Ch 7.1-7.2: Polynomial regression and Step Functions

Lecture 20 - CMSE 381

Prof. Elizabeth Munch

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

Fri, Nov 4, 2022

Announcements

Last time:

- Exam

This lecture:

- 7.1 Polynomial regression
- 7.2 Step functions

Announcements:

-

Section 1

Last time

Low-Dimensions

$n \gg p$

- Low here means p is low, or at least small relative to n
- Can do all the stuff we've talked about so far

High-Dimensions

$n \ll p$

- Issues show up even if p is close to or slightly smaller than n
- Classical approaches not appropriate since lots of overfitting

What to do about it?

Be less flexible....

Key points

- regularization or shrinkage plays a key role in high-dimensional problems,
- appropriate tuning parameter selection is crucial for good predictive performance, and
- the test error tends to increase as the dimensionality of the problem increases, unless the additional features are truly associated with the response.

- Curse of dimensionality
- Report results on an independent test set, or cross-validation errors.

Section 2

Polynomial Regression

Polynomial regression

Replace linear model

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

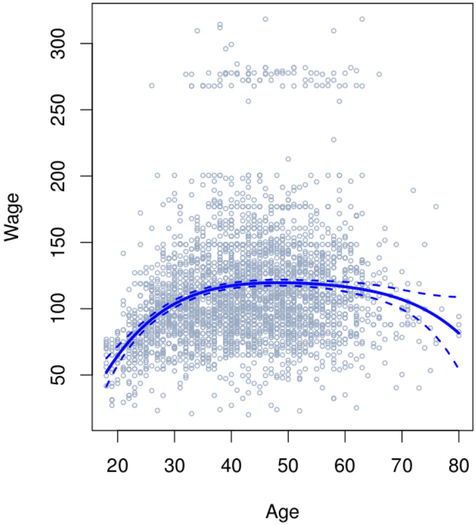
with

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

Coding bit

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

Example with wage data



Section 3

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

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

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				
7				
3				
5				
4				
9				

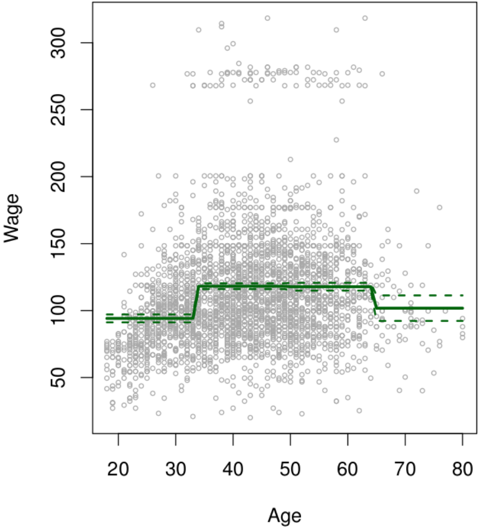
Step function: Learned model

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

Coding bit

Back to the `wage` data set

Step function example

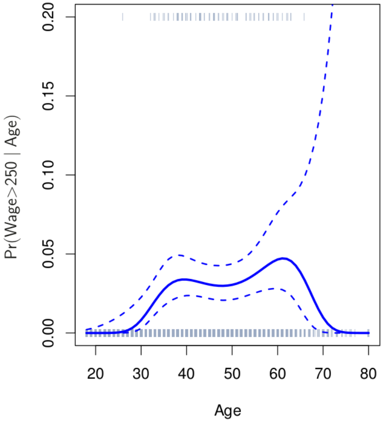


Section 4

Classification versions

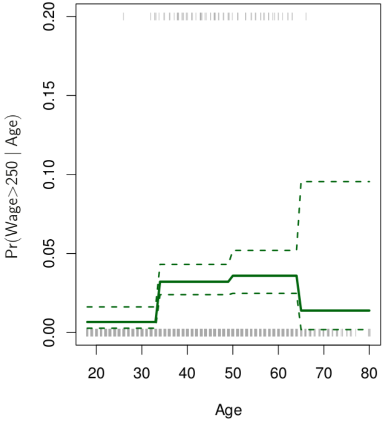
Classification version

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



Step function classification example

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



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	Basis functions, Regression Splines	7.3,7.4	
22	W	Nov 9	Smoothing Splines; Local regression; GAMs	7.5-7.7	
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	