PROBLEMS

Due Fri. Jan. 20, 2023 @ 4pm

1. A soccer team in the A league wants to recruit a player from the B league, which is at a lower level.

   (a) For this supervised learning problem, identify a possible target variable. (There are many "correct" answers.) Explain why they have access to data on this variable.

   (b) What scalar feature would work well to predict this target variable? Explain why they have access to data on this variable.

   (c) For the above target and feature, would a linear model be reasonable? Assuming a linear model was used, would you expect the slope to be positive or negative?

2. Suppose that we are given data samples (x_i, y_i):

x_i	0	1	2	3	4
y_i	0	0.5	1	0.5	3

   (a) What are the sample means, $\bar{x}$ and $\bar{y}$?

   (b) What are the sample variances and co-variance, s_x^2, s_y^2 and s_{xy}?

   (c) What are the least-squares parameters β_0, β_1 of the linear model

$$y \approx \beta_0 + \beta_1 x.$$

   (d) Plot the scatterplot and the regression line for the LS linear model.

   (e) What is the value of R^2 for the LS linear model?

   (f) Is linear regression effective on this dataset?

3. In seismology, the Gutenberg-Richter law expresses the relationship between the magnitude M and the total number of earthquakes N with that magnitude:

$$N = 10^{\gamma - \alpha M} \tag{1}$$

   where γ and α are constants. To estimate the parameters γ and α, a researcher collects a large number of pairs $\{(N_i, M_i)\}_{i=1}^n$ for a given region. Unfortunately, the model (1) is nonlinear, so she can't directly apply the linear regression formula.

(a) Show that we can transform (1) into a linear model of the form

$$y = \beta_0 + \beta_1 x, \tag{2}$$

where the target y is some function of N, the feature x is some function of M, and β_0 and β_1 are some functions of γ and α, respectively. Note that we are asking for a transformation, not an approximation; if (1) holds with equality then (2) should hold with equality.

(b) Using the result of part (a), describe all the steps in computing the least-squares estimates of the linear-model parameters $\{\beta_0, \beta_1\}$ from the given data.

(c) Now describe how the parameter estimates from part (b) can be transformed to get estimates of the original parameters γ and α.

4. Consider a linear model of the form,

$$y \approx \beta x,$$

which is a linear model, but with the intercept forced to zero. This occurs in applications where we want to force the predicted value $\widehat{y} = 0$ when $x = 0$. For example, if we are modeling $y =$ output power of a motor vs. $x =$ the input power, we would expect $x = 0 \Rightarrow y = 0$.

(a) Given data (x_i, y_i), write a cost function representing the residual sum of squares (RSS) between y_i and the predicted value $\widehat{y}_i$ as a function of β.

(b) Taking the derivative with respect to β, find the β that minimizes the RSS.