21. Piecewise Polynomials

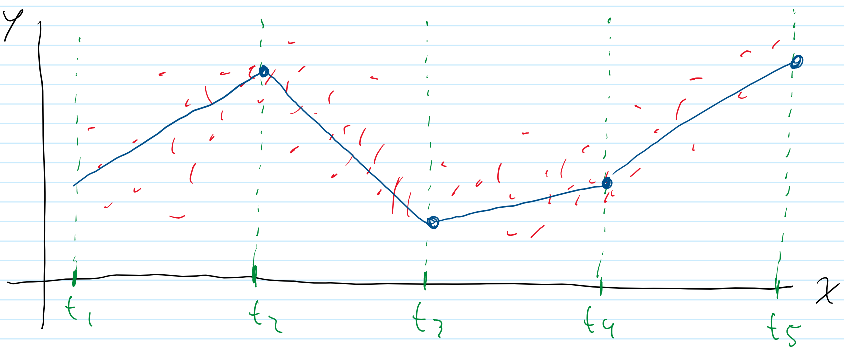
October 27, 2025 8:40 AM

Last time: Polynomial Regression

- instead of just using x as an input we can also use $x^2, x^3, \ldots, x^p$
- More power to model the data
- High order polynomial are troublesome
  - → over fitting → problems extrapolating

One solution: Not use 1 High order polynomial but use many low order polynomials.

Example: piecewise linear model



Break domain into sub intervals

Do a regression on each piece

piecewise linear but forcing the lines to connect.

* One of the most popular is piecewise cubic

→ smooth curve with 2 derivatives!

General Formula: $t_1 < t_2 < \ldots < t_{K+1}$ (K pieces)

$$y = \beta_{0,1} + \sum_{j=2}^{K} \beta_{0,j} \mathbb{1}[x \geq t_j] + \sum_{i=1}^{p} \sum_{j=1}^{K} \beta_{i,j} (x - t_j)_+^i + \varepsilon$$

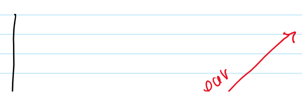
power poly   subinterval

where $(x-t_j)_+^i = (x-t_j)^i \mathbb{1}[x > t_j]$

Once we hit the point t_j, activate this polynomial.

Piece-wise linear + Connected (above drawing)

$$y = \beta_{0,1} + \sum_{j=1}^{K} \beta_{1,j} (x - t_j)_+ + \varepsilon$$



$$y = \beta_{0,1} + \sum_{j=1}^{K} \beta_{1,j}(x - t_j)_+ + \varepsilon$$

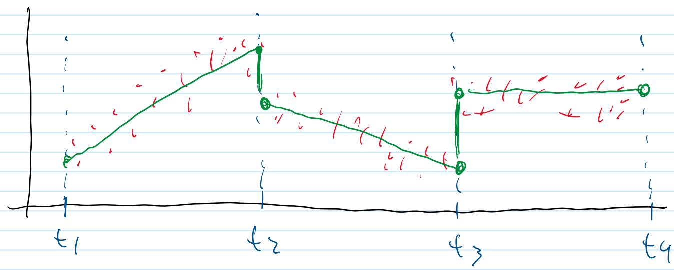
one intercept — K linear pieces

Piece-wise linear + Not connected

$$y = \beta_{0,1} + \underbrace{\sum_{j=2}^{K} \beta_{0,j}\,\mathbb{1}[x > t_j]}_{} + \sum_{j=1}^{K} \beta_{1,j}(x - t_j)_+ + \varepsilon$$

a vertical shift / new intercept For each piece / sub interval

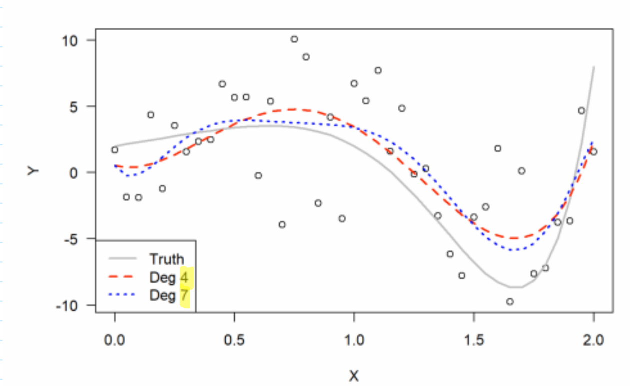
→ Fit a separate linear regression to each sub interval

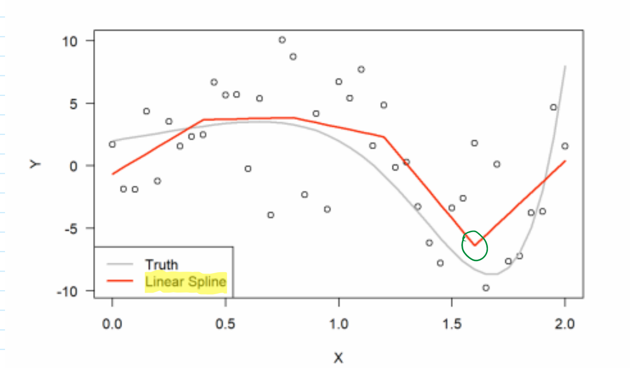


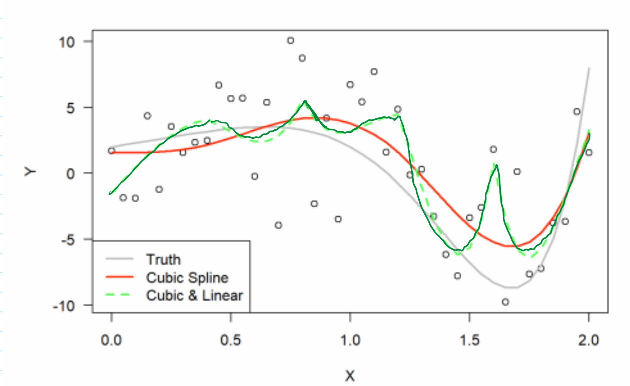
Piecewise Constant

$$y = \beta_{0,1} + \sum_{j=2}^{K} \beta_{0,j}\,\mathbb{1}[x > t_j] + \varepsilon$$

Example with Data from a degree 7 polynomial







← just cubic terms :)

← cubic + linear terms :(

Piecewise linear : Continuous!

Piecewise Quadratic : 1 Derivative!

Piecewise Cubic : 2 Derivatives!

Hypothesis Testing :

Hypothesis Testing:

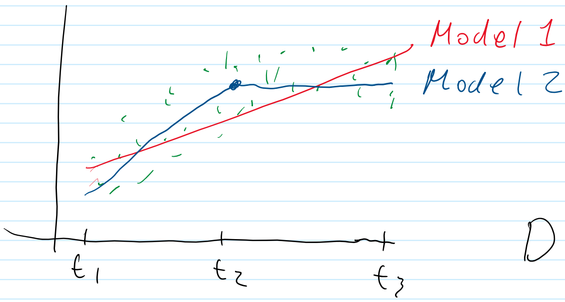
- If I have two Models, one nested in the other, then I can do a partial F-test.

Example: One line or Two lines?

Is there a change at t_2?

Model 1: $y = \beta_{0,1} + \beta_{1,1} x + \varepsilon$

Model 2: $y = \beta_{0,1} + \beta_{1,1} x + \beta_{1,2}(x - t_2)_+ + \varepsilon$



$H_0: \beta_{1,2} = 0$

$H_1: \beta_{1,2} \neq 0$

$DoF = (1, n-3)$

- Here, we have to pick a partition $t_1 < \ldots < t_{k+1}$ for our data.

Finding a "change point" in a statistical Model from the data can be a hard research problem.