Machine Learning & Big Data
ECON590
Problem Set 5

1. (8 points) This exercise shows that a function of the form

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

where $(x-\xi)^3_+ = (x-\xi)^3$ if $x > \xi$ and equals 0 otherwise, is indeed a cubic regression spline with one knot at ξ. That is, $f(x)$ is a piecewise cubic polynomial, continuous at ξ, with continuous first and second derivatives at ξ.

   (a) First, we show that $f(x)$ is a piecewise polynomial. To that end, find a cubic polynomial
   $$f_1(x) = a_1 + b_1 x + c_1 x^2 + d_1 x^3$$
   such that $f(x) = f_1(x)$ for all $x \leqslant \xi$. Express a_1, b_1, c_1, d_1 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.
   Next, find a cubic polynomial
   $$f_2(x) = a_2 + b_2 x + c_2 x^2 + d_2 x^3$$
   such that $f(x) = f_2(x)$ for $x > \xi$. Express a_2, b_2, c_2, d_2 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

   (b) Show that $f(x)$ is continuous at ξ, i.e., $f_1(\xi) = f_2(\xi)$.

   (c) Show that the first derivative of $f(x)$ is continuous at ξ, i.e., $f_1'(\xi) = f_2'(\xi)$.

   (d) Show that the second derivative of $f(x)$ is continuous at ξ, i.e., $f_1''(\xi) = f_2''(\xi)$.

2. (5 points) Suppose $\widehat{g}$ is computed to smoothly fit a set of n points by solving

$$\widehat{g} = \underset{g}{\operatorname{argmin}} \left(\sum_{i=1}^{n} (y_i - g(x_i))^2 + \lambda \int [g^{(m)}(x)]^2 \mathrm{d}x \right),$$

where $g^{(m)}$ is the mth derivative of g with $g^{(0)} = g$. For each of the following scenarios, describe $\widehat{g}$ or provide an example sketch. Be sure to explain your answer.

   (a) $\lambda = \infty, m = 0$.

   (b) $\lambda = \infty, m = 1$.

   (c) $\lambda = \infty, m = 2$.

   (d) $\lambda = \infty, m = 3$.

   (e) $\lambda = 0, m = 2$.

3. (7 points) This question involves the `Wage` dataset, which can be downloaded from the ISLR website.

(a) Fit `wage` to `age` using a (cubic) regression spline with knots at ages 30, 50 and 60. Plot the resulting curve. For a 30-year-old, what wage does the model predict?

(b) Fit `wage` to `age` using a smoothing spline, where the tuning parameter λ is chosen by cross-validation. Plot the resulting curve. For a 30-year-old, what wage does the model predict?

(c) Now fit a GAM to predict `wage` using `age` and `maritl` (marital status). Use a natural spline with 4 degrees of freedom for `age` and a separate constant for each category for `maritl`. Plot the resulting curve. For a married, 30-year-old, what wage does the model predict?