Johns Hopkins Engineering

625.464 Computational Statistics

Kernel Smoothers and Spline Smoothers

Module 12 Lecture 12C

Kernel Smoothers

Let K be a symmetric kernel centered at 0.

Ex/ $N(0,1)$ $\quad K(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

Let h be our smoothing parameter and bandwidth of the kernel.

$$\hat{S}_h(x) = \sum_{i=1}^{n} Y_i \frac{K\left(\frac{x - x_i}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{x - x_j}{h}\right)}$$

Comments On Kernel Smoothers

$$\hat{S}_h(x) = \sum_{i=1}^{n} Y_i \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{x-x_j}{h}\right)}$$

① K can be viewed as a weighting function

② K can be chosen so that only some of the observed data points are in a neighborhood of x or all data points can be used.

③ Retains concept of local averaging since proximity determines weight.

④ Large h → smoother

Small h → wiggly

Comments On Kernel Smoothers

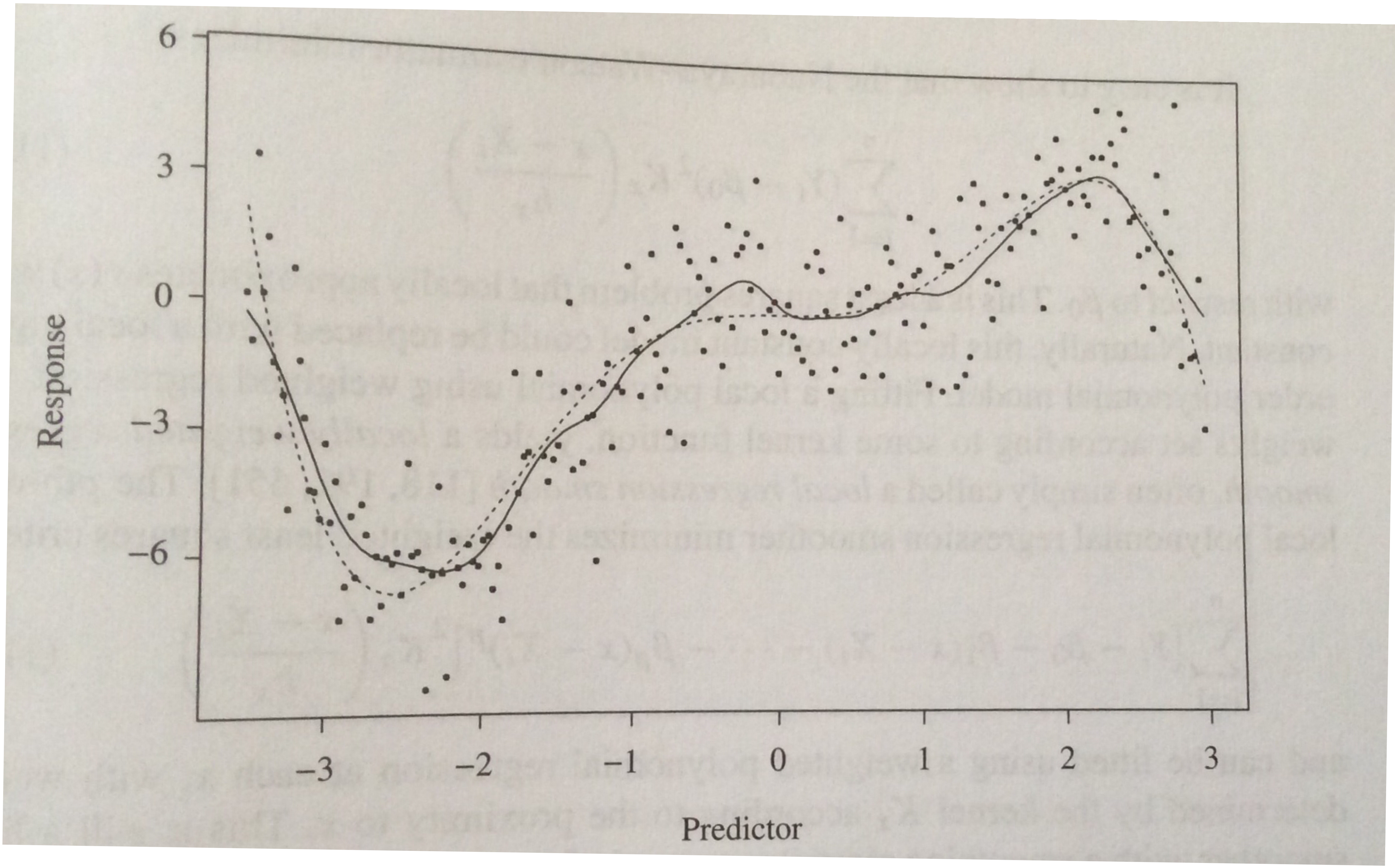
$$\hat{S}_h(x) = \sum_{i=1}^{n} Y_i \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{x-x_j}{h}\right)}$$

5. Choice of K is not as important as choice of h $\rightarrow$ no real reason not to use $N(0,1)$.

6. These are linear Smoothers.

7. Can use CV to optimize bandwidth.

Example of a Kernel Smoother



$K = N(0,1)$ $\qquad$ $h = .16$

Spline Smoothing

Assume the obs are sorted $x_1 < x_2 < \cdots < x_n$

define the metric

$$Q_\lambda(\hat{s}) = \underbrace{\sum_{i=1}^{n} \left(y_i - \hat{s}(x_i)\right)^2}_{①} + \underbrace{\lambda \int_{x_1}^{x_n} \hat{s}''(x)^2 \, dx}_{②}$$

① penalty for misfitting the data

② penalty for wiggliness

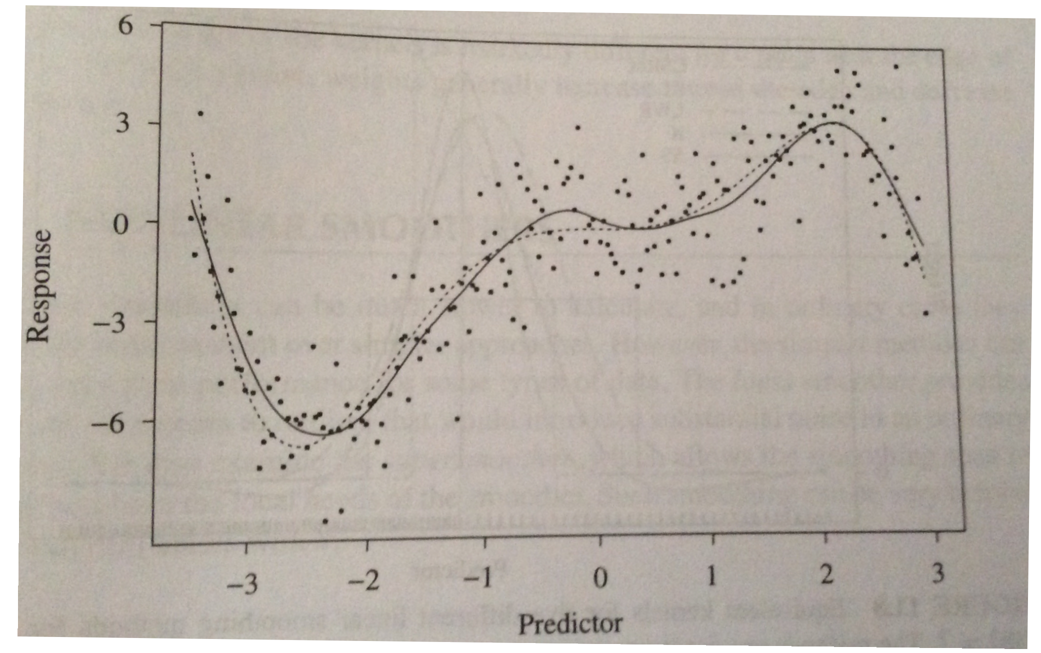
③ λ controls the weighting of the penalties

We minimize Q_λ by using a cubic smoothing spline with knots at $x_1, \ldots, x_n$

Comments on Spline Smoothing

① They are linear smoothers. and can be computed efficiently.

② Example $\lambda = .066$



③ As $\lambda \to \infty$ $\hat{S}_\lambda$ approaches a least squares line.

If $\lambda = 0$, $\hat{S}_\lambda$ is an interpolating spline connecting $x_1, \ldots, x_n$

④ Choose λ using CVRSS