Homework 3

due date: 11:59pm (Eastern time), Friday Oct 23, 2020

Questions 12 and 14 in Chapter 3.7 Exercises in An Introduction to Statistical Learning
(http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf).

Questions 9.1 and 9.4 in Applied Linear Regression book (4th Edition); see below for descriptions of these questions.

9.1 (Data file: `Rpdata`) The data in this file has a response `y` and six regressors `x1`, . . ., `x6`. The data are artificial, to make a few points.

Data link: http://users.stat.umn.edu/~sandy/alr4ed/data/

9.1.1 First draw a scatterplot matrix of all data and comment. Is there anything strange?

9.1.2 Fit the OLS regression `y ~ x1 + x2 + x3 + x4 + x5 + x6`. Is there anything strange?

9.1.3 Draw a plot of residuals versus fitted values. Is there anything strange? See Stefanski (2007) if you want to find out how this data set came about.

9.4 Simple regression Consider the simple regression model, $\mathrm{E}(Y|X = x) = \beta_0 + \beta_1 x$, $\mathrm{Var}(Y|X = x) = \sigma^2$.

9.4.1 Find a formula for the h_{ij} and for the leverages h_{ii}.

9.4.2 In a 2D plot of the response versus the predictor in a simple regression problem, explain how high-leverage points can be identified.

9.4.3 Make up a predictor X so that the value of the leverage in simple regression for one of the cases is equal to 1.
