

Instruction:

- Type your answers using Word (equation editor / MathType) or LaTeX. You may also use Jupyter Notebook for typing the document with math equations (see [\[1\]](#) and [\[2\]](#)). Handwritten solution will not be accepted.
- Save or print your submission as a PDF file and submit on Canvas.
- You may discuss among yourselves and ask hints from the TAs. However, you may not share your typed solution with others. The final work must be individual. Do not post the HW assignment or your solution to other websites such as Course Hero or GitHub.

1. Show that the following relationship on the simple linear regression class notebook is true:

$$\frac{\sum_{i=1}^n x_i y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n x_i^2 - n \bar{X}^2} = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{X})^2}$$

2. We showed that in a simple linear regression, the MLE estimates for β_1 and β_0 are the same as the OLS estimates. Show that the MLE estimate for σ^2 is $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$. (Hint: take derivative of the LL w.r.t σ^2).
3. In a simple linear regression, show that the OLS regression line always passes through the mean (average) of both x and y .
4. In the following estimated multiple regression model, we study the amount of sleep (in minutes) an average person gets every night. Each observation is a day, and if the day is a holiday, $Holiday = 1$; if the day is Monday, $Monday = 1$. In other words, all independent variables are dummy variables.

$$\widehat{sleep} = 458.5 + 8.6Holiday - 4.9Monday - 7.7Tuesday - 7.4Wednesday - 4.7Thursday + 23.4Friday + 30.6Saturday$$
 - a. Explain why Sunday is not included in the model as an independent variable.
 - b. Explain in plain English the meaning of the intercept 458.5. Does this number make sense?
 - c. According to the model, how much sleep does an average person get on a day that is both a Saturday and a Holiday?
5. We estimate the following regression model to predict wage of employees in a firm. Exp is years of experience, *male* is a dummy variable that equals 1 if the employee is male.

$$\widehat{wage} = \beta_0 + \beta_1 exp + \beta_2 male + \beta_3 exp \times male$$

If the true relationship between the variables are as the following plot suggests, what do you expect the signs (positive or negative) of β_1 , β_2 , and β_3 to be? Explain.

