ECE 492-45 Homework 2 (Fall 2021)
Instructor: Dr. Chau-Wai Wong
Material Covered: Linear Regression, Least Squares, Basic Linear Algebra, Off-the-Shelf Deep Learning Tools

Problem 1 (20 points) [Deriving Least-Squares Estimator]

a) Write model $y_i = \beta_0 + \beta_1 x_i + e_i,\ i = 1, \ldots, n$ into the matrix–vector form. Clearly indicate what are $\underset{\sim}{y}, \mathbf{X}, \underset{\sim}{\beta}$, and $\underset{\sim}{e}$. Set up the cost function $J(\underset{\sim}{\beta})$ to be minimized. Take gradient with respect to $\underset{\sim}{\beta}$ and set it to zero. Express $\hat{\underset{\sim}{\beta}}$ in terms of $\mathbf{X}$ and $\underset{\sim}{y}$. What are the closed-form solution for $\hat{\beta}_0$ and $\hat{\beta}_1$. Is $\hat{\beta}_1$ consistent with that in b)?

b) We have derived in class the least-squares estimator for p predictors. Now, show through partial differentiation that the least-squares estimator for β_1 of model $y_i = \beta_0 + \beta_1 x_i + e_i$ is

$$\hat{\beta}_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}. \tag{1}$$

Prove that the above expression is equivalent to

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}. \tag{2}$$

Problem 2 (20 points) [Least-Squares Estimator in Matrix–Vector Form] An ECE student John is doing an electronic circuits lab during which he needs to determine the conductance of a resistor using a voltage meter, a current meter, and a DC power source. The voltage meter is connected in parallel with the resistor and the current meter is in series with the resistor. Both meters are analog devices so the readings recorded by John have errors. The power source is tunable and has a range of 1 to 5 V. Each time John will try a uniformly random input voltage level and record the readings of both voltage and current meters. Denote the voltage reading as x_i and the current reading as y_i for the ith measurement. Assume the true conductance $G = 2\ \mathrm{m}\Omega^{-1}$.

a) Using a linear model $y_i = Gx_i + e_i$, where e_i is a zero-mean noise with standard deviation $\sigma_e = 0.1$ mA, simulate a dataset of $n = 10$ measurements. [Hint: You may use `rand()` to generate a uniformly random value in $[0, 1]$ and `0.1*randn()` to generate a zero-mean Gaussian noise with standard deviation 0.1. Note that the exact function names are dependent on your programming language. Throughout this problem, you may ignore the units such as $\mathrm{m}\Omega^{-1}$ and mA and focus only on the numbers.]

b) Express the linear model in a matrix-vector form. Clearly indicate the matrices/vectors $\underset{\sim}{y}, \mathbf{X}, \boldsymbol{\beta}$, and $\underset{\sim}{e}$. Directly implement the formula of the least-squares (LS) estimator, $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\underset{\sim}{y}$, into a computer function that takes as two input vectors $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, and output a number $\hat{G}$. Apply your function to the simulated data. What is the value of $\hat{G}$? [Hint: $\mathbf{X}$ is a n-by-1 "matrix," and $\boldsymbol{\beta}$ is a 1-by-1 "vector." If you are using Python, please use NumPy objects to store matrices and vectors. In R, you may generate a matrix using the `matrix()` function and carry out a matrix multiplication using operator `%*%`.]

c) John's friend, Tom, proposed a more intuitive estimator for the conductance: $\tilde{G} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i}$. Let's call it "Tom's estimator" for convenience. Write a computer function that takes as two input vectors $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, and output a number $\tilde{G}$. Apply your function to the simulated data. What is the value of $\tilde{G}$?

d) Generate 1,000 datasets. Repeatedly apply function written in (b) and collect 1,000 LS estimates and calculate the sample variance of these 1,000 values.

e) Use the 1,000 datasets generated in (d). Repeatedly apply function written in (c) and collect 1,000 Tom's estimates and calculate the sample variance of these 1,000 values. You should find that the LS estimator has a smaller variance than Tom's estimator.

Problem 3 (20 points) [Deep Learning with Matlab] In recent updates, Matlab has put together well-guided tutorials for deep learning. This is one set of tutorials on "Deep Learning with Images":

`https://www.mathworks.com/help/deeplearning/deep-learning-with-images.html`

Complete the following tutorials by running the example code. Write a concise report consisting of key source code, images, and your explanations.

a) Tutorial "Classify Webcam Images Using Deep Learning."

b) Tutorial "Create Simple Deep Learning Network for Classification."

c) (Bonus, 5') Tutorial "Transfer Learning with Deep Network Designer."

For more tutorials, see the left menu on this page: `https://www.mathworks.com/help/deeplearning/getting-started-with-deep-learning-toolbox.html`

These tutorials may give you ideas about your term project.

Problem 4 (20 points) [Linearly independence, Basis, and Vector Space]

a) Are vectors $[1\ \ 2], [4\ \ 5]$, and $[7\ \ 8]$ linearly independent? What about $[1\ \ 2\ \ 0], [1\ \ -1\ \ 1]$, and $[0\ \ 0\ \ 1]$? Justify your answers.

b) (Try it after 8/30 lecture) You are given a vector space $V = \text{span}\,\{[-1\ \ 0\ \ 0], [0\ \ -1\ \ 0]\}$.

(i) Express V in a set representation.

(ii) Can you find a basis for V?

(iii) Are $[5\ \ 8\ \ 0]$, $[8\ \ 0\ \ 5]$, and $[0\ \ 5\ \ 8]$ in vector space V? Is yes, what are the coefficient for each vector of the basis you found in (ii)?

(iv) Draw all points of (iii) in a 3D coordinate. Illustrate vector space V using a plane formed by the vectors of the basis.

c) (Try it after 8/30 lecture) Let

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & -1 & 0 \\ -1 & -1 & 2 & -3 \\ 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

What is the dimension of the column vector space of $\mathbf{A}$? What is the rank of $\mathbf{A}$?