Ch 7.2-7.3: Step Functions and Basis Functions

Lecture 21 - CMSE 381

Prof. Elizabeth Munch

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

Mon, Nov 7, 2022

Announcements

Last time:

- 7.1 Polynomial regression
- 7.2 Step functions

This lecture:

- 7.2 Step functions
- 7.3 Basis functions

Announcements:

- Vote tomorrow if you're elligible!

Section 1

Last time

Polynomial regression

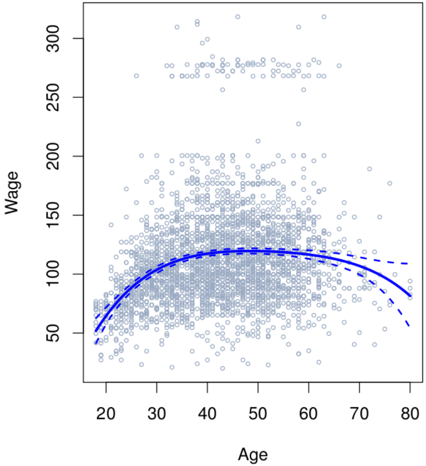
Replace linear model

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

with

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

Example with wage data



$$-184.1542 + 21.24552 * age + -0.56386 * age^2 + 0.00681 * age^3 + -3e - 05 * age^4$$

Section 2

Step Functions

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

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

Step function: Learned model

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

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				

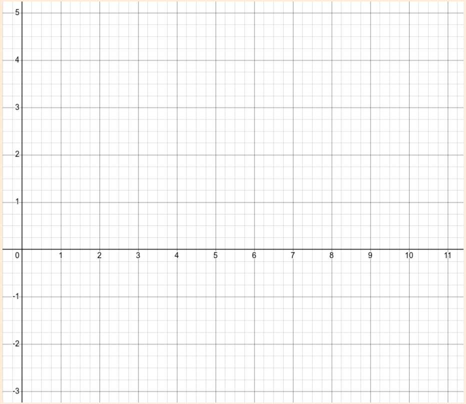
My code doing regression on the step function input returned the function.

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

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	



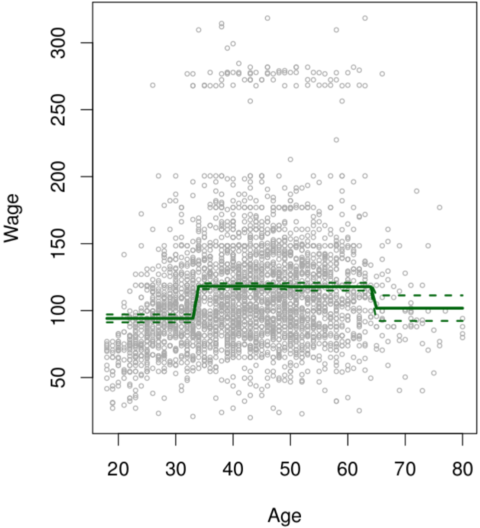
Step function: Learned model

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

Draw the function

Back to the `wage` data set

Step function example

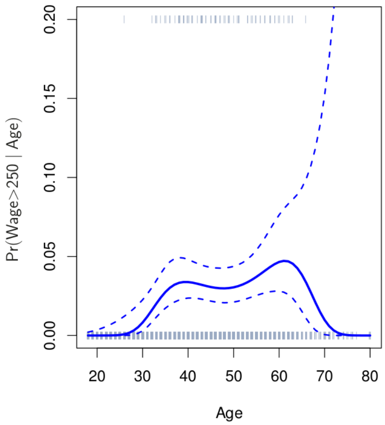


Section 3

Classification versions

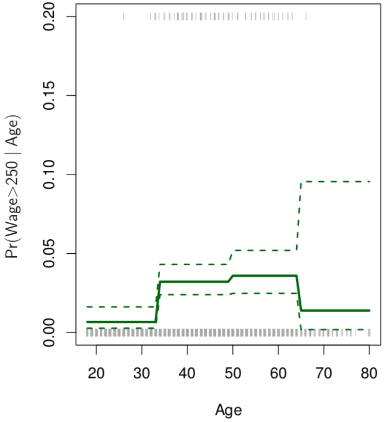
Classification version: Polynomial regression

$$\Pr(y_i > 250 \mid x_i) = \frac{\exp(\beta_0 + \beta_1 x_i + \cdots + \beta_d x_i^d)}{1 + \exp(\beta_0 + \beta_1 x_i + \cdots + \beta_d x_i^d)}$$

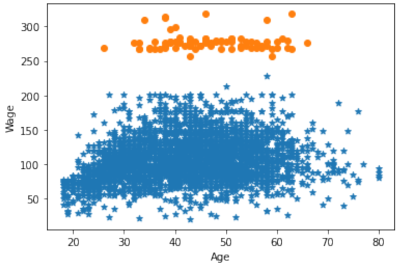


Classification version: Step functions

$$\Pr(y_i > 250 \mid x_i) = \frac{\exp(\beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \cdots + \beta_K C_K(x_i))}{1 + \exp(\beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \cdots + \beta_K C_K(x_i))}$$



Computation in the jupyter notebook



In [58]:
```python
from sklearn.linear_model import LogisticRegression
```

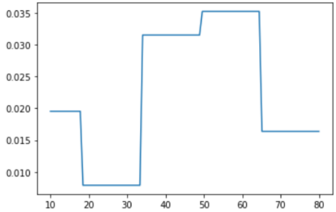
In [59]:
```python
y = np.array(df.wage>250) #<--- this makes sure I
                          #     just have true/false input
clf = LogisticRegression(random_state=48824)
clf.fit(df_steps_dummies,y)
```

Out[59]: `LogisticRegression(random_state=48824)`

In [56]:
```python
f = clf.predict_proba(t_dummies)
plt.plot(t,f[:,1])
```

(100, 2)

Out[56]: `[<matplotlib.lines.Line2D at 0x7fc2b938bb50>]`



A few more comments on step functions

Basis Functions Setup

Polynomial and piecewise-constant regression models are special cases of a *basis function* approach.

$$y_i = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \cdots + \beta_K b_K(x_i) + \varepsilon_i$$

Next time

20	F	Nov 4	Polynomial & Step Functions.	7.1,7.2	
21	M	Nov 7	Step Functions	7.2	
22	W	Nov 9	Basis functions, Regression Splines	7.3,7.4	
23	F	Nov 11	Decision Trees	8.1	HW #7 Due
24	M	Nov 14	Ensemble methods	8.2	
25	W	Nov 16	Maximal Margin Classifier	9.1	
26	F	Nov 18	SVC	9.2	HW #8 Due
27	M	Nov 21	SVM	9.3, 9.4, 9.5	
28	W	Nov 23	Single layer NN	10.1	
	F	Nov 25	No class - Thanksgiving		
29	M	Nov 28	Multi Layer NN	10.2	HW #9 Due
30	W	Nov 30	CNN	10.3	
31	F	Dec 2	Unsupervised Learning & Clustering	12.1, 12.4	
32	M	Dec 5	More Clustering	12.4	HW #10 Due
	W	Dec 7	Review		
	F	Dec 9	Midterm #3	Bring your cheat sheet and a non-internet-connected calculator	