2025 春季本科计量经济学

第 1 次作业参考答案

1. 对 2 元随机变量 X, Y，利用 $\mathrm{cov}(X, Y)$ 的对称双线性特征，证明相关系数 $\rho_{XY} \in [-1, 1]$。提示：对任意 $a \in \mathbb{R}$，考察 $\mathrm{var}(aX + Y) = \mathrm{cov}(aX + Y, aX + Y)$ 的展开式，并利用关于 a 的二次函数判别式，证明相应结论。

   证明：利用协方差的对称双线性特征，展开 $\mathrm{var}(aX + Y)$

$$\begin{aligned}\mathrm{var}(aX+Y) &= \mathrm{cov}(aX+Y, aX+Y) \\ &= \mathrm{cov}(aX, aX) + \mathrm{cov}(aX, Y) + \mathrm{cov}(Y, aX) + \mathrm{cov}(Y, Y) \\ &= a^2\mathrm{var}(X) + 2a\mathrm{cov}(X, Y) + \mathrm{var}(Y)\end{aligned}$$

   将其视作关于 a 的函数，令

$$f(a) = a^2\mathrm{var}(X) + 2a\mathrm{cov}(X, Y) + \mathrm{var}(Y)$$

   由于对于任意实数 a，方差总是非负的，即 $\mathrm{var}(aX + Y) \geq 0$，所以函数 $f(a) \geq 0$ 对所有 $a \in \mathbb{R}$ 成立，即二次函数 $f(a)$ 的判别式为

$$\Delta = [2\mathrm{cov}(X, Y)]^2 - 4\mathrm{var}(X)\mathrm{var}(Y) = 4\left[\mathrm{cov}(X, Y)^2 - \mathrm{var}(X)\mathrm{var}(Y)\right] \leq 0$$

   即

$$\mathrm{cov}(X, Y)^2 \leq \mathrm{var}(X)\mathrm{var}(Y)$$

   相关系数的定义为

$$\rho_{XY} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}}$$

   所以

$$\rho_{XY}^2 = \left(\frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}}\right)^2 = \frac{\mathrm{cov}(X, Y)^2}{\mathrm{var}(X)\mathrm{var}(Y)} \leq 1$$

   即

$$-1 \leq \rho_{XY} \leq 1$$

2. 对样本 $\{(X_i, Y_i)\}_{i=1}^N$，请证明样本相关系数

$$\hat{\rho}_{XY} = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_x\hat{\sigma}_Y} \in [-1, 1],$$

   其中样本协方差 $\hat{\sigma}_{XY} = \frac{1}{N-1}\sum_{i=1}^N (X_i - \hat{\mu}_X)(Y_i - \hat{\mu}_Y)$，样本方差 $\hat{\sigma}_X^2 = \frac{1}{N-1}\sum_{i=1}^N (X_i - \hat{\mu}_X)^2$，$\hat{\sigma}_Y^2$ 类似，$\hat{\mu}_X, \hat{\mu}_Y$ 为样本均值。提示，与上题类似，将样本协方差看做向量 $\boldsymbol{X} =$

$[X_1,...,X_N]^\intercal, \boldsymbol{Y}=[Y_1,...,Y_N]^\intercal$ 的对称双线性函数，然后对任意 $a\in\mathbb{R}$，考察 $a\boldsymbol{X}+\boldsymbol{Y}$ 的样本方差，并进行展开，得到 $\boldsymbol{X},\boldsymbol{Y}$ 的样本协方差、方差等表达式，再行推导。

证明：定义中心化的样本向量 $\mathbf{x}$,$\mathbf{y}$ 如下

$$\mathbf{x}=\begin{bmatrix} X_1-\hat{\mu}_X \\ X_2-\hat{\mu}_X \\ \vdots \\ X_N-\hat{\mu}_X \end{bmatrix}$$

$$\mathbf{y}=\begin{bmatrix} Y_1-\hat{\mu}_Y \\ Y_2-\hat{\mu}_Y \\ \vdots \\ Y_N-\hat{\mu}_Y \end{bmatrix}$$

因此，样本方差和协方差可以表示为向量内积的形式

$$\begin{aligned}\hat{\sigma}_X^2 &= \frac{1}{N-1}\mathbf{x}^\intercal\mathbf{x}\\ \hat{\sigma}_Y^2 &= \frac{1}{N-1}\mathbf{y}^\intercal\mathbf{y}\\ \hat{\sigma}_{XY} &= \frac{1}{N-1}\mathbf{x}^\intercal\mathbf{y}\\ &= \frac{1}{N-1}\mathbf{y}^\intercal\mathbf{x}\end{aligned}$$

与上题类似，考虑任意实数 a，考察线性组合 $a\boldsymbol{X}+\boldsymbol{Y}$ 的样本方差

$$\begin{aligned}\hat{\sigma}_{aX+Y}^2 &= \frac{1}{N-1}\sum_{i=1}^{N}(aX_i+Y_i-a\hat{\mu}_X-\hat{\mu}_Y)^2\\ &= \frac{1}{N-1}(a\mathbf{x}+\mathbf{y})^\intercal(a\mathbf{x}+\mathbf{y})\\ &= \frac{1}{N-1}\left(a^2\mathbf{x}^\intercal\mathbf{x}+2a\mathbf{x}^\intercal\mathbf{y}+\mathbf{y}^\intercal\mathbf{y}\right)\\ &= a^2\hat{\sigma}_X^2+2a\hat{\sigma}_{XY}+\hat{\sigma}_Y^2\end{aligned}$$

由于样本方差总是非负的，即

$$\hat{\sigma}_{aX+Y}^2 \geq 0, \quad \forall a\in\mathbb{R}$$

因此，关于 a 的二次函数

$$f(a)=a^2\hat{\sigma}_X^2+2a\hat{\sigma}_{XY}+\hat{\sigma}_Y^2\geq 0, \quad \forall a\in\mathbb{R}$$

计算判别式 Δ

$$\begin{aligned}\Delta &= [2\hat{\sigma}_{XY}]^2-4\hat{\sigma}_X^2\hat{\sigma}_Y^2\\ &= 4\hat{\sigma}_{XY}^2-4\hat{\sigma}_X^2\hat{\sigma}_Y^2\\ &= 4(\hat{\sigma}_{XY}^2-\hat{\sigma}_X^2\hat{\sigma}_Y^2)\leq 0\end{aligned}$$

即

$$\left(\frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X\hat{\sigma}_Y}\right)^2\leq 1 \Rightarrow -1\leq\frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X\hat{\sigma}_Y}\leq 1 \Leftrightarrow \hat{\rho}_{XY}\in[-1,1]$$

3. 假设总体分布为 $U([0,a])$，$a>0$，请计算该分布下的期望与方差表达式。进一步假设有 iid 样本 $\{X_i\}_{i=1}^N \overset{\text{iid}}{\sim} U([0,a])$，请利用样本均值与样本方差表达式，分别构造 a 的一个矩估计，记为 $\hat{a}_{1,N}, \hat{a}_{2,N}$，请说明这两个矩估计是否具有无偏性和一致性。

   解：由 $\mathbb{E}[X_i]=\frac{a}{2}, \mathbb{E}[X_i^2]=\frac{a^2}{3}$，可构造

   $$\hat{a}_{1,N}=\frac{2}{N}\sum_{i=1}^N X_i, \qquad \hat{a}_{2,N}=\sqrt{\frac{3}{N-1}\sum_{i=1}^N X_i^2}$$

   直接计算可知 $\mathbb{E}[\hat{a}_{1,N}]=\frac{2}{N}\sum_{i=1}^N \mathbb{E}[X_i]=a$，故 $\hat{a}_{1,N}$ 是无偏估计量。又由大数定律，知 $\hat{a}_{1,N}\xrightarrow{\text{a.s.}} 2\mathbb{E}[X_i]=2\times\frac{a}{2}=a, N\to\infty$，故是一致估计。对 $\hat{a}_{2,N}$，同样由大数定律可知 $\hat{a}_{1,N}\xrightarrow{\text{a.s.}}\sqrt{3\mathbb{E}[X_i^2]}=\sqrt{3\times\frac{a^2}{3}}=a, N\to\infty$，故是一致估计。但由于 $f(x)=\sqrt{x}$ 是严格凹函数，故由 Jensen 不等式可知，$\mathbb{E}[\hat{a}_{2,N}]<\sqrt{3\times\mathbb{E}\left[\frac{1}{N-1}\sum_{i=1}^N X_i^2\right]}=a$，为有篇估计量。

4. 假设 $\{X_i\}_{i=1}^N$ 为 iid 样本，X_i 期望为 μ，方差为 σ^2。

   (a) 定义基于总体均值的样本方差估计为

   $$\hat{s}_N^2=\frac{1}{N}\sum_{i=1}^N (X_i-\mu)^2,$$

   请证明 $\mathbb{E}\hat{s}_N^2=\sigma^2$，且 $\hat{s}_N^2\xrightarrow{\text{a.s.}}\sigma^2$。

   证明：直接计算可知

   $$\mathbb{E}(\hat{s}_N^2)=\frac{1}{N}\mathbb{E}\left[\sum_{i=1}^N (X_i-\mu)^2)\right]=\frac{1}{N}\sum_{i=1}^N \mathbb{E}(X_i-\mu)^2=\frac{1}{N}\times N\times\sigma^2=\sigma^2$$

   故 $\mathbb{E}[\hat{s}_N^2]=\sigma^2$。进一步计算可知

   $$\lim_{N\to\infty}\hat{s}_N^2=\lim_{N\to\infty}\left[\frac{1}{N}\sum_{i=1}^N (X_i-\mu)^2\right]=\lim_{N\to\infty}\left[\frac{\sum_{i=1}^N X_i^2}{N}-2\mu\frac{\sum_{i=1}^N X_i}{N}+\mu^2\right]$$

   由大数定律可以得出

   $$\frac{1}{N}\sum_{i=1}^N X_i\xrightarrow{\text{a.s.}}\mathbb{E}(X)=\mu, \quad \frac{1}{N}\sum_{i=1}^N X_i^2\xrightarrow{\text{a.s.}}\mathbb{E}(X^2)\Longrightarrow \hat{s}_N^2\xrightarrow{\text{a.s.}}\mathbb{E}(X^2)-\mu^2=\sigma^2$$

   (b) 定义基于样本均值的样本方差估计为

   $$\hat{\sigma}_N^2=\frac{1}{N-1}\sum_{i=1}^N (X_i-\hat{\mu}_N)^2,$$

   其中 $\hat{\mu}_N=\frac{1}{N}\sum_{i=1}^N X_i$，请证明 $\mathbb{E}\hat{\sigma}_N^2=\sigma^2$，且 $\hat{\sigma}_N^2\xrightarrow{\text{a.s.}}\sigma^2$。

   证明：计算可知

   $$\begin{aligned}\mathbb{E}(\hat{\sigma}_N^2)&=\frac{1}{N-1}\mathbb{E}\left[\sum_{i=1}^N\left(X_i-\frac{1}{N}\sum_{i=1}^N X_i\right)^2\right]\\&=\frac{1}{N-1}\mathbb{E}\left[\sum_{i=1}^N X_i^2-2\hat{\mu}_N\sum_{i=1}^N X_i+N\hat{\mu}_N^2\right]\end{aligned}$$

并且 $\mathbb{E}\left(\hat{\mu}_N\right) = \mathbb{E}\left(\frac{\sum X_i}{N}\right) = \frac{N\times\mu}{N} = \mu$，$\mathbb{E}(\hat{\mu}_N^2) = \mathrm{var}(\hat{\mu}_N) + [\mathbb{E}(\hat{\mu}_N)]^2 = \frac{1}{N^2}\sum_{i=1}^N \mathrm{var}(X_i) + \mu^2 = \frac{1}{N}\sigma^2 + \mu^2$ 因此有

$$
\begin{aligned}
\mathbb{E}(\hat{\sigma}_N^2) &= \frac{1}{N-1}\left[\mathbb{E}\left(\sum_{i=1}^N X_i^2\right) - 2N\mathbb{E}(\hat{\mu}_N^2) + N\mathbb{E}(\hat{\mu}_N^2)\right] \\
&= \frac{1}{N-1}\left[\sum_{i=1}^N \mathbb{E}(X_i^2) - N\mathbb{E}(\hat{\mu}_N^2)\right] \\
&= \frac{1}{N-1}\left[N(\sigma^2+\mu^2) - \sigma^2 - N\mu^2\right] \\
&= \frac{1}{N-1}\times(N-1)\times\sigma^2 \\
&= \sigma^2
\end{aligned}
$$

对一致性，利用上述计算和大数定律有

$$
\begin{aligned}
\lim_{N\to\infty}\hat{\sigma}_N^2 &= \lim_{N\to\infty}\frac{1}{N-1}\left(\sum_{i=1}^N X_i^2 - 2\hat{\mu}_N\sum_{i=1}^N X_i + N\hat{\mu}_N^2\right) \\
&= \lim_{N\to\infty}\frac{1}{N-1}\sum_{i=1}^N\left(X_i^2 - N\hat{\mu}_N^2\right) \\
&= \lim_{N\to\infty}\frac{N}{N-1}\frac{\sum_{i=1}^N X_i^2}{N} - \lim_{N\to\infty}\frac{N}{N-1}\hat{\mu}_N^2
\end{aligned}
$$

由

$$
\frac{\sum_{i=1}^N X_i^2}{N} \xrightarrow{\text{a.s.}} \mathbb{E}(X^2), \quad \hat{\mu}_N^2 \xrightarrow{\text{a.s.}} \mu
$$

可得

$$
\lim_{N\to\infty}\hat{\sigma}_N^2 = \mathbb{E}(X^2) - \mu^2 = \sigma^2
$$

5. 假设 $\boldsymbol{X} = [X_1,\dots,X_N] \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I})$ 为 N 维标准正态分布，即 $X_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$，请证明，对任意的 N 阶正交矩阵 $\boldsymbol{C}$，即 $\boldsymbol{C}\boldsymbol{C}^\intercal = \boldsymbol{I}_N$，$\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{X}$ 依然服从 N 维标准正态分布。

   证明：注意到 $\boldsymbol{Y}$ 是 $\boldsymbol{X}$ 的仿射（线性）变换，故 $\boldsymbol{Y}$ 的协方差矩阵可写为

$$
\mathrm{var}(\boldsymbol{Y}) = \boldsymbol{C}\mathrm{var}(\boldsymbol{X})\boldsymbol{C}^\intercal = \boldsymbol{C}\boldsymbol{I}\boldsymbol{C}^\intercal = \boldsymbol{C}\boldsymbol{C}^\intercal = \boldsymbol{I}
$$

   且 $\mathbb{E}[\boldsymbol{Y}] = \mathbb{E}[\boldsymbol{C}\boldsymbol{X}] = \boldsymbol{C}\mathbb{E}[\boldsymbol{X}] = \boldsymbol{0}$，再由多元正态分布随机向量仿射变换后依然为多元正态分布，可知 $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I})$

6. 课件 7 中定义的矩阵 $\boldsymbol{M}$ 称为幂等矩阵 (idempotent matrix)。该类矩阵的一般定义如下：若 N 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = \boldsymbol{A}$，则称为幂等矩阵。

   (a) 假设 λ 为 $\boldsymbol{A}$ 的特征值，请说明 $\lambda = 0$ 或 1。提示：请用特征值的等价定义，即存在非零向量 $\boldsymbol{c}$，使得 $\boldsymbol{A}\boldsymbol{c} = \lambda\boldsymbol{c}$，进行计算证明。

   证明：假设特征值 λ 对应的特征向量为 $\boldsymbol{c}$，则 $\boldsymbol{A}\boldsymbol{c} = \lambda\boldsymbol{c}$，两边同时乘以 $\boldsymbol{A}$，可得 $\boldsymbol{A}^2\boldsymbol{c} = \lambda\boldsymbol{A}\boldsymbol{c} = \lambda^2\boldsymbol{c}$。而利用幂等矩阵性质，左端等于 $\boldsymbol{A}\boldsymbol{c} = \lambda\boldsymbol{c}$，故有 $\lambda\boldsymbol{c} = \lambda^2\boldsymbol{c}$，重新整理可得 $\lambda(\lambda-1)\boldsymbol{c} = \boldsymbol{0}$。由 $\boldsymbol{c} \neq \boldsymbol{0}$，可知 $\lambda = 0$ 或 1。

(b) 假设 $\boldsymbol{A}$ 还是一个实数对称矩阵，请利用实对称矩阵特征值分解（第 4 讲课件），即存在正交矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\intercal$，其中 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ 为对角阵，来证明 $\mathrm{tr}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A})$。

证明：若 $\boldsymbol{A}$ 是一个实数对称矩阵，则存在正交阵 $\boldsymbol{C}$，使得 $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\intercal$，其中 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$。而 $\boldsymbol{A}$ 又是幂等矩阵，故由 (a) 可知 $\lambda_1, \ldots, \lambda_N = 0$ 或 1。因此，$\mathrm{tr}(\boldsymbol{\Lambda}) = \mathrm{rank}(\boldsymbol{\Lambda})$。最后，一方面由矩阵在可逆变换（左乘或右乘可逆矩阵）下保持秩不变，可知 $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{\Lambda})$；另一方面由乘积矩阵迹的基本性质，知 $\mathrm{tr}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\intercal) = \mathrm{tr}(\boldsymbol{\Lambda}\boldsymbol{C}^\intercal\boldsymbol{C}) = \mathrm{tr}(\boldsymbol{\Lambda})$。如此可得 $\mathrm{tr}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A})$。