EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-
Naturwissenschaftliche
Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Dr. Abhishek Chaudhary

Statistisches Lernen

Winter-Semester 2026/27 **Tübingen, 21.10.2026**

Homework 2

Problem 1. This exercise generalizes the problem from pattern recognition from the lecture. Suppose that $Y \in \{1,\dots,M\}$, where $M \geq 2$. The aim is to find a function $f^* : \mathbb{R}^d \to \{1,\dots,M\}$ such that

$$\mathbb{P}\big[f^*(\mathbf{X}) \neq Y\big] = \min_{f:\mathbb{R}^d \to \{1,\dots,M\}} \mathbb{P}\big[f(\mathbf{X}) \neq Y\big]\,. \tag{1}$$

(a) Show that (1) is met by the Bayes decision

$$f^*(\mathbf{x}) = \mathrm{argmax}_{1\leq k\leq M} \mathbb{P}\big[Y = k \big| \mathbf{X} = \mathbf{x}\big] \qquad \forall\, \mathbf{x} \in \mathbb{R}^d\,.$$

Remark: Let $\mathbf{x} \in \mathbb{R}^d$. The *a posteriori probabilities* are the functions

$$\mathbb{P}\big[Y = k \big| \mathbf{X} = \mathbf{x}\big] = \mathbb{E}\big[\mathbb{1}_{\{Y=k\}} \big| \mathbf{X} = \mathbf{x}\big] = m^{(k)}(\mathbf{x}) \qquad (1 \leq k \leq M)\,.$$

(b) Given a data set $D_n = \{(\mathbf{X}_j, Y_j)\}_{j=1}^n$, estimators $m_n^{(k)}$ of $m^{(k)}$ can be constructed from the data set

$$D_n^{(k)} = \big\{(\mathbf{X}_1, \mathbb{1}_{\{Y_1=k\}}), \dots, (\mathbf{X}_n, \mathbb{1}_{\{Y_n=k\}})\big\} \qquad (1 \leq k \leq M)\,,$$

and one can use the *plug-in* estimator

$$\forall\, \mathbf{x} \in \mathbb{R}^d : \qquad f_n(\mathbf{x}) = \mathrm{argmax}_{1\leq k\leq M} m_n^{(k)}(\mathbf{x}) \qquad (1 \leq k \leq M)$$

to estimate f^*. Show that the error probability of the *plug-in* estimators satisfies

$$\begin{aligned} 0 \;&\leq\; \mathbb{P}\big[f_n(\mathbf{X}) \neq Y \big| D_n\big] - \mathbb{P}\big[f^*(\mathbf{X}) \neq Y\big] \\ &\leq\; \sum_{k=1}^M \int_{\mathbb{R}^d} \big|m_n^{(k)}(\mathbf{x}) - m^{(k)}(\mathbf{x})\big|\, \mu[\mathrm{d}\mathbf{x}] \\ &\leq\; \sum_{k=1}^M \Big(\int_{\mathbb{R}^d} \big|m_n^{(k)}(\mathbf{x}) - m^{(k)}(\mathbf{x})\big|^2\, \mu[\mathrm{d}\mathbf{x}]\Big)^{1/2}\,. \end{aligned}$$

Problem 2. Let $n \in \mathbb{N}$. For $z_1, \dots, z_n \in \mathbb{R}$ we consider $\overline{z} := \frac{1}{n}\sum_{i=1}^n z_i$.

(a) Show that

$$\frac{1}{n}\sum_{i=1}^n |c - z_i|^2 = |c - \overline{z}|^2 + \frac{1}{n}\sum_{i=1}^n |\overline{z} - z_i|^2 \qquad \forall\, c \in \mathbb{R}\,.$$

(b) Conclude from (a) that

$$\frac{1}{n}\sum_{i=1}^{n}|\overline{z}-z_i|^2=\min_{c\in\mathbb{R}}\frac{1}{n}\sum_{i=1}^{n}|c-z_i|^2\,.$$

Problem 3. Let the data set $D_n=\{(\mathbf{X}_i,Y_i)\}_{i=1}^n$ be given. Show that the kernel estimator defined by

$$m_n(\mathbf{x})=\frac{\sum_{i=1}^n K(\frac{\mathbf{x}-\mathbf{X}_i}{h})Y_i}{\sum_{i=1}^n K(\frac{\mathbf{x}-\mathbf{X}_i}{h})}\,,$$

with $K(\mathbf{x})=\mathbb{1}_{\{\|\mathbf{x}\|\leq 1\}}$ the naive (or: window) kernel satisfies

$$m_n(\mathbf{x})=\arg\min_{c\in\mathbb{R}}\frac{1}{n}\sum_{i=1}^{n}K\big(\frac{\mathbf{x}-\mathbf{X}_i}{h}\big)|Y_i-c|^2\,.$$

Hint: Show that for any $c\in\mathbb{R}$,

$$\sum_{i=1}^{n}K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h}\Big)(Y_i-c)^2=\sum_{i=1}^{n}K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h}\Big)\big(Y_i-m_n(\mathbf{x})\big)^2+\sum_{i=1}^{n}K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h}\Big)\big(m_n(\mathbf{x})-c\big)^2\,.$$

Problem 4. Let $\mathscr{P}=\bigcup_{j\in\mathbb{N}}\mathscr{A}_j$ be a partitioning of $\mathbb{R}^d$, and $D_n=\{(\mathbf{X}_i,Y_i)\}_{i=1}^n$ a data set. Show that the partitioning estimate defined by

$$m_n(\mathbf{x})=\frac{\sum_{i=1}^n\mathbb{1}_{\{\mathbf{X}_i\in\mathscr{A}_j\}}Y_i}{\sum_{i=1}^n\mathbb{1}_{\{\mathbf{X}_i\in\mathscr{A}_j\}}}\qquad\forall\,\mathbf{x}\in\mathscr{A}_j \tag{2}$$

satisfies

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

where $\mathcal{F}_{\mathscr{P}}:=\big\{\sum_{j\in\mathbb{N}}a_j\mathbb{1}_{\mathscr{A}_j}:\ a_j\in\mathbb{R}\big\}$ denotes the space of all piecewise constant functions on $\mathscr{P}$.

Hint: Let m_h be the partitioning estimate in (2). Show by the aid of **Problem 1** that

$$\sum_{i=1}^{n}\big|f(\mathbf{X}_i)-Y_i\big|^2=\sum_{i=1}^{n}\big|f(\mathbf{X}_i)-m_n(\mathbf{X}_i)\big|^2+\sum_{i=1}^{n}\big|m_n(\mathbf{X}_i)-Y_i\big|^2\qquad\forall\,f\in\mathcal{F}_{\mathscr{P}}\,.$$

Date of Submission: 12.00 on 28.10.2026.