Nonparametrics homework problem

Assume we have iid data $\{(x_i, y_i)\}_{i=1}^n$, where $y_i = g(x_i) + \epsilon_i$. Suppose that $g(x)$ is approximated locally by a linear polynomial with kernel weight function $K_h(x - x_i)$.

- fitting criterion: in a small region around x_0, $g(x) \approx a_0 + b_0(x - x_0)$,

$$(\hat{a}_0, \hat{b}_0) = \arg\min_{a_0, b_0} \sum_{i=1}^n (y_i - a_0 - b_0(x_i - x_0))^2 w_i$$

- Kernel weights: $w_i = K_h(x_0 - x_i) / \sum_{i=1}^n K_h(x_0 - x_i)$.

Derive the local linear estimator $\hat{g}(x)$.

Solution: Let $\mathcal{L} = \sum_{i=1}^n (y_i - a_0 - b_0(x_i - x_0))^2 w_i$.

$$\text{Set} \qquad \partial \mathcal{L} / \partial a_0 = 2 \sum_{i=1}^n (y_i - a_0 - b_0(x_i - x_0)) w_i = 0.$$

$$\text{Then} \qquad a_0 = \sum_{i=1}^n (y_i w_i - b_0 w_i (x_i - x_0)) = \sum_{i=1}^n y_i w_i - b_0 S_{n1}, \tag{1}$$

where $S_{n1} = \sum_{i=1}^n w_i (x_i - x_0)$.

$$\text{Set} \qquad \partial \mathcal{L} / \partial b_0 = 2 \sum_{i=1}^n (y_i - a_0 - b_0(x_i - x_0)) w_i (x_i - x_0) = 0.$$

$$\text{Then} \qquad \sum_{i=1}^n y_i w_i (x_i - x_0) - a_0 S_{n1} - b_0 S_{n2} = 0, \tag{2}$$

where $S_{n2} = \sum_{i=1}^n w_i (x_i - x_0)^2$. Plug a_0 in (1) into (2) and solve for b_0. We get

$$\sum_{i=1}^n y_i w_i (x_i - x_0) - (\sum_{i=1}^n y_i w_i - b_0 S_{n1}) S_{n1} - b_0 S_{n2} = 0$$

$$\sum_{i=1}^n y_i w_i (x_i - x_0) - \sum_{i=1}^n y_i w_i S_{n1} = b_0 (S_{n2} - S_{n1}^2)$$

$$b_0 = \frac{\sum_{i=1}^n y_i w_i (x_i - x_0) - \sum_{i=1}^n y_i w_i S_{n1}}{S_{n2} - S_{n1}^2}. \tag{3}$$

Plug b_0 in (3) back into (1) and get

$$\begin{aligned} \hat{a}_0 &= \frac{\sum_{i=1}^n y_i w_i (S_{n2} - S_{n1}^2) - \sum_{i=1}^n y_i w_i (x_i - x_0) S_{n1} + \sum_{i=1}^n y_i w_i S_{n1}^2}{S_{n2} - S_{n1}^2} \\ &= \frac{\sum_{i=1}^n y_i w_i S_{n2} - \sum_{i=1}^n y_i w_i (x_i - x_0) S_{n1}}{S_{n2} - S_{n1}^2} \\ &= \frac{\sum_{i=1}^n y_i (S_{n2} - (x_i - x_0) S_{n1}) w_i}{S_{n2} - S_{n1}^2}. \end{aligned} \tag{4}$$

Since $g(x) \approx a_0 + b_0(x - x_0)$ in a small region around x_0, we use (4) for $\hat{g}(x_0)$. For a general $\hat{g}(x)$, replace x_0 by a general point x in all the above expressions, and we get

$$\hat{g}(x) = \frac{\sum_{i=1}^n y_i (S_{n2} - (x_i - x) S_{n1}) w_i}{S_{n2} - S_{n1}^2} \tag{5}$$

with $S_{n1} = \sum_{i=1}^n w_i (x_i - x)$, $S_{n2} = \sum_{i=1}^n w_i (x_i - x)^2$ and $w_i = K_h(x - x_i) / \sum_{i=1}^n K_h(x - x_i)$. □