HW3: More Linear Regression

Stat 154, Fall 2019

Problem 1

We examine a response variable Y in terms of two predictors X and Z. There are n observations. Let $\mathbf{X}$ be a matrix formed by a constant term of $\mathbf{1}$, and the vectors $\mathbf{x}$ and $\mathbf{z}$. Consider the cross-product matrix $\mathbf{X}^\mathsf{T}\mathbf{X}$ given below:

$$\mathbf{X}^\mathsf{T}\mathbf{X} = \begin{bmatrix} 30 & 0 & 0 \\ ? & 10 & 7 \\ ? & ? & 15 \end{bmatrix}$$

a. Complete the missing values denoted by "?", and determine the value of n?

b. Calculate the linear correlation coefficient between X and Z.

c. If the OLS regression equation is: $\hat{y}_i = -2 + x_i + 2z_i$, What is the value of $\bar{y}$?

d. If the residual sum of squares (RSS) is 12, What is the value of R^2? Recall that $\text{RSS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

Problem 2

Consider a linear regression model:

$$y_i = w_0 x_{i0} + w_1 x_{i1} + w_p x_{ip} + \epsilon_i = \mathbf{w}^\mathsf{T}\mathbf{x_i} + \epsilon_i$$

and assume that the noise terms ϵ_i are independent and have a Gaussian distribution with mean zero and constant variance σ^2:

$$\epsilon_i \sim N(0, \sigma^2)$$

In lecture, we discussed how to obtain the parameters $\mathbf{w} = (w_0, w_1, \dots, w_p)$ of a linear regression model via Maximum Likelihood (ML), which are given by: $\mathbf{w} = (\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}\mathbf{y}$

Determine the ML estimate of the other model paramater: σ^2 (the constant variance).

Problem 3

Multicollinearity is easy to detect in a linear regression model with two predictors; we need only look at the value of $r_{12} = cor(X_1, X_2)$. When there are more than two regressors, however, inspection of the r_{ij} is not sufficient.

For example, assume that we have four predictors X_1, X_2, X_3 and X_4, and correlation coefficients, r_{ij} are $r_{12} = r_{13} = r_{23} = 0$, with variances $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$, and $X_4 = X_1 + X_2 + X_3$.

Show that $r_{14} = r_{24} = r_{34} = 0.577$

Problem 4

Consider minimizing a quadratic function $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\mathsf{T}\mathbf{x}$, where $\mathbf{b}$ is a vector and $\mathbf{A}$ is a positive semidefinite matrix.

Suppose that $\mathbf{A}$ is invertible. Show that the minimum of $f(\mathbf{x})$ is given by $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$.

Problem 5

Let

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Let $f_1(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}_1\mathbf{x} - \mathbf{b}^\top\mathbf{x}$ and $f_2(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}_2\mathbf{x} - \mathbf{b}^\top\mathbf{x}$.

Implement Gradient Descent to minimize both $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. For each function, run gradient descent with 5 different random initializations, and print the solutions. Are they the same for each random initialization? Explain.