EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl

Dr. Abhishek Chaudhary

Statistical Learning 2

Summer-Semester 2023 **Tübingen, 21.06.2023**

Homework 6

Problem 1. Let the n-sample $D_n = \{(\mathbf{X}_i, Y_i)\}_{i=1}^n$ be given. Consider a partition $\mathscr{P}_n = \{\mathscr{A}_{n,j}\}_{j\in\mathbb{N}}$ of $\mathbb{R}^d$, and denote by $\mathscr{A}_n(\mathbf{x}) \in \mathscr{P}_n$ the cell which contains $\mathbf{x} \in \mathbb{R}^d$. The well-known partitioning estimator is

$$m_n(\mathbf{x}) = \frac{\sum_{i=1}^n Y_i \cdot \mathbb{1}_{\{\mathbf{X}_i \in \mathscr{A}_n(\mathbf{x})\}}}{\sum_{i=1}^n \mathbb{1}_{\{\mathbf{X}_i \in \mathscr{A}_n(\mathbf{x})\}}} \qquad \forall \mathbf{x} \in \mathbb{R}^d .$$

Show that $\{m_n\}_n$ is a LS estimator, *i.e.*, it satisfies

$$m_n = \operatorname{argmin}_{f\in\mathcal{F}_n} \frac{1}{n}\sum_{i=1}^n \big|f(\mathbf{X}_i) - Y_i\big|^2 ,$$

where $\mathcal{F}_n$ ensembles $\mathscr{P}_n$-locally **constant** functions.

Problem 2. We expect improved convergence properties for regular regression functions $m : \mathbb{R}^d \to \mathbb{R}$ if the piecewise constant estimator m_n in **Problem 1.** is replaced by a piecewise polynomial estimator, *i.e.*,

$$\mathcal{F}_n = \Big\{ \sum_{j=1}^{K_n} p_j \mathbb{1}_{\{\mathscr{A}_{n,j}\}}; \quad p_j \equiv \mathscr{P}_\ell(\mathbb{R}^d) \Big\} ,$$

where $\mathscr{P}_\ell(\mathbb{R}^d)$ is the space of polynoms of degree $\ell \in \mathbb{N}_0$, and $K_n \equiv \#\{\mathscr{A}_{n,j};\ \mathscr{A}_{n,j} \in \mathscr{P}_n\}$.

Suppose $\mathbf{X} \in [0,1]$, and $\mathbb{E}[Y^2] < \infty$. Choose $\ell \in \mathbb{N}_0$. Show that the corresponding **LS estimator** $\{m_n\}_n$ is strongly consistent, once (for some $\delta > 0$)

$$K_n \uparrow \infty , \qquad \beta_n \uparrow \infty , \qquad \frac{K_n \beta_n^4 \log \beta_n}{n} \downarrow 0 , \qquad \frac{\beta_n^4}{n^{1-\delta}} \downarrow 0 , \qquad \max_{1\le j\le K_n} \operatorname{diam}(\mathscr{A}_{n,j}) \downarrow 0 \qquad (n \uparrow 0) .$$

Date of Submission: 12.00 on 28.06.2023.