52805

2303

26-27.4.22 תרגול #6 מודלים ב'

- המשך: אלגו' גרעין
- אלגו' פולינומיאליים מקומיים.

פרק 4 - שאלה 8:

אומד K-NN לרגרסיה הוא

$$\hat{f}(x)=\frac{1}{K}\sum_{j\in N_k(x)} Y_j$$

כאשר $N_k(x)$ היא קבוצת האינדקס של k הנקודות design הקרובות ביותר ל-x.

(1) האם האומד $\hat{f}(x)$ שייך ל-$\Sigma(\beta,L)$ עבור סדר כלשהו?

פתרון:

__תזכורת:__ $f\in\Sigma(\beta,L)$ אם ל-f יש $l=\max\{k\in\mathbb{Z}_+ : k<\beta\}$ נגזרות ו-

$\forall x,y\in\mathbb{R}\quad |f^{(l)}(x)-f^{(l)}(y)|\le L|x-y|^{\beta-l}$ (הנגזרת היא $\beta-l$ Hölder)

[דוגמה: $f(x)=ax+b$ אז $f\in\Sigma(2,L)$ לכל $L>0$ כי $f'(x)=a$
וגם $f^{(k)}(x)=0\ \forall k>1$]

אם כן, נשים לב ש-$\hat{f}(x)$ היא פונקציה קבועה למקוטעין עם קפיצות בנקודות בהן $N_k(x)$ משתנה. דהיינו הפונקציה בכלל לא רציפה ולכן אין מה להסתכל על חלק של נגזרות.

(2) הניחו כי $X_j=j/n$ עבור $j=1,\dots,n$. מצאו את ערך k האופטימלי עבורו האומד הוא מינימקסי ביחס לסיכון MSE, מעל $\Sigma(1,L)$.

פתרון:

- נחשב את סכום ה-MSE, שונות ~~אמא~~ והטיה בנפרד, ונחסום את הביטויים ע"י ביטויים שלא תלויים בפונ' f הלא ידועה.
- נזכור שהמודל הוא

$$Y_i=f(x_i)+\varepsilon_i \quad ; \varepsilon_i \text{ iid with } \mathbb{E}\varepsilon_i=0,\ Var(\varepsilon_i)=\sigma^2$$

- נמזער על פני k את החסם.

חסם לשונות:

$$\mathrm{Var}_f(\hat f(x_0)) = \mathbb{E}_f\left(\hat f(x_0) - \mathbb{E}_f \hat f(x_0)\right)^2 = \mathbb{E}_f\left(\frac{1}{k}\sum_{j\in N_k(x_0)} Y_j - \frac{1}{k}\sum_{j\in N_k(x_0)} f(x_j)\right)^2$$

$$\left(\mathbb{E}_f \hat f(x_0) = \mathbb{E}_f \frac{1}{k}\sum_{j\in N_k(x_0)} Y_j = \frac{1}{k}\sum_{j\in N_k(x_0)} \mathbb{E}_f Y_j = \frac{1}{k}\sum_{j\in N_k(x_0)} f(x_j)\right)$$

$Y_j = f(x_j) + \varepsilon_j$ והסכומה (ε) על פני אלמ' אינדקסים.

$$= \mathbb{E}_f\left(\frac{1}{k}\sum_{j\in N_k(x_0)} \varepsilon_j\right)^2 = \frac{1}{k^2}\cdot \underbrace{|N_k(x_0)|}_{k}\cdot \sigma^2 = \frac{\sigma^2}{k}$$

(k האברים הקרובים ביותר)

חסם להטיה:

חישבנו למעלה.

$$\left|\mathbb{E}_f \hat f(x_0) - f(x_0)\right| = \left|\frac{1}{k}\sum_{j\in N_k(x_0)} f(x_j) - f(x_0)\right|$$

לפי ה-design

$$= \left|\frac{1}{k}\sum_{j\in N_k(x_0)}\left(f(j/n) - f(x_0)\right)\right|$$

אי"ש המשולש

$$\le \frac{1}{k}\sum_{j\in N_k(x_0)} \left|f(j/n) - f(x_0)\right|$$

$f\in\Sigma(1,L)$ והורדנו ה-ס

היא ספוס הפונקציה.

עבור n גדול מספיק: בסביבה $N_k(x_0)$ יש $\frac{k}{2}$ נקודות מימין ל-x_0 ו-$\frac{k}{2}$ משמאל.

עבור $j\in N_k(x_0)$,

$|j/n - x_0| \le |j/n|$

$$\le \frac{1}{k}\sum_{j\in N_k(x_0)} L\cdot |j/n - x_0| = \cancel{\frac{L}{kn}\cdot\sum_{j\in N_k(x_0)} |j - nx_0|}$$

$$\le \frac{2L}{kn}\sum_{j=1}^{k/2} |j| \le 2L\cdot\frac{k}{n}$$

לכן בסה"כ:

$$\mathbb{E}_f\left(\hat f(x_0) - f(x_0)\right)^2 \le \frac{\sigma^2}{k} + \left(2L\frac{k}{n}\right)^2 =: g(k)$$

$$g'(k) = -\frac{\sigma^2}{k^2} + 2\cdot 2L\frac{k}{n}\cdot\frac{2L}{n} = -\frac{\sigma^2}{k^2} + \frac{8L^2}{n^2}\cdot k \stackrel{!}{=} 0$$

$$\Leftrightarrow \frac{8L^2}{n^2}k = \frac{\sigma^2}{k^2} \Leftrightarrow k^3 = \sigma^2\cdot\frac{n^2}{8L^2} \Rightarrow \boxed{k^* = C(L,\sigma)\cdot n^{2/3}}$$

Local Polynomial Estimators - פרק 5

- הסתכלנו על הבעיה של רגרסיה: f עבור $(X_1,Y_1),\dots,(X_n,Y_n)$ מדגם i.i.d. עם

$$Y_j = f(X_j) + \varepsilon_j \qquad \varepsilon_j \text{ - i.i.d.} \quad \mathbb{E}\varepsilon_j = 0 \quad \mathbb{E}\varepsilon_j^2 = 1$$

- במקרה הזה, אומד Nadaraya-Watson הוא

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

Kernel

Bandwidth

- אם K חיובי, אז ניתן לכתוב

$$\hat{f}(x) = \arg\min_{\theta\in\mathbb{R}} \sum_{j=1}^n (Y_j-\theta)^2 K\left(\frac{x-X_j}{h}\right)$$

- רצינו להכליל את האומד. נניח כי f גזירה l פעמים. אז לכל z בסביבה של x,

$$f(z) = f(x) + f'(x)(z-x) + \dots + \frac{1}{l!} f^{(l)}(x)(z-x)^l + o((z-x)^l)$$
$$= \theta(x)^T U\left(\frac{z-x}{h}\right) + o((z-x)^l)$$

$$\theta(x)^T = \left(f(x), f'(x)h, \dots, f^{(l)}(x)h^l\right) \;;\; U^T(t) = \left(1, t, \frac{1}{2!}t^2, \dots, \frac{1}{l!}t^l\right)$$

⇐ עבור X_j-ים קרובים ל-x,

אז
$$Y_j = \theta(x)^T U\left(\frac{X_j-x}{h}\right) + o\left((X_j-x)^l\right) + \varepsilon_j$$

אז
$$\hat{\theta}(x) = \arg\min_{\theta\in\mathbb{R}^{l+1}} \sum_{j=1}^n \left(Y_j - \theta^T U\left(\frac{X_j-x}{h}\right)\right)^2 K\left(\frac{x-X_j}{h}\right)$$

Local Polynomial Estimator of order l → $\hat{f}(x) := \hat{\theta}(x)^T U(0)$

הערה 1: אומד N-W הוא אומד LP מסדר 0.

הערה 2: שאר הרכיבים של θ יהוו אומדים לנגזרות של f, אחרי נרמול מתאים.

• אם $B_n(x)$ הפיכה, אז הפתרון הוא:

נסמן $$\hat{\Theta}(x) = B_n(x)^{-1} a_n(x)$$ ★

$$a_n(x) = \frac{1}{nh} \sum_{j=1}^{n} Y_j U\left(\frac{X_j - x}{h}\right) K\left(\frac{x - X_j}{h}\right)$$

$$B_n(x) = \frac{1}{nh} \sum_{j=1}^{n} U\left(\frac{X_j - x}{h}\right) U\left(\frac{X_j - x}{h}\right)^T K\left(\frac{x - X_j}{h}\right)$$

• לכן, האומד LP ל- f הוא:

בשלב הבא ננסה לחקור B המטריצה של אומדי LP.

$$\hat{f}(x) = \sum_{j=1}^{n} Y_j W_{nj}(x)$$

$$W_{n,j}(x) = \frac{1}{nh} U^T(0) B_n(x)^{-1} U\left(\frac{X_j - x}{h}\right) K\left(\frac{x - X_j}{h}\right)$$

תרגילים:

שאלה 1: הראו כי הנוסחה ★ נכונה.

פתרון:

נגדיר $$g(\theta) = \sum_{i=1}^{n} (y_i - \theta^T v_i)^2 K_i \quad , \quad v_i \in \mathbb{R}^{l+1},\ K_i \in \mathbb{R}$$

בהמשך נבין את v_i, K_i המתאימים.

$$f_i(\theta) = (y_i - \theta^T v_i)^2$$

$$\nabla f_i(\theta) = 2(y_i - \theta^T v_i) v_i^T$$

$$\Rightarrow \quad \nabla g(\theta) = \sum_{i=1}^{n} \nabla f_i(\theta) K_i = \sum_{i=1}^{n} 2(y_i - \theta^T v_i) v_i^T K_i$$

$$= \sum_{i=1}^{n} 2 y_i v_i^T K_i - 2\sum_{i=1}^{n} \theta^T v_i v_i^T K_i \overset{!}{=} 0$$

$$\Rightarrow \quad \theta^T \underbrace{\sum_{i=1}^{n} v_i v_i^T K_i}_{\text{מטריצה}} = \underbrace{\sum_{i=1}^{n} y_i v_i^T K_i}_{\text{וקטור}}$$

בהנחה שהמטריצה שלנו הפיכה.

$$\theta^T = \left(\sum_{i=1}^{n} y_i v_i^T K_i\right) \cdot \left(\sum_{i=1}^{n} v_i v_i^T K_i\right)^{-1}$$

$$\Rightarrow \quad \hat{\theta} = \left(\sum_{i=1}^{n} v_i v_i^T K_i\right)^{-1} \left(\sum_{i=1}^{n} y_i v_i^T K_i\right)$$

והנוסחה מתקבלת זהה:

(וכפלה ב- $\frac{nh}{nh}$)

$$y_i = Y_i \in \mathbb{R} \ ; \quad v_i = U\left(\frac{X_i - x}{h}\right) \in \mathbb{R}^{l+1} \ ; \quad K_i = K\left(\frac{x - X_i}{h}\right) \in \mathbb{R}$$

ב:

נמצא את האומד LP לרגרסיה מסדר $l=1$. נמצא נוסחאות סגורות ל- f ול- f'.

פתרון:

עבור $l=1$:

$$\theta(x)^T = (f(x), f'(x)h)$$
$$U^T(t) = (1, t)$$

צריך למצוא: $\hat{\theta}, \hat{f}, \hat{f}'$.

$$\hat{\theta}(x) = B_n^{-1}(x)\, a_n(x) \quad ; \quad \hat{f}(x) = \hat{\theta}(x)^T U(0) \quad ; \quad \hat{f}'(x) = \frac{1}{h}\hat{\theta}(x)^T U'(0)$$

$$\ast\ B_n(x) = \frac{1}{nh}\sum_{i=1}^{n} U\left(\frac{X_j - x}{h}\right) U\left(\frac{X_j - x}{h}\right)^T K\left(\frac{x - X_j}{h}\right)$$

$$\ast\ U(t)U(t)^T = \begin{pmatrix} 1 \\ t \end{pmatrix}(1, t) = \begin{pmatrix} 1 & t \\ t & t^2 \end{pmatrix}$$

$$= \frac{1}{nh}\sum_{i=1}^{n} \begin{pmatrix} 1 & \frac{X_j - x}{h} \\ \frac{X_j - x}{h} & \left(\frac{X_j - x}{h}\right)^2 \end{pmatrix} K\left(\frac{x - X_j}{h}\right)$$

$$= \frac{1}{nh}\begin{pmatrix} \sum_{i=1}^{n} K\left(\frac{x - X_j}{h}\right) & \sum_{i=1}^{n} \frac{X_j - x}{h}\cdot K\left(\frac{x - X_j}{h}\right) \\ \sum_{i=1}^{n} \frac{X_j - x}{h} K\left(\frac{x - X_j}{h}\right) & \sum_{i=1}^{n} \left(\frac{X_j - x}{h}\right)^2 K\left(\frac{x - X_j}{h}\right) \end{pmatrix} =: \frac{1}{nh}\cdot\begin{pmatrix} S_0 & S_1 \\ S_1 & S_2 \end{pmatrix}$$

$$\ast\ a_n(x) = \frac{1}{nh}\sum_{j=1}^{n} Y_j \cdot U\left(\frac{X_j - x}{h}\right) K\left(\frac{x - X_j}{h}\right) = \frac{1}{nh}\begin{pmatrix} \sum_{j=1}^{n} Y_j \cdot K\left(\frac{x - X_j}{h}\right) \\ \sum_{j=1}^{n} Y_j \cdot \frac{X_j - x}{h} \cdot K\left(\frac{x - X_j}{h}\right) \end{pmatrix} =: \frac{1}{nh}\cdot\begin{pmatrix} T_0 \\ T_1 \end{pmatrix}$$

$$\Rightarrow \hat{\theta}(x) = B_n^{-1}(x)\, a_n(x) = \frac{1}{S_0 S_2 - S_1^2}\cdot\begin{pmatrix} S_2 & -S_1 \\ -S_1 & S_0 \end{pmatrix}\cdot nh \cdot \frac{1}{nh}\cdot\begin{pmatrix} T_0 \\ T_1 \end{pmatrix} = \frac{1}{S_0 S_2 - S_1^2}\cdot\begin{pmatrix} S_2 T_0 - S_1 T_1 \\ -S_1 T_0 + S_0 T_1 \end{pmatrix}$$

$$\Rightarrow \hat{f}(x) = \hat{\theta}(x)^T U(0) = \frac{S_2 T_0 - S_1 T_1}{S_0 S_2 - S_1^2}$$

$$\Rightarrow \hat{f}'(x) = \frac{1}{h}\cdot\hat{\theta}(x)^T U'(0) = \cancel{\frac{1}{h}\hat{\theta}(x)^T \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}}\ \frac{1}{h}\cdot\frac{S_0 T_1 - S_1 T_0}{S_0 S_2 - S_1^2}$$

$$U'(t) = (0, 1)$$

//