

# Regression Splines

Following the concept of [Basis Functions](#), Regression Splines is a non-parametric method which an extension to the linear models falls under the non-linear regression functions similar to [Polynomial Regression](#)

But unlike [Polynomial Regression](#) that transform the design matrix  $X$  globally, **Splines** fit a [Basis Functions](#) **locally** within each region and fitted using [Ordinary Least Squares](#) :

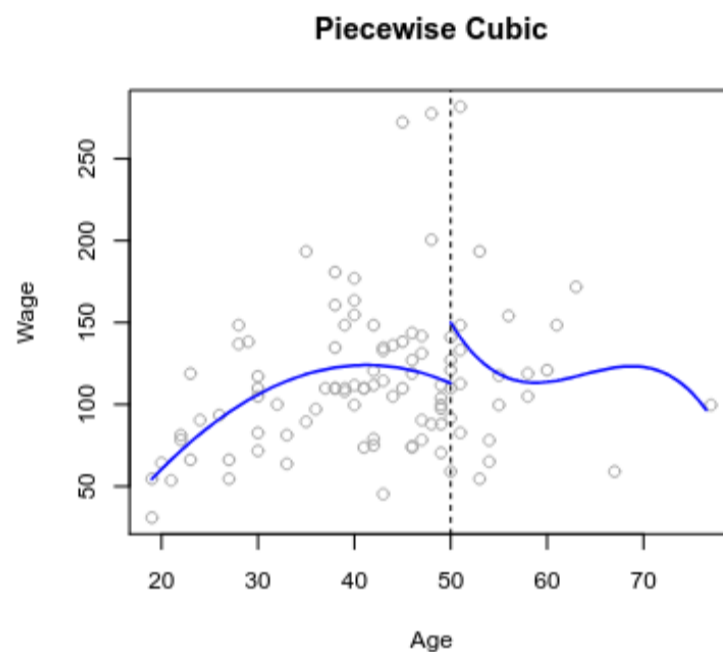
$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \epsilon_i$$

- This is a **Cubic Polynomial Regression** applied globally

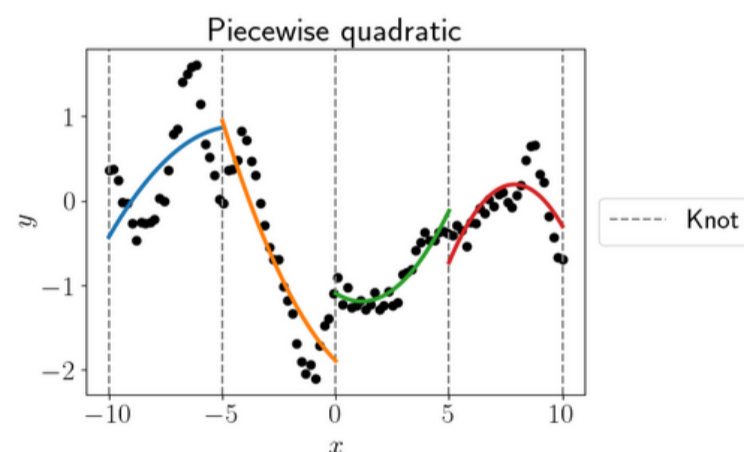
## Piecewise Polynomials

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

- Here we fitting the **Cubic Polynomial Regression** two times depending on the region  $\kappa$  which called **Knot**

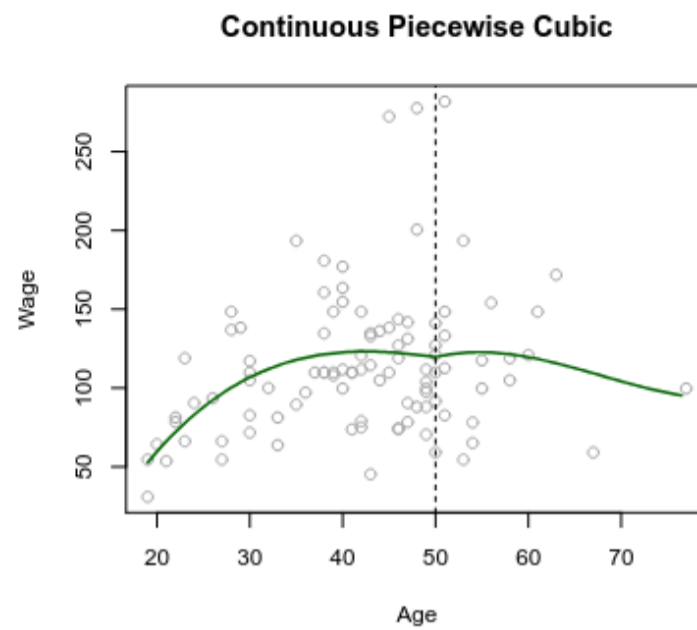


Using more knots results in more flexible piecewise polynomial, generally when fitting **piecewise polynomials splines** we will end up fitting  $K + 1$



The thing to notice is that the fit is discontinuous at the **knots** to fix that we add a **constraint** to reinforce **continuity**

Let's take for example this **piecewise poly spline** and  $\tau_K$  as the **knots** location :



**Truncated Power Basis Function :**

$$y_i = \beta_0 + \beta_1 h_1(x_i) + \cdots + \beta_{K+d} h_{K+d}(x_i) + \epsilon_i$$

- $h_K(x_i) = x^{K-1}$  represent the **polynomial basis function**
- $h_{K+d}(x_i)$  is a **basis function** to adjust the *intercept* to fix the **discontinuity**

With :

$$h_{K+d}(x) = (x - \tau_k)_+^{d-1} = \begin{cases} (x - \tau)^{d-1} & \text{if } x > \tau \\ 0 & \text{otherwise} \end{cases}$$

for **Cubic Polynomial** :

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 (x - \tau_1)^3 + \epsilon_i$$

- One **knot**  $\tau$
- The polynomial of an order  $d = 4$

if  $x < \tau_1$  : **Region 1**

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$$

if  $x \geq \tau_1$  : **Region 2**

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 (x - \tau_1)^3$$

By doing some algebra we get :

$$y_i = (\beta_0 + \beta_4 \tau_1^3) + x_i(\beta_1 + 3\beta_4 \tau_1^2) + x_i^2(\beta_2 - 3\beta_4 \tau_1) + x_i^3(\beta_3 + \beta_4) + \epsilon_i$$

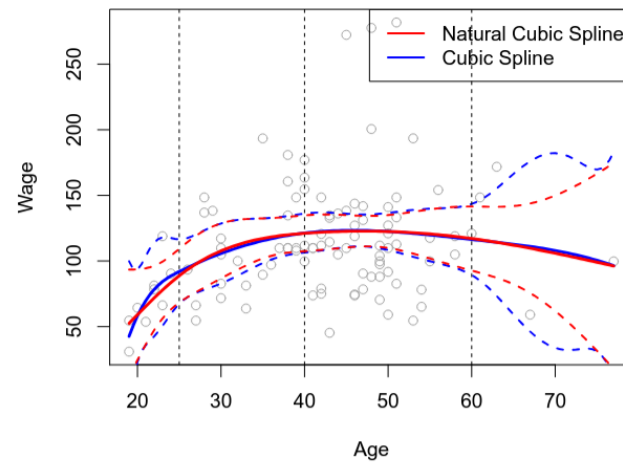
- The coefficient  $\beta_4$  from the **truncated basis function** act as an **adjustment factor** across all local coefficients  $\beta_0, \beta_1, \beta_2, \beta_3$

At  $x = \tau_1$  both **Region 1** and **Region 2** share the same value which ensures **continuity**

**Note** : This mean we will always fit  $K + d$  coefficients , with  $K$  being the number of **knots**

## Limits of Piecewise Polynomials

They can have very high variance at the outer range of the predictors , where  $X$  takes very large/small values



B-spline basis are piece-wise polynomial functions of order  $k$ , they overlap each other on knots which results in support and **continuity** , like the adjacent ones overlaps

$$B_{i,1}(x) = \begin{cases} 1 & \text{if } \tau_i \leq x < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

for  $i = 1, \dots, K + 2M - 1$ . (By convention,  $B_{i,1} = 0$  if  $\tau_i = \tau_{i+1}$ ).

$$B_{i,m} = \frac{x - \tau_i}{\tau_{i+m-1} - \tau_i} B_{i,m-1}(x) + \frac{\tau_{i+m} - x}{\tau_{i+m} - \tau_{i+1}} B_{i+1,m-1}(x)$$

every B-spline is built recursively(cox-de Boor) following the formula above it either take the left or the right support

Changing a coefficient  $\beta_j$  effects only the spline on the small region

$$y_i = \sum_{j=1}^M \beta_j B_j(x)$$

- local control
- low variance

