**Հետևանք** Թեորեմ 4-ից և լեմմա 2-ից հետևում է, որ

$$\sup_{f^* \in \mathcal{H}(\beta, L)} \mathbb{E}\left[L(\hat{f}_n, f^*)\right] \leq \underbrace{L^2 \cdot h^{2\beta}}_{\substack{\text{մոտարկման} \\ \text{սխալանք} \\ = \text{bias}}} + \underbrace{\frac{1}{nh}}_{\substack{\text{ստոխաստիկ} \\ \text{սխալանք} \\ = \text{variance}}} :$$

**Դիտողություն** Մի կողմից՝ ստոխաստիկ սխալանքը կարելի է գրել $\frac{1}{nh} = \frac{K}{n}$ տեսքով, որտեղ $K \in \mathbb{N}^*$ «սյուների» քանակն է: Մյուս կողմից՝ սյունապատկերը հետևյալ բազմության տարր է

$$\mathcal{F}_K = \left\{ f : [0,1] \to [0, +\infty) \text{ s.t. } f(x) = \sum_{j=1}^{K} \theta_j \mathbb{1}_{C_j}(x) \text{ for some } \theta = (\theta_1, \dots, \theta_K) \in \mathbb{R}_+ \right\}$$

Եթե մի պահ ենթադրենք, որ $f^* \in \mathcal{F}_K$, ապա $f^*$-ի գնահատման խնդիրը հանգում է $\theta$-պարամետրի գնահատման խնդրին: Քանի որ վերջինս $K$-չափանի է, այդ իսկ պատճառով առաջանում է $\frac{K}{n}$ ստոխաստիկ սխալանք: Սա ընդհանուր երևույթ է մեծաչափականությունում:

**Դիտողություն** Երբ $h$-ը փոքրացնում ենք, մոտարկման սխալանքը փոքրանում է, իսկ ստոխաստիկ սխալանքը՝ մեծանում: Այս երևույթն անվանում են գերմարզում (overfitting) կամ թերողորկում (undersmoothing):

# 8. Օպտիմալ կամ մինիմաքս ռիսկ

Փորձենք գտնել լավագույն լայնությունը՝ $h$-ը: Նշան.

$$\varphi(h) = L^2 h^{2\beta} + \frac{1}{nh}$$

Հաշվենք $\varphi$-ի առաջին և երկրորդ ածանցյալները՝

$$\varphi'(h) = 2\beta L^2 h^{2\beta-1} - \frac{1}{nh^2}$$

$$\varphi''(h) = 2\beta(2\beta-1) L^2 h^{2\beta-2} + \frac{2}{nh^3}$$

Նկատենք, որ $\varphi''(h) \geqslant 0 \ \forall h>0$: Հետևաբար $\varphi$ ֆունկցիան ուռուցիկ է $[0;+\infty)$-ում: Կամայական կետ, որտեղ $\varphi'$-ը հավասար է $0$-ի, կլինի գլոբալ մինիմումի կետ:

$$\varphi'(h) = 0 \iff (2\beta L^2)\, h^{2\beta-1} = \frac{1}{nh^2}$$

$$\iff (2\beta L^2)\cdot h^{2\beta+1} = n^{-1}$$

$$\iff h = \underbrace{(2\beta L^2)^{-\frac{1}{2\beta+1}}}_{C} \times n^{-\frac{1}{2\beta+1}}$$

$$\iff K = C \times n^{\frac{1}{2\beta+1}}$$

Քանի որ $K$-ն պետք է լինի ամբողջ թիվ, վերցնենք

$K = \left[n^{\frac{1}{2\beta+1}}\right] + 1$ , որտեղ $[x]$-ը $x$-ի ամբողջ մասն է:

**Թեորեմ 5** Եթե ընտրենք $K = [n^{\frac{1}{2\beta+1}}] + 1$, ապա

$$\sup_{f^* \in \mathcal{H}_{\beta,L}} \mathbb{E}\left[L(\hat{f}_n, f^*)\right] \leq (L^2+2)\, n^{-\frac{2\beta}{2\beta+1}} :$$

Մասնավորապես, $\beta = 1$ (Լիպշից) դեպքում ստանում ենք արդեն ծանոթ $n^{-2/3}$ արագությունը:

**Ապացույց** Հաշվի առնելով որ $x-1 \leq [x] \leq x \;\; \forall x \in \mathbb{R}$, և $n \geq 1$,

$$n^{\frac{1}{2\beta+1}} - 1 \leq \left[n^{\frac{1}{2\beta+1}}\right] \leq n^{\frac{1}{2\beta+1}}$$

$$n^{\frac{1}{2\beta+1}} \leq K \leq n^{\frac{1}{2\beta+1}} + 1 \leq 2n^{\frac{1}{2\beta+1}} :$$

Հետևաբար,

$$\varphi(h) = L^2 h^{2\beta} + \frac{1}{nh} = \frac{L^2}{K^{2\beta}} + \frac{K}{n}$$

$$\leq L^2 n^{-2\beta/2\beta+1} + 2\, n^{\frac{1}{2\beta+1}-1}$$

$$= (L^2+2)\, n^{-2\beta/2\beta+1} :$$

∎

**Դիտարկումներ**

1) Կարելի է ապացուցել, որ $n^{-2\beta/2\beta+1}$ արագությունը օպտիմալ է: Այսինքն $\exists\, C>0$ s.t.

$$\inf_{\hat{f}} \sup_{f^* \in \mathcal{H}(\beta,L)} \mathbb{E}\left[L(\hat{f}, f^*)\right] \geq C \cdot n^{-2\beta/2\beta+1} .$$

որտեղ inf-ը բոլոր հնարավոր գնահատականների վրա է:

Դաս 4

2) Եթե դիտարկենք նույն խնդիրը $d$-չափանի դեպքում, այսինքն $f^*: [0,1]^d \longrightarrow [0,+\infty)$, կստանանք $n^{-2\beta/2\beta+d}$ արագությունը: Այս արագությունը դանդաղում է, երբ $d$-ն մեծանում է:

3) Բացի որ գործնականում $f^*$-ը անհայտ է, անհայտ է նաև $\beta$-ն: Հետևաբար չենք կարող կիրառել $K=\left[n^{\frac{1}{2\beta+1}}\right]+1$ արժեքը: Բայց սա չի նշանակում, որ ամբողջ կատարված աշխատանքը անիմաստ էր: Այժմ մենք գիտենք, որ լավագույն հնարավոր արագությունը $n^{-2\beta/2\beta+1}$-ն է: Գիտենք նաև, որ բոլոր պարամետրերի դասում կա այդ արագությունն ունեցող գնահատական: Այս 2 դիտարկումները շատ կարևոր են օպտիմալ գնահատական կառուցելու համար: