Dr. Stefan Richter
Statistische Analyse von maschinellen Lernalgorithmen
Wintersemester 2020/2021

3. Übungsblatt

Aufgabe 1 (Diskussion: Übertragung der Beweistechnik von Satz 3.12).
In dieser Aufgabe wollen wir die Beweistechnik von Satz 3.12 auf den KQ-Schätzer im linearen Modell mit deterministischen Design anwenden, d.h. $X_1, ..., X_n$ sind hier deterministisch und $\hat{\Sigma} = \frac{1}{n}\mathbb{X}^T\mathbb{X} = \Sigma$. Außerdem seien $\varepsilon_i \sim N(0, \sigma^2)$.
Erinnerung: $\hat{\beta} = \mathrm{argmin}_{\beta \in \mathbb{R}^d} \hat{R}_n(\beta)$ mit $\hat{R}_n(\beta) = \frac{1}{n}\|\mathbb{Y} - \mathbb{X}\beta\|_2^2$, $R(\beta) = \|\Sigma^{1/2}(\beta - \beta^*)\|_2^2 + \sigma^2$.

(a) Zeigen Sie:
$$R(\hat{\beta}) - R(\beta^*) \leq \big|(\hat{R}_n(\hat{\beta}) - R(\hat{\beta})) - (\hat{R}_n(\beta^*) - R(\beta^*))\big|.$$

(b) Sei $\gamma > 0$, $T := \frac{\gamma}{\gamma + \|\Sigma^{1/2}(\hat{\beta} - \beta^*)\|_2}$ und $\tilde{\beta} := T\hat{\beta} + (1-T)\beta^*$. Zeigen Sie, dass $\|\Sigma^{1/2}(\tilde{\beta} - \beta^*)\|_2 \leq \gamma$ und
$$R(\tilde{\beta}) - R(\beta^*) \leq \sup_{\beta: \|\Sigma^{1/2}(\beta - \beta^*)\|_2 \leq \gamma} \big|(\hat{R}_n(\beta) - R(\beta)) - (\hat{R}_n(\beta^*) - R(\beta^*))\big| =: Z_\gamma$$

(c) Sei $a > 0$. Zeigen Sie: Gilt $(a\gamma)^{1/2} \leq \frac{\gamma}{2}$, so gilt auf $A = \{Z_\gamma \leq a \cdot \gamma\}$
$$\|\Sigma^{1/2}(\hat{\beta} - \beta)\|_2 \leq \gamma,$$
und dann wiederum $R(\hat{\beta}) - R(\beta^*) \leq a \cdot \gamma$.

(d) Folgern Sie:
$$\mathbb{P}(R(\hat{\beta}) - R(\beta^*) > a \cdot \gamma) \leq \mathbb{P}(A^c).$$

Ab jetzt geht es um die Wahl von a, γ.

(e) Zeigen Sie, dass
$$(\hat{R}_n(\beta) - R(\beta)) - (\hat{R}_n(\beta^*) - R(\beta^*)) = -\frac{2}{n}\mathbb{e}^T\mathbb{X}(\beta - \beta^*)$$
und weiter $|Z_\gamma| \leq \frac{2\gamma}{n}\|\mathbb{e}^T\mathbb{X}\Sigma^{-1/2}\|_2$.
Hinweis: Fügen Sie $\Sigma^{-1/2}\Sigma$ *ein und nutzen Sie* $|v^T w| \leq \|v\|_2\|w\|_2$.

(f) Zeigen Sie: $\mathbb{E}|Z_\gamma| \leq 2\gamma\sigma(\frac{d}{n})^{1/2}$.

(g) Wählen Sie a, γ geeignet, um folgende Aussagen zu zeigen:
$$\mathbb{P}(R(\hat{\beta}) - R(\beta^*) \geq 16\sigma^2\frac{d}{n} \cdot t^2) \leq \frac{1}{t}.$$

Nun ermitteln wir eine bessere Abschätzung für $\mathbb{P}(Z_\gamma > a\gamma)$.

(h) Zeigen Sie, dass $\frac{1}{\sigma\sqrt{n}}\|\mathbb{e}^T\mathbb{X}\Sigma^{-1/2}\|_2 \overset{d}{=} \|W\|_2$ mit $W \sim N(0, I_{d\times d})$.

(i) Zeigen Sie, dass

$$\mathbb{P}(\|W\|_2^2 - 2\mathbb{E}\|W\|_2^2 \geq t) \leq e^{-t/4}.$$

Hinweis: Berechnen Sie $\mathbb{E}\|W\|_2^2$, schieben Sie es auf die rechte Seite und wenden Sie dann Markov-Ungleichung mit $g(x) = e^{-x/4}$ an. Für $S \sim N(0,1)$ gilt $\mathbb{E}[\exp(\frac{S^2}{4})] = \sqrt{2}$, und es ist $\sqrt{2}e^{-1/2} \leq 1$.

(j) Zeigen Sie mit der Abschätzung von Z_γ aus (e), dass mit $a = \frac{\sigma}{\sqrt{n}}\sqrt{2d+t}$ gilt:

$$\mathbb{P}(A^c) = \mathbb{P}(Z_\gamma > a\gamma) \leq e^{-t}.$$

(k) Folgern Sie wie in (g):

$$\mathbb{P}\Big(R(\hat{\beta}) - R(\beta^*) \geq 4\frac{\sigma^2}{n} \cdot (2d+t)\Big) \leq e^{-t}$$

Homepage der Vorlesung:
`https://ssp.math.uni-heidelberg.de/sam-ws2020/`