CS 5565, HW6(Moving Beyond Linearity) 85pts (65pts Undergrad)

Name ____________________________________

1. (15 points total) Suppose that a curve $\hat{g}$ is computed to smoothly fit a set of n points using the following formula:

$$\hat{g} = \arg\min_g \left(\sum_{i=1}^{n} (y_i - g(x_i))^2 + \lambda \int \left[g^{(m)}(x) \right]^2 dx \right)$$

where $g^{(m)}$ represents the mth derivative of g (and $g^{(0)} = g$). Provide example sketches of $\hat{g}$ in each of the following scenarios.

   (a) (3 points) $\lambda = \infty,\ m = 0$

   (b) (3 points) $\lambda = \infty,\ m = 1$

   (c) (3 points) $\lambda = \infty,\ m = 2$

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

   (e) (3 points) $\lambda = 0,\ m = 3$

2. (10 points total) Suppose we fit a curve with basis functions $b_1(X) = X$, $b_2(X) = (X-1)^2 I(X \geq 1)$. (Note that $I(X \geq 1)$ equals 1 for $X \geq 1$ and 0 otherwise.) We fit the linear regression model

$$Y = \beta_0 + \beta_1 b_1(X) + \beta_2 b_2(X) + \epsilon,$$

and obtain coefficient estimates $\hat{\beta}_0 = 1,\ \hat{\beta}_1 = 1,\ \hat{\beta}_2 = -2$.

Sketch the estimated curve between $X = -2$ and $X = 2$. Note the intercepts, slopes, and other relevant information.

3. (10 points total) Suppose we fit a curve with basis functions $b_1(X) = I(0 \leq X \leq 2) - (X-1)I(1 \leq X \leq 2)$, $b_2(X) = (X-3)I(3 \leq X \leq 4) + I(4 < X \leq 5)$. We fit the linear regression model

$$Y = \beta_0 + \beta_1 b_1(X) + \beta_2 b_2(X) + \epsilon,$$

and obtain coefficient estimates $\hat{\beta}_0 = 1,\ \hat{\beta}_1 = 1,\ \hat{\beta}_2 = 3$.

Sketch the estimated curve between $X = -2$ and $X = 2$. Note the intercepts, slopes, and other relevant information.

4. (15 points total) Consider two curves, $\hat{g}_1$ and $\hat{g}_2$ defined by

$$\begin{aligned}
\hat{g}_1 &= \arg\min_g \left(\sum_{i=1}^{n} (y_i - g(x_i))^2 + \lambda \int \left[g^{(3)}(x) \right]^2 dx \right), \\
\hat{g}_2 &= \arg\min_g \left(\sum_{i=1}^{n} (y_i - g(x_i))^2 + \lambda \int \left[g^{(4)}(x) \right]^2 dx \right),
\end{aligned}$$

where $g^{(m)}$ represents the mth derivative of g.

(a) (5 points) As $\lambda \longrightarrow \infty$, will $\hat{g}_1$ or $\hat{g}_2$ have the smaller training RSS?

(b) (5 points) As $\lambda \longrightarrow \infty$, will $\hat{g}_1$ or $\hat{g}_2$ have the smaller test RSS?

(c) (5 points) For $\lambda = 0$, will $\hat{g}_1$ or $\hat{g}_2$ have the smaller training and test RSS?

5. (20 points total (extra credit undergraduate)) It was mentioned in the chapter that a cubic regression spline with one knot at ξ can be obtained using a basis of the form x, x^2, x^3, $(x-\xi)^3_+$, where $(x-\xi)^3_+ = (x-\xi)^3$ if $x > \xi$ and equals 0 otherwise. We will now show that a function of the form

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

is indeed a cubic regression spline, regardless of the values of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

(a) (5 points) 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 $x \leq \xi$. Express a_1, b_1, c_1, d_1 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

(b) (5 points) 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 all $x > \xi$. Express a_2, b_2, c_2, d_2 in terms of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$.

(c) (5 points) Show that $f_1(\xi) = f_2(\xi)$. That is, $f(x)$ is continuous at ξ.

(d) (5 points) Show that $f_1'(\xi) = f_2'(\xi)$. That is, $f'(x)$ is continuous at ξ.

(e) (5 points) Show that $f_1''(\xi) = f_2''(\xi)$ i.e. $f(x)$ is indeed a cubic spline.

6. (15 points) Construct a natural cubic spline that passes through the points (1,2),(2,4), and (3,3).