Spring 2020

Quantitative Analysis
Chung-Ming Kuan

Problem Set 9

Due: 5/18

Part One: Hand-Written Exercise

1. Let

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \xi)^3_+,$$

where $(x-\xi)^3_+ = (x-\xi)^3$ if $x > \xi$ and equals 0 otherwise. We will now show that $f(x)$ is a polynomial continuous at ξ, up to second derivatives, regardless of the values of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

   (a) Find a cubic polynomial $f_1(x)$ such that $f(x) = f_1(x)$ for all $x \leq \xi$. Find another cubic polynomial $f_2(x)$ such that $f(x) = f_2(x)$ for all $x > \xi$. We have now established that $f(x)$ is a piecewise polynomial.

   (b) Show that $f_1(\xi) = f_2(\xi)$, $f_1'(\xi) = f_2'(\xi)$, and $f_1''(\xi) = f_2''(\xi)$. Therefore, $f(x)$, $f'(x)$, and $f''(x)$ are all continuous at ξ.

   (c) Show that $\frac{\partial^3 f_1(\xi)}{\partial x} \neq \frac{\partial^3 f_2(\xi)}{\partial x}$. Therefore, $\frac{\partial^3 f(x)}{\partial x}$ is not continuous at ξ.

Part Two: Computer Exercise

1. Please load the data set `Wage` from the package `ISLR`. Suppose we want to fit the variable of interest `wage` on a single predictor `age`, using cubic spline.

   (a) Please use 10-fold CV to determine the optimal interior knots (from 1 to 5) for the cubic spline to fit our data.

   (b) Please draw the scatter plot with the cubic spline(blue line) with the optimal interior knots you just found.

2. Continue with question (1), instead of a cubic spline, suppose now we wish to fit a natural cubic spline using 4 combinations of boundary knots, $(q_{.05}, q_{.95}), (q_{.1}, q_{.9}), (q_{.15}, q_{.85})$ $, (q_{.2}, q_{.8})$, where q_k is the $100k^{th}$ quantile of `age`.

   (a) Please use 10-fold CV to determine the optimal combination of boundary knots and use `df`=4 for the natural cubic spline to fit our data. Please also show the location of the optimal boundary knots.

(b) Please draw the scatter plot with the nature cubic spline(red line) with the optimal boundary knots you just found together with the cubic spline(blue line) you found in question (1).