EECS553 HW1

Lingqi Huang

September 2024

1 Problem 1

(a): It is the responsibility of faculty members to specify their policies in writing at the beginning of each semester. Students are responsible for understanding these policies and should consult the instructor if they are unclear. [**True**]

(b): Students who are not members of the College of Engineering and who take a course offered by the College are bound by the policies of the Engineering Honor Code.[**True**]

(c):If a student is accused of academic misconduct, they may simply withdraw from the class to avoid any blemish on their academic record.[**False**]

2 Problem 2

(a): Notice that the covariance matrix Σ could be written as

$$\Sigma = Cov(\mathbf{X}, \mathbf{X}) = \mathbb{E}[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^T)]$$

Notice that for any vector $\mathbf{a} \in \mathbb{R}^n$, we have that

$$\mathbf{a}^T \Sigma \mathbf{a} = \mathbf{a}^T \mathbb{E}[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^T)]\mathbf{a}$$

Since $\mathbf{a}$ is a constant vector, define a new random variable $\mathbf{Y} = [\mathbf{X} - \mathbb{E}(\mathbf{X})^T]\mathbf{a}$, we notice that

$$\mathbf{a}^T \Sigma \mathbf{a} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T] \geq 0$$

Thus, by definition, we conclude that covariance matrices are positive semi-definite.

(b): "$\Rightarrow$":Suppose that the matrix $\mathbf{A}$ is invertible and symmetric, then by Spectral Theorem, $\mathbf{A}$ could be decomposed to

$$\mathbf{A} = \mu \Lambda \mu^T$$

where Λ is a diagonal matrix and $\mu\mu^T = \mu^T\mu = \mathbf{I}$.Now by property of determinant of matrix, we have that

$$det(\mathbf{A}) = (det(\mu))^2 \det(\Lambda) \neq 0$$

because $det(\Lambda) \neq 0$ for $\mathbf{A}$ is invertible and Λ contain all eigenvalue of $\mathbf{A}$. Thus, we conclude all eigenvalue of $\mathbf{A}$ are non-zero.

"$\Leftarrow$":Now suppose all eigenvalues of $\mathbf{A}$ are non-zero, meaning $\det(A) \neq 0$, we can easily conclude that $\mathbf{A}$ is invertible. To find the spectral decomposition of $\mathbf{A}^{-1}$, we notice that

$$\mathbf{A}^{-1} = (\mu \Lambda \mu^T)^{-1} = (\mu^T)^{-1} \Lambda^{-1} \mu^{-1} = \mu \Lambda \mu^T$$

Now if $\mathbf{A}$ is also positive definite, then we first notice that $\mathbf{A}^{-1}$ is also symmetric because we notice

$$\mathbf{A}^{-1} = (\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$$

Now by computation of determinant, as $det(\mathbf{A}) > 0$, we must have $det(\mathbf{A}^{-1}) > 0$ since $det(\mathbf{A})det(\mathbf{A}^{-1}) = det(\mathbf{I}) = 1$. BY theorem, we show $\mathbf{A}^{-1}$ is positive definite.

3 Problem 3

(a): We first notice that the event of $[X = x]$ can be written as $\bigcup_{i=1}^{n} A_i$, where $A_i = [X = x, Y = y_i]$, and each A_i, $A_j$are disjoint for $i \neq j$. By property of probability measure, we notice that

$$P(X = x) = P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{n} P(X = x, Y = y_i) = \sum_{y} P(x, y)$$

This completes the proof.

(b):Notice that by definition:

$$\mathbb{E}(X) = \sum_{y}\sum_{x} xp(x, y) = \sum_{y}\sum_{x} xp(x|y)p(y) = \sum_{y}\left(\sum_{x} xp(x|y)\right)p(y)$$

Notice that $\sum_{x} xp(x|y) = \mathbb{E}(X|y)$, and so $\mathbb{E}(X) = \sum_{y} \mathbb{E}(X|y)p(y) = \mathbb{E}_y[\mathbb{E}(\mathbb{X}|\mathbb{Y})]$. This completes the proof.

4 Problem 4

(a): We notice that $\Sigma^{-1} = U\Lambda^{-1}U^T$ from problem 2, now as we know $\theta = \frac{\pi}{6}$, the matrix U can be written as:

$$U = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

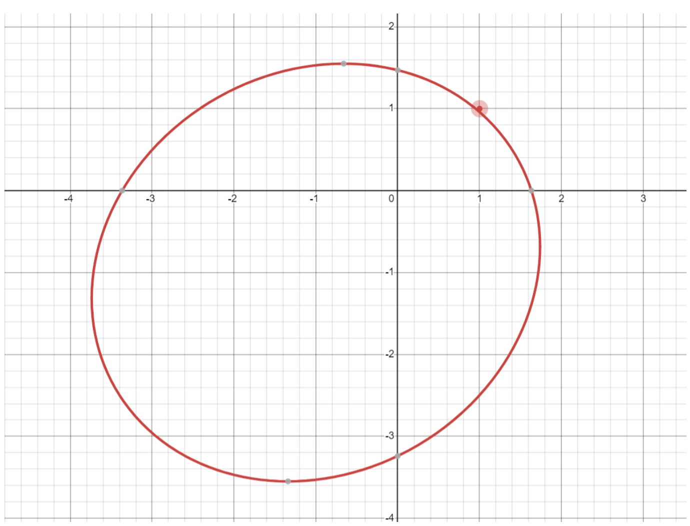
Now by some matrix computation, we find that

$$\Sigma^{-1} = \begin{pmatrix} \frac{13}{48} & -\frac{\sqrt{3}}{48} \\ -\frac{\sqrt{3}}{48} & \frac{5}{16} \end{pmatrix}$$

So we can get that the region can be written as

$$(\mathbf{x} - \mu)^T\Sigma^{-1}(\mathbf{x} - \mu) \leq 2 \iff 13(x_1 + 1)^2 + 15(x_2 + 1)^2 - 2\sqrt{3}(x_1 + 1)(x_2 + 1) \leq 96$$

where $\mathbf{x} = [x_1, x_2]$. Now notice that the degree of rotation is exactly the angle between major axis and x-axis, which is just $\theta = \frac{\pi}{6}$. Since we know $r = \sqrt{2}$, and two eigenvalues of Σ are 4 and 3, so the length of two axes are $2\sqrt{2}$, $\sqrt{6}$.(The below skecth is drawn by Desmos)



(b):To solve this problem, I would use a Theorem from the textbook *Linear Models in Statitics, 2nd Edition* written by *Alvin C. Rencher and G. Bruce Schaalje.*

Theorem 5.5: Let $\mathbf{y}$ be distributed as $N_p(\mu, \Sigma)$, let A be a symmetric matrix of constants of rank, and let $\lambda = \frac{1}{2}\mu^T A\mu$.Then $y^T Ay$ is $\chi^2(r, \lambda)$, if and only if $A\Sigma$ is idempotent.

Now for this problem, it is straightforward to see that $\Sigma^{-1}\Sigma = (\Sigma^{-1}\Sigma)^2 = I$, so the condition of idempotent is satisfied. Also, notice that $det(\Sigma^{-1}) \neq 0$, thus we conclude that Σ^{-1} has full rank of 2. Now since $\mathbf{x} - \mu \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma\right)$, we get $\lambda = 0$ that the non-centrality is 0. So we conclude that

$$(\mathbf{X} - \mu)^T \Sigma^T (\mathbf{X} - \mu) \sim \chi^2_2$$

which is a Chi-squared distribution with degree of 2. Now by conmand pchisq(2, 2) in R language, we conclude that $P(\mathbf{X} \in C) \approx 0.63$.

5 Problem 5

(a): It is suffice to show that $\forall \mathbf{x_1}, \mathbf{x_2}$, and $t \in [0, 1]$, we have

$$g(t\mathbf{x}_1 + (1 - t\mathbf{x}_2)) \leq tg(\mathbf{x}_1) + (1 - t)g(\mathbf{x}_2)$$

Now notice that

$$\begin{aligned}
g(t\mathbf{x}_1 + (1 - t\mathbf{x}_2)) &= f(A(t\mathbf{x}_1 + (1 - t)\mathbf{x}_2) + b) && (1) \\
&= f(At\mathbf{x}_1 + A(1 - t)\mathbf{x}_2 + bt + (1 - t)b) && (2) \\
&= f(t(A\mathbf{x}_1 + b) + (1 - t)(A\mathbf{x}_2 + b)) && (3) \\
&\leq tf(A\mathbf{x_1} + b) + (1 - t)f(A\mathbf{x}_2 + b) \quad \text{(by convexity of } f\text{)} && (4) \\
&= tg(\mathbf{x}_1) + (1 - t)g(\mathbf{x}_2) && (5)
\end{aligned}$$

Thus, we have shown that $g(\mathbf{x})$ is convex.

(b): we first compute the gradient of f, which is

$$\nabla f(\mathbf{x}) = \frac{1}{2}(A + A^T)\mathbf{x} + b = A\mathbf{x} + b$$

The Hessian would be

$$\nabla^2 f(\mathbf{x}) = A$$

Thus, we conclude that f is convex if A is positive semi-definite, where we need eigenvalues of A are non-negative. If f is strictly convex, we require all eigenvalues of A are all positive.