EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Dr. Abhishek Chaudhary

Statistical Learning 2

Summer-Semester 2023 **Tübingen, 06.08.2023**

Homework 8

We know about two main assumptions **(Ass)**$_{1,2}$ which, if they hold for a general LS estimator, validate its universal strongly consistency. A specific example covered by this theorem is a *data-dependent partitioning estimator* $\{m_n\}_n$, for which we verified in the lecture the 'best approximation property' **(Ass)**$_1$, *i.e.*,

$$\inf_{f \in T_{\beta_n}\mathcal{G}_c \circ \mathscr{P}_n} \int_{\mathbb{R}^d} \big|f(\mathbf{x}) - m(\mathbf{x})\big|^2 \mu[\mathrm{d}\mathbf{x}] \downarrow 0 \quad (n \uparrow \infty) \qquad \mathbb{P}\text{-a.s.}$$

The aim of this homework sheet is to complete this program, by verifying **(Ass)**$_2$ for $\{m_n\}_n$, which is

$$\sup_{f \in T_{\beta_n}\mathcal{G}_c \circ \mathscr{P}_n} \Big|\frac{1}{n}\sum_{i=1}^n \big|f(\mathbf{X}_i) - Y_{i,L}\big|^2 - \mathbb{E}\big[|f(\mathbf{X}) - Y_L|^2\big]\Big| \downarrow 0 \quad (n \uparrow \infty) \qquad \mathbb{P}\text{-a.s.} \tag{1}$$

In the lecture, we already discussed relevant steps which are needed here to settle (1) at earlier places for different LS estimators.[1]

Problem 1. Use **Lemma 13.1.** in the book by [Györfi et al., p. 238] to prove (1).

Problem 2. Detail the proof for **Lemma 13.1.** given in the book by [Györfi et al., p. 238].

Date of Submission: 12.00 on 12.07.2023.

[1] One place where we used related arguments was when we when we showed (1) in the case of the linear LS series estimator.