Stats 203V: Introduction to Regression Models and ANOVA Summer 2020

Homework 2

Due date : July 12, 11:59 PM (PT)

For full credit, solutions (.pdf file, either typed or handwritten and scanned) must be uploaded to `Gradescope` by **11:59 pm PT on on Sunday July 12**. Any changes to the Homework will be posted on `Canvas`.

Collaboration on homework problems is fine, but your write up should be your own and the write up should mention the names of your collaborators.

From now on, we request that solutions to questions using `R` be prepared using `R Markdown`, e.g. within `R Studio`. Please convert your `.Rmd` file to a `.pdf` file before uploading.

1. (10 points) [$R^2 = r^2$ for simple linear regression.] Consider the least squares fit of $a + bx_i$ to Y_i for $i = 1, \ldots, n$, which results in least squares fitted values $\hat{Y}_i$ and estimates $\hat{a}, \hat{b}$ related by $\hat{Y}_i = \hat{a} + \hat{b}x_i$. Let Y, $\hat{Y}$ and x be vectors with components Y_i, $\hat{Y}_i$ and x_i respectively.
   (a) (5 points) Show that $Cov(\hat{b}x, Y) = Cov(\hat{Y}, Y)$.
   (b) (5 points) Hence show that for this model $R^2 = r^2$.

2. (20 points) [Teenage gambling dataset `teengamb` in `library(faraway)`] Faraway, Exercise 2.1 (i.e. Chapter 2 Problem 1). Also explain why the answers to (c),(d),(e) come out the way they do.

3. (20 points) [Weekly wages of US workers dataset `uswages` in `library(faraway)`] Faraway, Exercise 2.2.

4. (20 points) [F test for added variable is equivalent to (two-sided) t.] Suppose that $Y = X\beta + \epsilon$ with X fixed and $n \times p$ of full column rank and with $\epsilon \sim N_n(0, \sigma^2 I_n)$. Consider the t and F tests of $H_0 : \beta_p = 0$ versus the alternative $H_A : \beta_p \neq 0$. This exercise outlines the demonstration that in this case $F = t^2$.
   (a) (3 points) Write the design matrix $X = [X_\omega \;\; \boldsymbol{x}]$ where X_ω is $n \times (p-1)$ and consists of the first $(p-1)$ columns of X and $\boldsymbol{x} = X^{(p)}$ is the p-th column. Write $X'X$ as a 2×2 block matrix.
   (b) (7 points) Look up the formula for the inverse of a symmetric 2×2 block matrix and use it to show that
   $$1/(X'X)^{-1}_{pp} = \boldsymbol{x}'(I_n - H_\omega)\boldsymbol{x},$$
   where H_ω is the hat matrix of X_ω.
   (c) (10 points) [may be easier after Lecture 7.] Let $H_\Delta = H - H_\omega$. Show that $H_\Delta Y = \hat{\beta}_p H_\Delta \boldsymbol{x} = \hat{\beta}_p (I_n - H_\omega)\boldsymbol{x}$ and hence show that $F = ||H_\Delta Y||_2^2 / \hat{\sigma}^2 = t^2$.

5. (30 points, 3 points for each part) [Generalized inverse and projection matrix] For any square matrix A of dimension $m \times m$, we say that G (of dimension $m \times m$) is **a generalized inverse** (in short g-inverse) of A if $AGA = A$. G is called **Reflexive g-inverse** of A if $AGA = A, GAG = G$.
   (a) Show that if A is invertible, then the only g-inverse of A is A^{-1}.

(b) Consider the 2×2 matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Demonstrate that A has infinitely many symmetric reflexive g-inverses.

(c) Take X matrix of dimension $n \times p$ (We do not assume $n \geq p$ or X is of full rank). Note that $X'X$ need not be invertible. Let G be a symmetric reflexive g-inverse of $X'X$. Define

$$P_X(G) := XGX',$$

where the G in $P_X(G)$ signifies the possible dependence of $P_X(G)$ on the choice of G. Show that $P_X(G)$ is symmetric and idempotent. (P is called idempotent if $P^2 = P$).

[Hint : Use the definition of g-inverse.]

(d) Show that $X'P_X(G)X = X'X$ and hence conclude that $X'(I_n - P_X(G))X = 0$.

(e) Using (c), show that $I_n - P_X(G)$ is also idempotent. Hence, conclude using (d) that $P_X(G)X = X$. [Hint : Use the fact that $A'A = 0$ implies $A = 0$.]

(f) Recall the fact that a vector $\boldsymbol{v} \in \mathbb{R}^n$ is in $Col(X)^{\perp}$ if any only if $X'\boldsymbol{v} = 0$. Show that for any $\boldsymbol{v} \in Col(X)^{\perp}$, we have $P_X(G)\boldsymbol{v} = 0$.

(g) Recall the fact that for any subspace $V \subset \mathbb{R}^n$ and any $\boldsymbol{u} \in \mathbb{R}^n$, there exist unique $\boldsymbol{u}_V \in V$ and $\boldsymbol{u}_{V^{\perp}} \in V^{\perp}$ such that $\boldsymbol{u} = \boldsymbol{u}_V + \boldsymbol{u}_{V^{\perp}}$. Combining (e) and (f) show that

$$P_X(G)\boldsymbol{u} = \boldsymbol{u}_{Col(X)}, \ \ \forall\, \boldsymbol{u} \in \mathbb{R}^n.$$

(h) Using (g), show that $P_X(G)$ does not depend on the choice of G. Therefore we shall write it as P_X.

[Hint : Use the fact that $A\boldsymbol{u} = 0$, for all $\boldsymbol{u}$ implies that $A = 0$.]

(i) Using (g), show that $Col(P_X) = Col(X)$.

(j) For any subspace $V \subset \mathbb{R}^n$, if Q is a symmetric, idempotent matrix such that $Q\boldsymbol{u} = \boldsymbol{u}_V$, for all $\boldsymbol{u} \in \mathbb{R}^n$, then Q is called the **Orthogonal Projection matrix onto V**. Combining previous parts show that P_X is the orthogonal projection matrix onto $Col(X)$.