Smoothing Splines

- Unlike regression splines and natural splines, there are no knots!
- The discrete problem of selecting a number of knots into a continuous penalization problem
- We seek a function g among all possible functions (linear + non-linear) which minimizes

$$\underbrace{\text{model fit}}_{\text{not the usual RSS}} + \text{penalty term} \;=\; \sum_{i=1}^{n} (y_i - g(x_i))^2 + \lambda \int \underbrace{(g''(t))^2}_{\text{catches wiggles or non-linearities}} dt$$

- The function g that minimizes the above quantity is called a **smoothing spline**
- $\lambda \geq 0$ is the tuning penalty parameter, also called **roughness penalty**
  - when $\lambda = 0$ we get an extremely wiggly non-linear function g (completely useless)
  - as λ increases, the function becomes smoother
  - theoretically: when $\lambda \to \infty$, g'' is zero everywhere $\Longrightarrow g(X) = \beta_0 + \beta_2 X$ i.e. linear model
- the solution for any finite and non-zero λ is that the function g is a natural cubic spline but with knots placed on each individual sample point $x_1, x_2, x_3, \ldots, x_n$

Cubic vs. Natural vs. Smoothing Splines

Example: Wage (ISLR2)

Training data = 50