Ch 7.1-7.2: Polynomial regression and Step Functions

Lecture 21 - CMSE 381

Prof. Elizabeth Munch

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Mon, Oct 28, 2024

Announcements

Last time:

- Test 2

This lecture:

- 7.1 Polynomial regression
- 7.2 Step functions

Announcements:

- HW #6 due Sunday 11/3
- Project stuff

Lec #	Date			Reading	HW
21	Mon	10/28	Polynomial & Step Functions	7.1,7.2	
22	Wed	10/30	Step Functions; Basis functions; Start Splines	7.2 - 7.4	
23	Fri	11/1	Regression Splines	7.4	HW #7 Due Sun 11/3
24	Mon	11/4	Decision Trees	8.1	
25	Wed	11/6	Class Cancelled (Dr Munch out of town)		
26	Fri	11/8	Random Forests	8.2.1, 8.2.2	HW #8 Due Sun 11/10
27	Mon	11/11	Maximal Margin Classifier	9.1	
28	Wed	11/13	SVC	9.2	
29	Fri	11/15	SVM	9.3, 9.4	HW #9 Due Sun 11/17
30	Mon	11/18	Single layer NN	10.1	
31	Wed	11/20	Multi Layer NN	10.2	
32	Fri	11/22	CNN	10.3	HW #10 Due Sun 11/24
33	Mon	11/25	TBD: Unsupervised learning/clustering	12.1, 12.4?	
	Wed	11/27	Virtual: Project office hours		
	Fri	11/29	No class - Thanksgiving		
	Mon	12/2	Review		

Section 1

Polynomial Regression

Polynomial regression

Replace linear model

$$y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i$$

with

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_i^2 + \cdots + \beta_d x_i^d + \varepsilon_i$$

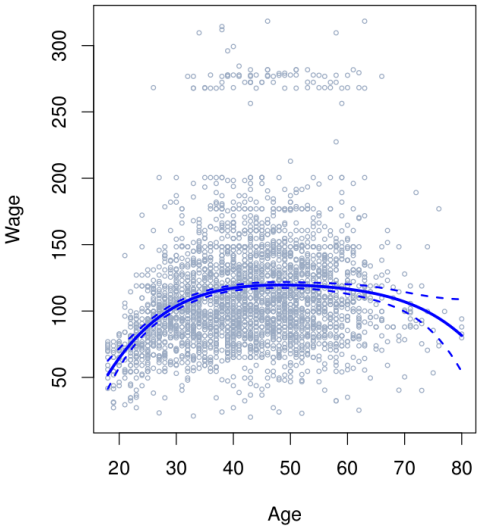
Example: Wage Data

$$\texttt{wage} = \beta_0 + \beta_1 \texttt{age} + \beta_2 \texttt{age}^2 + \cdots + \beta_p \texttt{age}^p + \varepsilon.$$

My code learned:

$$-184.1542 + 21.24552 \cdot age + -0.56386 \cdot age^2 + 0.00681 \cdot age^3 + (-3 \cdot 10^{-5}) \cdot age^4$$

Example with wage data



Section 2

Step function

Step functions

$$\mathrm{I}(X < c) \qquad \mathrm{I}(c_1 \leq X < c_2) \qquad \mathrm{I}(c \leq X)$$

More on step function setup

$$
\begin{aligned}
C_0(X) &= I(X < c_1), \\
C_1(X) &= I(c_1 \leq X < c_2), \\
C_2(X) &= I(c_2 \leq X < c_3), \\
&\vdots \\
C_{K-1}(X) &= I(c_{K-1} \leq X < c_K), \\
C_K(X) &= I(c_K \leq X),
\end{aligned}
$$

Example

Given knots $c_1 = 3$, $c_2 = 5$, $c_3 = 7$, determine the entries in the columns for $C_i(X)$ in the below matrix.

X	$C_0(X)$	$C_1(X)$	$C_2(X)$	$C_3(X)$
1				
2				
3				
4				
5				

X	$C_0(X)$	$C_1(X)$	$C_2(X)$	$C_3(X)$
6				
7				
8				
9				
10				

Draw the function

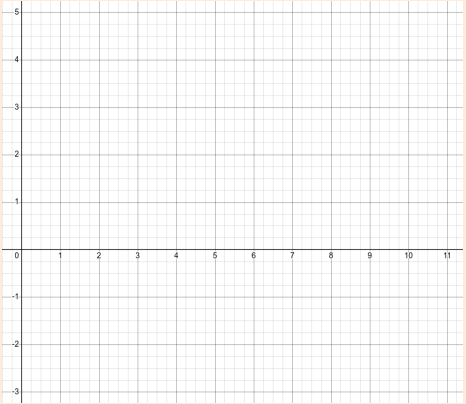
My code doing regression on the step function input returned the function.

$$f(X) = -1 + 3C_1(X) + 4C_2(X) - 2C_3(X).$$

Fill in the table of values, then draw this function below.

X	$F(X)$
1	
2	
3	
4	
5	

X	$F(X)$
6	
7	
8	
9	
10	



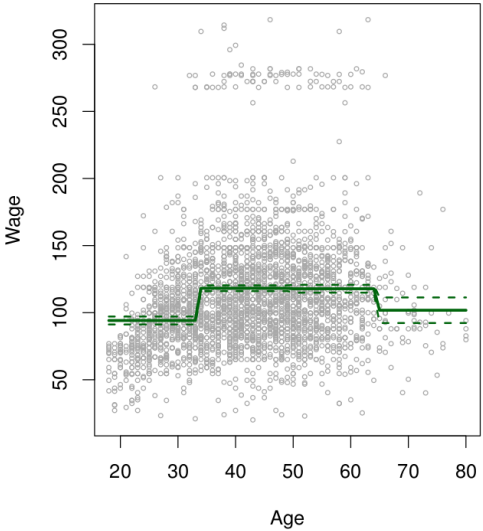
Step function: Learned model

$$y_i = \beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \cdots + \beta_K C_K(x_i) + \varepsilon_i$$

Coding bit

Back to the `wage` data set

Step function example



Next time

Lec #	Date			Reading	HW
21	Mon	10/28	Polynomial & Step Functions	7.1,7.2	
22	Wed	10/30	Step Functions; Basis functions; Start Splines	7.2 - 7.4	
23	Fri	11/1	Regression Splines	7.4	HW #7 Due Sun 11/3
24	Mon	11/4	Decision Trees	8.1	
25	Wed	11/6	Class Cancelled (Dr Munch out of town)		
26	Fri	11/8	Random Forests	8.2.1, 8.2.2	HW #8 Due Sun 11/10
27	Mon	11/11	Maximal Margin Classifier	9.1	
28	Wed	11/13	SVC	9.2	
29	Fri	11/15	SVM	9.3, 9.4	HW #9 Due Sun 11/17
30	Mon	11/18	Single layer NN	10.1	
31	Wed	11/20	Multi Layer NN	10.2	
32	Fri	11/22	CNN	10.3	HW #10 Due Sun 11/24
33	Mon	11/25	TBD: Unsupervised learning/clustering	12.1, 12.4?	
	Wed	11/27	Virtual: Project office hours		
	Fri	11/29	No class - Thanksgiving		
	Mon	12/2	Review		
	Wed	12/4	**Midterm #3**		