Spring 2022

Quantitative Analysis
Chung-Ming Kuan
Jui-Chung Yang

Problem Set 9

Due: 5/9

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

Part Two: Computer Exercise

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`.

1. Suppose we use a **polynomial regression**. Please use 50-fold CV to determine the optimal degree (from 1 to 5) for the polynomial to fit our data.

2. Suppose we use a **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) Draw the scatter plot with the cubic spline (blue line) with the optimal interior knots you just found.

3. Suppose we use a **natural cubic spline** with 4 different groups of boundary knots, $(q_{.05}, q_{.95})$, $(q_{.1}, q_{.9})$, $(q_{.15}, q_{.85})$, and $(q_{.2}, q_{.8})$, where q_k is the $100k^{th}$ quantile of `age`.

(a) Please use 10-fold CV to determine the optimal group 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 of the nature cubic spline (red line) with the optimal boundary knots together with the cubic spline (blue line) that you found in the previous question.