

ISLR | Chapter 7 Exercises

Marshall McQuillen

9/12/2018

Conceptual

1

- **A.** The cubic piecewise polynomial:

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \xi)_+^3 \quad \text{where} \quad (x - \xi)_+^3 = \begin{cases} 0, & x \leq \xi \\ (x - \xi)^3, & \text{otherwise} \end{cases}$$

...can be broken up and rewritten to be:

$$f(x) = \begin{cases} f_1(x) = a_1 + b_1 x + c_1 x^2 + d_1 x^3, & x \leq \xi \\ f_2(x) = a_2 + b_2 x + c_2 x^2 + d_2 x^3, & \text{otherwise} \end{cases}$$

In $f_1(x)$, since $(x - \xi)_+^3 = 0$ (because $x \leq \xi$), the fifth term (of $f(x)$) zeroes out and the coefficients can be expressed as $a_1 = \beta_0$, $b_1 = \beta_1$, $c_1 = \beta_2$ and $d_1 = \beta_3$.

- **B.** Expanding the fifth term in $f(x)$ allows for the various powers of x to be grouped together and then recondensed. a_2 , b_2 , c_2 and d_2 are expressed in terms of the coefficients below.

$$f_2(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \xi)^3 \quad (1)$$

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

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

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x^3 - x^2\xi - 2x^2\xi + 2x\xi^2 + \xi^2 x - \xi^3) \quad (4)$$

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x^3 - 3x^2\xi + 3x\xi^2 - \xi^3) \quad (5)$$

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^3 - \beta_4 3x^2\xi + \beta_4 3x\xi^2 - \beta_4 \xi^3 \quad (6)$$

$$= (\beta_0 - \beta_4 \xi^3) + (\beta_1 x + \beta_4 3x\xi^2) + (\beta_2 x^2 - \beta_4 3x^2\xi) + (\beta_3 x^3 + \beta_4 x^3) \quad (7)$$

$$= (\beta_0 - \beta_4 \xi^3) + (\beta_1 + 3\beta_4 \xi^2)x + (\beta_2 - 3\beta_4 \xi)x^2 + (\beta_3 + \beta_4)x^3 \quad (8)$$

$$f_2(x) = a_2 + b_2 x + c_2 x^2 + d_2 x^3 \quad \text{where} \quad \begin{cases} a_2 = \beta_0 - \beta_4 \xi^3 \\ b_2 = \beta_1 + 3\beta_4 \xi^2 \\ c_2 = \beta_2 - 3\beta_4 \xi \\ d_2 = \beta_3 + \beta_4 \end{cases} \quad (9)$$

- **C.** Showing that $f(x)$ is continuous at ξ is illustrated by showing that $f(\xi)_1 = f(\xi)_2$.

$$f_1(\xi) = a_1 + b_1(\xi) + c_1(\xi)^2 + d_1(\xi)^3 \quad (10)$$

$$= \beta_0 + \beta_1(\xi) + \beta_2(\xi)^2 + \beta_3(\xi)^3 \quad (11)$$

$$(12)$$

$$f_2(\xi) = a_2 + b_2(\xi) + c_2(\xi)^2 + d_2(\xi)^3 \quad (13)$$

$$= (\beta_0 - \beta_4\xi^3) + (\beta_1 + 3\beta_4\xi^2)(\xi) + (\beta_2 - 3\beta_4\xi)(\xi)^2 + (\beta_3 + \beta_4)(\xi)^3 \quad (14)$$

$$= (\beta_0 - \beta_4\xi^3) + (\beta_1\xi + 3\beta_4\xi^3) + (\beta_2\xi^2 - 3\beta_4\xi^3) + (\beta_3\xi^3 + \beta_4\xi^3) \quad (15)$$

$$= \beta_0 - \beta_4\xi^3 + \beta_1\xi + 3\beta_4\xi^3 + \beta_2\xi^2 - 3\beta_4\xi^3 + \beta_3\xi^3 + \beta_4\xi^3 \quad (16)$$

$$= \beta_0 + \beta_1\xi + \beta_2\xi^2 + \beta_3\xi^3 + 3\beta_4\xi^3 - 3\beta_4\xi^3 + \beta_4\xi^3 - \beta_4\xi^3 \quad (17)$$

$$= \beta_0 + \beta_1\xi + \beta_2\xi^2 + \beta_3\xi^3 + (3\beta_4\xi^3 - 3\beta_4\xi^3) + (\beta_4\xi^3 - \beta_4\xi^3) \quad (18)$$

$$f_2(\xi) = \beta_0 + \beta_1\xi + \beta_2\xi^2 + \beta_3\xi^3 \quad (19)$$

$$f_2(\xi) = \beta_0 + \beta_1\xi + \beta_2\xi^2 + \beta_3\xi^3 = f_1(\xi)$$

- **D.** In order to show that $f'_1(\xi) = f'_2(\xi)$, we must first find $f'(x)$ with respect to x and then simplify both $f'_1(\xi)$ and $f'_2(\xi)$.

$$f(x) = a_1 + b_1x + c_1x^2 + d_1x^3 \quad (20)$$

$$f'(x) = b_1 + 2c_1x + 3d_1x^2 \quad (21)$$

Therefore, substituting the necessary coefficients in for b_1 , c_1 and d_1 in both $f'_1(\xi)$ and $f'_2(\xi)$, we get:

$$f'(x) = b_1 + 2c_1x + 3d_1x^2 \quad \text{then} \quad \begin{cases} f'_1(\xi) = \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 \\ f'_2(\xi) = (\beta_1 + 3\beta_4\xi^2) + 2(\beta_2 - 3\beta_4\xi)\xi + 3(\beta_3 + \beta_4)\xi^2 \end{cases} \quad (22)$$

$$f'_2(\xi) = (\beta_1 + 3\beta_4\xi^2) + 2(\beta_2 - 3\beta_4\xi)\xi + 3(\beta_3 + \beta_4)\xi^2 \quad (23)$$

$$= \beta_1 + 3\beta_4\xi^2 + 2\beta_2\xi - 6\beta_4\xi^2 + 3\beta_3\xi^2 + 3\beta_4\xi^2 \quad (24)$$

$$= \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 + (3\beta_4\xi^2 + 3\beta_4\xi^2 - 6\beta_4\xi^2) \quad (25)$$

$$= \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 + (6\beta_4\xi^2 - 6\beta_4\xi^2) \quad (26)$$

$$f'_2(\xi) = \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 \quad (27)$$

We now see that the derivative $f'(x)$ is continuous at knot ξ , which is to say $f'_1(\xi) = f'_2(\xi)$:

$$f'_2(\xi) = \beta_1 + 2\beta_2\xi + 3\beta_3\xi^2 = f'_1(\xi)$$

- **E.** In order to show that $f_1''(\xi) = f_2''(\xi)$, we must first find $f''(x)$ with respect to x and then simplify both $f_1''(\xi)$ and $f_2''(\xi)$.

$$f(x) = a_1 + b_1x + c_1x^2 + d_1x^3 \quad (28)$$

$$f'(x) = b_1 + 2c_1x + 3d_1x^2 \quad (29)$$

$$f''(x) = 2c_1 + 6d_1x \quad (30)$$

Therefore, substituting the necessary coefficients in for c_1 and d_1 in both $f_1''(\xi)$ and $f_2''(\xi)$, we come to:

$$f''(x) = 2c_1 + 6d_1x \quad \text{then} \quad \begin{cases} f_1''(\xi) = 2\beta_2 + 6\beta_3\xi \\ f_2''(\xi) = 2(\beta_2 - 3\beta_4\xi) + 6(\beta_3 + \beta_4)\xi \end{cases} \quad (31)$$

$$f_2''(\xi) = 2(\beta_2 - 3\beta_4\xi) + 6(\beta_3 + \beta_4)\xi \quad (32)$$

$$= 2\beta_2 - 6\beta_4\xi + 6\beta_3\xi + 6\beta_4\xi \quad (33)$$

$$= 2\beta_2 + 6\beta_3\xi + (6\beta_4\xi - 6\beta_4\xi) \quad (34)$$

$$f_2''(\xi) = 2\beta_2 + 6\beta_3\xi \quad (35)$$

We now see that the second derivative $f''(x)$ is continuous at knot ξ , which is to say $f_1''(\xi) = f_2''(\xi)$:

$$f_2''(\xi) = 2\beta_2 + 6\beta_3\xi = f_1''(\xi)$$

2

(sketches on following page)

- **A.** With $\lambda = \infty$, the second term will dominate the above equation and the RSS will be ignored. Since $g^0 = g$, this comes out to finding $g(x)$ that minimizes the integral of $g(x)$. Therefore, $g(x) = 0$.
- **B.** With $\lambda = \infty$ and $m = 1$, the second term will dominate the above equation and the RSS will be ignored. This then becomes a problem of finding a function $g(x)$ where $\int g'(x)$ is minimized. Therefore, $g(x) = c$ (a flat line) where c is a constant, ensuring that $g'(x) = 0$.
- **C.** With $\lambda = \infty$ and $m = 2$, the second term will dominate the above equation and the RSS will be ignored. This then becomes a problem of finding a function $g(x)$ where $\int g''(x)$ is minimized.

If we work backwards conceptually, we will see that $g(x) = \beta_0 + \beta_1x$. Since $\int g''(x)$ must be minimized, $g''(x) = 0$. Therefore, $g'(x) = c$ where c is some constant. This implies that $g(x)$ must have a constant slope, c aka β_1 . Therefore, $g(x) = \beta_0 + \beta_1x$

- **D.** With $\lambda = \infty$ and $m = 3$, the second term will dominate the above equation and the RSS will be ignored. This then becomes a problem of finding a function $g(x)$ where $\int g'''(x)$ is minimized. Therefore, $g(x) = \beta_0 + \beta_1x + \beta_2x^2$, $g(x)$ will be quadratic in some sense

Once again, working backwards conceptually, if the goal is to minimize $\int g'''(x)$, then $g'''(x) = 0$. Therefore, $g''(x) = c$, where c is some constant. This implies that $g'(x)$ must have a constant slope, c . if $g'(x)$ has a constant slope, then $g(x) = \beta_0 + \beta_1x + \beta_2x^2$. Having a quadratic equation means that the slope of $g(x)$ is changing at a fixed rate, which satisfies our condition that $g'(x) = c$.

- **E.** With $\lambda = 0$ and $m = 3$, the second term in the equation is completely ignored, and $g(x)$ becomes the line that interpolates all data points.

Introduction to Statistical Learning - Chapter 7 #2

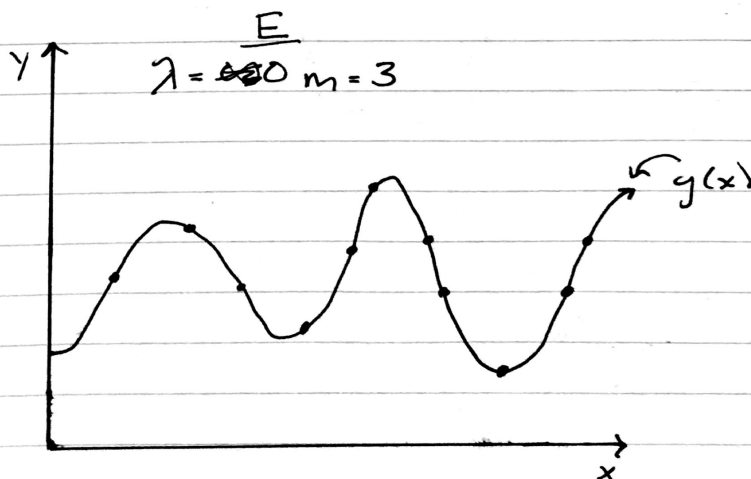
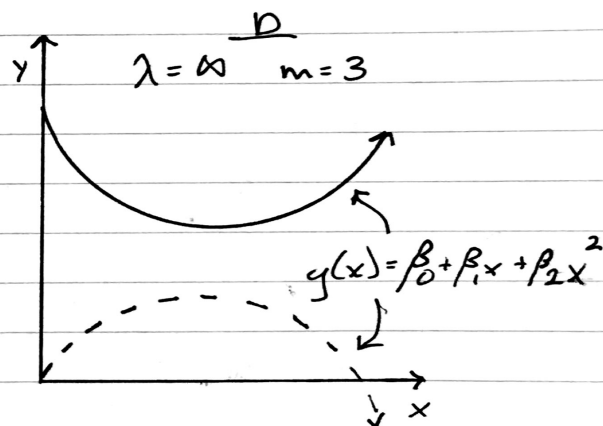
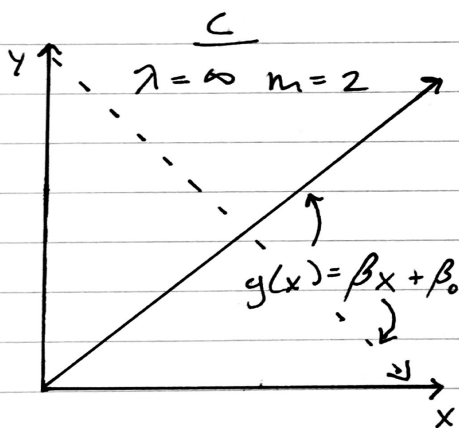
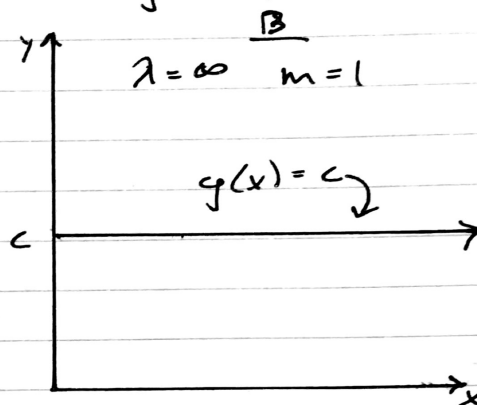
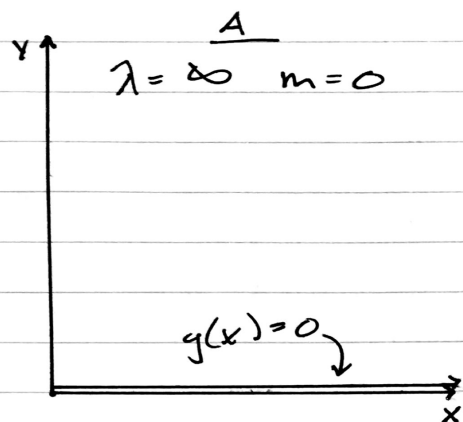


Figure 1: "Conceptual Exercise 2"