Homework 6: Linear Regression

Instructor: Daniel L. Pimentel-Alarcón Due 04/16/2019

Problem 6.1. In the lecture notes we saw that the goal in linear regression is to find

$$\boldsymbol{\beta}^\star \;=\; \underset{\boldsymbol{\beta}\in\mathbb{R}^{D+1}}{\arg\min}\; \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2.$$

To this end we used matrix derivatives to conclude that

$$\boldsymbol{\beta}^\star \;=\; (\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}\mathbf{y}. \tag{6.1}$$

In this problem you will use scalar derivatives to show that this is correct. First observe that

$$\begin{aligned}
g(\boldsymbol{\beta}) &= \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 \\
&= \sum_{i=1}^{N}(y_i - \mathbf{x}_i^\mathsf{T}\boldsymbol{\beta})^2 \\
&= \sum_{i=1}^{N}\left(y_i^2 - 2y_i\mathbf{x}_i^\mathsf{T}\boldsymbol{\beta} + (\mathbf{x}_i^\mathsf{T}\boldsymbol{\beta})^2\right) \\
&= \sum_{i=1}^{N}\left[y_i^2 - 2y_i\sum_{j=1}^{D}\beta_j x_{ij} + \left(\sum_{j=1}^{D}\beta_j x_{ij}\right)^2\right].
\end{aligned}$$

(a) Compute the derivative of $g(\boldsymbol{\beta})$ with respect to β_k.

(b) Set the derivative of $g(\boldsymbol{\beta})$ with respect to β_k to zero, and solve for β_k.

Problem 6.2. Consider the following vector $\mathbf{y}$, containing information about glucose level of three individuals, and the following data matrix $\mathbf{X}$ containing information about height and weight of the corresponding individuals:

$$\mathbf{y} = \begin{bmatrix}110\\140\\180\end{bmatrix}, \qquad \mathbf{X} \;=\; \begin{bmatrix}180 & 150\\150 & 175\\170 & 165\end{bmatrix}.$$

(a) Use (6.1) to find the vector $\boldsymbol{\beta}$ that best explains $\mathbf{y}$ as a linear function of $\mathbf{X}$.

(b) Use your answer from Part (b) in Problem 6.1 to find the vector $\boldsymbol{\beta}$ that best explains $\mathbf{y}$ as a linear function of $\mathbf{X}$.

(c) Do your answers from parts (a) and (b) coincide?

(d) Which option do you prefer?

(e) Given a new sample with feature vector $\mathbf{x} = [175\;\;170]^\mathsf{T}$, what would be your prediction for y?