2025-2026-1 数据分析基础复习资料

November 17, 2025

1. 设三维随机向量 $X \sim N_3(\mu, 2I_3)$, 已知

$$\mu = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0.5 & -1 & 0.5 \\ -0.5 & 0 & -0.5 \end{bmatrix}, \quad d = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

试求 $Y = AX + d$ 的分布.

解. 本题考察多元正态的线性变换性质:

若 $X \sim N_p(\mu, \Sigma)$，则

$$Y = AX + b \quad \sim \quad N_q(A\mu + b,\ A\Sigma A^T),$$

其中 A 是 $q \times p$ 矩阵，b 是 $q \times 1$ 向量。

$Y \sim N(A\mu + d, A(2I_3)A^{\mathrm{T}})$.

$$A\mu + d = \begin{bmatrix} 0.5 & -1 & 0.5 \\ -0.5 & 0 & -0.5 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

$$A(2I_3)A^{\mathrm{T}} = 2AA^{\mathrm{T}} = 2\begin{bmatrix} 0.5 & -1 & 0.5 \\ -0.5 & 0 & -0.5 \end{bmatrix} \begin{bmatrix} 0.5 & -0.5 \\ -1 & 0 \\ 0.5 & -0.5 \end{bmatrix} = 2\begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}.$$

2. 已知

$$X = (X_1, X_2, X_3)^{\mathrm{T}} \sim N_3(\mu, \Sigma), \quad \mu = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

(1) 求 $3X_1 - 2X_2 + X_3$ 的分布

(2) 求 $a = (a_1, a_2)'$，使 X_3 与 $X_3 - a' \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ 相互独立

解. (1) 令 $Y = 3X_1 - 2X_2 + X_3$，即

$$Y = [3, -2, 1]X.$$

根据多为正态分布线性变换的性质，若 $X \sim N_3(\mu, \Sigma)$，则 $a^T X \sim N(a^T\mu, a^T\Sigma a)$。这里 $a = (3, -2, 1)^T$。

$$\mathbb{E}(Y) = a^T\mu = 3 \times 2 + (-2) \times (-3) + 1 \times 1 = 6 + 6 + 1 = 13.$$

$$\mathrm{Var}(Y) = a^T \Sigma a = [3, -2, 1] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$$

$$\Sigma a = \begin{bmatrix} 1 \cdot 3 + 1 \cdot (-2) + 1 \cdot 1 \\ 1 \cdot 3 + 3 \cdot (-2) + 2 \cdot 1 \\ 1 \cdot 3 + 2 \cdot (-2) + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 3 - 2 + 1 \\ 3 - 6 + 2 \\ 3 - 4 + 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

于是，

$$\mathrm{Var}(Y) = 3 \times 2 + (-2) \times (-1) + 1 \times 1 = 6 + 2 + 1 = 9.$$

所以方差为 9。

因此，

$$3X_1 - 2X_2 + X_3 \sim N(13, 9).$$

(2) 记 $Z = X_3 - a_1 X_1 - a_2 X_2$。要求 X_3 与 Z 独立。

对于多元正态分布，不相关性与独立性等价。所以只需

$$\mathrm{Cov}(X_3, Z) = 0.$$

利用协方差的性质，

$$\mathrm{Cov}(X_3, Z) = \mathrm{Cov}(X_3, X_3 - a_1 X_1 - a_2 X_2) = \mathrm{Var}(X_3) - a_1 \mathrm{Cov}(X_3, X_1) - a_2 \mathrm{Cov}(X_3, X_2).$$

根据 Σ 的信息，有

$$\mathrm{Var}(X_3) = 2, \quad \mathrm{Cov}(X_3, X_1) = 1, \quad \mathrm{Cov}(X_3, X_2) = 2.$$

代入，得

$$2 - a_1 \cdot 1 - a_2 \cdot 2 = 0$$

即

$$a_1 + 2a_2 = 2.$$

所以解不唯一，例如取 $a_2 = 0$，则 $a_1 = 2$，即 $a = (2, 0)^T$。

易知，一般解为

$$a_1 = 2 - 2t, \quad a_2 = t, \quad t \in \mathbb{R}.$$

取最简单的一个，

$$a = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

此时 $Z = X_3 - 2X_1$，与 X_3 独立。

3. 设

$$\begin{cases} y_1 = a + \varepsilon_1, \\ y_2 = 2a - b + \varepsilon_2, \\ y_3 = a + 2b + \varepsilon_3, \end{cases} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \sim N_3(0, \sigma^2 I_3).$$

试求参数 a, b 的最小二乘估计;

解. 方法一:

最小化残差的平方和

$$S=\sum_{i=1}^{3}(y_i-\hat{y}_i)^2=(y_1-a)^2+(y_2-2a+b)^2+(y_3-a-2b)^2.$$

分别对 a,b 求偏导, 并令其为 0, 得到如下方程:

$$\frac{\partial S}{\partial a}=-2(y_1-a)-4(y_2-2a+b)-2(y_3-a-2b)=0,$$
$$\frac{\partial S}{\partial b}=2(y_2-2a+b)-4(y_3-a-2b)=0.$$

解得 $\hat{a}=\frac{1}{6}(y_1+2y_2+y_3),\hat{b}=\frac{1}{5}(-y_2+2y_3)$.

方法二:

将模型转化为矩阵表示:

$$\begin{cases} y_1=a+\varepsilon_1, \\ y_2=2a-b+\varepsilon_2, \\ y_3=a+2b+\varepsilon_3, \end{cases} \quad 其中\ \varepsilon\sim N_3(0,\sigma^2 I_3).$$

写成矩阵形式:

$$Y=X\beta+\varepsilon,\ 其中\ Y=\begin{bmatrix} y_1\\ y_2\\ y_3 \end{bmatrix},\beta=\begin{bmatrix} a\\ b \end{bmatrix},X=\begin{bmatrix} 1 & 0\\ 2 & -1\\ 1 & 2 \end{bmatrix}.$$

即:

$$Y=\begin{bmatrix} 1 & 0\\ 2 & -1\\ 1 & 2 \end{bmatrix}\begin{bmatrix} a\\ b \end{bmatrix}+\varepsilon.$$

利用课本最小二乘估计的结论，有

$$\hat{\beta}=(X^TX)^{-1}X^TY.$$

于是,

$$X^TX=\begin{bmatrix} 1 & 2 & 1\\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0\\ 2 & -1\\ 1 & 2 \end{bmatrix}=\begin{bmatrix} 1+4+1 & 0-2+2\\ 0-2+2 & 0+1+4 \end{bmatrix}=\begin{bmatrix} 6 & 0\\ 0 & 5 \end{bmatrix}.$$

所以,

$$(X^TX)^{-1}=\begin{bmatrix} \frac{1}{6} & 0\\ 0 & \frac{1}{5} \end{bmatrix}.$$

再由,

$$X^TY=\begin{bmatrix} 1 & 2 & 1\\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} y_1\\ y_2\\ y_3 \end{bmatrix}=\begin{bmatrix} y_1+2y_2+y_3\\ -y_2+2y_3 \end{bmatrix}.$$

故,

$$\hat{a}=\frac{y_1+2y_2+y_3}{6},$$
$$\hat{b}=\frac{-y_2+2y_3}{5}.$$

4. 设 $y_i = \beta_0 + \beta_1 x_i + \beta_2(3x_i^2 - 2) + \varepsilon_i (i = 1, 2, 3), x_1 = -1, x_2 = 0, x_3 = 1$, 其中, $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 相互独立, 服从 $N(0, \sigma^2)$ 正态分布.

(1) 写出矩阵 $Y = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ 中的 $\boldsymbol{X}$.

(2) 求 $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^{\mathrm{T}}$ 的最小二乘估计.

(3) 证明当 $\beta_2 = 0$ 时, β_0 与 β_1 的最小二乘估计不变.

解. (1) 把 $x_1 = -1, \quad x_2 = 0, \quad x_3 = 1$ 代入下式得,

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_1 & 3x_1^2 - 2 \\ 1 & x_2 & 3x_2^2 - 2 \\ 1 & x_3 & 3x_3^2 - 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}.$$

(2) 设 $\boldsymbol{\beta}$ 的最小二乘估计为 $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^{\mathrm{T}}$, 易得, $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$, 于是

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 \\ -\frac{1}{2}y_1 + \frac{1}{2}y_3 \\ \frac{1}{6}y_1 - \frac{1}{3}y_2 + \frac{1}{6}y_3 \end{bmatrix}.$$

(3) 当 $\beta_2 = 0$ 时, 模型变为: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, 2, 3$, 设 $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)^{\mathrm{T}}$, $\boldsymbol{X} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$, $\boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, $(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$, 得

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \begin{bmatrix} \frac{1}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 \\ -\frac{1}{2}y_1 + \frac{1}{2}y_3 \end{bmatrix},$$

即 β_0 与 β_1 的最小二乘估计不变.

5. 对于只有一个自变量的线性回归模型 $Y = \beta_0 + \beta_1 X + \varepsilon$, 利用观测值 $(y_i; x_i)(i = 1, 2, \cdots, n)$ 求 β_0, β_1 的最小二乘估计及 $\sigma^2 = \mathrm{Var}(\varepsilon)$ 的估计, 其中 $x_i (i = 1, 2, \cdots, n)$ 不全相同.

解.

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix},$$

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} = \frac{1}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix}$$

$$= \frac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix},$$

其中 $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$.

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \begin{bmatrix} 1 & 1 & & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ \sum_{i=1}^n x_i y_i \end{bmatrix},$$

其中 $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i$. 故由

$$\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)^{\mathrm{T}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y},$$

式可得 β_0 和 β_1 的最小二乘估计为

$$\begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{bmatrix} = \frac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix} \begin{bmatrix} n\bar{y} \\ \sum_{i=1}^n x_i y_i \end{bmatrix}$$

$$= \frac{1}{n\sum_{i=1}^n (x_i - \bar{x})^2} \begin{bmatrix} n\bar{y}\sum_{i=1}^n x_i^2 - n\bar{x}\sum_{i=1}^n x_i y_i \\ n\sum_{i=1}^n x_i y_i - n^2\bar{x}\bar{y} \end{bmatrix},$$

即

$$\widehat{\beta}_0 = \frac{\bar{y}\sum_{i=1}^n x_i^2 - \bar{x}\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \widehat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

σ^2 的无偏估计为

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2.$$

6. 设随机向量 $\boldsymbol{X} = (X_1, X_2, X_3)^{\mathrm{T}}$ 的协方差矩阵为

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

求 X 的各主成分.

解. 计算 $\boldsymbol{\Sigma}$ 的特征值

$$|\lambda E - \boldsymbol{\Sigma}| = \begin{vmatrix} \lambda - 1 & 2 & 0 \\ 2 & \lambda - 5 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 2)\begin{vmatrix} \lambda - 1 & 2 \\ 2 & \lambda - 5 \end{vmatrix} = (\lambda - 2)(\lambda^2 - 6\lambda + 1) = 0,$$

得 $\lambda_1 = 3 + 2\sqrt{2}, \lambda_2 = 2, \lambda_3 = 3 - 2\sqrt{2}$.

对应于特征值 $\lambda_2 = 2$ 的特征向量为如下方程组的非零解

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = 0$$

于是解得, $\boldsymbol{\xi_2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. 同理可以计算, 对应于 $\lambda_1 = 3 + 2\sqrt{2}$ 的特征向量为 $\boldsymbol{\xi}_1 = \begin{bmatrix} 1-\sqrt{2} \\ 1 \\ 0 \end{bmatrix}$, 对应于 $\lambda_3 = 3 - 2\sqrt{2}$ 的特征向量为 $\boldsymbol{\xi}_3 = \begin{bmatrix} 1 \\ \sqrt{2}-1 \\ 0 \end{bmatrix}$.

这些特征向量显然是正交的, 对上述特征向量单位化处理, 于是矩阵 $\boldsymbol{\Sigma}$ 的特征值及对应的正交单位化特征向量分别为

$$\lambda_1 = 3 + 2\sqrt{2}, \quad \boldsymbol{e}_1 = \begin{bmatrix} \frac{1-\sqrt{2}}{4-2\sqrt{2}} \\ \frac{1}{4-2\sqrt{2}} \\ 0 \end{bmatrix};$$

$$\lambda_2 = 2, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$

$$\lambda_3 = 3 - 2\sqrt{2} \quad \boldsymbol{e}_3 = \begin{bmatrix} \frac{1}{4-2\sqrt{2}} \\ \frac{\sqrt{2}-1}{4-2\sqrt{2}} \\ 0 \end{bmatrix}.$$

因此 $\boldsymbol{X}$ 的三个主成分分别为

$$\begin{aligned} Y_1 &= \boldsymbol{e}_1^{\mathrm{T}} \boldsymbol{X} = \frac{1-\sqrt{2}}{4-2\sqrt{2}} X_1 + \frac{1}{4-2\sqrt{2}} X_2; \\ Y_2 &= \boldsymbol{e}_2^{\mathrm{T}} \boldsymbol{X} = X_3; \\ Y_3 &= \boldsymbol{e}_3^{\mathrm{T}} \boldsymbol{X} = \frac{1}{4-2\sqrt{2}} X_1 + \frac{\sqrt{2}-1}{4-2\sqrt{2}} X_2. \end{aligned}$$

由上述结果, 第一主成分的贡献率为

$$\frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \approx 73\%,$$

前两个主成分的累计贡献率为

$$\frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \approx 98\%.$$

7. 设三元总体 X 的协方差阵为

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

试求总体主成分.

解. $\boldsymbol{\Sigma}$ 的特征值为 $\lambda_1 = 4, \lambda_2 = 4, \lambda_3 = 2$, 对应的正交单位特征向量为 $\boldsymbol{\xi}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\boldsymbol{\xi}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\boldsymbol{\xi}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. 总体主成分 $Y_i = X_i, i = 1, 2, 3$.

8. 设 $\boldsymbol{X}=(X_1,X_2)^T$ 的协方差矩阵为

$$\boldsymbol{\Sigma}=\begin{bmatrix}1 & 4\\ 4 & 100\end{bmatrix},$$

相应的相关系数矩阵为

$$\boldsymbol{\rho}=\begin{bmatrix}1 & 0.4\\ 0.4 & 1\end{bmatrix},$$

分别从 $\boldsymbol{\Sigma}$ 和 $\boldsymbol{\rho}$ 出发进行主成分分析.

9. 设总体 $\boldsymbol{X}=(X_1,X_2,X_3)^{\mathrm{T}}$ 的协方差矩阵为

$$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma^2 & \sigma^2\rho & 0\\ \sigma^2\rho & \sigma^2 & \sigma^2\rho\\ 0 & \sigma^2\rho & \sigma^2\end{bmatrix},\quad 0<\rho<\frac{1}{\sqrt{2}},$$

求 $\boldsymbol{X}$ 的主成分以及各主成分的贡献率.

解.

$$\Sigma=\sigma^2\begin{bmatrix}1 & \rho & 0\\ \rho & 1 & \rho\\ 0 & \rho & 1\end{bmatrix},\quad 0<\rho<\frac{1}{\sqrt{2}}.$$

令 $A=\begin{bmatrix}1 & \rho & 0\\ \rho & 1 & \rho\\ 0 & \rho & 1\end{bmatrix}$. 于是,矩阵 A 的特征多项式为 $\det(A-\lambda I)=\begin{vmatrix}1-\lambda & \rho & 0\\ \rho & 1-\lambda & \rho\\ 0 & \rho & 1-\lambda\end{vmatrix}$.
按第一行展开, 得

$$\begin{aligned}\det(A-\lambda I)&=(1-\lambda)\left[(1-\lambda)^2-\rho^2\right]-\rho\left[\rho(1-\lambda)-0\right]+0\\ &=(1-\lambda)\left[(1-\lambda)^2-\rho^2\right]-\rho^2(1-\lambda)\\ &=(1-\lambda)\left[(1-\lambda)^2-\rho^2-\rho^2\right]\\ &=(1-\lambda)\left[(1-\lambda)^2-2\rho^2\right].\end{aligned}$$

令上式为 0，得

$$(1-\lambda)\left[(1-\lambda)^2-2\rho^2\right]=0.$$

解得

$$\lambda=1,\quad \lambda=1\pm\sqrt{2}\,\rho.$$

因为 $0<\rho<\frac{1}{\sqrt{2}}$，所以

$$1+\sqrt{2}\rho>1>1-\sqrt{2}\rho>0.$$

于是特征值从大到小为:

$$\lambda_1=1+\sqrt{2}\rho,\quad \lambda_2=1,\quad \lambda_3=1-\sqrt{2}\rho.$$

于是，Σ 的特征值为

$$l_1=\sigma^2(1+\sqrt{2}\rho),\quad l_2=\sigma^2,\quad l_3=\sigma^2(1-\sqrt{2}\rho).$$

接下来，分别求所有特征值对应的特征向量。

（1）$\lambda = 1+\sqrt{2}\rho$ 解 $(A-(1+\sqrt{2}\rho)I)v=0$，即

$$\begin{bmatrix} -\sqrt{2}\rho & \rho & 0 \\ \rho & -\sqrt{2}\rho & \rho \\ 0 & \rho & -\sqrt{2}\rho \end{bmatrix} v = 0.$$

第一行：$-\sqrt{2}\rho v_1 + \rho v_2 = 0 \implies v_2 = \sqrt{2}v_1$。第三行：$\rho v_2 - \sqrt{2}\rho v_3 = 0 \implies v_2 = \sqrt{2}v_3 \implies v_3 = v_1$。第二行自动满足（代 v_2, v_3 可得 0）。取 $v_1 = 1$，则 $v = (1, \sqrt{2}, 1)^T$，模长 $\sqrt{1+2+1} = 2$，单位化，得

$$\mathbf{u}_1 = \frac{1}{2}(1, \sqrt{2}, 1)^T.$$

（2）$\lambda = 1$ 解 $(A-I)v = 0$，即

$$\begin{bmatrix} 0 & \rho & 0 \\ \rho & 0 & \rho \\ 0 & \rho & 0 \end{bmatrix} v = 0.$$

第一行：$\rho v_2 = 0 \implies v_2 = 0$。第三行：$\rho v_2 = 0$ 自动满足。第二行：$\rho v_1 + \rho v_3 = 0 \implies v_1 + v_3 = 0 \implies v_3 = -v_1$。取 $v_1 = 1$，则 $v = (1, 0, -1)^T$，模长 $\sqrt{2}$，单位化，得

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 0, -1)^T.$$

（3）$\lambda = 1-\sqrt{2}\rho$ 解 $(A-(1-\sqrt{2}\rho)I)v=0$，即

$$\begin{bmatrix} \sqrt{2}\rho & \rho & 0 \\ \rho & \sqrt{2}\rho & \rho \\ 0 & \rho & \sqrt{2}\rho \end{bmatrix} v = 0.$$

第一行：$\sqrt{2}\rho v_1 + \rho v_2 = 0 \implies v_2 = -\sqrt{2}v_1$。第三行：$\rho v_2 + \sqrt{2}\rho v_3 = 0 \implies v_2 + \sqrt{2}v_3 = 0 \implies v_3 = -v_2/\sqrt{2} = v_1$。第二行自动满足。取 $v_1 = 1$，则 $v = (1, -\sqrt{2}, 1)^T$，模长 2，单位化，得

$$\mathbf{u}_3 = \frac{1}{2}(1, -\sqrt{2}, 1)^T.$$

于是，主成分 $Y_k = \mathbf{u}_k^T X$，即，

$$Y_1 = \frac{1}{2}(X_1 + \sqrt{2}X_2 + X_3), \quad \text{方差} = \sigma^2(1+\sqrt{2}\rho),$$

$$Y_2 = \frac{1}{\sqrt{2}}(X_1 - X_3), \quad \text{方差} = \sigma^2,$$

$$Y_3 = \frac{1}{2}(X_1 - \sqrt{2}X_2 + X_3), \quad \text{方差} = \sigma^2(1-\sqrt{2}\rho).$$

总方差：

$$\mathrm{tr}(\Sigma) = 3\sigma^2.$$

各主成分贡献率 = 对应特征值 / 总方差：

$$\text{贡献率}(Y_1) = \frac{\sigma^2(1+\sqrt{2}\rho)}{3\sigma^2} = \frac{1+\sqrt{2}\rho}{3},$$

$$\text{贡献率}(Y_2) = \frac{\sigma^2}{3\sigma^2} = \frac{1}{3},$$

$$\text{贡献率}(Y_3) = \frac{\sigma^2(1-\sqrt{2}\rho)}{3\sigma^2} = \frac{1-\sqrt{2}\rho}{3}.$$

10. 设总体 $X=(X_1,X_2,X_3)^{\mathrm{T}}$ 的相关系数矩阵为

$$\boldsymbol{\rho}=\begin{bmatrix}1&\rho&\rho\\ \rho&1&\rho\\ \rho&\rho&1\end{bmatrix},\quad \rho>0,$$

求 $\boldsymbol{X}$ 的标准化变量的主成分及其贡献率.

解.

$$\begin{aligned}|\lambda E-\boldsymbol{\Sigma}|&=\begin{vmatrix}\lambda-1&-\rho&-\rho\\ -\rho&\lambda-1&-\rho\\ -\rho&-\rho&\lambda-1\end{vmatrix}=\begin{vmatrix}-2\rho+\lambda-1&-\rho&-\rho\\ -2\rho+\lambda-1&\lambda-1&-\rho\\ -2\rho+\lambda-1&-\rho&\lambda-1\end{vmatrix}\\ &=(-2\rho+\lambda-1)\begin{vmatrix}1&-\rho&-\rho\\ 1&\lambda-1&-\rho\\ 1&-\rho&\lambda-1\end{vmatrix}\\ &=(-2\rho+\lambda-1)\begin{vmatrix}1&-\rho&-\rho\\ 0&\lambda-1+\rho&-\rho\\ 0&0&\lambda-1+\rho\end{vmatrix}\\ &=(-2\rho+\lambda-1)(\lambda-1+\rho)^2=0.\end{aligned}$$

于是, $\lambda_1=2\rho+1,\lambda_2=\lambda_3=1-\rho$. 可求得一组正交化单位特征向量:

$$\boldsymbol{\xi}_1=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}});$$

$$\boldsymbol{\xi}_2=(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0);$$

$$\boldsymbol{\xi}_3=(-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}}).$$

主成分分别为 $Y_i=\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{X}(i=1,2,3)$, Y_1 的贡献率为 $\frac{2\rho+1}{3}$, Y_2 和 Y_3 的贡献率为 $\frac{1-\rho}{3}$.

11. 已知 $\boldsymbol{X}=(X_1,X_2,X_3)^{\mathrm{T}}$ 的协方差矩阵为 $\begin{pmatrix}11&\frac{\sqrt{3}}{2}&\frac{3}{2}\\ \frac{\sqrt{3}}{2}&\frac{21}{4}&\frac{5\sqrt{3}}{4}\\ \frac{3}{2}&\frac{5\sqrt{3}}{4}&\frac{31}{4}\end{pmatrix}$, 试求出其主成分 Y_1,Y_2,Y_3, 并求它们的贡献率.

参考答案: $Y_1=\frac{\sqrt{3}}{2}X_1+\frac{1}{4}X_2+\frac{\sqrt{3}}{4}X_3,\mathrm{Var}(Y_1)=12$, 贡献率为 $\frac{12}{12+8+4}=50\%$;

$Y_2=\frac{-1}{2}X_1+\frac{\sqrt{3}}{4}X_2+\frac{\sqrt{3}}{4}X_3,\mathrm{Var}(Y_2)=8$, 贡献率为 $\frac{8}{12+8+4}=33.33\%$;

$Y_3=-\frac{\sqrt{3}}{2}X_2+\frac{1}{2}X_3,\mathrm{Var}(Y_3)=4$, 贡献率为 $\frac{4}{12+8+4}=16.67\%$;

12. 设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_p)^{\mathrm{T}}$ 的协方差矩阵为 $\boldsymbol{\Sigma}=(\sigma_{ij})_{p\times p},\boldsymbol{P}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_p)$ 为由 $\boldsymbol{\Sigma}$ 的 p 个正交单位化特征向量为列所构成的矩阵,$\boldsymbol{Y}=(Y_1,Y_2,\cdots,Y_p)^{\mathrm{T}}$ 为 $\boldsymbol{X}$ 的 p 个主成分所构成的向量.

(1) 证明 $\boldsymbol{Y}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X}$.

(2) 利用 (1) 的结果求第 k 个主成分 Y_k 与各个 $X_j(j=1,2,\cdots,p)$ 的相关系数, 它对主成分的解释有何作用?

(3) 在样本主成分下的情况又如何?

解. (1) 证明 $Y = P^T X$

已知，Σ 是 X 的协方差矩阵。$P = (e_1, e_2, \ldots, e_p)$，其中 e_k 是 Σ 的单位正交特征向量，对应特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$。由主成分的定义，有 $Y_k = e_k^T X$。

于是，

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} e_1^T X \\ e_2^T X \\ \vdots \\ e_p^T X \end{bmatrix} = \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_p^T \end{bmatrix} X = P^T X.$$

因为 P 是正交矩阵（$P^T P = I$），所以 $P^T = P^{-1}$。

(2) 第 k 个主成分 Y_k 与 X_j 的相关系数

记 $Y_k = e_k^T X = \sum_{t=1}^p e_{kt} X_t$. 已知 $\mathrm{Var}(Y_k) = \lambda_k, \quad \mathrm{Var}(X_j) = \sigma_{jj}$. 于是，$Y_k$ 和 X_j 的协方差为

$$\mathrm{Cov}(Y_k, X_j) = \mathrm{Cov}\left(\sum_{t=1}^p e_{kt} X_t,\ X_j\right) = \sum_{t=1}^p e_{kt}\, \mathrm{Cov}(X_t, X_j) = \sum_{t=1}^p e_{kt} \sigma_{tj}.$$

注意 $\Sigma e_k = \lambda_k e_k$，即 $\sum_{t=1}^p \sigma_{jt} e_{kt} = \lambda_k e_{kj}$（这里 t 是求和下标，注意 $\sigma_{jt} = \sigma_{tj}$ 因为对称性）。所以，

$$\mathrm{Cov}(Y_k, X_j) = \sum_{t=1}^p e_{kt} \sigma_{tj} = \sum_{t=1}^p \sigma_{jt} e_{kt} = \lambda_k e_{kj}.$$

因此相关系数，

$$\rho_{Y_k, X_j} = \frac{\mathrm{Cov}(Y_k, X_j)}{\sqrt{\mathrm{Var}(Y_k) \cdot \mathrm{Var}(X_j)}} = \frac{\lambda_k e_{kj}}{\sqrt{\lambda_k \cdot \sigma_{jj}}} = \frac{\sqrt{\lambda_k}\, e_{kj}}{\sqrt{\sigma_{jj}}}.$$

这个相关系数反映了 X_j 对主成分 Y_k 的"贡献"大小，可用于解释主成分的实际意义。如果某个 Y_k 与某些 X_j 的相关系数绝对值很大，说明 Y_k 主要代表了这些原始变量的信息。这比只看主成分载荷 e_{kj} 多考虑了 X_j 的方差和主成分的方差，更标准化。

(3) 样本主成分的情况

在样本情况下，用样本协方差矩阵 S 代替 Σ。S 的特征值 $\hat{\lambda}_k$，特征向量 $\hat{e}_k$。于是，样本主成分：$\hat{Y}_k = \hat{e}_k^T X$（对中心化的数据）。样本相关系数公式类似，用 $\hat{\lambda}_k$、$\hat{e}_{kj}$、s_{jj}（样本方差）代替相应量：

$$\hat{\rho}_{Y_k, X_j} = \frac{\sqrt{\hat{\lambda}_k}\, \hat{e}_{kj}}{\sqrt{s_{jj}}}.$$

13. 已知总体 $G_i(m = 1)$ 的分布为 $N(\mu^{(i)}, \sigma_i^2)(i = 1, 2)$, 按距离判别准则为 (不妨设 $\mu^{(1)} > \mu^{(2)}$)

$$\begin{cases} x \in G_1, & 若x > \mu^*, \\ x \in G_2, & 若x \leqslant \mu^*, \end{cases}$$

其中 $\mu^* = \frac{\sigma_1 \mu^{(2)} + \sigma_2 \mu^{(1)}}{\sigma_1 + \sigma_2}$. 试求错判概率 $P(2|1)$ 和 $P(1|2)$.

解. 错判概率 $P(2|1)$ 表示来自 G_1 的个体被判为 G_2 的概率。来自 G_1 时，$X \sim N(\mu^{(1)}, \sigma_1^2)$。判为 G_2 当且仅当 $X \leq \mu^*$。所以，

$$P(2|1) = P\left(X \leq \mu^* \mid X \sim N(\mu^{(1)}, \sigma_1^2)\right) = \Phi\left(\frac{\mu^* - \mu^{(1)}}{\sigma_1}\right),$$

其中 Φ 是标准正态分布函数。

代入 μ^*，得

$$\mu^* - \mu^{(1)} = \frac{\sigma_1\mu^{(2)} + \sigma_2\mu^{(1)}}{\sigma_1+\sigma_2} - \mu^{(1)} = \frac{\sigma_1\mu^{(2)} + \sigma_2\mu^{(1)} - \sigma_1\mu^{(1)} - \sigma_2\mu^{(1)}}{\sigma_1+\sigma_2}$$
$$= \frac{\sigma_1(\mu^{(2)} - \mu^{(1)})}{\sigma_1+\sigma_2}.$$

因此,

$$\frac{\mu^* - \mu^{(1)}}{\sigma_1} = \frac{\mu^{(2)} - \mu^{(1)}}{\sigma_1+\sigma_2}.$$

记 $d = \mu^{(1)} - \mu^{(2)} > 0$, 则,

$$\frac{\mu^* - \mu^{(1)}}{\sigma_1} = -\frac{d}{\sigma_1+\sigma_2}.$$

所以,

$$P(2|1) = \Phi\left(-\frac{d}{\sigma_1+\sigma_2}\right) = 1 - \Phi\left(\frac{d}{\sigma_1+\sigma_2}\right).$$

错判概率 $P(1|2)$ 表示来自 G_2 的个体被判为 G_1 的概率。来自 G_2 时, $X \sim N(\mu^{(2)}, \sigma_2^2)$。判为 G_1 当且仅当 $X > \mu^*$。所以,

$$P(1|2) = P\left(X > \mu^* \mid X \sim N(\mu^{(2)}, \sigma_2^2)\right) = 1 - \Phi\left(\frac{\mu^* - \mu^{(2)}}{\sigma_2}\right).$$

$$\begin{aligned}\mu^* - \mu^{(2)} &= \frac{\sigma_1\mu^{(2)} + \sigma_2\mu^{(1)}}{\sigma_1+\sigma_2} - \mu^{(2)} \\ &= \frac{\sigma_1\mu^{(2)} + \sigma_2\mu^{(1)} - \sigma_1\mu^{(2)} - \sigma_2\mu^{(2)}}{\sigma_1+\sigma_2} \\ &= \frac{\sigma_2(\mu^{(1)} - \mu^{(2)})}{\sigma_1+\sigma_2} \\ &= \frac{\sigma_2 d}{\sigma_1+\sigma_2}.\end{aligned}$$

因此,

$$\frac{\mu^* - \mu^{(2)}}{\sigma_2} = \frac{d}{\sigma_1+\sigma_2}.$$

于是,

$$P(1|2) = 1 - \Phi\left(\frac{d}{\sigma_1+\sigma_2}\right).$$

故 $P(2 \mid 1) = P(1 \mid 2) = 1 - \Phi\left(\frac{\mu_1-\mu_2}{\sigma_1+\sigma_2}\right)$.

14. 已知两个总体 $G_1 : X \sim N_p(\mu^{(1)}, \Sigma), G_2 : X \sim N_p(\mu^{(2)}, \Sigma)$,$\mu^{(1)}$, $\mu^{(2)}$, 协方差矩阵 Σ 均已知。又设线性判别函数为

$$W(X) = (X - \bar{\mu})^T \Sigma^{-1} (\mu^{(1)} - \mu^{(2)}),$$

其中 $\bar{\mu} = \frac{\mu^{(1)}+\mu^{(2)}}{2}$. 判别准则为:

$$\begin{cases} W(X) > 0 & \Rightarrow \quad \text{判为} G_1, \\ W(X) \le 0 & \Rightarrow \quad \text{判为} G_2. \end{cases}$$

试求错判概率 $P(2|1)$ 和 $P(1|2)$。

解. 由于 $P(2|1) = P(W(X) \le 0 \mid G_1)$。故需要求当 $X \in G_1$ 时 $W(X)$ 的分布。

设 $X \sim N_p(\mu^{(1)}, \Sigma)$。令 $\delta = \mu^{(1)} - \mu^{(2)}$，则

$$W(X) = (X - \bar{\mu})^T \Sigma^{-1} \delta.$$

这是 X 的一个线性函数，所以 $W(X)$ 服从一元正态分布。

先求均值 $E[W(X) \mid G_1]$，

$$E[X \mid G_1] = \mu^{(1)},$$

$$E[W(X) \mid G_1] = (\mu^{(1)} - \bar{\mu})^T \Sigma^{-1} \delta.$$

又因为，

$$\mu^{(1)} - \bar{\mu} = \mu^{(1)} - \frac{\mu^{(1)} + \mu^{(2)}}{2} = \frac{\mu^{(1)} - \mu^{(2)}}{2} = \frac{\delta}{2}.$$

所以，

$$E[W(X) \mid G_1] = \left(\frac{\delta}{2}\right)^T \Sigma^{-1} \delta = \frac{1}{2} \delta^T \Sigma^{-1} \delta.$$

记马氏距离平方，

$$\Delta^2 = \delta^T \Sigma^{-1} \delta.$$

则，

$$E[W(X) \mid G_1] = \frac{\Delta^2}{2}.$$

再求方差 $\mathrm{Var}[W(X) \mid G_1]$，

$$W(X) = \delta^T \Sigma^{-1} X - \delta^T \Sigma^{-1} \bar{\mu}.$$

常数部分不影响方差，所以

$$\mathrm{Var}[W(X) \mid G_1] = \mathrm{Var}(\delta^T \Sigma^{-1} X) = \delta^T \Sigma^{-1} \cdot \mathrm{Cov}(X) \cdot \Sigma^{-1} \delta.$$

由于 $\mathrm{Cov}(X) = \Sigma$，所以

$$\mathrm{Var}[W(X) \mid G_1] = \delta^T \Sigma^{-1} \Sigma \Sigma^{-1} \delta = \delta^T \Sigma^{-1} \delta = \Delta^2.$$

因此

$$W(X) \mid G_1 \sim N\left(\frac{\Delta^2}{2},\ \Delta^2\right).$$

由于 $P(1|2) = P(W(X) > 0 \mid G_2)$。故需要计算当 $X \in G_2$ 时 $W(X)$ 的分布。类似地，$X \sim N_p(\mu^{(2)}, \Sigma)$ 时，

$$E[X \mid G_2] = \mu^{(2)},$$

$$\mu^{(2)} - \bar{\mu} = \mu^{(2)} - \frac{\mu^{(1)} + \mu^{(2)}}{2} = -\frac{\delta}{2}.$$

所以，

$$E[W(X) \mid G_2] = \left(-\frac{\delta}{2}\right)^T \Sigma^{-1} \delta = -\frac{1}{2} \Delta^2.$$

方差与 G_1 情况相同（因为方差与均值无关），

$$\mathrm{Var}[W(X) \mid G_2] = \Delta^2.$$

因此，

$$W(X) \mid G_2 \sim N\left(-\frac{\Delta^2}{2},\ \Delta^2\right).$$

下面计算错判概率 $P(2|1)$，其中 $P(2|1) = P(W(X) \le 0 \mid G_1)$。

令

$$Z = \frac{W(X) - \frac{\Delta^2}{2}}{\Delta} \sim N(0,1) \quad (\text{在}G_1\text{下}).$$

当 $W(X) = 0$ 时，

$$Z = \frac{0 - \frac{\Delta^2}{2}}{\Delta} = -\frac{\Delta}{2}.$$

所以，

$$P(2|1) = \Phi\left(-\frac{\Delta}{2}\right) = 1 - \Phi\left(\frac{\Delta}{2}\right).$$

下面计算错判概率 $P(1|2)$，其中 $P(1|2) = P(W(X) > 0 \mid G_2)$。

标准化（在 G_2 下），

$$Z = \frac{W(X) + \frac{\Delta^2}{2}}{\Delta} \sim N(0,1).$$

当 $W(X) = 0$ 时，

$$Z = \frac{0 + \frac{\Delta^2}{2}}{\Delta} = \frac{\Delta}{2}.$$

所以，

$$P(1|2) = 1 - \Phi\left(\frac{\Delta}{2}\right).$$

于是，

$$P(2|1) = 1 - \Phi\left(\frac{\Delta}{2}\right),$$

$$P(1|2) = 1 - \Phi\left(\frac{\Delta}{2}\right),$$

其中 $\Delta^2 = (\mu^{(1)} - \mu^{(2)})^T \Sigma^{-1} (\mu^{(1)} - \mu^{(2)})$。

15. 设有 3 个组 G_1, G_2 和 G_3, 欲判别某样品 x_0 属于何组, 已知先验概率 $p_1 = 0.05, p_2 = 0.65, p_3 = 0.30$, 且 $f_1(x_0) = 0.10, f_2(x_0) = 0.63, f_3(x_0) = 2.4$. 求 x_0 属于各组的后验概率, 并判别 x_0 应该属于哪一组.

解.

$$P(G_1|x_0) = \frac{p_1 f_1(x_0)}{\sum_{i=1}^3 p_i f_i(x_0)} = \frac{0.05 \times 0.10}{0.05 \times 0.10 + 0.65 \times 0.63 + 0.30 \times 2.4} = \frac{0.005}{1.1345} = 0.004;$$

同理可得

$$P(G_2|x_0) = \frac{p_2 f_2(x_0)}{\sum_{i=1}^3 p_i f_i(x_0)} = \frac{0.65 \times 0.63}{1.1345} = 0.361;$$

$$P(G_3|x_0) = \frac{p_3 f_3(x_0)}{\sum_{i=1}^3 p_i f_i(x_0)} = \frac{0.30 \times 2.4}{1.1345} = 0.635.$$

因此, 应将 x_0 判给 G_3.

16. 设三个总体 G_1, G_2 和 G_3 的分布分别为：$N(2, 0.5^2)$、$N(0, 2^2)$ 和 $N(3, 1^2)$。试问样品 $x = 2.5$ 应判归哪一类?

（1）按距离判别准则；

（2）按贝叶斯判别准则

$$取 q_1 = q_2 = q_3 = \frac{1}{3}, \quad c(j|i) = \begin{cases} 1, & i \neq j \\ 0, & i = j \end{cases}$$

解. (1) 距离判别准则

距离判别：把 x 判给马氏距离最小的总体（对于一维正态，马氏距离 $= |x - \mu_i|/\sigma_i$ 的平方）。

计算每个总体的平方马氏距离：

$$d_i^2(x) = \frac{(x - \mu_i)^2}{\sigma_i^2}.$$

$G_1 : \mu_1 = 2, \sigma_1 = 0.5$

$$d_1^2 = \frac{(2.5 - 2)^2}{0.5^2} = \frac{0.25}{0.25} = 1.$$

$G_2 : \mu_2 = 0, \sigma_2 = 2$

$$d_2^2 = \frac{(2.5 - 0)^2}{2^2} = \frac{6.25}{4} \approx 1.5625.$$

$G_3 : \mu_3 = 3, \sigma_3 = 1$

$$d_3^2 = \frac{(2.5 - 3)^2}{1^2} = \frac{0.25}{1} = 0.25.$$

最小的是 $d_3^2 = 0.25$，所以判给 G_3。

(2) 贝叶斯判别准则

已知：

$$q_1 = q_2 = q_3 = \frac{1}{3}, \quad c(j|i) = \begin{cases} 1, & i \neq j, \\ 0, & i = j. \end{cases}$$

贝叶斯判别：把 x 判给后验概率最大的总体。

后验概率：

$$P(G_i \mid x) = \frac{q_i f_i(x)}{\sum_{k=1}^{3} q_k f_k(x)}.$$

因为 q_i 都相等，所以只需比较 $f_i(x)$。

计算概率密度（忽略常数 $\frac{1}{\sqrt{2\pi}}$，比较核心部分）：

$$f_i(x) \propto \frac{1}{\sigma_i} \exp\left[-\frac{1}{2}\frac{(x - \mu_i)^2}{\sigma_i^2}\right].$$

比较时可以用对数似然（加常数不影响大小顺序）。

令

$$g_i(x) = \ln f_i(x) = -\ln \sigma_i - \frac{1}{2}\frac{(x - \mu_i)^2}{\sigma_i^2} + 常数.$$

去掉公共常数，比较：

$$h_i(x) = -\ln \sigma_i - \frac{1}{2}\frac{(x - \mu_i)^2}{\sigma_i^2}.$$

$G_1 : \sigma_1 = 0.5$

$$h_1 = -\ln 0.5 - \frac{1}{2} \cdot 1 = 0.6931 - 0.5 = 0.1931.$$

$G_2 : \sigma_2 = 2$

$$h_2 = -\ln 2 - \frac{1}{2} \cdot 1.5625 = -0.6931 - 0.78125 = -1.47435.$$

$G_3 : \sigma_3 = 1$

$$h_3 = -\ln 1 - \frac{1}{2} \cdot 0.25 = 0 - 0.125 = -0.125.$$

最大的是 $h_1 \approx 0.1931$，所以判给 G_1。