EXERCISE 13

WEIYU LI

1. Consider the bandwidth selection method of Nadaraya-Waston estimator $\hat{m}$. If the bandwidth is chosen without leave-one-out as LSCV does, that is,

$$h_o = \arg\min \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{m}(X_i))^2,$$

Prove that $h_0 = 0$.

Remark 1. *The claim is informal. In fact, there may be some other minimizers of the optimization problem, based on specific structure of X_i, Y_i. For example, if $Y_i = Y_j$ for all i, j, then any h minimizes the objective. And $h_0 = 0$ is not well-defined. Therefore, there're two alternatives to rigorously state the proposition:*

(1) *Letting $h_0 \to 0+$ minimizes $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{m}(X_i))^2$ for arbitrary X_i, Y_i.*
(2) *Letting $h_0 \to 0+$ is a minimizor of $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{m}(X_i))^2$.*

Remark 2. *We will need the assumption that $K(x) \to 0 (x \to \infty)$, which cannot be guaranteed by the integrability of K.*

Proof. Since we can regard $\hat{m}(X_i)$ as $\sum_{j=1}^{n} W_{i,j} Y_j$, where $W_{i,j}(\mathbf{X}) \in [0,1]$ and $\sum_{j=1}^{n} W_{i,j} = 1$, we can rewrite the objective as

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{m}(X_i))^2 &= \frac{1}{n} \sum_{i=1}^{n} (Y_i - \sum_{j=1}^{n} W_{i,j} Y_j)^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} (\sum_{j=1}^{n} \tilde{W}_{i,j} Y_j)^2 \\
&\geq 0,
\end{aligned}$$

where $\tilde{W}_{i,j} = \delta_{ij} - W_{ij}$ and the inequality becomes equal when $\sum_{j=1}^{n} \tilde{W}_{i,j} Y_j = 0, \ \forall i$.

Notice that

$$W_{i,j} = \frac{K(\frac{X_i - X_j}{h_0})}{K(0) + \sum_{k \neq j} K(\frac{X_i - X_k}{h_0})},$$

then if $X_i \neq X_j$, $\tilde{W}_{i,j} = W_{i,j} \to 0$ as $h_0 \to 0$. From $\sum_{j=1}^{n} \tilde{W}_{i,j} = 0$, we further have

$$\sum_{j=1}^{n} \tilde{W}_{i,j} Y_j = \sum_{j: X_j \neq X_i} \tilde{W}_{i,j} Y_j + \sum_{j: X_j = X_i} \tilde{W}_{i,j} Y_i = \sum_{j: X_j \neq X_i} \tilde{W}_{i,j} Y_j - \sum_{j: X_j \neq X_i} \tilde{W}_{i,j} Y_i \to 0,$$

in which case $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{m}(X_i))^2$ is minimized at its lower bound 0. □

Date: 2019/11/18.

liweiyu@mail.ustc.edu.cn.

2. Prove that for Nadaraya-Waston estimator $\hat{m}$, the objective function of leave-one-out bandwidth selection satisfies

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{m}_{-i}(X_i))^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i - \hat{m}(X_i)}{1 - W_i(X_i)}\right)^2,$$

where $W_i(x) = \mathcal{K}_h(X_i - x) / \sum_{j=1}^{n} \mathcal{K}_h(X_j - x)$**.**

Proof. Notice that

$$\hat{m}_{-i}(X_i) = \frac{\sum_{j\neq i}\mathcal{K}_h(X_j - X_i)Y_j}{\sum_{j\neq i}\mathcal{K}_h(X_j - X_i)} = \frac{\sum_{j\neq i} W_j(X_i)\left(\sum_{j'=1}^{n}\mathcal{K}_h(X'_j - X_i)\right)Y_j}{[1 - W_i(X_i)]\sum_{j=1}^{n}\mathcal{K}_h(X_j - X_i)} = \frac{\sum_{j\neq i} W_j(X_i)Y_j}{1 - W_i(X_i)}.$$

From the definition of $\hat{m}(X_i) = \sum_{j=1}^{n} W_j(X_i)Y_j$, we have

$$\begin{aligned}\frac{Y_i - \hat{m}(X_i)}{1 - W_i(X_i)} &= \frac{Y_i - W_i(X_i)Y_i - \sum_{j\neq i} W_j(X_i)Y_j}{1 - W_i(X_i)} \\ &= Y_i - \frac{\sum_{j\neq i} W_j(X_i)Y_j}{1 - W_i(X_i)} \\ &= Y_i - \hat{m}_{-i}(X_i),\end{aligned}$$

which gives that $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{m}_{-i}(X_i))^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i - \hat{m}(X_i)}{1 - W_i(X_i)}\right)^2$. $\square$