Johns Hopkins Engineering

625.464 Computational Statistics

Linear Smoothers

Module 12 Lecture 12B

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Linear Smoothers

Given p-R data (x_i, y_i) we want to find est.

$$\hat{S}_x(x) = \text{ave}\{y_i \mid x_i \in N(x)\}$$

For a linear smoother, the prediction at any point x, $\hat{S}_a(x)$, will be a linear comb. of the response values. We focus on est the smooth at obs. values x_i and then we will obtain the smooth at all x values by using interpolation.

Linear Smoothers

Given $x = (x_1, \ldots, x_n)^T$ and $y = (y_1, \ldots, y_n)^T$, then $\hat{s} = (\hat{s}(x_1), \ldots, \hat{s}(x_n))^T$ can be expressed as

$$\hat{s} = \mathcal{S} y$$

where $\mathcal{S}$ is an $n \times n$ smoothing matrix that does not depend on y.

Constant Span Running Mean

Basic Idea: Take the sample mean of of k nearby points.

$$\hat{S}_k(x_i) = \sum_{\{j : x_j \in N(x_i)\}} y_j / k$$

$$S^{\perp} = (0 \cdots 0 \; 1/k \cdots 1/k \; 0 \cdots\cdots 0)$$

where k is odd and $N(x_i)$ is x_i along with the $(k-1)/2$ values nearest above and below

if x_i are sorted: $\hat{S}_k(x_i) = \text{mean}\left[y_j \text{ for } \max\left(i - \frac{k-1}{2}, 1\right) \leq j \leq \min\left(i + \frac{k-1}{2}, n\right) \right]$

$$\hat{S}_k(x_{i+1}) = \hat{S}_k(x_i) - \frac{y_{i-(k-1)/2}}{k} + \frac{y_{i+(k+1)/2}}{k}$$

Smoothing Matrix and Example

How to deal w/ the edges. $K=5$

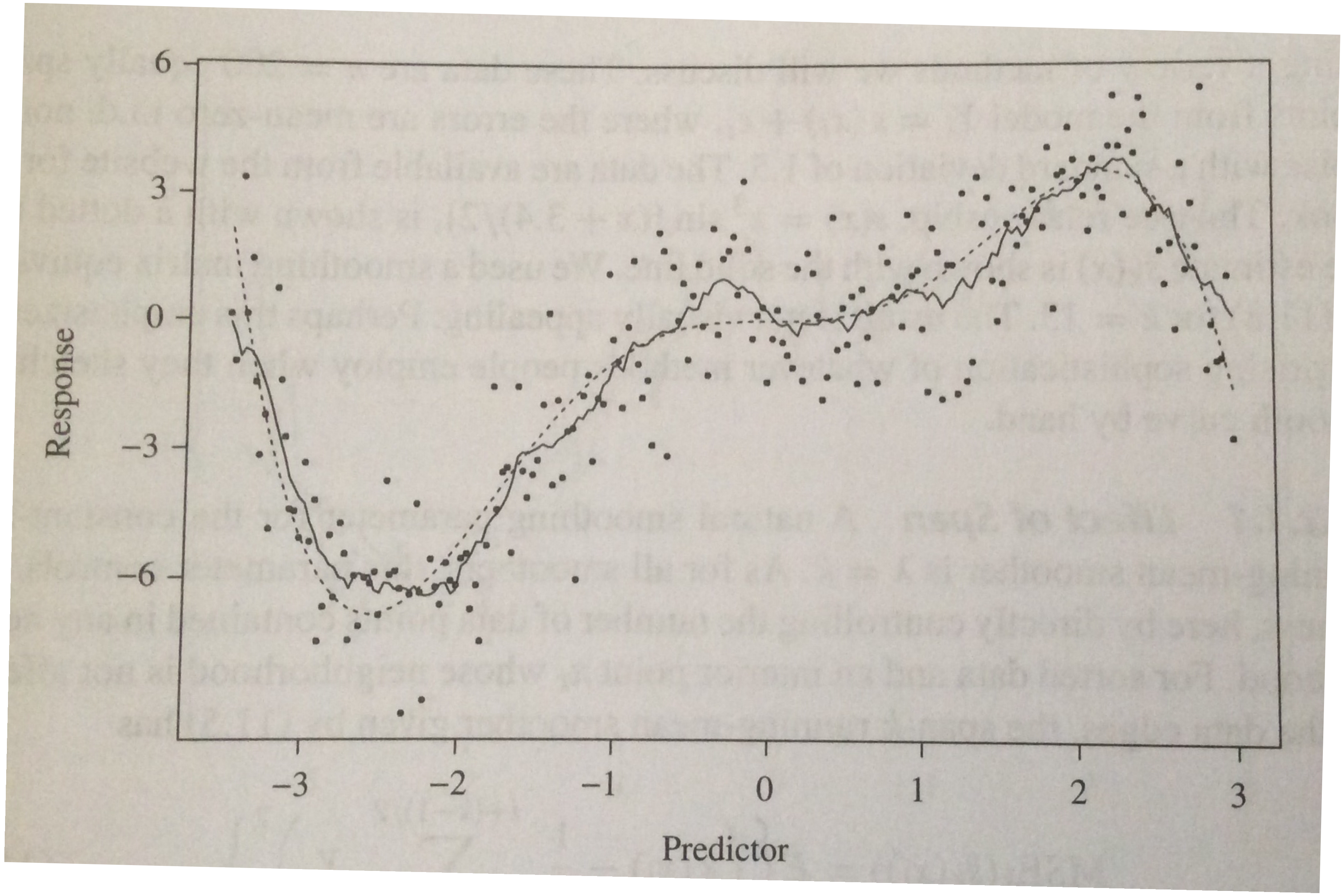
① Shrink the neighborhood

$$S = \begin{pmatrix} 1 & 0 & \cdots & & & & & \\ 1/3 & 1/3 & 1/3 & 0 & \cdots & & & \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \cdots & \\ 0 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \cdots \\ 0 & 0 & & & & & & \end{pmatrix}$$

② Truncate the neighborhood

$$S = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 & \cdots & & & \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & \cdots & & \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \cdots & \\ 0 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \cdots \\ 0 & 0 & & \cdot & \cdot & & & \end{pmatrix}$$

Constant Span Running Mean Example



$n=200 \quad Y_i = s(x_i)+\varepsilon_i \quad \varepsilon_i \sim N(0, 1.5^2)$

$s(x) = x^3 \sin\left(\frac{x+3.4}{2}\right) \quad k=13$

Effect of Span on the Smooth

$\lambda = k$

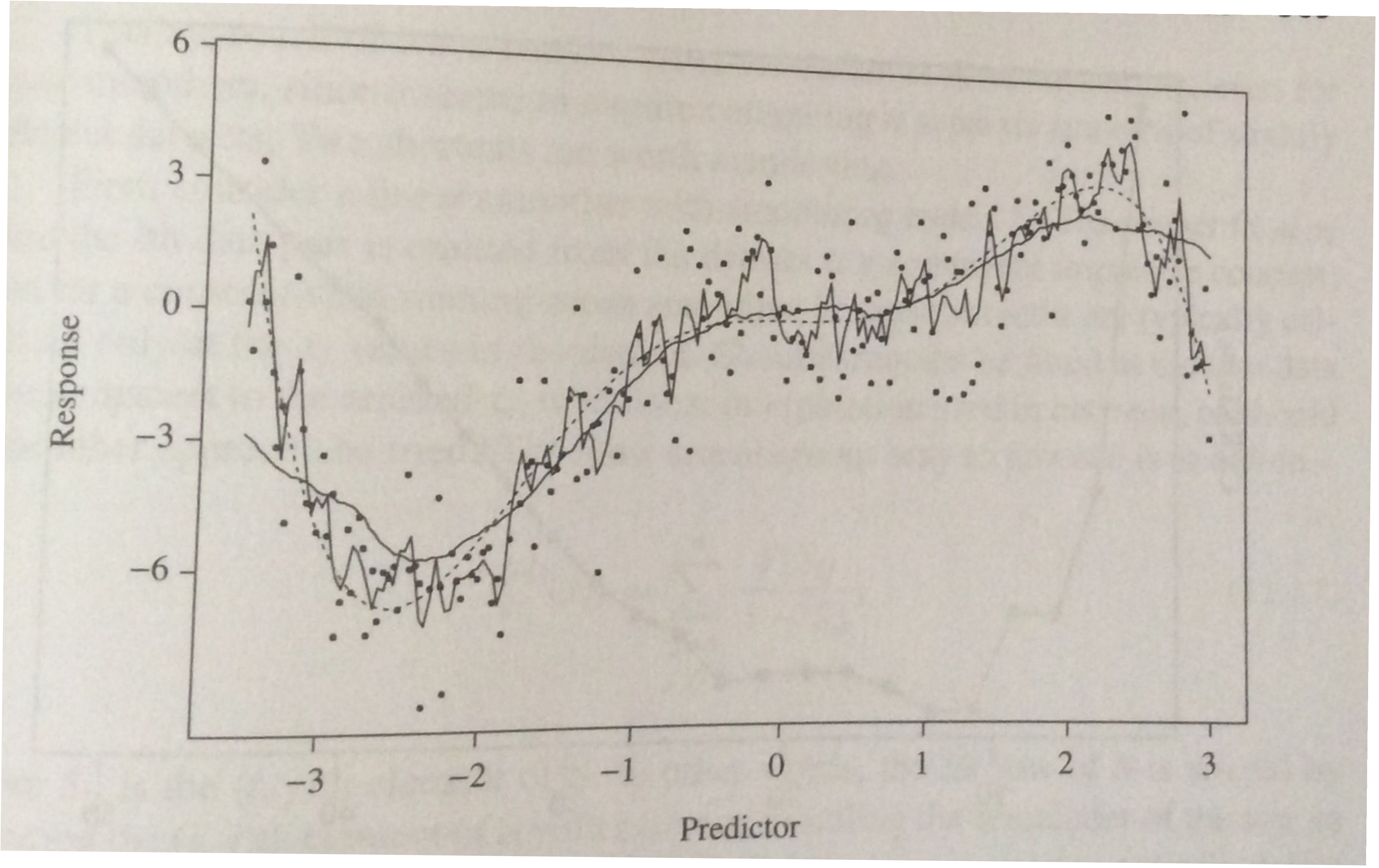
$$MSPE_k(\hat{S}_k(x_i)) = Var(Y \mid X = x_i) + MSE_k(\hat{S}_k(x_i))$$

$$= (1 + 1/k)\sigma^2 + (bias(\hat{S}_k(x_i)))^2$$

$\uparrow$ decreases $\quad\quad$ $\leftarrow$ increases

As span k goes up

Effect of Span



K=3

K=43

How to select the span for linear smoothers?

minimize w.r.t K the residual MSE

$$RSS_K(\hat{S}_K)/n = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{S}_K(x_i))^2$$

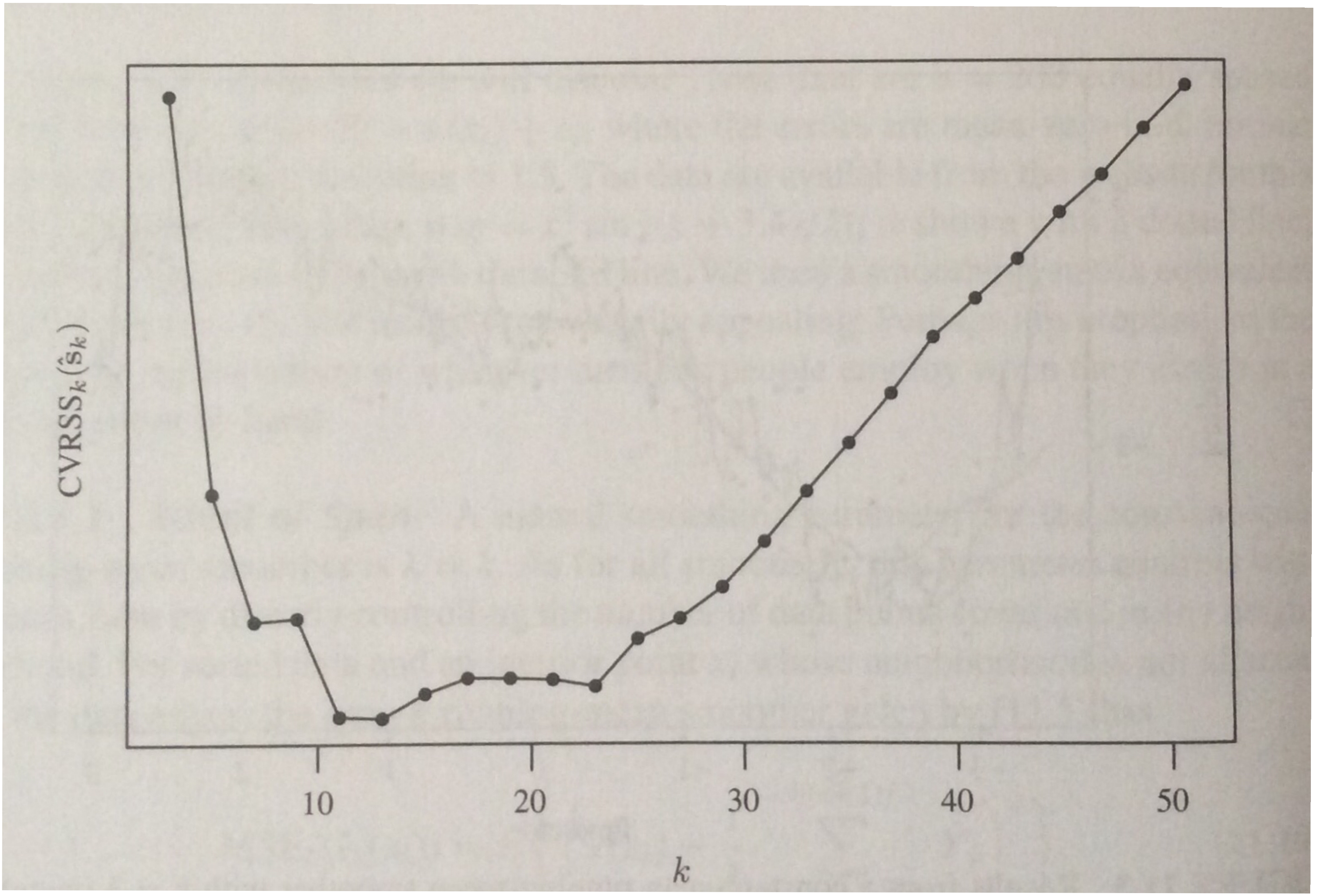
$$E[RSS_K(\hat{S}_K)/n] = \overline{MSPE_K}(\hat{S}_K) - \frac{1}{n}\sum_{i \neq j} cov(Y_i, \hat{f}(x_j))$$

To find K you can minimize

$$\frac{\underline{CVRSS_K(\hat{S}_K)}}{n} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{S}_K^{(-i)}(x_i))^2$$

where $\hat{S}_K^{(-i)}(x_i)$ is the value of smooth omitting (x_i, y_i).

CVRSS Example



Speeding Up Cross Validation

① Leave out groups of data not just 1 point.

② Recall that in the original smoother

$$\hat{S}_K(X_i) = \sum_{j=1}^{n} Y_j S_{ij} \quad \text{where } S_{ij} \nearrow 0, \searrow \frac{1}{K}$$

- Smooth with (X_i, y_i) eliminated

$$\hat{S}_K^{(-i)}(X_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{Y_j S_{ij}}{1 - S_{ii}}$$

In this case the CVRSS simplifies to

$$\frac{CVRSS_K(\hat{S}_K)}{n} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{Y_i - \hat{S}_K(X_i)}{1 - S_{ii}} \right)^2$$