1 Basic concept

1.1 Bias-variance decomposition

Define $e(x) = f(x) - \hat{f}(x)$ and $p(e)$ as its pdf.

$$\text{MSE} = E(e^2) = \int e^2 pde = (\int epde)^2 + (\int e^2 pde - (\int epde)^2) = E(e)^2 + Var(e) \tag{1}$$

1.2 A simple example

(A) Y follows binomal distribution $B(n, \pi_h)$. $E(Y) = n\pi_h$, $Var(Y) = n\pi_h(1 - \pi_h)$. One can estimate $f(0)$ by

$$f(0)_{\text{estimate}} = \frac{Y}{nh} \tag{2}$$

(B)

$$\begin{aligned} \pi_h &\approx hf(0) + \frac{f''(0)}{2}\int x^2 dx \\ &= hf(0) + \frac{f''(0)h^3}{24} \end{aligned} \tag{3}$$

$$\begin{aligned} \text{MSE}(0) &= (E(\hat{f}(0)) - f(0))^2 + Var(\hat{f}(0)) \\ &\approx (\frac{\pi_h}{h} - f(0))^2 + \frac{\pi_h}{nh^2}(1 - \pi_h) \\ &= (\frac{f''(0)}{24})^2 h^4 + \frac{1}{nh}(f(0) + \frac{f''(0)h^2}{24})(1 - hf(0) + \frac{f''(0)h^3}{24}) \\ &\approx (\frac{f''(0)}{24})^2 h^4 + \frac{1}{nh} f(0)(1 - hf(0)) \end{aligned} \tag{4}$$

(C)

$$\frac{\partial \text{MSE}(0)}{\partial h} = 4Ah^3 - \frac{f(0)}{nh^2} \tag{5}$$

In order to minimize the MSE, $h = (\frac{f(0)}{4An})^{\frac{1}{5}}$

2 Curve Fitting by linear smoothing

(A)

$$\begin{aligned} y_{\text{estimate}} &= \beta_{\text{estimate}} x \\ &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} x \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sum_i x_i y_i}{\sum_i x_i^2} x \\
&= \sum_i \frac{x_i x}{\vec{x}^2} y_i
\end{aligned} \tag{6}
$$

So we have $\omega_i = \frac{x_i x}{\vec{x}^2}$