Review Session 4 – Solutions

1 Solutions

1.1 Previous Core Competency Problems

Problem 1. [2018 Summer Practice, # 10] Suppose that $X_1, \dots, X_n \overset{i.i.d.}{\sim} N(0,1)$, and A is an $n \times n$ matrix which is symmetric (i.e., $A^T = A$) and idempotent (i.e., $A^2 = A$). Find the distribution of $\sum_{i,j=1}^n X_i X_j A(i,j)$. Assume if necessary that $\sum_{i=1}^n A(i,i) = s$.

Solution

Since A is symmetric, it is orthonormally diagonalizable $A = QDQ^T$. Then, $Y = \sum X_i X_j A(i,j)$ can be expressed as $X^T AX$ where $X = (X_1, \dots, X_n)^T$. Then, let $Y := (Q^T X)^T D (Q^T X)$. Letting $V = Q^T X$, we have

$$Y = \sum_{i=1}^n D_{ii} V_i^2$$

where $D_{ii} \in \{0,1\}$ since the eigenvalues of A are all 0 or 1 by virtue of A being idempotent (to see this, compare the eigendecompositions of A and A^2). If $V = (V_1, \dots, V_n)^T$, then $V_i \sim N(0,1)$ since Q is orthonormal.

Then, Y is a sum of standard normals squared meaning $Y \sim \chi^2_d$ where d is the multiplicity of eigenvalue 1 in A, or $\mathrm{rank}(A)$.

Problem 2 (2018 Summer Practice, # 17)**.** Let X and Y be i.i.d. $\mathcal{N}(0,1)$ random variables. Consider

$$Z := \text{sign}(Y) \cdot X$$

where sign$(y) := 1$ if $y > 0$ and sign$(y) := -1$ if $y \leq 0$.

(a) Find the distribution of Z.

(b) Compute the covariance of X and Z.

(c) Determine $P[X + Z = 0]$.

(d) Are X and Z independent? (Give a precise mathematical argument).

Solution

(a) We have

$$\begin{aligned}
\mathbb{P}(Z \leq z_0) &= \mathbb{P}(X \leq z_0 | \operatorname{sgn}(Y) = 1) \cdot \mathbb{P}(\operatorname{sgn}(Y) = 1) + \mathbb{P}(-X \leq z_0 | \operatorname{sgn}(Y) = -1) \cdot \mathbb{P}(\operatorname{sgn}(Y) = -1) \\
&= \frac{1}{2} \cdot \mathbb{P}(X \leq z_0) + \mathbb{P}(X \geq -z_0) \cdot \frac{1}{2} \\
&= \mathbb{P}(X \leq z_0) \implies Z \sim N(0,1)
\end{aligned}$$

(b) The covariance is

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[XZ | \operatorname{sgn}(Y)]] = \frac{1}{2}\mathbb{E}[XZ | \operatorname{sgn}(Y) = 1] + \frac{1}{2}\mathbb{E}[XZ | \operatorname{sgn}(Y) = -1] = \frac{1}{2}\mathbb{E}[X^2] - \frac{1}{2}\mathbb{E}[X^2] = 0$$

(c) We have

$$\mathbb{P}(X + Z = 0) = \frac{1}{2}\mathbb{P}(X + X = 0) + \frac{1}{2}\mathbb{P}(X - X = 0) = \frac{1}{2}$$

(d) No, X, Z are not independent. If they were independent then $X + Z$ would be Gaussian as well, being a sum of independent Gaussians, but $\mathbb{P}(X + Z = 0) > 0$ from (c), a contradiction. We could also cite the following observation:

$$\mathbb{P}(Z > 1 | X \in [-1/2, 1/2]) = 0 \neq \mathbb{P}(Z > 1)$$

Problem 3 (2018 September, # 8)**.** Suppose $(\mathbf{X}, \mathbf{Y})$ have a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix

$$\Sigma = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where A is $m \times m$, B is $m \times n$, and C is $n \times n$, and A and C are non-singular. Define a vector $\mathbf{Z} := \mathbf{Y} - B^T A^{-1} \mathbf{X}$.

(i) Find the $m \times n$ covariance matrix of $\mathbf{X}$ and $\mathbf{Z}$.

(ii) Express $\mathbf{Y}$ as $\mathbf{Z} + B^T A^{-1} \mathbf{X}$, and, hence deduce the conditional distribution of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$.

Solution

(i) We have $\mathbb{E}[\mathbf{X}\mathbf{Z}^T] = \mathbb{E}[\mathbf{X}(\mathbf{Y}^T - \mathbf{X}^T (A^T)^{-1} B)] = B - A(A^T)^{-1} B = \mathbf{0}_{m \times n}$, where we use the fact that A is symmetric, being a covariance matrix.

(ii) The covariance of $\mathbf{Z}$ is

$$\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] = \mathbb{E}[(\mathbf{Y} - B^T A^{-1} \mathbf{X})(\mathbf{Y}^T - \mathbf{X}^T A^{-1} B)] = C - B^T A^{-1} B.$$

Now, $\mathbf{Z}$ is a Gaussian by Cramér-Wold device. We also have by (i) that $\mathbf{Z}$ is independent of $\mathbf{Z}$. Thus, $\mathbf{Y} | \mathbf{X} = \mathbf{x} \sim N(B^T A^{-1} \mathbf{x}, C - B^T A^{-1} B)$.

Problem 4. [2018 September, # 9] Let $X \in \mathbb{R}^d$ be a centered normal random vector and $A \in \mathbb{R}^{d \times d}$ a fixed symmetric matrix. Denote by Y an independent copy of X. Show that

$$X^T A X - Y^T A Y \overset{\mathsf{d}}{=} 2 X^T A Y.$$

Hint: $(X \pm Y)/\sqrt{2}$ are i.i.d. random vectors following the same distribution as X.

Solution

By considering the eigendecomposition of A, WLOG assume A is a diagonal matrix with eigenvalues $\{\lambda_i\}_{i=1}^d$. Observe then that

$$X^T A X - Y^T A Y = \sum_{i=1}^n \lambda_i (X_i + Y_i)(X_i - Y_i),$$

so that (using the hint) this is equal to

$$2 \sum_i \lambda_i \left(\frac{X_i + Y_i}{\sqrt{2}} \right) \left(\frac{X_i - Y_i}{\sqrt{2}} \right) \overset{\mathsf{d}}{=} 2 \sum_i \lambda_i X_i Y_i = 2 X^T A Y$$

Problem 5 (2020 September, # 7)**.** Suppose X_1, X_2 are i.i.d. $N(0, 1)$.

(a) Find the joint distribution of $X_1 + X_2$ and $X_1 - X_2$.

(b) Show that $2X_1 X_2$ has the same distribution as $X_1^2 - X_2^2$.

Solution

(a) The transformation $(X_1, X_2) \mapsto (X_1 + X_2, X_1 - X_2)$ has Jacobian

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

which has determinant 2, meaning the inverse mapping has determinant $-1/2$. Next, we note

$$x_1^2 + x_2^2 = \frac{1}{2}\left((x_1 + x_2)^2 + (x_1 - x_2)^2\right).$$

Thus, the joint density of $(Y_1, Y_2) \sim (X_1 + X_2, X_1 - X_2)$ is

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{4\pi} e^{-\frac{1}{4}(y_1^2 + y_2^2)}.$$

(b) This is the same idea as Problem 4:

$$2X_1X_2 = \frac{1}{2}\left((X_1 + X_2)^2 - (X_1 - X_2)^2\right) = \frac{1}{2}(Y_1^2 - Y_2^2) = \left(\frac{Y_1}{\sqrt{2}}\right)^2 - \left(\frac{Y_2}{\sqrt{2}}\right)^2 \overset{d}{=} X_1^2 - X_2^2.$$

1.2 Additional Practice

Problem 6 (Casella & Berger, Exercise 4.20)**.** Suppose X_1, X_2 are independent $\mathcal{N}(0, \sigma^2)$ random variables.

1. Find the joint distribution of Y_1 and Y_2, where

$$Y_1 = X_1^2 + X_2^2 \text{ and } Y_2 = \frac{X_1}{\sqrt{Y_1}}.$$

2. Show that Y_1 and Y_2 are independent, and interpret this result geometrically.

Solution

The transformation $(X_1, X_2) \mapsto (Y_1, Y_2)$ is not one-to-one because we cannot determine the sign of X_2 from Y_1 and Y_2. Thus, to use a pdf transformation law, we need to split up the support of (X_1, X_2) into regions where this transformation is one-to-one. Intuitively, these regions should be where the sign of X_2 is determined. Consider the partition $\mathcal{A}_0 \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_2 = \mathbb{R}^2$ such that

$$\begin{aligned} \mathcal{A}_0 &:= \{(x_1, x_2) : x_2 = 0\} \\ \mathcal{A}_1 &:= \{(x_1, x_2) : x_2 > 0\} \\ \mathcal{A}_2 &:= \{(x_1, x_2) : x_2 < 0\} \end{aligned}$$

Meanwhile, the support of (Y_1, Y_2) is the set $\{(y_1, y_2) : y_1 > 0, y_2 \in (-1, 1)\}$. Then, the inverse transformation $(Y_1, Y_2) \mapsto (X_1, X_2)$ from sets $\mathcal{B}$ to $\mathcal{A}_1$ is given by $(y_1, y_2) \mapsto (y_2\sqrt{y_1}, \sqrt{y_1 - y_1 y_2^2})$. We can verify this has Jacobian

$$J_1 := \frac{1}{2\sqrt{1 - y_2^2}}.$$

Similarly, the inverse transformation from sets $\mathcal{B}$ to $\mathcal{A}_2$ is given by $(y_1, y_2) \mapsto (y_2\sqrt{y_1}, -\sqrt{y_1 - y_1 y_2^2})$ with Jacobian $J_2 = -J_1$. Meanwhile, $\mathcal{A}_0$ is a measure-zero set with respect to the distribution of X_1, X_2 so we don't need to worry about value of the pdf there. Thus, the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ is the sum of two terms:

$$f_{Y_1,Y_2}(y_1, y_2) = 2\left(\frac{1}{2\pi\sigma^2} e^{-y_1/(2\sigma^2)} \cdot \frac{1}{2\sqrt{1 - y_2^2}}\right) = \frac{1}{2\pi\sigma^2} e^{-y_1/(2\sigma^2)} \cdot \frac{1}{\sqrt{1 - y_2^2}} \text{ for } y_1 > 0, y_2 \in (-1, 1).$$

Next, the joint pdf of Y_1 and Y_2 factors into a function of y_1 and a function of y_2. So, Y_1 and Y_2 are independent. Geometrically, Y_1 is the squared distance of (X_1, X_2) to the origin. Meanwhile, Y_2 is the cosine of the angle between the positive x_1-axis and the line from (X_1, X_2) to the origin. So, independence says the distance from the origin is independent of the orientation.

Problem 7 (marginal normality does not imply bivariate normality). [Casella & Berger, Exercise 4.47] Let X and Y be independent $\mathcal{N}(0,1)$ random variables, and define a new random variable Z by

$$Z = \begin{cases} X & XY > 0 \\ -X & XY < 0 \end{cases}.$$

1. Show that Z has a normal distribution.
2. Show that the joint distribution of Z and Y is not bivariate normal. Hint: show that Z and Y always have the same sign.

Solution

We have using the definition of Z, for $z < 0$:

$$\begin{aligned}
\mathbb{P}(Z \le z) &= \mathbb{P}(X \le z, XY > 0) + \mathbb{P}(-X \le z, XY < 0) \\
&= \mathbb{P}(X \le z, Y < 0) + \mathbb{P}(X \ge -z, Y < 0) \text{ (since } z < 0) \\
&= \mathbb{P}(X \le z)\mathbb{P}(Y < 0) + \mathbb{P}(X \ge -z)\mathbb{P}(Y > 0) \text{ (independence)} \\
&= \mathbb{P}(X \le z)\mathbb{P}(Y < 0) + \mathbb{P}(X \le z)\mathbb{P}(Y > 0) \text{ (symmetry of } X \text{ and } Y) \\
&= \mathbb{P}(X \le z) \cdot (\mathbb{P}(Y < 0) + \mathbb{P}(Y > 0)) \\
&= \mathbb{P}(X \le z).
\end{aligned}$$

By a similar argument, for $z > 0$, we have $\mathbb{P}(Z > z) = \mathbb{P}(X > z)$. Thus, $\mathbb{P}(Z \le z) = \mathbb{P}(X \le z)$ meaning Z and X have the same cdf's and thus $Z \sim \mathcal{N}(0,1)$.

Following the hint, $Z > 0$ iff either (i) $X < 0$ and $Y > 0$ or (ii) $X > 0$ and $Y > 0$. Thus, Z and Y always have the same sign and hence the support of (Z, Y) is not all of $\mathbb{R}^2$. Thus, (Z, Y) cannot be bivariate normal.

Problem 8 (Casella & Berger, Exercise 5.22). Let X and Y be iid $\mathcal{N}(0,1)$ random variables, and define $Z = \min(X, Y)$. Prove that $Z^2 \sim \chi^2_1$.

Solution

We claim the cdf of Z^2 as $F_{Z^2}(z) = 1 - 2F_X(-\sqrt{z})$. We have

$$\begin{aligned}
F_{Z^2}(z) &= \mathbb{P}(\min(X,Y)^2 \le z) \\
&= \mathbb{P}(-\sqrt{z} \le \min(X,Y) \le \sqrt{z}) \\
&= \mathbb{P}(\min(X,Y) \le \sqrt{z}) - \mathbb{P}(\min(X,Y) \le -\sqrt{z}) \\
&= (1 - \mathbb{P}(\min(X,Y) > \sqrt{z})) - (1 - \mathbb{P}(\min(X,Y) > -\sqrt{z})) \\
&= \mathbb{P}(\min(X,Y) > -\sqrt{z}) - \mathbb{P}(\min(X,Y) > \sqrt{z}) \\
&= \mathbb{P}(X > -\sqrt{z})\mathbb{P}(Y > -\sqrt{z}) - \mathbb{P}(X > \sqrt{z})\mathbb{P}(Y > \sqrt{z}),
\end{aligned}$$

where we use the independence of X and Y to establish the last equality. Since X and Y are identically distributed, $\mathbb{P}(X > a) = \mathbb{P}(Y > a) = 1 - F_X(a)$. So,

$$F_{Z^2}(z) = (1 - F_X(-\sqrt{z}))^2 - (1 - F_X(\sqrt{z}))^2 = 1 - 2F_X(-\sqrt{z}).$$

Differentiating then gives pdf

$$f_{Z^2}(z) = \frac{1}{\sqrt{2\pi}} e^{-z/2} \cdot z^{-1/2}.$$

Indeed, this is the pdf of a χ^2_1 random variable.