

## DSC 241 – Homework 3

**Problem 1.** (Is polynomial regression well-conditioned?) Consider the canonical design matrix for fitting a polynomial of degree  $p$  based on  $x_1, \dots, x_n$ , meaning,  $\mathbf{X} = (x_i^j)$  for  $i = 1, \dots, n$  and  $j = 0, \dots, p$ . Suppose the  $x_i$ 's are evenly distributed in  $(0, 1)$ , for example,  $x_i = i/(n+1)$  for  $i = 1, \dots, n$ . For each  $p \in \{1, \dots, 20\}$  and each  $n \in \{30, 50, 100, 200, 500, 1000\}$ , compute the condition number of the design matrix. Produce a useful plot for visualizing the result of these computations. Offer brief comments.

**Problem 2.** (Piecewise constant fit.)

- a. Write a function `piecewiseConstant(x, y, L, plot = TRUE)` taking in a one dimensional predictor variable  $x$  with values in  $[0, 1]$  and a response  $y$ , and fits a piecewise constant model (by least squares) on  $2^L$  intervals of equal length partitioning the unit interval ( $L$  is a nonnegative integer) in the form of a numerical vector of length  $2^L$ , with the option of producing a scatterplot with the fit overlaid.
- b. Apply your function to explaining City Mpg as a piecewise constant function of Horsepower in the `04cars` dataset. Produce a single scatterplot, with lines corresponding to the fit with  $L = 2$  (blue),  $L = 3$  (green), and  $L = 4$  (red). Add a legend, etc, so it looks 'nice'.