EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-
Naturwissenschaftliche
Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Armin Beck

Statistical Learning 1

Summer semester 2024 **Tübingen, 24.06.2024**

Assignment 10

Problem 1
Let (X, Y) fulfill the following assumptions

(1) $\mathcal{S} := \operatorname{supp}(\mu)$ is bounded.

(2) $|Y| \leq L$ $\mathbb{P}$ a.s. for some $L > 0$.

(3) $\exists p \in (0,1],\ \exists C_{H\ddot{o}l} > 0 :\ |m(x) - m(z)| \leq C_{H\ddot{o}l}\|x - z\|^p$ für alle $x, z \in \mathcal{S}$.

Let m_n be the kernel estimate with naive kernel, where the bandwidth is chosen from the set

$$\mathcal{P}_n \equiv \mathcal{P}_n^{Ker} := \Big\{ 2^k :\ k \in \{-n, ...n\} \Big\}$$

using the Algorithm 4.1. Furthermore, let

$$n_L = \lfloor \frac{n}{2} \rfloor.$$

Then

$$\mathbb{E}\Big[\int_{\mathbb{R}^d} |m_n(x) - m(x)|^2 \mu[dx] \Big] = \mathcal{O}(n^{-\frac{2p}{2p+d}})$$

Problem 2
The proof of strong consistency of local averaging estimates uses the following special version of the Banach-Steinhaus theorem for integral operators in $L_1(\mathbb{R}^d; \mu)$.
Let $K_n(x, z)$ be functions on $\mathbb{R}^d \times \mathbb{R}^d$ satisfying the following conditions:

(i) There is a constant $c > 0$ such that, for all n,

$$\int_{\mathbb{R}^d} |K_n(x, z)| \mu[dx] \leq c$$

for $\mu-$almost all z.

(ii) There is a constant $D \geq 1$ such that

$$\int_{\mathbb{R}^d} |K_n(x, z)| \mu[dz] \leq D$$

for all x and n.

(iii) For all $a > 0$,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |K_n(x,z)| \mathbb{1}_{\{\|x-z\|>a\}} \mu[dz]\mu[dx] = 0.$$

(iv)

$$\lim_{n\to\infty} \operatorname{esssup}_{x\in\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} |K_n(x,z)|\mu(dz) - 1 \Big| = 0.$$

Then for all $m \in L_1(\mathbb{R}^d;\mu)$

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \Big| m(x) - K_n(x,z)m(z)\mu[dz] \Big| \mu[dx] = 0.$$

Date of Submission: 1.07.2024 in the mailbox at 12 noon.