6.2

记$b(x) = (1, x, x^2, x^3, \cdots, x^k)^T$,$\mathbf{B} = (b_1(x), b_2(x), b_3(x), \cdots, b_N(x))^T$。有

$$b(x_0)^T = b(x_0)^T(\mathbf{B}^T\mathbf{W}(x_0)\mathbf{B})^{-1}\mathbf{B}^T\mathbf{W}(x_0)\mathbf{B}$$

根据$l_i(x_0)$的定义，有

$$\begin{aligned}1 &= b(x_0)^T(\mathbf{B}^T\mathbf{W}(x_0)\mathbf{B})^{-1}\mathbf{B}^T\mathbf{W}(x_0)\mathbf{1} = \sum_{i=1}^N l_i(x_0)\\ x_0^k &= b(x_0)^T(\mathbf{B}^T\mathbf{W}(x_0)\mathbf{B})^{-1}\mathbf{B}^T\mathbf{W}(x_0)\mathbf{B}_{k+1} = \sum_{i=1}^N l_i(x_0)x_i^k\end{aligned}$$

于是

$$\sum_{i=1}^N l_i(x_0)(x_i - x_0) = \sum_{i=1}^N l_i(x_0)x_i - x_0\sum_{i=1}^N l_i(x_0) = 0$$

对于$j \geq 2$的情况

$$\begin{aligned}k_j(x_0) &= \sum_{i=1}^N (x_i - x_0)^j l_i(x_0)\\ &= \sum_{i=1}^N \left(\sum_{k=0}^j (-1)^k C_j^k x_i^{j-k} x_0^k\right) l_i(x_0)\\ &= \sum_{k=0}^j (-1)^k C_j^k x_0^k \left(\sum_{i=1}^N l_i(x_0) x_i^{j-k}\right)\\ &= \sum_{k=0}^j (-1)^k C_j^k x_0^k x_0^{j-k}\\ &= x_0^j \sum_{k=0}^j C_j^k (-1)^k\\ &= 0\end{aligned}$$

使用泰勒展开

$$\begin{aligned}E\hat{f}(x_0) = f(x_0)&\sum_{i=1}^N l_i(x_0) + f'(x_0)\sum_{i=1}^N (x_i - x_0) l_i(x_0)\\ &+ \frac{f''(x_0)}{2}\sum_{i=1}^N (x_i - x_0)^2 l_i(x_0)\\ &+ \cdots\\ &+ (-1)^k \frac{f^{(k)}}{k!}\sum_{i=1}^N (x_i - x_0)^k l_i(x_0) + R\\ &= f(x_0) + R\end{aligned}$$

也就是说，k次局部多项式回归的偏倚是一个$k+1$阶小量。

6.7

类似于平滑样条，定义$\mathbf{S}_\lambda$满足$\{\mathbf{S}_\lambda\}_{ij} = l_i(x_j)$，有$\hat{\mathbf{f}} = \mathbf{S}_\lambda\mathbf{y}$。

于是

$$y_i - \hat{f}^{-i}(x_i) = \frac{y_i - \hat{f}(x_i)}{1 - \{\mathbf{S}_\lambda\}_{ii}} = \frac{y_i - \hat{f}(x_i)}{1 - l_i(x_i)}$$

6.10

$$
\begin{aligned}
&E[\mathrm{ASR}(\lambda) - \mathrm{PE}(\lambda)] \\
&\quad = E\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{f}_\lambda(x_i))^2 - \frac{1}{N}\sum_{i=1}^{N}(y_i^* - \hat{f}_\lambda(x_i))^2 \\
&\quad = \frac{1}{N}E\sum_{i=1}^{N}(y_i - y_i^*)(y_i + y_i^* - 2\hat{f}_\lambda(x_i)) \\
&\quad = \frac{1}{N}E_{\mathbf{y}}\sum_{i=1}^{N} y_i^2 + y_i y_i^* - 2\hat{f}_\lambda(x_i)y_i - y_i^* y_i - {y_i^*}^2 + 2y_i^*\hat{f}_\lambda(x_i) \\
&\quad = \frac{1}{N}E_{\mathbf{y}}\sum_{i=1}^{N} -2\hat{f}_\lambda(x_i)y_i + 2y_i^*\hat{f}_\lambda(x_i) \\
&\quad = -\frac{2}{N}\sum_{i=1}^{N}\mathrm{Cov}(\hat{f}_\lambda(x_i), y_i)
\end{aligned}
$$

所以

$$
\mathrm{PE}(\lambda) = \mathrm{ASR}(\lambda) + \frac{2}{N}\sum_{i=1}^{N}\mathrm{Cov}\left(\hat{f}_\lambda(x_i), y_i\right)
$$

其中

$$
\begin{aligned}
\sum_{i=1}^{N}\mathrm{Cov}(\hat{f}_\lambda(x_i), y_i) &= \mathrm{trace}(\mathrm{Cov}(\mathbf{S}\mathbf{y}, \mathbf{y})) \\
&= \mathrm{trace}(\mathbf{S}\mathrm{Var}(\mathbf{y})) \\
&= \mathrm{trace}(\mathbf{S})\sigma_\epsilon^2
\end{aligned}
$$

所以

$$
C_\lambda = \mathrm{ASR}(\lambda) + \frac{2\sigma^2}{N}\mathrm{trace}(\mathbf{S}_\lambda)
$$