Parte teórica TP final

viernes, 5 de julio de 2024 20:52

$(x_i, y_i) \sim (x, y)$ $i = 1 \cdots n$ donde $X, Y \in \mathbb{R}$ y $m(x) = E(Y|X=x)$

Definimos $\hat{m}_h(x) = \sum_{i=1}^{n} y_i\, m_{i,h}(x)$ con Y el vector de respuestas compuesto por y_i. Probar que $\hat{Y} = SY$ es un t.l en Y y hallar la expresión de S definiendo cada una de sus componentes.

- Para ver que SY es una t.l, ~~sug~~ ① $S(Y_1+Y_2) = SY_1 + SY_2$

$\hat{Y} \in \mathbb{R}^{n\times 1}$, $Y \in \mathbb{R}^{n\times 1} \Rightarrow S \in \mathbb{R}^{n\times n}$ ② $S(\alpha Y) = \alpha S Y$

①
$$S(Y_1+Y_2) = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} Y_{11}+Y_{12} \\ \vdots \\ Y_{n1}+Y_{n2} \end{pmatrix} = \begin{pmatrix} S_{11}(Y_{11}+Y_{12}) + S_{12}(Y_{21}+Y_{22}) \cdots + S_{1n}(Y_{n1}+Y_{n2}) \\ \vdots \\ S_{n1}(Y_{11}+Y_{12}) + \cdots + S_{nn}(Y_{n1}+Y_{n2}) \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{n} S_{1k} Y_{k1} + S_{1k} Y_{k2} \\ \vdots \\ \sum_{k=1}^{n} S_{nk} Y_{k1} + S_{nk} Y_{k2} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} S_{1k} Y_{k1} \\ \vdots \\ \sum_{k=1}^{n} S_{nk} Y_{k1} \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^{n} S_{1k} Y_{k2} \\ \vdots \\ \sum_{k=1}^{n} S_{nk} Y_{k2} \end{pmatrix} = SY_1 + SY_2 \quad \checkmark$$

②
$$S\cdot(\alpha Y) \underset{\text{asociatividad}}{=} \Big(S\,\underbrace{(\alpha, \ldots, \alpha)}_{n}^{t}\Big) Y \underset{\alpha\in\mathbb{R}}{=} \alpha S \cdot Y \quad \checkmark$$

Luego, SY es una t.l en Y, veamos quién es S.

Sabemos que podemos estimar $\hat{Y}_i = \hat{m}_h(x_i)$ donde

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} Y_i\, K\left(\frac{X_i - x}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)}$$

es el estimador de Nadaraya Watson para $m(x) = E(Y|X=x)$ (para el caso continuo $\left|\frac{X-x}{h}\right| \le 1$)

obs: $E(Y|X=x) \xrightarrow{LGN} \frac{1}{n}\sum Y_i\, \mathbb{1}\{|\frac{X_i - x}{h}| \le 1\} \ldots$

$\Rightarrow$ Para la muestra X_i
$$\hat{Y}_i = \hat{m}_h(x_i) = \frac{\sum_{j=1}^{n} Y_j\, K\left(\frac{X_j - X_i}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{X_k - X_i}{h}\right)} = \sum_{j=1}^{n} Y_i\, \omega_{j,h}(x_i)$$

metemos el denominador en la sumatoria

$$\Rightarrow \hat{Y}_i \underset{S \text{ es t.l}}{=} S\cdot Y_i = \sum_{j=1}^{n} Y_i\, \omega_{j,h}(x_i)$$

$\Rightarrow$ Cada componente $S_{ij} = \omega_{j,h}(x_i) = \dfrac{K\left(\frac{X_j - X_i}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{X_k - X_i}{h}\right)}$

Por lo tanto

$$S = \begin{pmatrix} \omega_{1,h}(x_1) & \cdots & \omega_{n,h}(x_1) \\ \vdots & \ddots & \vdots \\ \omega_{1,h}(x_n) & \cdots & \omega_{n,h}(x_n) \end{pmatrix} = \begin{pmatrix} \frac{K\left(\frac{X_1 - X_1}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{X_k - X_1}{h}\right)} & \cdots & \frac{K\left(\frac{X_n - X_1}{h}\right)}{\sum_{k=1}^{n} K\left(\frac{X_k - X_1}{h}\right)} \\ \vdots & \ddots & \vdots \end{pmatrix}$$

dos: quedan 0 los elementos de la diagonal.