STA 437/2005

Midterm Exam 2, Section LEC0201

Date: March 14, 2025

STA 437/2005, Midterm 2

Student Name: ______________________________

(Please use capital letters)

Student Number: ______________________________

I am taking (please circle): **STA437** **STA2005**

Instructions:

- Fill out your name and student number on this page.
- Carefully edit your answer in the provided space. Use the last two pages for your own notes (this will not be graded).
- Show all your work to receive full credit.
- This exam has three problems. Each problem is on a separate page.
- No electronic devices, notes, or books are allowed during the exam.

Good Luck!

Problem 1 (20 points)

Let $X \sim N_3(\mathbf{0}, \Sigma)$, where the covariance matrix is given by

$$\Sigma = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

1. (10 pts) What is the PCA transformation Z of X in this case?
   Note that there is one eigenvactor $\mathbf{u}_3 := (0, 0, 1)$ with eigenvalue 1. All other eigenvectors are of the form $(*, *, 0)$ so enough to check the eigenvectors of the upper 2×2 submatrix. We have

   $$\det(\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}) = 0 \text{ if and only if } \lambda = 1, 3.$$

   The associated eigenvectors are $\mathbf{u}_2 := \frac{1}{\sqrt{2}}(0, 1, -1)$, $\mathbf{u}_1 := \frac{1}{\sqrt{2}}(0, 1, 1)$ respectively. So the transformation is $Z = U^\top X$ where U has columns $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.

2. (10 pts) Find the distribution of (Z_1, Z_2).
   Since $\Sigma = U \Lambda U^T$, the distribution of Z is $N_2(\mathbf{0}, \mathrm{diag}(3, 1))$.

Problem 2 (20 points)

Consider the random vector $X = (X_1, X_2)$, where X is uniformly distributed on the unit square in $\mathbb{R}^2$, i.e., the probability density function of X is

$$f(x_1, x_2) = \begin{cases} \frac{1}{4}, & \text{if } -1 \le x_1, x_2 \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Answer the following questions.

1. (6 pts) Argue whether X follows spherical distribution.

2. (6 pts) Determine which of the following vectors follow the same distribution as X:

$$Y_1 = \begin{pmatrix} X_2 \\ X_1 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} -X_1 \\ X_2 \end{pmatrix}, \quad Y_3 = \begin{pmatrix} \frac{\sqrt{3}}{2} X_1 - \frac{1}{2} X_2 \\ \frac{1}{2} X_1 + \frac{\sqrt{3}}{2} X_2 \end{pmatrix}.$$

   Choose from the following options and briefly justify your answer:

   A. only Y_2.

   B. only Y_1 and Y_2.

   C. only Y_2 and Y_3.

   D. All of them.

3. (8 pts) True or False (no need to provide justifications):

T / F: X_1, X_2 have the same distribution.

T / F: X_1, X_2 are independent.

T / F: $X_1 \sim U[-1, 1]$.

T / F: The copula of X_1 and X_2 is $C(u, v) = uv$.

Problem 3 (20 points)

Consider a vector $X = (X_1, X_2)$ whose distribution is a mixture of two Gaussian distributions with parameters: $\pi_1 = 0.6$, $\pi_2 = 0.4$ and

$$\mu_1 = (1, 2),\ \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \qquad \mu_2 = (0, 0),\ \Sigma_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

(i) (10 pts) Denote by $f_1(x)$, $f_2(x)$ the densities of the two Gaussian components. Suppose $f_1(3, 3) \approx 0.013$ and $f_2(3, 3) \approx 0.005$. Explain how you can use this information to compute the probability that the observation $(3, 3)$ comes from the first Gaussian component.

We have

$$p(z = 1|\mathbf{x} = (3, 3)) \ = \ \frac{\pi_1 p(\mathbf{x} = (3, 3)|z = 1)}{\pi_1 p(\mathbf{x} = (3, 3)|z = 1) + \pi_2 p(\mathbf{x} = (3, 3)|z = 2)}.$$

Since $p(\mathbf{x} = (3, 3)|z = 1) = 0.013$ and $p(\mathbf{x} = (3, 3)|z = 2) = 0.005$ we have everything we need to use this formula.

(ii) (10 pts) Compute $\mathbb{E}(X_1X_2)$.
We have $\mathbb{E}(X_1X_2) = 0.4 \cdot (0 + 1 \cdot 2) + 0.6 \cdot (1 + 0) = 1.4$

You can use this page for rough work. This page will not be evaluated.

You can use this page for rough work. This page will not be evaluated.