Ch 7.2-7.4: Step Functions, Basis Functions, Start Splines

Lecture 22 - CMSE 381

Prof. Mengsen Zhang

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

Mon, Oct 27, 2025

Announcements

Last time:

- 7.1 Polynomial regression
- 7.2 Step functions

This lecture:

- 7.2 Step functions
- 7.3 Basis functions
- 7.4 Regression Splines (Finish next lecture)

Announcements:

- HW #6 Due Sunday
- Exam 2 graded. What worked? PollEv

Section 1

Last time

Polynomial regression

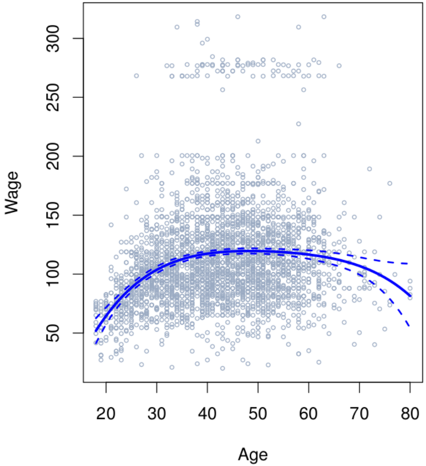
Replace linear model

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

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



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

Step functions

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

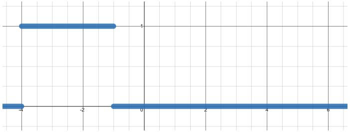
$$
\begin{aligned}
C_0(X) &= I(X < c_1), \\
C_1(X) &= I(c_1 \leq X < c_2), \\
C_2(X) &= I(c_2 \leq X < c_3), \\
&\vdots \\
C_{K-1}(X) &= I(c_{K-1} \leq X < c_K), \\
C_K(X) &= I(c_K \leq X),
\end{aligned}
$$

Learned model:

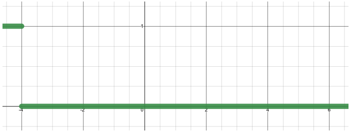
$$y_i = \beta_0 + \beta_1 C_1(x_i) + \beta_2 C_2(x_i) + \cdots + \beta_K C_K(x_i) + \varepsilon_i$$

Example: Cut points at -4, -1, 3, 6

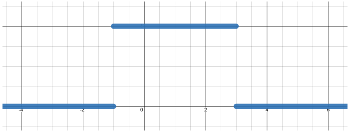
$C_1(X)$



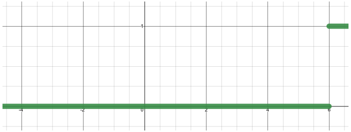
$C_0(X)$



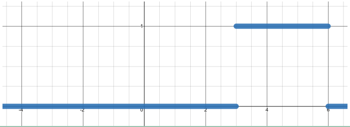
$C_2(X)$



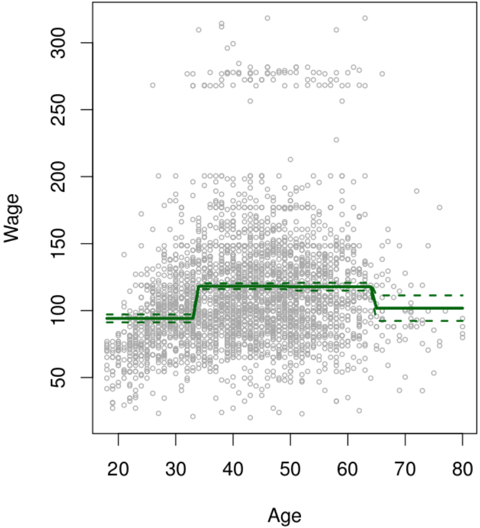
$C_4(X)$



$C_3(X)$



Step function example



What will you learn today?

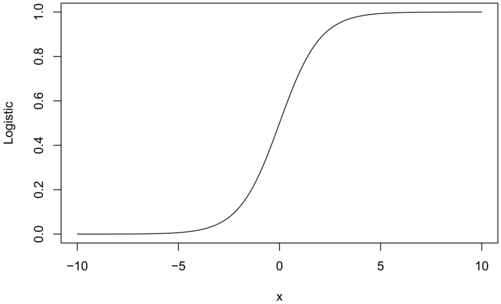
- How to fit step function models for classification problems?
  - You should also be able to implement them in `Python`.
- What are basis functions?
  - You should be able to articulate what the basis functions b_i are for different models that we have covered.
  - Examples include polynomial, step functions, and cubic splines.
- What is the purpose of using basis functions?
  - How do they influence model flexibility?
  - How does this purpose manifest in different models, such as regression splines?
- How to define cubic splines mathematically?
  - You should be able to write down the equations for the model between knots.
  - And the equations for constraints at the knots.
- How to calculate the cubic spline coefficients by hand by applying the above mathematical definitions?

Section 2

Classification versions

Remember logisitic regression?

$$y = \frac{e^x}{1 + e^x}$$



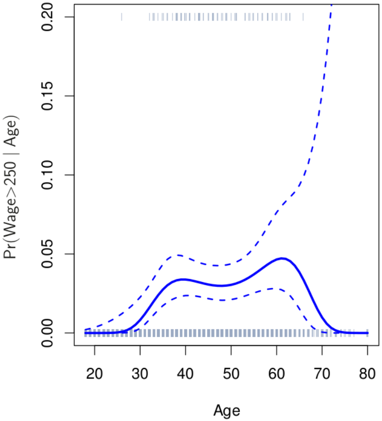
$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

Multiple features:

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}}$$

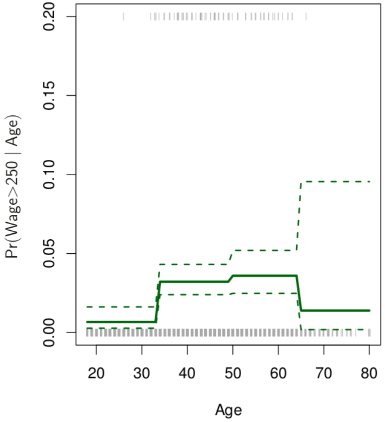
Classification version: Polynomial regression

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

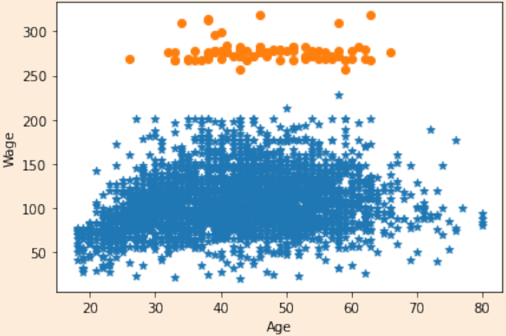


Classification version: Step functions

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



Coding bit: classification version



A few more comments on step functions

- Gives the chance to break up the domain, avoid forcing global structure
- Need to make decisions about the c_i. A bit arbitrary unless your data has natural breakpoints.
- Popular in biostats and epidemiology

Section 3

Basis functions

Basis Functions Setup

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

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

Section 4

Regression Splines

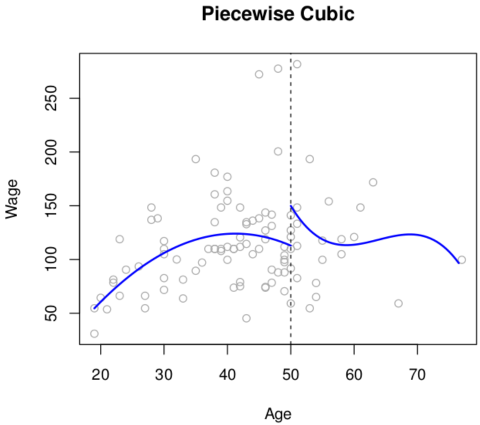
Piecewise polynomials

- Fit a polynomial regression

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

- Let the β_i's be different at different locations of the range.

Example of piecewise polynomial

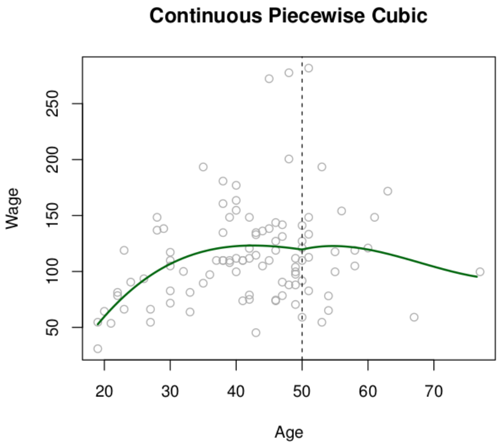


Example:

$$y_i = \begin{cases} \beta_{01} + \beta_{11}x_i + \beta_{21}x_i^2 + \beta_{31}x_i^3 + \epsilon_i & \text{if } x_i < c \\ \beta_{02} + \beta_{12}x_i + \beta_{22}x_i^2 + \beta_{32}x_i^3 + \epsilon_i & \text{if } x_i \geq c. \end{cases}$$

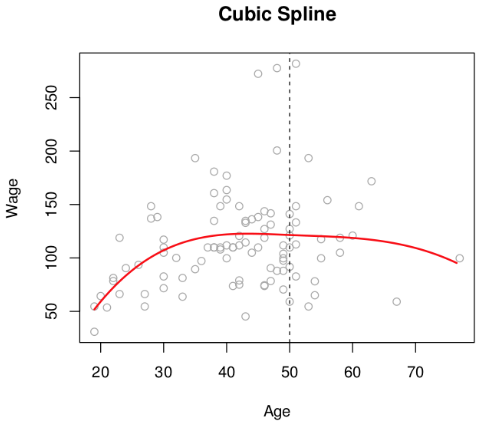
The fix

- Fit piecewise polynomial
- Require continuity at knots



The better fix: Cubic splines

- Fit piecewise polynomial
- Require continuity at knots
- Require the first and second derivatives to be continuous at knots



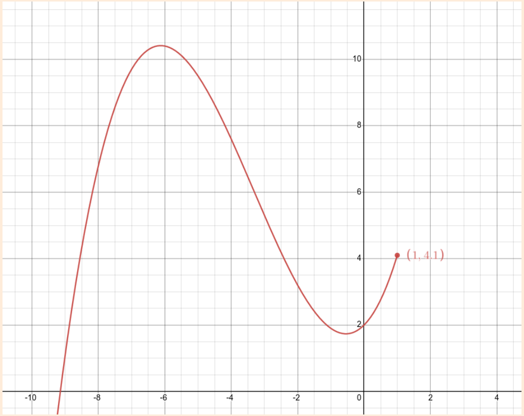
Test your understanding: PollEv

Example

We have the following piecewise cubic polynomial:

$$f(x) = \begin{cases} 2 + x + x^2 + 0.1x^3 & x \leq 1 \\ b_0 + b_1 x + b_2 x^2 - x^3 & x > 1 \end{cases}$$

What are b_1, b_1, and b_2 to make this a cubic spline?



Check your answers: www.desmos.com/calculator/kbm0zivqco

More space for work

$$f(x) = \begin{cases} 2 + x + x^2 + 0.1x^3 & x \leq 1 \\ b_0 + b_1 x + b_2 x^2 - x^3 & x > 1 \end{cases}$$

Next time

21	F	10/24	Polynomial & Step Functions	7.1-7.2	HW #5 Due Sun 10/26
22	M	10/27	Step Functions; Basis functions; Start Splines	7.2-7.4	
23	W	10/29	Regression Splines	7.4	
24	F	10/31	Decision Trees	8.1	HW #6 Due Sun 11/2
25	M	11/3	Random Forests	8.2.1, 8.2.2	
26	W	11/5	Maximal Margin Classifier	9.1	
27	F	11/7	SVC	9.2	HW #7 Due Sun 11/9
28	M	11/10	SVM	9.3, 9.4	
29	W	11/12	Single Layer NN	10.1	
30	F	11/13	Multi Layer NN	10.2	HW #8 Due Sun 11/16
31	M	11/17	CNN	10.3	
32	W	11/19	Unsupervised learning / clustering	12.1, 12.4	
33	F	11/21	***Review***		HW #9 Due Sun 11/23
	M	11/24	**Midterm #3**		
	W	11/26	Virtual: Project Office Hours		
	F	11/28	Thanksgiving		
	M	12/1	Virtual: Project Office Hours		
	W	12/3	Virtual: Project Office Hours		
	F	12/5			**Project Due**
	M	12/8			
	W	12/10			
	F	12/12	**No final exam**		**Honors Project Due**