$ 的特征值为 5, 0, 所以 A 的奇异值为 $\sqrt{5}, 0$

对于特征值 5, 解齐次线性方程组 $(5I - A^H A)X = 0 \rightarrow \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}X = 0 \rightarrow X = k \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

对于特征值 0, 解齐次线性方程组 $(-A^H A)X = 0 \rightarrow \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix}X = 0 \rightarrow X = k \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

将 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ 标准化得 $v_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, v_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$

$\because A^H A$ 的非零奇异值个数为 1

\therefore 取 $\Delta = [\sqrt{5}]$, 取 $V = [V_1, V_2]$, 其中 $V_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, V_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$

计算 $U_1 = AV_1\Delta^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

构造 $U_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, U_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, 得 $U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

于是可得奇异值分解式为 $A = U \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} V^H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$

另一种方法 (有时需要试凑?):

解: (1) 容易计算 AA^H 的特征值为 5, 0, 0, 所以 A 的奇异值为 $\sqrt{5}$

下面计算 AA^H 的标准正交特征向量

$AA^H = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ 对应的正交特征向量分别为 $\mu_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mu_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mu_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

由这三个标准正交特征向量组成矩阵 U , 所以有 $U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$A^H A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

再计算 $A^H A$ 的标准正交特征向量, 解得分别与 5, 0 对应的两个标准正交特征向量

对于特征值 5, 解齐次线性方程组 $(5I - A^H A)X = 0 \rightarrow \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} X = 0 \rightarrow X = k \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

对于特征值 0, 解齐次线性方程组 $(-A^H A)X = 0 \rightarrow \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} X = 0 \rightarrow X = k \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$\text{将 } \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ 标准化得 } v_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, \quad v_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

由这两个标准正交特征向量组成矩阵 U , 那么有 $V = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$

$$\text{于是可得奇异值分解式为 } A = U \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} V^H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

(2)

$A^H A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, 那么 A 的非零奇异值为 $\sqrt{5}, \sqrt{2}$, $A^H A$ 对应于特征值 5, 2 的标准特征向量为

$$v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ 由这两个标准正交特征向量组成矩阵 } V, \text{ 那么有 } V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\because A^H A \text{ 的非零奇异值个数为 } 2 \therefore \text{ 取 } \Delta = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \end{bmatrix}, V_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\text{计算 } U_1 = A V_1 \Delta^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{5}}{5} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \\ \frac{2\sqrt{5}}{5} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{bmatrix}, \text{ 再扩充成 } U$$

另一种方法 (有时需要试凑?):

再计算 AA^H 的标准正交特征向量, 解得分别与 5, 2, 0, 0 对应的两个标准正交特征向量

$$\mu_1 = \begin{bmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \\ 0 \end{bmatrix}, \mu_2 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \mu_3 = \begin{bmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \mu_4 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, U = \begin{bmatrix} 1/\sqrt{5} & 0 & -2/\sqrt{5} & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 2/\sqrt{5} & 0 & 1/\sqrt{5} & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$$

注: 这里定义 $A^H A$ 对应的..是 V , AA^H 对应的..是 U

$$\text{于是可得 } A = U \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} V^H = \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

伪逆矩阵求法:

- ①首先求 SVD 分解。 $(A = U S V^H)$
- ②将 S 矩阵中的非零元素取倒数, 再转置, 得到 Σ^+
- ③ $A^+ = V \Sigma^+ U^H$

注: 这里定义 $A^H A$ 对应的...是 V , AA^H 对应的...是 U 。

5. 会求 $\begin{cases} \frac{dx}{dt} = Ax + f(t) \\ x(t_0) = x_0 \end{cases}$ 的解

$$x = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-\tau)}f(\tau)d\tau, f(A) = Pf(J)P^{-1}$$

例: $A = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$, 求微分方程组 $\frac{dx(t)}{dt} = Ax(t)$ 满足初始条件 $x(0) = [1 \ 1 \ 1]^T$ 的解。

$$\phi(t) = L^{-1}[(sI - A)^{-1}], \quad L^{-1}\left[\frac{1}{s+a}\right] = e^{-at}, L^{-1}\left[\frac{1}{(s+a)^2}\right] = te^{-at}$$

$$(sI - A) = \begin{bmatrix} s+1 & 2 & -6 \\ 1 & s & -3 \\ 1 & 1 & s-4 \end{bmatrix}, (sI - A)^{-1} = \begin{bmatrix} \frac{s-3}{s^2-2s+1} & -\frac{2}{s^2-2s+1} & \frac{6}{s^2-2s+1} \\ -\frac{1}{s^2-2s+1} & \frac{s-2}{s^2-2s+1} & \frac{3}{s^2-2s+1} \\ -\frac{1}{s^2-2s+1} & -\frac{1}{s^2-2s+1} & \frac{s+2}{s^2-2s+1} \end{bmatrix}$$

$$e^{At} = \phi(t) = L^{-1}[(sI - A)^{-1}] = \begin{bmatrix} -(2t-1)e^t & -2te^t & 6te^t \\ -te^t & -(t-1)e^t & 3te^t \\ -te^t & -te^t & (3t+1)e^t \end{bmatrix}$$

$$x(t) = \phi(t)x(0) = \begin{bmatrix} -(2t-1)e^t & -2te^t & 6te^t \\ -te^t & -(t-1)e^t & 3te^t \\ -te^t & -te^t & (3t+1)e^t \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} (2t+1)e^t \\ (t+1)e^t \\ (t+1)e^t \end{bmatrix}$$

$$\text{例: } A = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}, f(t) = \begin{bmatrix} -e^t \\ 0 \\ e^t \end{bmatrix}, \text{ 求 } \frac{dx(t)}{dt} = Ax(t) + f(t) \text{ 满足 } x(0) = [1 \ 1 \ 1]^T \text{ 的解。}$$

满足所给初始条件的解为 $x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}f(\tau)d\tau$

$$e^{A(t-\tau)} = \begin{bmatrix} (1-2t+2\tau)e^{t-\tau} & -2(t-\tau)e^{t-\tau} & 6(t-\tau)e^{t-\tau} \\ -(t-\tau)e^{t-\tau} & (1-t+\tau)e^{t-\tau} & 3(t-\tau)e^{t-\tau} \\ -(t-\tau)e^{t-\tau} & -(t-\tau)e^{t-\tau} & (1+3t-3\tau)e^{t-\tau} \end{bmatrix}, e^{A(t-\tau)}f(\tau) = \begin{bmatrix} -1+8(t-\tau) \\ 4(t-\tau) \\ 1+4(t-\tau) \end{bmatrix} e^t$$

$$\int_0^t e^{A(t-\tau)}f(\tau)d\tau = e^t \begin{bmatrix} \int_0^t [-1+8(t-\tau)]d\tau \\ \int_0^t 4(t-\tau)d\tau \\ \int_0^t [1+4(t-\tau)]d\tau \end{bmatrix} = e^t \begin{bmatrix} 4t^2 - t \\ 2t^2 \\ 2t^2 + t \end{bmatrix}$$

二、数值计算

1. 误差的来源有哪些? 为了使某个数的近似值其误差不超过 ϵ , 至少需要几位有效数字?

- (1) 模型误差——数学模型与实际问题之间的误差.
- (2) 观测误差——由观测产生的误差.
- (3) 截断误差——由简化问题(公式)所引起的解的误差(也称方法误差).
- (4) 舍入误差——数字计算过程中产生的误差

有效数字定义: 如果 $|e| = |x^* - x| \leq 0.5 \times 10^{-k}$, 称近似数 x^* 准确到小数点后第 k 位, 从这小数点后第 k 位数字直到最左边非零数字之间的所有数字都称为有效数字。

(k 加上小数点左边的数字个数, 为有效数字个数)

例: $x_1 = 3.142, x_2 = 3.141, x_3 = \frac{22}{7}$ 分别作为 π 的近似值各具有几位有效数字?

解: $\pi = 3.14159265358979323846\dots$

$$|\pi - x_1| = 0.00040\dots \leq \frac{1}{2} \times 10^{-3} = 0.0005, \quad 4\text{位}$$

$$|\pi - x_2| = 0.00059\dots \leq \frac{1}{2} \times 10^{-2} = 0.005, \quad 3\text{位}$$

$$|\pi - x_3| = 0.00126\dots \leq \frac{1}{2} \times 10^{-2} = 0.005, \quad 3\text{位}$$

有效数位与相对误差的关系:

设 x 的近似值为: $x^* = \pm 0.a_1 a_2 \dots a_n \times 10^m \quad (a_1 \neq 0)$,

(1) 若 x 具有 n 位有效数字, 则 $\varepsilon_r \leq \frac{1}{2a_1} \times 10^{-n+1}$

(2) 若 $\varepsilon_r \leq \frac{1}{2(a_1+1)} \times 10^{-n+1}$, 则 x 至少有 n 位有效数字

例: 要使 $\sqrt{20}$ 的近似值有小于0.1%的相对误差, 要取几位有效数字。

解: $\varepsilon_r \leq \frac{1}{2a_1} \times 10^{-n+1} \quad \because \sqrt{20} = 4.4\dots, \therefore a_1 = 4$

$$\varepsilon_r \leq \frac{1}{2a_1} \times 10^{-n+1} = \frac{1}{2 \times 4} \times 10^{-n+1} = 0.125 \times 10^{-n+1}$$

故取 $n = 4$, 则 $\varepsilon_r \leq 0.125 \times 10^{-n+1} < 10^{-3} = 0.1\%$

2. 熟悉 Lagrange 与 Newton 插值多项式, 会利用分段线性、抛物插值求函数值, 积分的近似值, 并能估计误差。

Lagrange 插值:

$$\text{基函数 } l_i(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n)}{(x_i - x_0)(x_i - x_1)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$$

$$\text{Lagrange 插值多项式 } L_n(x) = \sum_{i=0}^n y_i l_i(x) = \sum_{k=0}^n y_k \frac{\omega_{n+1}(x)}{(x - x_k) \omega_{n+1}'(x_k)}$$

$$\omega_{n+1}(x) = (x - x_0)(x - x_1)\cdots(x - x_n) \quad (\text{全部乘满})$$

$$\omega_{n+1}'(x_k) = (x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n) \quad (\text{除去 } (x_k - x_k))$$

$$\text{插值余项 } R_n(x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \xi \in (a, b)$$

$$\text{误差估计 } |R_n(x)| = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^n |x - x_j| \leq \frac{\max_{a < x < b} |f^{(n+1)}(x)|}{(n+1)!} \prod_{j=0}^n |x - x_j|$$

$$\max_{a < x < b} |f^{(n+1)}(x)| = M_{n+1}$$

注: n 代表插值多项式的次数! n=1, 一次, 线性; n=2, 二次, 抛物。n+1 点。

例: 已给 $\sin 0.32 = 0.314567, \sin 0.34 = 0.333487, \sin 0.36 = 0.3522787$

用线性插值及抛物插值计算 $\sin 0.3367$ 的值, 并估计截断误差。

解: 取 $x_0 = 0.32, y_0 = 0.314567, x_1 = 0.34, y_1 = 0.333487, x_2 = 0.36, y_2 = 0.352274$

用线性插值计算, 取 $x_0 = 0.32$ 及 $x_1 = 0.34$, 得

$$\begin{aligned} \sin 0.3367 &\approx L_1(0.3367) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(0.3367 - x_0) \\ &= 0.314567 + \frac{0.01892}{0.02} \times 0.0167 = 0.330365 \end{aligned}$$

用抛物插值计算:

$$\begin{aligned} \sin 0.3367 &\approx y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} \\ &= L_2(0.3367) = 0.330374 \end{aligned}$$

$$|R_2(x)| \leq \frac{M_3}{6} |(x - x_0)(x - x_1)(x - x_2)|$$

$$\text{其中 } M_3 = \max_{x_0 \leq x \leq x_2} |f'''(x)| = \cos x_0 < 0.828$$

$$\therefore |R_2(x)| \leq \frac{1}{6} (0.828)(0.0167)(0.033)(0.0233) < 0.178 \times 10^{-6}$$

Newton 插值:

$$N_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0)\dots(x - x_{n-1})$$

注: n 代表插值多项式的次数! n=1, 一次, 线性; n=2, 二次, 抛物。n+1 点。

$$a_0 = f[x_0]$$

$$a_1 = f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}$$

$$a_2 = f[x_0, x_1, x_2] = \frac{f[x_0, x_2] - f[x_0, x_1]}{x_2 - x_1} = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

$$f[x_0, x_1, \dots, x_k] = \frac{f[x_0, x_1, \dots, x_{k-2}, x_k] - f[x_0, x_1, \dots, x_{k-1}]}{x_k - x_{k-1}} = \frac{f[x_1, \dots, x_k] - f[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0}$$

$$(k = 0, 1, \dots, n)$$

x_k	$f(x_k)$	一阶差商	二阶差商	三阶差商	四阶差商
x_0	$f(x_0)$				
x_1	$f(x_1)$	$f[x_0, x_1]$			
x_2	$f(x_2)$	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$		
x_3	$f(x_3)$	$f[x_2, x_3]$	$f[x_1, x_2, x_3]$	$f[x_0, x_1, x_2, x_3]$	
x_4	$f(x_4)$	$f[x_3, x_4]$	$f[x_2, x_3, x_4]$	$f[x_1, x_2, x_3, x_4]$	$f[x_0, x_1, x_2, x_3, x_4]$
x_5	$f(x_5)$	$f[x_4, x_5]$	$f[x_3, x_4, x_5]$	$f[x_2, x_3, x_4, x_5]$	$f[x_1, x_2, x_3, x_4, x_5]$

线性插值误差估计:

$$|R(x)| = |f(x) - \varphi(x)| \leq \frac{h^2}{8} M, h = \max_{0 \leq i \leq n-1} (x_{i+1} - x_i), M = \max_{a \leq x \leq b} |f''(x)|$$

牛顿插值的余项与误差估计? ? ?

求积分近似值:

$$\text{梯形公式: } \int_a^b f(x) dx \approx \frac{b-a}{2} [f(a) + f(b)]$$

$$\text{误差估计: } R_1(f) = -\frac{(b-a)^3}{12} f''(\eta)$$

$$\text{Simpson求积公式: } \int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

$$\text{误差估计: } R_2(f) = -\frac{b-a}{180} \left(\frac{b-a}{2} \right)^4 f^{(4)}(\eta)$$

例 1 分别用梯形公式与辛普森公式计算积分 $I = \int_0^1 e^x dx$ 的近似值并估计误差。

解: 由梯形公式 $I = \frac{1}{2}(e^0 + e^1) = 1.85914$

截断误差为: $|R_1| = \frac{1}{12} |f''(\xi)| \leq \frac{e}{12} = 0.22652$

由Simpson公式 $I \approx \frac{1}{6} \left(e^0 + 4e^{\frac{1}{2}} + e^1 \right) = 1.71886$

截断误差为: $|R_2| = \frac{1}{2880} |f^{(4)}(\xi)| \leq \frac{e}{2880} = 0.00095$

复化求积? ? ?

3.会建立非线性方程 $f(x) = 0$ 求近似根的迭代格式, 比如 Newton 切线法, 弦截法并能用你熟知的语言编程。建立求非线性常微分方程 Cauchy 问题数值解的 Euler 折线法的迭代格式

Newton 切线法:

切线法:

设已知方程 $f(x)=0$ 有近似根 x_k (假定 $f'(x_k) \neq 0$) , 将函数 $f(x)$ 在点 x_k 展开, 有

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k)$$

$$f(x) = 0 \text{ 可近似表示为 } f(x_k) + f'(x_k)(x - x_k) = 0$$

$$\text{则迭代形式为: } x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (k = 0, 1, \dots)$$

$$\text{弦截法: } x_{k+1} = x_k - \frac{(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} f(x_k)$$

非线性常微分方程的 Euler 法:

$$\begin{cases} \frac{dy}{dx} = f(x, y) \\ y(x_0) = y_0 \end{cases} \rightarrow \begin{cases} y_{n+1} = y_n + hf(x_n, y_n) \\ y_0 = y(x_0) \end{cases}$$

4.会判断一数值积分公式的代数精度。

如果对于次数不超过 m 的多项式均能准确地成立, 但对于 $m+1$ 次多项式不准确成立, 则称该求积公式具有 m 次代数精度。

此处的“多项式”, 是指 $\{1, x, x^2, x^3, \dots\}$ 。

$n+1$ 个节点的插值型求积公式至少具有 n 次代数精确度。

梯形公式代数精确度是 1, Simpson 公式代数精度为 3.

三、数理统计

1.会求参数估计: 矩估计与极大似然估计。

一、矩估计

方法: 用样本 k 阶矩 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ 估计总体 k 阶矩 $\mu_k = E(X^k)$, 建立含有

待估参数的方程, 从而解出待估参数。

例 1. 设总体 $X \sim B(m, p)$, 其中 p 未知, X_1, X_2, \dots, X_n 为总体的样本, 求 p 的矩估计量。

解: $E(X) = mp$, 令 $\bar{X} = mp$, 得 $\hat{p} = \bar{X}/m$.

例 2 设总体 X 的概率密度为 $f(x) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}$, X_1, \dots, X_n 为样本, 求参数 σ 的矩估计

$$\text{解: } E(X) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} dx = 0$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}} dx = \int_0^{+\infty} x^2 \cdot \frac{1}{\sigma} e^{-\frac{x}{\sigma}} dx = \int_0^{+\infty} (-x^2) de^{-\frac{x}{\sigma}} \quad (\text{用两次分部积分法})$$

$$= (-x^2) e^{-\frac{x}{\sigma}} \Big|_0^{+\infty} + \int_0^{+\infty} 2xe^{-\frac{x}{\sigma}} dx = 2 \int_0^{+\infty} (-x) d\sigma e^{-\frac{x}{\sigma}} = 2\sigma \int_0^{+\infty} e^{-\frac{x}{\sigma}} dx = 2\sigma^2$$

$$\text{令 } A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2 = 2\sigma^2 \text{ 得 } \hat{\sigma} = \sqrt{\frac{1}{2n} \sum_{i=1}^n X_i^2}$$

注: 分部积分法: $\int u dv = uv - \int v du$

例 3 设 X_1, \dots, X_n 是取自总体 X 的一个样本

$$X \sim f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}}, & x \geq \mu, \theta, \mu \text{ 为未知参数, 其中 } \theta > 0, \\ 0, & \text{其它} \end{cases}$$

$$\text{解: } E(X) = \int_{\mu}^{+\infty} x \cdot \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} dx = \int_{\mu}^{+\infty} (-x) \cdot de^{-\frac{x-\mu}{\theta}} = (-x) e^{-\frac{x-\mu}{\theta}} \Big|_{\mu}^{+\infty} + \int_{\mu}^{+\infty} e^{-\frac{x-\mu}{\theta}} dx$$

$$= \mu - \theta e^{-\frac{x-\mu}{\theta}} \Big|_{\mu}^{+\infty} = \mu + \theta$$

$$E(X^2) = \int_{\mu}^{+\infty} x^2 \cdot \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} dx = \int_{\mu}^{+\infty} (-x^2) \cdot de^{-\frac{x-\mu}{\theta}} = (-x^2) e^{-\frac{x-\mu}{\theta}} \Big|_{\mu}^{+\infty} + \int_{\mu}^{+\infty} 2xe^{-\frac{x-\mu}{\theta}} dx$$

$$= \mu^2 + 2\theta \int_{\mu}^{+\infty} x \cdot \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}} dx = \mu^2 + 2\theta E(X) = \mu^2 + 2\mu\theta + 2\theta^2$$

$$\text{令 } \begin{cases} \bar{X} = \mu + \theta, \\ \frac{1}{n} \sum_{i=1}^n X_i^2 = \mu^2 + 2\mu\theta + 2\theta^2 \end{cases} \text{ 解得 } \begin{cases} \hat{\mu} = \bar{X} - \sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2} \\ \hat{\theta} = \sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2} \end{cases}$$

$$\text{方差计算公式: } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

二、极大似然估计

例 1. 设总体 X 服从参数为 $\lambda (\lambda > 0)$ 的泊松分布, 求 λ 的极大似然估计。

解: 设 (X_1, X_2, \dots, X_n) 为从总体 X 中随机抽取的样本, 样本观察值为 (x_1, x_2, \dots, x_n)

由 X 服从泊松分布, 得 X 的分布律为 $p(x; \lambda) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0, 1, \dots$

$$\text{似然函数为 } L(\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{x_1! x_2! \cdots x_n!}$$

$$\text{两边取对数, 得 } \ln L(\lambda) = -n\lambda + \left(\sum_{i=1}^n x_i \right) \ln \lambda - \sum_{i=1}^n \ln x_i!$$

$$\text{对 } \lambda \text{ 求导, 并令其为 } 0, \frac{d \ln L(\lambda)}{d \lambda} = -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0$$

$$\text{得 } \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x} \quad \text{所以 } \lambda \text{ 的极大似然估计为 } \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

例 2 设总体 X 的概率密度为 $f(x) = \begin{cases} \lambda a x^{a-1} e^{-\lambda x^a}, & x \geq 0 \\ 0, & x < 0 \end{cases}$, 其中 $\lambda > 0$ 为待估参数,

$a > 0$ 是已知常数, (x_1, x_2, \dots, x_n) 是取自总体 X 的样本值, 求 λ 的极大似然估计值。

$$\text{解: } L(\lambda) = \prod_{i=1}^n f(x_i, \lambda) = \prod_{i=1}^n \lambda a x_i^{a-1} e^{-\lambda x_i^a} = \lambda^n a^n \left(\prod_{i=1}^n x_i \right)^{a-1} e^{-\lambda \sum_{i=1}^n x_i^a}$$

$$\text{两边取对数, 得 } \ln L(\lambda) = n \ln \lambda + n \ln a + (a-1) \ln \left(\prod_{i=1}^n x_i \right) - \lambda \sum_{i=1}^n x_i^a$$

$$\text{对 } \lambda \text{ 求导, 并令其为 } 0, \frac{d \ln L(\lambda)}{d \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i^a = 0$$

$$\text{得 } \hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i^a}, \text{ 为 } \lambda \text{ 的极大似然估计值.}$$

例 3. 设总体 X 的分布律

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 θ ($0 < \theta < 1/2$) 是未知参数, 3, 1, 3, 0, 3, 1, 2, 3, 是来自总体 X 的样本观察值, 求参数 θ 的极大似然估计值.

解: $L(\theta) = (1-2\theta) \cdot 2\theta \cdot (1-\theta) \cdot (1-2\theta) \cdot \theta^2 \cdot (1-2\theta) \cdot 2\theta \cdot (1-\theta) \cdot \theta^2 \cdot (1-2\theta)$
 $= 4\theta^6 (1-2\theta)^4 (1-\theta)^2$

两边取对数, 得 $\ln L(\theta) = \ln 4 + 6 \ln \theta + 4 \ln(1-2\theta) + 2 \ln(1-\theta)$

对 θ 求导, 并令其为 0, $\frac{d \ln L(\theta)}{d\theta} = \frac{6}{\theta} - \frac{8}{1-2\theta} - \frac{2}{1-\theta} = 0$

得 $\theta = \frac{7+\sqrt{13}}{12}$ 和 $\theta = \frac{7-\sqrt{13}}{12}$, 因为 $\theta = \frac{7+\sqrt{13}}{12} > \frac{1}{2}$ 不合题意,

所以 θ 的极大似然估计值为 $\hat{\theta} = \frac{7-\sqrt{13}}{12}$

例 4. 设 $X \sim U(a, b)$, x_1, x_2, \dots, x_n 是 X 的一个样本值, 求 a, b 的极大似然估计值与极大似然估计量.

解: 由 $X \sim U(a, b)$ 知, X 的密度函数为 $f(x; a, b) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其它} \end{cases}$

似然函数为 $L(a, b) = \prod_{i=1}^n f(x_i; a, b) = \begin{cases} \frac{1}{(b-a)^n}, & a < x_i < b \\ 0, & \text{其它} \end{cases}$

似然函数只有当 $a < x_i < b, i = 1, 2, \dots, n$ 时才能获得最大值, 且 a 越大, b 越小, $L(a, b)$ 越大

令 $x_{\min} = \min\{x_1, x_2, \dots, x_n\}$, $x_{\max} = \max\{x_1, x_2, \dots, x_n\}$, 取 $\hat{a} = x_{\min}$, $\hat{b} = x_{\max}$

则对满足 $a \leq x_{\min} \leq x_{\max} \leq b$ 的一切 $a < b$, 都有 $\frac{1}{(b-a)^n} \leq \frac{1}{(x_{\max}-x_{\min})^n}$

故有 $\hat{a} = x_{\min}$, $\hat{b} = x_{\max}$ 是 a, b 的极大似然估计值.

$\hat{a} = X_{\min} = \min\{X_1, X_2, \dots, X_n\}$, $\hat{b} = X_{\max} = \max\{X_1, X_2, \dots, X_n\}$

分别为 a, b 的极大似然估计量

2. 假设检验的思想、原理是什么？假设检验中 p —值是什么含义？你知道哪几种非参数检验方法？会列联表的独立性检验。

假设检验的思想原理：

假设检验的理论依据：实际推断原理（小概率原理）：

小概率事件在一次试验中几乎是不可能发生的。

假设检验是概率意义下的反证法，即：

首先假定原假设 H_0 成立，依照事先给定的概率 α （称为显著性水平），构造一个小概率事件。然后根据抽样的结果，观察此小概率事件是否发生。若此小概率事件发生了，则认为原假设是不真的，从而作出拒绝 H_0 的判断。否则，就接受 H_0 。

拒绝原假设是有说服力的，而接受原假设是没有说服力的，因此应把希望否定的假设作为原假设。

P 值的含义：

一般在一个假设检验中，利用观测值能够做出的拒绝原假设 H_0 的最小显著性水平称为该检验的 p 值。按 p 值的定义，对于任意指定的显著性水平 α ，有以下结论：

- (1) 若 $\alpha < p$ 值，则在显著性水平 α 下接受 H_0 。
- (2) 若 $\alpha \geq p$ 值，则在显著性水平 α 下拒绝 H_0 。

p 值反映了样本信息中所包含的反对原假设 H_0 的依据的强度， p 值是已经观测到的一个小概率事件的概率， p 值越小， H_0 越有可能不成立，说明样本信息中反对 H_0 的依据的强度越强、越充分。

非参数检验方法：

χ^2 拟合优度检验：检验总体 X 的分布函数是否为 $F(x)$

柯尔莫哥洛夫检验：检验经验分布是否服从某种理论分布

斯米尔诺夫检验：检验两个样本是否服从同一分布

独立性检验：

独立性检验列联表：

$X \backslash Y$		Y				$n_{i\bullet} = \sum_{j=1}^k n_{ij}$
		B_1	B_2	\dots	B_k	
X	A_1	n_{11}	n_{12}	\dots	n_{1k}	$n_{1\bullet}$
	A_2	n_{21}	n_{22}	\dots	n_{2k}	$n_{2\bullet}$
	\vdots	\vdots	\vdots	\dots	\vdots	\vdots
	A_m	n_{m1}	n_{m2}	\dots	n_{mk}	$n_{m\bullet}$
$n_{\bullet j} = \sum_{i=1}^m n_{ij}$		$n_{\bullet 1}$	$n_{\bullet 2}$	\dots	$n_{\bullet k}$	

例 1.

$X \backslash Y$	患慢性气管炎者	未患慢性气管炎者	合计	患病率 %
吸烟	43	162	205	21.0
不吸烟	13	121	134	9.7
合计	56	283	339	16.5

解: 检验的问题为: H_0 : 总体的两个指标X 和Y 是相互独立的。

$$\text{统计量为: } \chi^2 = n \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - \frac{n_i \cdot n_{\bullet j}}{n})^2}{n_i n_{\bullet j}}$$

$$\text{观察值为: } \hat{\chi}^2 = n \sum_{i=1}^m \sum_{j=1}^k \frac{(n_{ij} - \frac{n_i \cdot n_{\bullet j}}{n})^2}{n_i n_{\bullet j}} = 7.48$$

$$[\text{上侧}] \text{ 分位数: } \chi^2_{0.01}((m-1)(k-1)) = \chi^2_{0.01}(1) = 6.635$$

$$\text{显然: } \chi^2_{0.01}(1) = 6.635 < \hat{\chi}^2 = 7.48$$

因而拒绝原假设, 即认为慢性气管炎与吸烟有关。

3. 掌握方差分析表格结构, 会填出空格 (单因素、双因素无交互作用、有交互作用)

SST : 全部观察值 x_{ij} 与总平均值 \bar{x} 的离差平方和

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

SSE : 组内离差平方和

每个水平或组的各样本数据与对应组平均值作差, 平方, 再求和

$$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

SSA : 组间平方和

各组平均值 \bar{x}_i ($i = 1, 2, \dots, k$) 与总平均值 \bar{x} 的离差平方和

$$SSA = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2, n_i \text{ 为对应组观察值个数}$$

$$SST = SSE + SSA$$

数据的简化: 数据经过变换 $X'_{ij} = b(X_{ij} - a)$ 后, 对 F 值的计算没有影响

单因素方差分析表: (其中, k 为水平的个数, n 为总的数据个数)

方差来源	平方和 SS	自由度 df	均方 MS	F 值
组间 (因素影响)	SSA	$k-1$	$MSA = SSA/(k-1)$	$F = \frac{MSA}{MSE}$
组内 (误差)	SSE	$n-k$	$MSE = SSE/(n-k)$	
总和	SST	$n-1$		

双因素无重复 (无交互作用):

因素 A \ 因素 B	B_1	B_2	\dots	B_b	$T_{i.} = \sum_{j=1}^b X_{ij}$	$\bar{X}_{i.} = T_{i.}/b$
A_1	X_{11}	X_{12}	\dots	X_{1b}	$T_{1.}$	$\bar{X}_{1.}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots
A_a	X_{a1}	X_{a2}	\dots	X_{ab}	$T_{a.}$	$\bar{X}_{a.}$
$T_{.j} = \sum_{i=1}^a X_{ij}$	$T_{.1}$	$T_{.2}$	\dots	$T_{.b}$	$T = \sum_{i=1}^a \sum_{j=1}^b X_{ij}$	
$\bar{X}_{.j} = T_{.j}/a$	$\bar{X}_{.1}$	$\bar{X}_{.2}$	\dots	$\bar{X}_{.b}$		$\bar{X} = \frac{1}{ab} T$

总离差平方和: $SS_T = \sum_{i=1}^a \sum_{j=1}^b (X_{ij} - \bar{X})^2, SS_T = SS_A + SS_B + SS_E$

因素 A/B 的离差平方和: $SS_A = b \sum_{i=1}^a (\bar{X}_{i.} - \bar{X})^2, SS_B = a \sum_{j=1}^b (\bar{X}_{.j} - \bar{X})^2$

误差平方和: $SS_E = \sum_{i=1}^a \sum_{j=1}^b (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X})^2$

双因素（无交互作用）试验的方差分析表

方差来源	平方和 SS	自由度 df	均方和 MS	F 值	F 值临界值
因素 A	SS_A	df_A	$MS_A = \frac{SS_A}{df_A}$	$F_A = \frac{MS_A}{MS_E}$	$F_\alpha((a-1), (a-1)(b-1))$
因素 B	SS_B	df_B	$MS_B = \frac{SS_B}{df_B}$	$F_B = \frac{MS_B}{MS_E}$	$F_\alpha((b-1), (a-1)(b-1))$
误差	SS_E	df_E	$MS_E = \frac{SS_E}{df_E}$		
总和	SS_T	df_T			

$$df_E = df_T - df_A - df_B, \quad SS_E = SS_T - SS_A - SS_B$$

各因素离差平方和的自由度 df_A 、 df_B 为水平数减一，总平方和的自由度 df_T 为试验总次数减一。

双因素有重复（有交互作用）试验资料表

因素 B		B_1	B_2	...	B_b
因素 A	A_1	X_{111}	X_{121}	...	X_{1b1}

		X_{11n}	X_{12n}	...	X_{1bn}
...		• • • •			
		X_{a11}	X_{a21}	...	X_{ab1}
A_a
	X_{a1n}	X_{a2n}	...	X_{abn}	

所有观测值的平均值: $\bar{X} = \frac{1}{abn} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n X_{ijk}$

在条件 (A_i, B_j) 下的平均值: $\bar{X}_{ij\bullet} = \frac{1}{n} \sum_{k=1}^n X_{ijk}$

在条件 A_i 下的平均值: $\bar{X}_{i\bullet\bullet} = \frac{1}{bn} \sum_{j=1}^b \sum_{k=1}^n X_{ijk} = \frac{1}{b} \sum_{j=1}^b \bar{X}_{ij\bullet}$

在条件 B_i 下的平均值: $\bar{X}_{\bullet j\bullet} = \frac{1}{an} \sum_{i=1}^a \sum_{k=1}^n X_{ijk} = \frac{1}{a} \sum_{i=1}^a \bar{X}_{ij\bullet}$

总离差平方和: $SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (X_{ijk} - \bar{X})^2$

可分解为: $SS_T = SS_A + SS_B + SS_{A \times B} + SS_E$

因素 A 的离差平方和:

$$SS_A = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{X}_{i\bullet\bullet} - \bar{X})^2 = bn \sum_{i=1}^a (\bar{X}_{i\bullet\bullet} - \bar{X})^2$$

因素 B 的离差平方和:

$$SS_B = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{X}_{\bullet j\bullet} - \bar{X})^2 = an \sum_{j=1}^b (\bar{X}_{\bullet j\bullet} - \bar{X})^2$$

交互作用的离差平方和:

$$SS_{A \times B} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (\bar{X}_{ij\bullet} - \bar{X}_{i\bullet\bullet} - \bar{X}_{\bullet j\bullet} + \bar{X})^2 = n \sum_{i=1}^a \sum_{j=1}^b (\bar{X}_{ij\bullet} - \bar{X}_{i\bullet\bullet} - \bar{X}_{\bullet j\bullet} + \bar{X})^2$$

双因素 (有重复) 试验的方差分析表

方差来源	平方和	自由度	均方和	F 值	F 值临介值
因素 A	SS_A	df_A	$MS_A = \frac{SS_A}{df_A}$	$F_A = \frac{MS_A}{MS_E}$	$F_\alpha((a-1), ab(n-1))$
因素 B	SS_B	df_B	$MS_B = \frac{SS_B}{df_B}$	$F_B = \frac{MS_B}{MS_E}$	$F_\alpha((b-1), ab(n-1))$
$A \times B$	$SS_{A \times B}$	$df_{A \times B}$	$MS_{A \times B} = \frac{SS_{A \times B}}{df_{A \times B}}$	$F_{A \times B} = \frac{MS_{A \times B}}{MS_E}$	$F_\alpha((a-1)(b-1), ab(n-1))$
误差	SS_E	df_E	$MS_E = \frac{SS_E}{df_E}$		
总和	SS_T	df_T			

$$df_A = a-1, df_B = b-1, df_{A \times B} = df_A \times df_B, df_E = ab(n-1), df_T = abn-1$$

-
- 4.会解读 Excel 的回归分析输出结果, 据此能写出自变量因变量以及回归方程, 判别回归效果 (拟合程度)【见其它文档】
 - 5.会正交设计的极差分析法, 找出各因素影响的主次顺序, 最佳生产工艺