STA 302f14 Assignment Four[1]

Except for Question 2(g)vii, these problems are preparation for the quiz in tutorial on Friday October 10th, and are not to be handed in. Please bring your printout from Question 2(g)vii to the quiz. Please look at the current formula sheet while you do these problems.

1. The general linear model is $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ is an $n \times (k+1)$ matrix of observable constants, $\boldsymbol{\beta}$ is a $(k+1) \times 1$ vector of unknown constants (parameters), and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of unobservable random variables with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $cov(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}_n$, where $\sigma^2 > 0$ is an unknown constant parameter.
   (a) Show that the matrix $\mathbf{X}'\mathbf{X}$ is symmetric.
   (b) Recall that the $p \times p$ matrix $\mathbf{A}$ is said to be *non-negative definite* if $\mathbf{v}'\mathbf{A}\mathbf{v} \geq 0$ for all constant vectors $\mathbf{v} \in \mathbb{R}^p$. Show that $\mathbf{X}'\mathbf{X}$ is non-negative definite.
   (c) Show that if the columns of $\mathbf{X}$ are linearly independent, then $\mathbf{X}'\mathbf{X}$ is positive definite.
   (d) Show that if $\mathbf{X}'\mathbf{X}$ is positive definite, then $(\mathbf{X}'\mathbf{X})^{-1}$ exists.
   (e) Show that if $(\mathbf{X}'\mathbf{X})^{-1}$ exists, then the columns of $\mathbf{X}$ are linearly independent.

   This is a good problem because it establishes that the least squares estimator $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ exists if and only if the columns of $\mathbf{X}$ are linearly independent, meaning that no independent variable is a linear combination of the other ones.

2. This question is an example of *simple regression*. "Simple" means one independent variable. Chapter 6 in the text is about simple regression. It covers testing as well as estimation. We'll get to testing later.

   Here is the model. Let $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ 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. The numbers $x_1, \ldots, x_n$ are known, observed constants, while the parameters β_0 β_1 and σ^2 are unknown constants (parameters).
   (a) What is $E(Y_i)$?
   (b) What is $Var(Y_i)$?
   (c) Find the Least Squares estimates of β_0 and β_1 by minimizing the function

   $$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 x_i)^2$$

   over all values of (β_0, β_1). Let $(\widehat{\beta}_0, \widehat{\beta}_1)$ denote the point at which $Q(\boldsymbol{\beta})$ is minimal. Your answer is a pair of formulas, one for $\widehat{\beta}_0$ and one for $\widehat{\beta}_1$.

[1]Copyright information is at the end of the last page.

(d) Give the equation of the least-squares line.

(e) Recall that a statistic is an *unbiased estimator* of a parameter if the expected value of the statistic is equal to the parameter.

   i. Is $\widehat{\beta}_0$ an unbiased estimator of β_0? Answer Yes or No and show your work.

   ii. Is $\widehat{\beta}_1$ an unbiased estimator of β_1? Answer Yes or No and show your work.

(f) Fitting this simple regression problem into the matrix framework of Question 1,

   i. What is $\mathbf{X}'\mathbf{X}$?

   ii. What is $\mathbf{X}'\mathbf{Y}$?

   iii. What is $(\mathbf{X}'\mathbf{X})^{-1}$?

   iv. Verify that your expression for $\widehat{\beta}_1$ agrees with $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. The formula for $\widehat{\beta}_0$ agrees also, but it's messy so don't bother.

(g) Please use this small data set for the following questions:

x	1	8	3	6	4	7
y	14	2	14	10	9	9

   i. What is $\widehat{\beta}_0$? Your answer is a number. Two decimal places of accuracy will be fine.

   ii. What is $\widehat{\beta}_1$? Your answer is a number. Two decimal places of accuracy will be fine.

   iii. What is $\widehat{Y}_3$? Your answer is a number. Again, two decimal places of accuracy will be fine.

   iv. What is $\widehat{\epsilon}_3$? Your answer is a number.

   v. Based on these data, what value of y would you predict for $x = 5$? Your answer is a number.

   vi. Plot the least-squares line. You can do it freehand; it does not need to be perfect.

   vii. Use R to estimate β_0 and β_1. Bring your printout (one page maximum) to the quiz. You may be asked to hand it in. You may write your name and student number on the printout (or put them in a comment statement), but don't write anything else on the printout.

3. In Question 2, the model had both an intercept and one independent variable. But suppose the model has no intercept. This is called simple *regression through the origin.* The model would be $Y_i = \beta_1 x_i + \epsilon_i$.

(a) Find the least squares estimator $\widehat{\beta}_1$ with calculus.

(b) What is $\mathbf{X}'\mathbf{X}$?

(c) What is $\mathbf{X}'\mathbf{Y}$?

(d) What is $(\mathbf{X}'\mathbf{X})^{-1}$?

(e) Verify that your expression for $\widehat{\beta}_1$ agrees with $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

4. There can even be a regression model with an intercept but no independent variable. In this case the model would be $Y_i = \beta_0 + \epsilon_i$.

   (a) Find the least squares estimator $\widehat{\beta}_0$ with calculus.

   (b) What is the $\mathbf{X}$ matrix?

   (c) What is $\mathbf{X}'\mathbf{X}$?

   (d) What is $\mathbf{X}'\mathbf{Y}$?

   (e) What is $(\mathbf{X}'\mathbf{X})^{-1}$?

   (f) Verify that your expression for $\widehat{\beta}_0$ agrees with $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

5. In scalar form, the model of Question 1 is

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i,$$

and we obtain least-squares estimates of the β values by minimizing the sum of squared differences between observed Y_i and $E(Y_i)$. That is, we choose $\beta_0, \ldots, \beta_k$ to make

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2$$

as small as possible.

   (a) Differentiate $Q(\boldsymbol{\beta})$ with respect to β_0 and set the derivative to zero, obtaining the first *normal equation*.

   (b) Noting that the quantities $\widehat{\beta}_0, \ldots, \widehat{\beta}_k$ must satisfy the first normal equation, show that the least squares plane must pass through the point $(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k, \overline{Y})$.

   (c) Defining "predicted" Y_i as $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \cdots + \widehat{\beta}_k x_{ik}$, show that $\sum_{i=1}^n \widehat{Y}_i = \sum_{i=1}^n Y_i$.

   (d) The *residual* for observation i is defined by $\widehat{\epsilon}_i = Y_i - \widehat{Y}_i$. Show that the sum of residuals equals exactly zero.

6. Referring to the matrix version of the linear model (see Question 1) and letting $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ (which implies that the columns of $\mathbf{X}$ must be linearly independent), show that $(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})'(\mathbf{X}\widehat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{0}$.

7. Using the result of the preceding question and writing $Q(\boldsymbol{\beta})$ as $Q(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$, show that $Q(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}) + (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\mathbf{X}'\mathbf{X})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Why does this imply that the minimum of $Q(\boldsymbol{\beta})$ occurs at $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$? How do you know that the minimum is unique?

8. The set of vectors $\mathcal{V} = \{\mathbf{v} = \mathbf{Xb} : \mathbf{b} \in \mathbb{R}^{k+1}\}$ is the subset of $\mathbb{R}^n$ consisting of linear combinations of the columns of $\mathbf{X}$. That is, $\mathcal{V}$ is the space *spanned* by the columns of $\mathbf{X}$. The least squares estimator $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ was obtained by minimizing $(\mathbf{Y}-\mathbf{Xb})'(\mathbf{Y}-\mathbf{Xb})$ over all $\mathbf{b} \in \mathbb{R}^{k+1}$. Thus, $\widehat{\mathbf{Y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$ is the point in $\mathcal{V}$ that is *closest* to the data vector $\mathbf{Y}$. Geometrically, $\widehat{\mathbf{Y}}$ is the *projection* (shadow) of $\mathbf{Y}$ onto $\mathcal{V}$.

   This means the vector of differences $\widehat{\boldsymbol{\epsilon}} = \mathbf{Y} - \widehat{\mathbf{Y}}$ should be perpendicular (at right angles) to each and every vector in $\mathcal{V}$. Prove it, by calculating the inner product $(\mathbf{Xb})'\widehat{\boldsymbol{\epsilon}}$ for general $\mathbf{b}$.

9. Is $\widehat{\boldsymbol{\beta}}$ an unbiased estimator of $\boldsymbol{\beta}$? Answer Yes or No and show your work.

10. Calculate $cov(\widehat{\boldsymbol{\beta}})$ and simplify. Show your work.
