EXERCISE 16

WEIYU LI

1. Consider the regression model

$$y_i = f(x_i) + e_i, \quad i = 1, \dots, n.$$

Suppose that $N_j(x), j = 1, \dots, n$ **are the basis functions of** 3**-rd order natural splines (with knots at** $x_i, i = 1, \dots, n$**),** $f(x) = \sum_{j=1}^n N_j(x)\beta_j$**,** $\Omega = (\Omega_{jk})$ **with** $\Omega_{jk} = \int N_j''(t)N_k''(t)dt$**, and** $N = (N_{ij})$ **with** $N_{ij} = N_j(x_i)$**.**

(1) Using the smooth splines method, solve the estimation of $\beta = (\beta_1, \dots, \beta_n)^\top$**.**

Solve. From Page 53 of *Lec15.pdf*,

$$\hat{\beta} = (N^\top N + \lambda\Omega)^{-1} N^\top y.$$

□

(2) Denote $S_\lambda = N(N^\top N + \lambda\Omega)^{-1}N^\top$**, prove that**

$$CV(\lambda) = \frac{1}{n}\sum_{i=1}^n \left(y_i - \hat{f}^{-i}(x_i)\right)^2 = \frac{1}{n}\sum_{i=1}^n \left[\frac{y_i - \hat{f}(x_i)}{1 - S_\lambda(i,i)}\right]^2,$$

where $\hat{f}^{-i}(x_i)$ **is the predict value at** x_i **given the leave-**i**-out sample.** *(Hint: the smoother matrix $S_\lambda = S_\lambda(x)$ such that $\hat{y} = S_\lambda y$ is independent of the value of y.)*

Proof. Denote $y^i = \left(y_1, \dots, y_{i-1}, \hat{f}^{-i}(x_i), y_i, \dots, y_n\right)^\top$. Notice that

$$\hat{f}(x_i) = e_i^\top S_\lambda y,$$

where $e_i \in \mathbb{R}^n$ is the identity vector with only the i-th entry nonzero and being 1. We first prove an important claim.

[CLAIM]

$$\hat{f}^{-i}(x_i) = e_i^\top S_\lambda y^i.$$

Proof of **[CLAIM]***:* Note that $\hat{f}$ is the twice differentiable function that minimizes $\sum_{j=1}^n \left(y_j - f(x_j)\right)^2 + \lambda \int \left(f''\right)^2$, *i.e.*,

$$\hat{f} = \arg\min_{f \in \mathcal{C}^2} \sum_{j=1}^n \left(y_j - f(x_j)\right)^2 + \lambda \int \left(f''\right)^2,$$

where $\mathcal{C}^2$ consists of twice differentiable function. Similarly,

$$\hat{f}^{-i} = \arg\min_{f \in \mathcal{C}^2} \sum_{j \neq i} \left(y_j - f(x_j)\right)^2 + \lambda \int \left(f''\right)^2.$$

Date: 2019/12/02.

liweiyu@mail.ustc.edu.cn.

On the one hand, we have that

$$\begin{aligned}\min_{f\in\mathcal{C}^2}\sum_{j=1}^{n}\left(y_j^i-f(x_j)\right)^2+\lambda\int\left(f''\right)^2 &\leq \sum_{j=1}^{n}\left(y_j^i-\hat{f}^{-i}(x_j)\right)^2+\lambda\int\left((\hat{f}^{-i})''\right)^2\\ &=\sum_{j\neq i}\left(y_j^i-\hat{f}^{-i}(x_j)\right)^2+\lambda\int\left((\hat{f}^{-i})''\right)^2,\end{aligned}$$

since $y_i^i=\hat{f}^{-i}(x_i)$. On the other hand, from the property that $\min(A+B)\geq\min A+\min B$,

$$\begin{aligned}&\min_{f\in\mathcal{C}^2}\sum_{j=1}^{n}\left(y_j^i-f(x_j)\right)^2+\lambda\int\left(f''\right)^2\\ \geq&\min_{f\in\mathcal{C}^2}\left[\sum_{j\neq i}\left(y_j^i-f(x_j)\right)^2+\lambda\int\left(f''\right)^2\right]+\min_{f\in\mathcal{C}^2}\left(y_i^i-f(x_i)\right)^2\\ =&\sum_{j\neq i}\left(y_j^i-\hat{f}^{-i}(x_j)\right)^2+\lambda\int\left((\hat{f}^{-i})''\right)^2+0.\end{aligned}$$

Consequently, we find that

$$\hat{f}^{-i}=\arg\min_{f\in\mathcal{C}^2}\sum_{j=1}^{n}\left(y_j^i-f(x_j)\right)^2+\lambda\int\left(f''\right)^2,$$

that is, $\hat{f}^{-i}$ is the fitted spline for the data (x,y^i), then $\hat{f}^{-i}(x_i)=e_i^\top S_\lambda y^i$.

Now we turn back to our main proof. Note that the right side can be rewritten as

$$\begin{aligned}\frac{y_i-\hat{f}(x_i)}{1-S_\lambda(i,i)}=&\frac{y_i-\hat{f}^{-i}(x_i)+\hat{f}^{-i}(x_i)-\hat{f}(x_i)}{1-S_\lambda(i,i)}\\ =&\frac{1}{1-S_\lambda(i,i)}\left(y_i-\hat{f}^{-i}(x_i)+e_i^\top S_\lambda(y^i-y)\right)\\ =&\frac{1}{1-S_\lambda(i,i)}\left(y_i-\hat{f}^{-i}(x_i)+e_i^\top S_\lambda e_i(\hat{f}^{-i}(x_i)-y_i)\right)\\ =&y_i-\hat{f}^{-i}(x_i).\end{aligned}$$

□