STA 2101 Assignment 2[1]

The questions on this assignment are not to be handed in. They are practice for Quiz Two on Friday September 27th. There is a posted formula sheet that will be provided with the quiz. The linear algebra questions are more review.

1. Let $\mathbf{A} = \left(\begin{array}{cc} 1 & 2 \\ 2 & 4 \end{array} \right)$ $\mathbf{B} = \left(\begin{array}{cc} 0 & 2 \\ 2 & 1 \end{array} \right)$ $\mathbf{C} = \left(\begin{array}{cc} 2 & 0 \\ 1 & 2 \end{array} \right)$

   (a) Calculate $\mathbf{AB}$ and $\mathbf{AC}$

   (b) Do we have $\mathbf{AB} = \mathbf{AC}$? Answer Yes or No.

   (c) Prove $\mathbf{B} = \mathbf{C}$. Show your work.

2. Let $\mathbf{X}$ be an n by p matrix with $n \neq p$. Why is it incorrect to say that $(\mathbf{X}^\top \mathbf{X})^{-1} = \mathbf{X}^{-1}\mathbf{X}^{\top -1}$?

3. Let $\mathbf{a}$ be an $n \times 1$ matrix of real constants. How do you know $\mathbf{a}^\top \mathbf{a} \geq 0$?

4. The $p \times p$ matrix $\boldsymbol{\Sigma}$ is said to be *positive definite* if $\mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a} > 0$ for all $p \times 1$ vectors $\mathbf{a} \neq \mathbf{0}$. Show that the eigenvalues of a positive definite matrix are all strictly positive. A good approach is to start with the definition of an eigenvalue and the corresponding eigenvalue: $\boldsymbol{\Sigma}\mathbf{v} = \lambda \mathbf{v}$. Eigenvectors are typically scaled to have length one, so you may assume $\mathbf{v}^\top \mathbf{v} = 1$.

5. Recall the *spectral decomposition* of a symmetric matrix (for example, a variance-covariance matrix). Any such matrix $\boldsymbol{\Sigma}$ can be written as $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^\top$, where $\mathbf{P}$ is a matrix whose columns are the (orthonormal) eigenvectors of $\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda}$ is a diagonal matrix of the corresponding eigenvalues, and $\mathbf{P}^\top\mathbf{P} = \mathbf{P}\mathbf{P}^\top = \mathbf{I}$. If $\boldsymbol{\Sigma}$ is real, the eigenvalues are real as well.

   (a) Let $\boldsymbol{\Sigma}$ be a square symmetric matrix with eigenvalues that are all strictly positive.

      i. What is $\boldsymbol{\Lambda}^{-1}$?

      ii. Show $\boldsymbol{\Sigma}^{-1} = \mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}^\top$

   (b) Let $\boldsymbol{\Sigma}$ be a square symmetric matrix, and this time the eigenvalues are non-negative.

      i. What do you think $\boldsymbol{\Lambda}^{1/2}$ might be?

      ii. Define $\boldsymbol{\Sigma}^{1/2}$ as $\mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{P}^\top$. Show $\boldsymbol{\Sigma}^{1/2}$ is symmetric.

      iii. Show $\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Sigma}$, justifying the notation.

(c) Now return to the situation where the eigenvalues of the square symmetric matrix $\mathbf{\Sigma}$ are all strictly positive. Define $\mathbf{\Sigma}^{-1/2}$ as $\mathbf{P}\mathbf{\Lambda}^{-1/2}\mathbf{P}^\top$, where the elements of the diagonal matrix $\mathbf{\Lambda}^{-1/2}$ are the reciprocals of the corresponding elements of $\mathbf{\Lambda}^{1/2}$.

   i. Show that the inverse of $\mathbf{\Sigma}^{1/2}$ is $\mathbf{\Sigma}^{-1/2}$, justifying the notation.

   ii. Show $\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{-1/2} = \mathbf{\Sigma}^{-1}$.

(d) Let $\mathbf{\Sigma}$ be a symmetric, positive definite matrix. How do you know that $\mathbf{\Sigma}^{-1}$ exists?

6. Let $\mathbf{X}$ be an $n \times p$ matrix of constants. The idea is that $\mathbf{X}$ is the "design matrix" in the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, so this problem is really about linear regression.

   (a) Recall that $\mathbf{A}$ symmetric means $\mathbf{A} = \mathbf{A}^\top$. Let $\mathbf{X}$ be an n by p matrix. Show that $\mathbf{X}^\top\mathbf{X}$ is symmetric.

   (b) Recall the definition of linear independence. The columns of $\mathbf{A}$ are said to be *linearly dependent* if there exists a column vector $\mathbf{v} \neq \mathbf{0}$ with $\mathbf{A}\mathbf{v} = \mathbf{0}$. If $\mathbf{A}\mathbf{v} = \mathbf{0}$ implies $\mathbf{v} = \mathbf{0}$, the columns of $\mathbf{A}$ are said to be linearly *independent*. Show that if the columns of $\mathbf{X}$ are linearly independent, then $\mathbf{X}^\top\mathbf{X}$ is positive definite.

   (c) Show that if $\mathbf{X}^\top\mathbf{X}$ is positive definite then $(\mathbf{X}^\top\mathbf{X})^{-1}$ exists.

   (d) Show that if $(\mathbf{X}^\top\mathbf{X})^{-1}$ exists then the columns of $\mathbf{X}$ are linearly independent.

   This is a good problem because it establishes that the least squares estimator $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$ exists if and only if the columns of $\mathbf{X}$ are linearly independent.

---

7. Women and men are coming into a store according to independent Poisson processes with rates λ_1 for women and λ_2 for men. You don't have to know anything about Poisson processes to do this question. We have that the number of women and the number of men entering the store in a given time period are independent Poisson random variables, with expected values λ_1 for women and λ_2 for men. Because the Poisson process is an independent increments process, we can treat the numbers from n time periods as a random sample.

   Management wants to know the expected number of male customers and the expected number of female customers. Unfortunately, the total numbers of customers were recorded, but not their sex. Let $y_1, \ldots, y_n$ denote the total numbers of customers who enter the store in n time periods. That's all the data we have.

   (a) What is the distribution of y_i? If you know the answer, just write it down without proof.

   (b) What is the parameter space?

   (c) Find the MLE of the parameter vector (λ_1, λ_2). Show your work.

   (d) How is this question related to the Zipper example?

8. Suppose $\sqrt{n}(T_n - \theta) \stackrel{d}{\rightarrow} T$. Show $T_n \stackrel{p}{\rightarrow} \theta$. Please use Slutsky lemmas rather than definitions. Hint: Think of the sequence of constants $\frac{1}{\sqrt{n}}$ as a sequence of degenerate random variables (variance zero) that converge almost surely and hence in probability to zero. Now you can use a Slutsky lemma.

9. Let $X_1, \ldots, X_n$ be a random sample from a Binomial distribution with parameters 3 and θ. That is,
$$P(X_i = x_i) = \binom{3}{x_i} \theta^{x_i}(1-\theta)^{3-x_i},$$
for $x_i = 0, 1, 2, 3$. Find the maximum likelihood estimator of θ, and show that it is strongly consistent.

10. Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution with density
$$f(x;\tau) = \frac{\tau^{1/2}}{\sqrt{2\pi}}\, e^{-\frac{\tau x^2}{2}},$$
where the parameter $\tau > 0$. Let
$$\widehat{\tau} = \frac{n}{\sum_{i=1}^n X_i^2}.$$
Is $\widehat{\tau}$ a consistent estimator of τ? Answer Yes or No and prove your answer. Hint: You can just write down $E(X^2)$ by inspection. This is a very familiar distribution.

11. Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ. Show that $T_n = \frac{1}{n+400}\sum_{i=1}^n X_i$ is a strongly consistent estimator of μ.

12. Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ and variance σ^2. Prove that the sample variance $S^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}$ is a strongly consistent estimator of σ^2.

13. Independently for $i = 1, \ldots, n$, let
$$Y_i = \beta X_i + \epsilon_i,$$
where $E(X_i) = E(\epsilon_i) = 0$, $Var(X_i) = \sigma_X^2$, $Var(\epsilon_i) = \sigma_\epsilon^2$, and ϵ_i is independent of X_i. Let
$$\widehat{\beta}_n = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}.$$
Is $\widehat{\beta}_n$ a consistent estimator of β? Answer Yes or No and prove your answer.

14. In this problem, you'll use (without proof) the *variance rule*, which says that if θ is a real constant and $T_1, T_2, \ldots$ is a sequence of random variables with

$$\lim_{n\to\infty} E(T_n) = \theta \text{ and } \lim_{n\to\infty} Var(T_n) = 0,$$

then $T_n \stackrel{P}{\to} \theta$.

In Problem 13, the independent variables are random. Here they are fixed constants, which is more standard (though a little strange if you think about it). Accordingly, let

$$Y_i = \beta x_i + \epsilon_i$$

for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a distribution with expected value zero and variance σ^2, and β and σ^2 are unknown constants.

(a) What is $E(Y_i)$?

(b) What is $Var(Y_i)$?

(c) Use the same estimator as in Problem 13. Is $\widehat{\beta}_n$ unbiased? Answer Yes or No and show your work.

(d) Suppose that the sequence of constants $\sum_{i=1}^n x_i^2 \to \infty$ as $n \to \infty$. Does this guarantee $\widehat{\beta}_n$ will be consistent? Answer Yes or No. Show your work.

(e) Let $\widehat{\beta}_{2,n} = \frac{\overline{Y}_n}{\overline{x}_n}$. Is $\widehat{\beta}_{2,n}$ unbiased? Consistent? Answer Yes or No to each question and show your work. Do you need a condition on the x_i values ?

(f) Prove that $\widehat{\beta}_n$ is a more accurate estimator than $\widehat{\beta}_{2,n}$ in the sense that it has smaller variance. Hint: The sample variance of the explanatory variable values cannot be negative.

15. Let X be a random variable with expected value μ and variance σ^2. Show $\frac{X}{n} \stackrel{p}{\to} 0$.

16. Let $X_1, \ldots, X_n$ be a random sample from a Gamma distribution with $\alpha = \beta = \theta > 0$. That is, the density is

$$f(x;\theta) = \frac{1}{\theta^\theta \Gamma(\theta)} e^{-x/\theta} x^{\theta-1},$$

for $x > 0$. Let $\widehat{\theta} = \overline{X}_n$. Is $\widehat{\theta}$ a consistent estimator of θ? Answer Yes or No and prove your answer.

17. Here is an integral you cannot do in closed form, and numerical integration is challenging. For example, R's `integrate` function fails.

$$\int_0^{1/2} e^{\cos(1/x)}\, dx$$

Using R, approximate the integral with Monte Carlo integration, and give a 99% confidence interval for your answer. You need to produce 3 numbers: the estimate, a lower confidence limit and an upper confidence limit. See lecture slides. **Bring your printout to the quiz.**