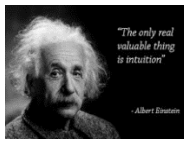
Piecewise Polynomial Models



Piecewise Polynomial: Intuition

- As we discussed earlier, **high-degree polynomials** cause o**ver-fitting**, are difficult to **interpret** and impose a particular **mathematical structure** on the fitted model.
- Piecewise linear models help in this regard, but fitting **polynomial rather than linear** models may provide a better fit
- A **piecewise polynomial** is a hybrid between piecewise linear and polynomial, and it **fits** separate **low-degree polynomials** over **different regions** of a predictor.
- Polynomials can be of various degrees but **quadratic** and **cubic** polynomials are the most popular because the whole objective is to use **low-degree** polynomials
- For example, a **cubic** piecewise polynomial with **3 sections** divided at $x = c_1$ and $x = c_2$ would look like this:

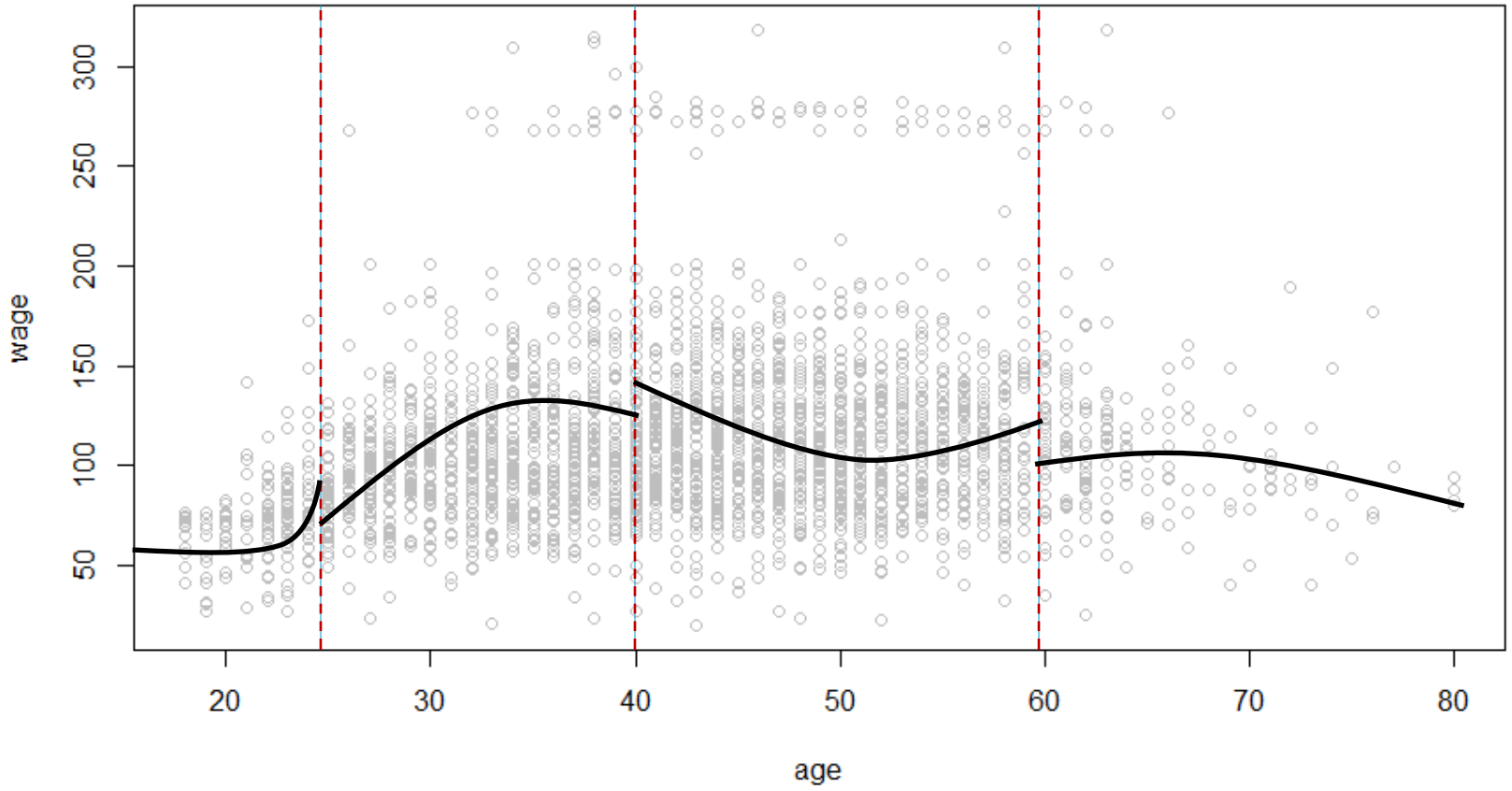
$$y = \beta_{01} + \beta_{11}x + \beta_{21}x^2 + \beta_{31}x^3 + \varepsilon \text{ if } x < c_1$$

$$y = \beta_{02} + \beta_{12}x + \beta_{22}x^2 + \beta_{32}x^3 + \varepsilon \text{ if } c_1 <= x < c_2$$

$$y = \beta_{03} + \beta_{13}x + \beta_{23}x^2 + \beta_{33}x^3 + \varepsilon \text{ if } x >= c_2$$

Piecewise Polynomial Illustration

- A **cubic piecewise polynomial** predicting **wage** as a function of **age**, fitting 4 different cubic polynomials around 3 knots
- Problem: there are 2 predicted points at each knot because the curves are **not continuous a the knots**



R Tips

```
fit.piecewise.cube=
  lm(y~
  x+I(x^2)+I(x^3)+
  I((x-25)*(x>25))+I((x-25)^2*(x>25))+I((x-25)^3*(x>25))+
  I((x-40)*(x>40))+I((x-40)^2*(x>40))+I((x-40)^3*(x>40))+
  I((x-60)*(x>60))+I((x-60)^2*(x>60))+I((x-60)^3*(x>60)),
  data=dataName)
```

→ Fits a stepwise cubic polynomial model with 4 segments separated by 3 knots set arbitrarily at `x=25, 40` and `60` respectively

Note: I do not recommend modeling a piecewise polynomial this way, but of course, it can be done. This model is not very interpretable. While it can improve predictive accuracy, it is better to fit Spline models (next section). Spline models are also difficult to interpret, but they are as accurate or more than piecewise polynomials, and much easier to model.

KOGOD SCHOOL
of
BUSINESS