STA 302f17 Assignment Four[1]

1. Independently for $i = 1, \ldots, n$, let $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$, where the β_j are unknown constants, the x_{ij} are known, observable constants, and the ϵ_i are unobservable random variables with expected value zero. If course, values of the dependent variable y_i are observable. Start deriving the least squares estimates of β_0, β_1 and β_2 by minimizing the sum of squared differences between the y_i and their expected values. I say *start* because you don't have to finish the job. Stop when you have three linear equations in three unknowns, arranged so they are clearly the so-called "normal" equations $X'X\boldsymbol{\beta} = X'\mathbf{y}$.

2. For the general linear regression model in matrix form,

   (a) Show (there is no difference beween "show" and "prove") that the matrix $X'X$ is symmetric.

   (b) Show that $X'X$ is non-negative definite.

   (c) Show that if the columns of X are linearly independent, then $X'X$ is positive definite.

   (d) Show that if $X'X$ is positive definite, then $(X'X)^{-1}$ exists.

   (e) Show that if $(X'X)^{-1}$ exists, then the columns of X are linearly independent.

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

3. In the matrix version of the general linear regression model, X is $n \times (k+1)$ and $\mathbf{y}$ is $n \times 1$.

   (a) What are the dimensions of the matrix H? Give the number of rows and the number of columns.

   (b) Assuming that the column of X are linearly independent (and we always do), what is the rank of H?

   (c) You know that the inverse of a square matrix exists if and only if its columns (and rows) are linearly independent. If $n > k+1$, can H have an inverse? Answer Yes or No.

   (d) Show that H is symmetric.

   (e) Show that H is idempotent, meaning $H = H^2$

   (f) Using $tr(AB) = tr(BA)$, find $tr(H)$.

   (g) Show that $\widehat{\mathbf{y}} = H\mathbf{y}$.

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

(h) Show that $\mathbf{e} = (I - H)\mathbf{y}$.

(i) Show that $M = I - H$ is symmetric.

(j) Show that M is idempotent.

(k) Find $tr(M)$.

4. Please read Chapter 2, pages 28-37 in the textbook.

   (a) Show that $M\boldsymbol{\epsilon} = \mathbf{e}$.

   (b) Prove Theorem 2.1 in the text.

   (c) Why does $X'\mathbf{e} = \mathbf{0}$ tell you that if a regression model has an intercept, the residuals must add up to zero?

   (d) Letting $\mathcal{S} = (\mathbf{y} - X\boldsymbol{\beta})'(\mathbf{y} - X\boldsymbol{\beta})$,

      i. Show that $\mathcal{S} = \mathbf{e}'\mathbf{e} + (\mathbf{b} - \boldsymbol{\beta})'(X'X)(\mathbf{b} - \boldsymbol{\beta})$.

      ii. Why does this imply that the minimum of $\mathcal{S}(\boldsymbol{\beta})$ occurs at $\boldsymbol{\beta} = \mathbf{b}$?

      iii. The columns of X are linearly independent. Why does linear independence guarantee that the minimum is unique?

   (e) What are the dimensions of the random vector $\mathbf{b}$ as defined in Expression (2.9)? Give the number of rows and the number of columns.

   (f) Is $\mathbf{b}$ an unbiased estimator of $\boldsymbol{\beta}$? Answer Yes or No and show your work.

   (g) Calculate $cov(\mathbf{b})$ and simplify. Show your work.

   (h) What are the dimensions of the random vector $\widehat{\mathbf{y}}$?

   (i) What is $E(\widehat{\mathbf{y}})$? Show your work.

   (j) What is $cov(\widehat{\mathbf{y}})$? Show your work. It is easier if you use H.

   (k) What are the dimensions of the random vector $\mathbf{e}$?

   (l) What is $E(\mathbf{e})$? Show your work. Is $\mathbf{e}$ an unbiased estimator of $\boldsymbol{\epsilon}$? This is a trick question, and requires thought.

   (m) What is $cov(\mathbf{e})$? Show your work. It is easier if you use $I - H$.

   (n) Let $s^2 = \mathbf{e}'\mathbf{e}/(n - k - 1)$ as in Expression (2.33). Show that s^2 is an unbiased estimator of σ^2. The way this was done in lecture is preferable to the way it is done in the text, in my opinion.

   (o) Do Exercises 2.1, 2.3 and 2.6 in the text. In 2.6, Γ orthogonal means $\Gamma' = \Gamma^{-1}$.

5. The scalar form of the general linear regression model is

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

where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a distribution with expected value zero and variance σ^2. The numbers x_{ij} are known, observed constants, while $\beta_0, \ldots, \beta_k$ and σ^2 are unknown constants (parameters). The term "random sample" means independent and identically distributed in this course, so the ϵ_i random variables have zero covariance with one another.

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

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

   (c) What is $Cov(y_i, y_j)$ for $i \neq j$?

   (d) Defining $SSTO = \sum_{i=1}^n (y_i - \overline{y})^2$, $SSR = \sum_{i=1}^n (\widehat{y}_i - \overline{y})^2$ and $SSE = \sum_{i=1}^n (y_i - \widehat{y}_i)^2$, show $SSTO = SSE + SSR$. I find it helpful to switch to matrix notation partway through the calculation.

6. "Simple" regression is just regression with a single independent variable. The model equation is $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. Fitting this simple regression problem into the matrix framework of the general linear regression model in matrix form (see formula sheet),

   (a) What is the X matrix?

   (b) What is $X'X$? Your answer is a 2×2 matrix with a formula in each cell.

   (c) What is $X'\mathbf{y}$? Again, your answer is a matrix with a formula in each cell.

7. Show that for simple regression, the proportion of explained sum of squares is the square of the correlation coefficient. That is, $R^2 = \frac{SSR}{SST} = r^2$.

8. In *simple regression through the origin*, there is one independent variable and no intercept. The model is $y_i = \beta_1 x_i + \epsilon_i$.

   (a) What is the X matrix?

   (b) What is $X'X$?

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

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

   (e) What is $b_1 = (X'X)^{-1}X'\mathbf{y}$? Compare your answer to (1.22) on page 11 in the textbook.

9. There can even be a regression model with an intercept and no independent variables. In this case the model would be $y_i = \beta_0 + \epsilon_i$.

   (a) Find the least squares estimator of β_0 with calculus.

   (b) What is the X matrix?

   (c) What is $X'X$?

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

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

   (f) What is $b_0 = (X'X)^{-1}X'\mathbf{y}$? Compare this with your answer to Question 9a.

10. The set of vectors $\mathcal{V} = \{\mathbf{v} = X\mathbf{a} : \mathbf{a} \in \mathbb{R}^{k+1}\}$ is the subset of $\mathbb{R}^n$ consisting of linear combinations of the columns of X. That is, $\mathcal{V}$ is the space *spanned* by the columns of X. The least squares estimator $\mathbf{b} = (X'X)^{-1}X'\mathbf{y}$ was obtained by minimizing $(\mathbf{y} - X\mathbf{a})'(\mathbf{y} - X\mathbf{a})$ over all $\mathbf{a} \in \mathbb{R}^{k+1}$. Thus, $\widehat{\mathbf{y}} = X\mathbf{b}$ 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 $\mathcal{V}$. The hat matrix H is a *projection matrix*. It projects the image on any point in $\mathbb{R}^n$ onto $\mathcal{V}$. Now we will test out several consequences of this idea.

    (a) The shadow of a point already in $\mathcal{V}$ should be right at the point itself. Show that if $\mathbf{v} \in \mathcal{V}$, then $H\mathbf{v} = \mathbf{v}$.

    (b) The vector of differences $\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}}$ should be perpendicular (at right angles) to each and every basis vector of $\mathcal{V}$. How is this related to Theorem 2.1?

    (c) Show that the vector of residuals $\mathbf{e}$ is perpendicular to any $\mathbf{v} \in \mathcal{V}$.
