MLDS 401: Homework 2
Due: October 1, 23:59
Professor Malthouse

You may work in homework groups. Turn in one copy per group, with all names on it. Turn in a single PDF including all explanations, code and output. Clearly indicate the problem number in the PDF. Please do not submit a zip file. Points will be deducted if you don't follow these instructions.

1. Define the following, where $\mathbf{A}$ is symmetrical:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

   (a) (2 points) Find $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$

   (b) (2 points) Show

$$\frac{\partial \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}$$

2. Suppose that we observe n data pairs: (x_i, y_i), $i = 1, \ldots, n$. Assume that $y_i = \beta_0 + \beta_1 x_i + e_i$, where $e_i \sim \mathcal{N}(0, \sigma^2)$ and the errors (e_i) are independent. This problem asks you to consider the matrix formulation of the regression problem.

   (a) (2 points) Identify $n \times 2$ matrix $\mathbf{X}$ for the model. Hint: the first column is for the intercept and the second for predictor variable x.

   (b) (2 points) Compute $\mathbf{X}^\mathsf{T}\mathbf{X}$. Write the answer in terms such as n and $\sum_i x_i$.

   (c) (2 points) Is your matrix $\mathbf{X}^\mathsf{T}\mathbf{X}$ symmetrical?

   (d) (4 points) Now suppose that you have p predictors instead of 1, so that $\mathbf{X}$ is now $n \times (p + 1)$. Show that $\mathbf{X}^\mathsf{T}\mathbf{X}$ is symmetrical. Hint: if $\mathbf{A} = \mathbf{X}^\mathsf{T}\mathbf{X}$, show that $a_{ij} = a_{ji}$.

3. Consider the regression model $y_i = \alpha + \beta x_i + e_i$, where e_i are independent random variables with $\mathbb{E}(e_i) = 0$ and $\mathbb{V}(e_i) = \sigma^2$ for all i.

   (a) (2 points) What is the implication for the regression function if $\beta = 0$, so that the model is $y_i = \alpha + e_i$? How would the regression function plot on a graph?

   (b) (3 points) Derive the least square estimator a of α for the model above (with $\beta = 0$) and show that it equals the sample mean $a = \bar{y}$.

   (c) (3 points) Prove that the estimate a in the previous part is an unbiased estimator of α.

   (d) (3 points) What is the variance of your estimate a?

   (e) (3 points) Discuss why your estimates are (at least approximately) normally distributed.

(f) The Gauss-Markov theorem states that OLS estimates are best linear unbiased estimates ("BLUE"), i.e., among all linear, unbiased estimates, the OLS estimates have the smallest variance. Show that your estimate from part (b) is BLUE. Hints: Let $\hat{\alpha} = \sum_{i=1}^{n} c_i y_i$ be another linear (it is a linear combination of y_i) unbiased estimate, where c_i are constants. Let $d_i = c_i - 1/n$ be the difference between the constants of the new estimator and those from OLS $(1/n)$. Show that $d_i = 0$ for all i, otherwise the variance will be greater than that of $\bar{y}$ from part (d). When $d_i = 0$ the new estimate is the same as the OLS one.

i. (2 points) What does the unbiased assumption imply about the sum of c_i?

ii. (2 points) Show $\sum_i d_i/n = 0$.

iii. (2 points) Evaluate $\mathbb{V}(\hat{\alpha})$ in terms of d_i and find when it is minimized over the d_i values.

iv. Or you can think geometrically.

4. (8 points) ACT problem 2.5: show $\overline{y}$ and $\widehat{\beta}_1$ are independent by showing

$$\mathsf{C}(\overline{y}, \hat{\beta}_1) = \frac{1}{S_{xx}} \sum_{i=1}^{n} \mathsf{C}(\overline{y}, y_i) = 0$$

Hint: it is useful to establish the following lemmas: $\mathsf{C}(aX, bY) = ab\mathsf{C}(X, Y)$ and $\mathsf{C}(X + Y, Z) = \mathsf{C}(X, Z) + \mathsf{C}(Y, Z)$.