Agenda 8.12.:

- Terminänderungen: Achtung, verlängerte Fristen!

  → Bearbeitung EA 5 11.12.-7.1. — Besprechung EA5 12.1.

  EA 6 8.1. - 21.1. — EA6 26.1.

  EA 7 21.1. - 4.2. — EA7 9.2.

  (wie im OÜS)

  → Aufgabenstellungen werden vorher schon hochgeladen (zB EA 6 ist jetzt schon online)

- Übungstermine: ~~15.12.23~~ 12.1.24

  26.1.24

  ~~16.2.24~~ 9.2.24

- Videoliste am Ende des Semesters (→ Forum)

  → gerne noch Vorschläge machen
  - spezielle Themen / Aufgaben
  - Kurseinheiten / Vorlesungen
  - Anhang / Vorbereitung auf die Vorlesung

- Besprechung EA 4
  - Aufgabe 4 (Orth. in hohen Dim.)
  - nach Bedarf auch andere Aufgaben

- Tipps EA5 (und ggf. EA 6)

EA 4, A4

4. Orthogonalität in hohen Dimensionen (10 Punkte). In dieser Aufgabe ist unser Ziel, zu zeigen, dass zufällige Vektoren in hohen Dimensionen mit großer Wahrscheinlichkeit orthogonal sind. Wir betrachten zwei Vektoren $X, Y \in B^d$, die unabhängig und jeweils gleichverteilt auf der Kugel B^d sind. Wir betrachten $d \geq 3$.

a) (2 Punkte) Es seien $\alpha, \beta \in (0,1)$. Beweisen Sie die Ungleichung

$$\mathbb{P}(|\langle X,Y\rangle| < \alpha) \geq \mathbb{P}(|\langle X,Y\rangle| < \alpha \mid \|Y\| > \beta)\cdot \mathbb{P}(\|Y\| > \beta).$$

b) (2 Punkte) Folgern Sie:

$$\mathbb{P}(|\langle X,Y\rangle| < \alpha) \geq \mathbb{P}(|\langle X,\tfrac{Y}{\|Y\|}\rangle| < \alpha \mid \|Y\| > \beta)\cdot \mathbb{P}(\|Y\| > \beta).$$

c) (3 Punkte) Nutzen Sie Theorem 4.3.3 um zu zeigen, dass für $c \geq 1$ und $\alpha = \frac{c}{\sqrt{d-1}}$ die Ungleichung

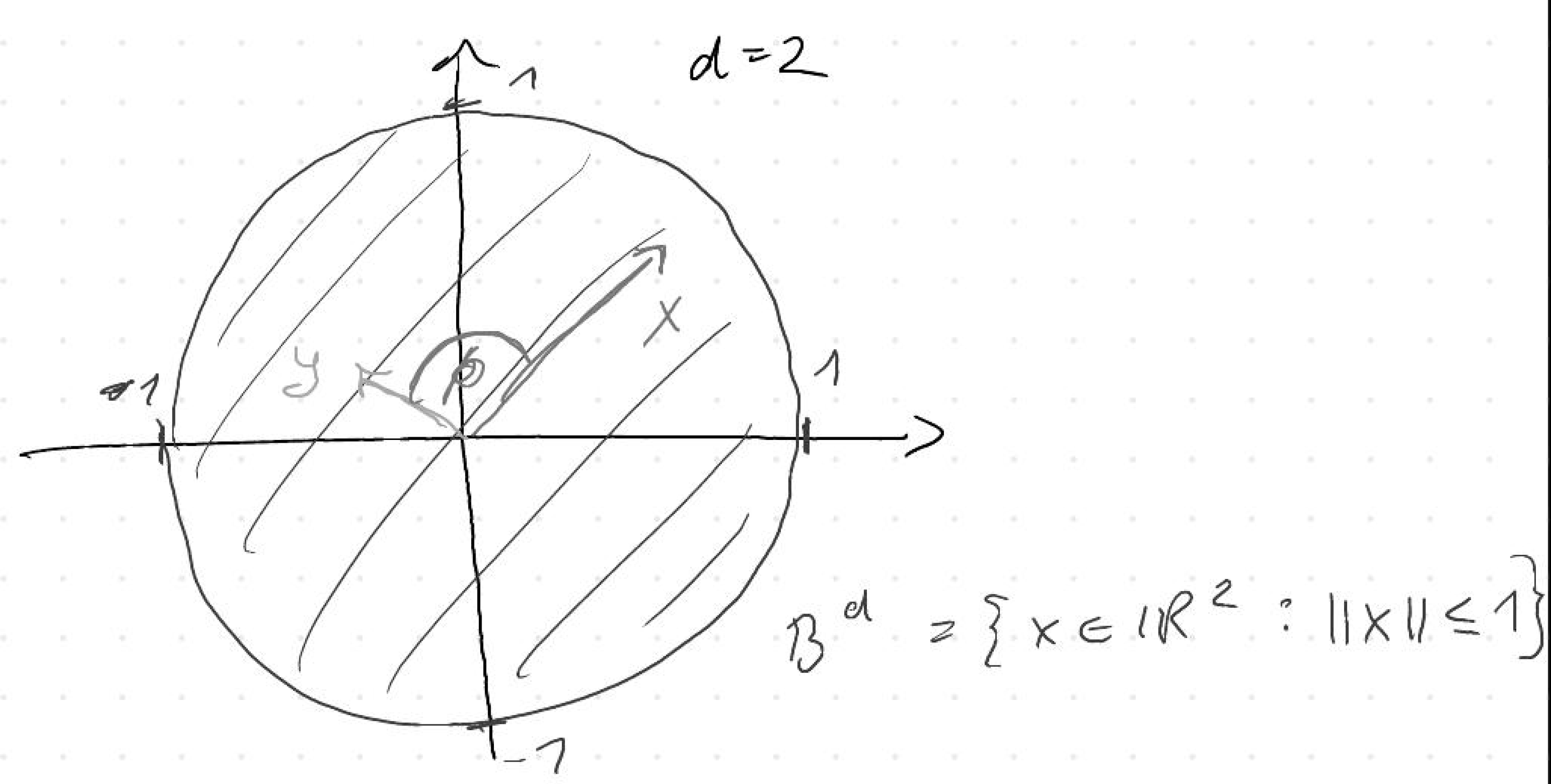
$$\mathbb{P}(|\langle X,\tfrac{Y}{\|Y\|}\rangle| < \alpha \mid \|Y\| > \beta) \geq 1 - \frac{2}{c}e^{-\frac{c^2}{2}}$$

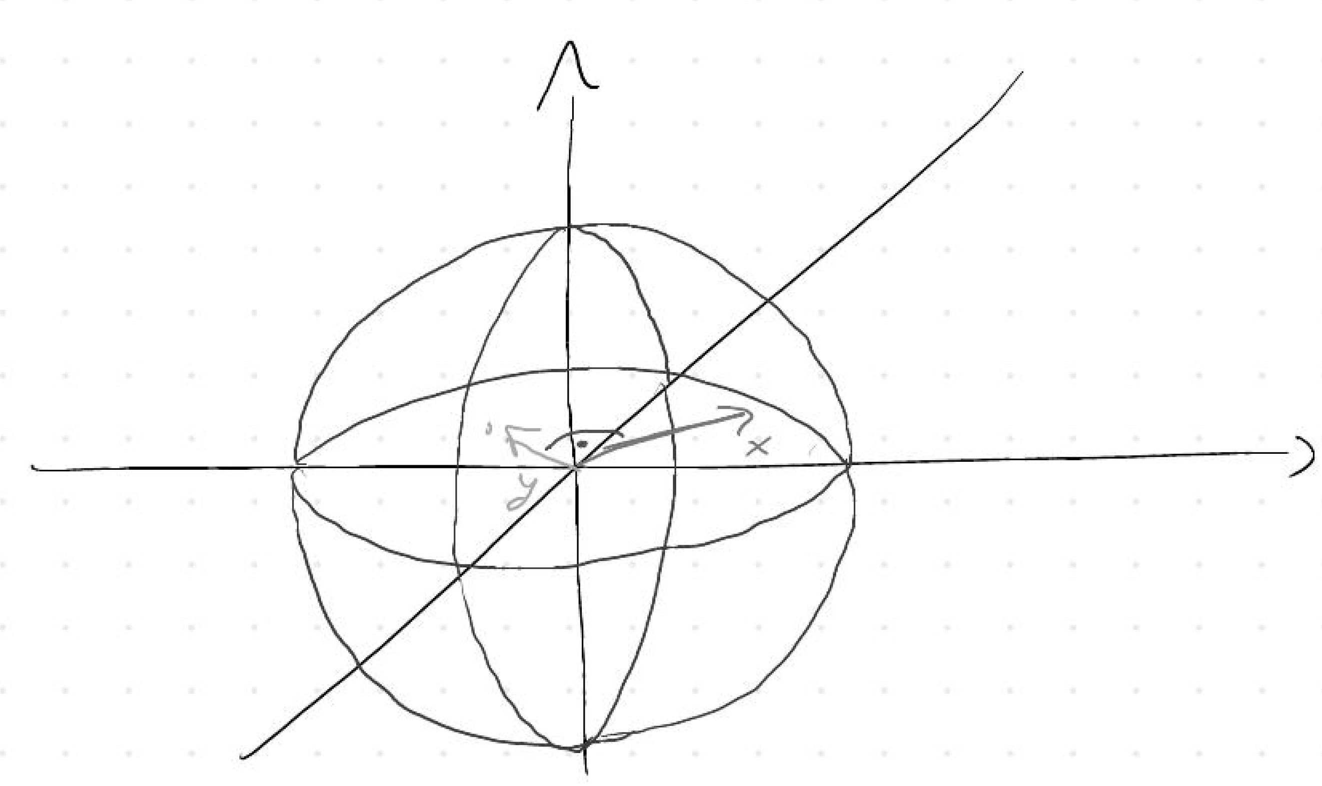
gilt.

d) (3 Punkte) Schließen Sie, dass

$$\mathbb{P}\left(|\langle X,Y\rangle| < \frac{c}{\sqrt{d-1}}\right) \geq 1 - \frac{2}{c}e^{-\frac{c^2}{2}}$$

gilt. Warum erreichen wir damit unser oben formuliertes Ziel?
Hinweis: $\mathbb{P}(X \in K_{z,\alpha}) = \mathrm{vol}(K_{z,\alpha})/\mathrm{vol}(B^d)$ für $K_{z,\alpha} \subset B^d$, wobei $K_{z,\alpha} = \{x \in B^d : |\langle x,z\rangle| < \alpha\}$, $\|z\| = 1$, $z \in B^d$.





a) $\alpha, \beta \in (0,1)$

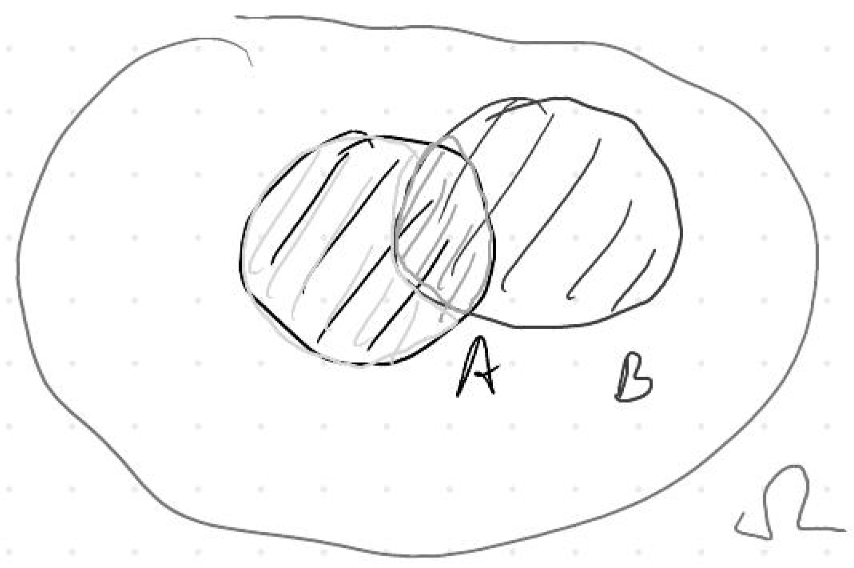
Allgemein: Seien A und B zwei Ereignisse

$$\mathbb{P}[A] = \mathbb{P}[A \cap B] + \mathbb{P}[A \cap B^c] = \mathbb{P}[A|B]\,\mathbb{P}[B] + \underbrace{\mathbb{P}[A|B^c]\,\mathbb{P}[B^c]}_{\geq 0}$$

A

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

(*) $\rightsquigarrow \mathbb{P}[A \cap B] = \mathbb{P}[A|B] \cdot \mathbb{P}[B]$



$$\geq \mathbb{P}[A|B]\,\mathbb{P}[B]$$

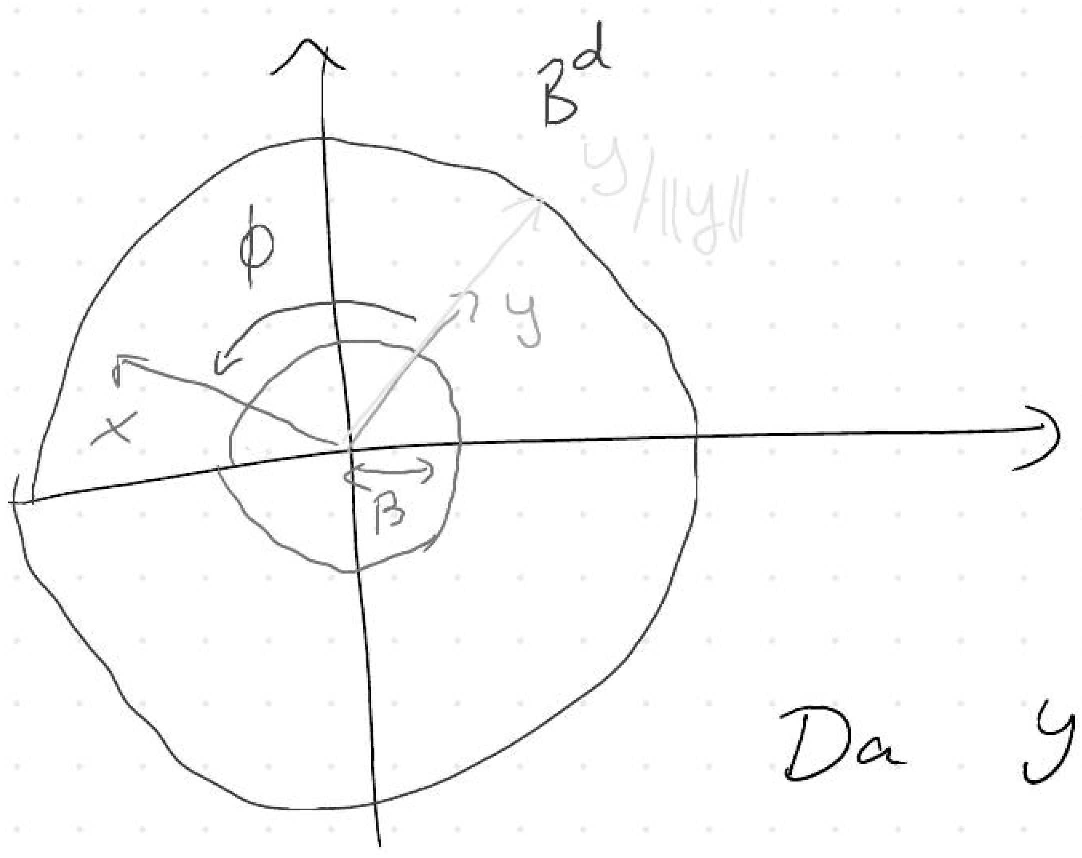
Setze ein: $\underline{A} = \{|\langle x,y\rangle| < \alpha\}$ $\underline{B} = \{\|y\| > \beta\}$

$$\rightarrow \underline{\mathbb{P}[\,|\langle x,y\rangle| < \alpha \mid \|y\| > \beta] \cdot \mathbb{P}[\|y\| > \beta] \leq \mathbb{P}[\,|\langle x,y\rangle| < \alpha]}. \checkmark$$

(*) Wir haben gezeigt

$$\mathbb{P}[A] \geq \mathbb{P}[A|B] \cdot \mathbb{P}[B] \qquad A, B \text{ wie oben wählen}$$

$$\underline{\mathbb{P}[\,|\langle x,y\rangle| < \alpha\,] \geq \mathbb{P}[\,|\langle x,y\rangle| < \alpha \mid \|y\| > \beta] \cdot \mathbb{P}[\|y\| > \beta]}$$



b) $\underline{\mathbb{P}[\,|\langle x,y\rangle| < \alpha\,]} \geq \underline{\mathbb{P}[\,|\langle x, \tfrac{y}{\|y\|}\rangle| < \alpha \mid \|y\| > \beta] \cdot \mathbb{P}[\|y\| > \beta]}$

Es gilt $|\langle x,y\rangle| < \alpha \quad \Big| \cdot \frac{1}{\|y\|}$

$$|\langle x, \tfrac{y}{\|y\|}\rangle| < \frac{\alpha}{\|y\|}$$

Da $y \in B_d \Rightarrow \|y\| \leq 1$, also $\frac{1}{\|y\|} \geq 1$ und $\frac{\alpha}{\|y\|} \geq \alpha$.

Daher

$$\mathbb{P}[\,\underline{|\langle x,y\rangle| < \alpha} \mid \|y\| > \beta] = \mathbb{P}[\,|\langle x, \tfrac{y}{\|y\|}\rangle| < \underbrace{\frac{\alpha}{\|y\|}}_{\geq \alpha} \mid \|y\| > \beta]$$

$$\geq \mathbb{P}[\,|\langle x, \tfrac{y}{\|y\|}\rangle| < \underline{\alpha} \mid \|y\| > \beta]$$

c) Zzg. $\mathbb{P}\left[\left|\left\langle X, \frac{y}{\|y\|} \right\rangle\right| < \alpha \;\middle|\; \|y\| > \beta \right] \geq 1 - \frac{2}{c} e^{-c^2/2}$

für $c \geq 1$ und $\alpha = \frac{c}{\sqrt{d-1}}$.

Wir wissen $\frac{y}{\|y\|} = 1$ (s.o.). Wir setzen α oben ein & erhalten

$$\mathbb{P}\left[\left|\left\langle X, \frac{y}{\|y\|} \right\rangle\right| < \frac{c}{\sqrt{d-1}} \;\middle|\; \|y\| > \beta \right] \geq \inf_{z: \|z\|=1} \mathbb{P}\left[|\langle X, z \rangle| < \frac{c}{\sqrt{d-1}} \right] = *$$

Idee: links ist eine Eigenschaft erfüllt ($\|z\| = 1$ für $z = \frac{y}{\|y\|}$) $\rightsquigarrow$ größer gleich dem kleinsten Term, sodass diese Eigenschaft erfüllt ist

Jetzt: Hinweis! $K_{z,\alpha} := \{x \in B^d : |\langle x, z \rangle| < \alpha\}$ für $\|z\| = 1$, $z \in B^d$.

d.h. $|\langle X, z \rangle| < \frac{c}{\sqrt{d-1}} \iff X \in K_{z, \frac{c}{\sqrt{d-1}}}$. Aus dem Hinweis folgt

$\mathbb{P}[X \in K_{z,\alpha}] = \frac{\mathrm{vol}(K_{z,\alpha})}{\mathrm{vol}(B^d)}$ (X gleichverteilt auf B^d)

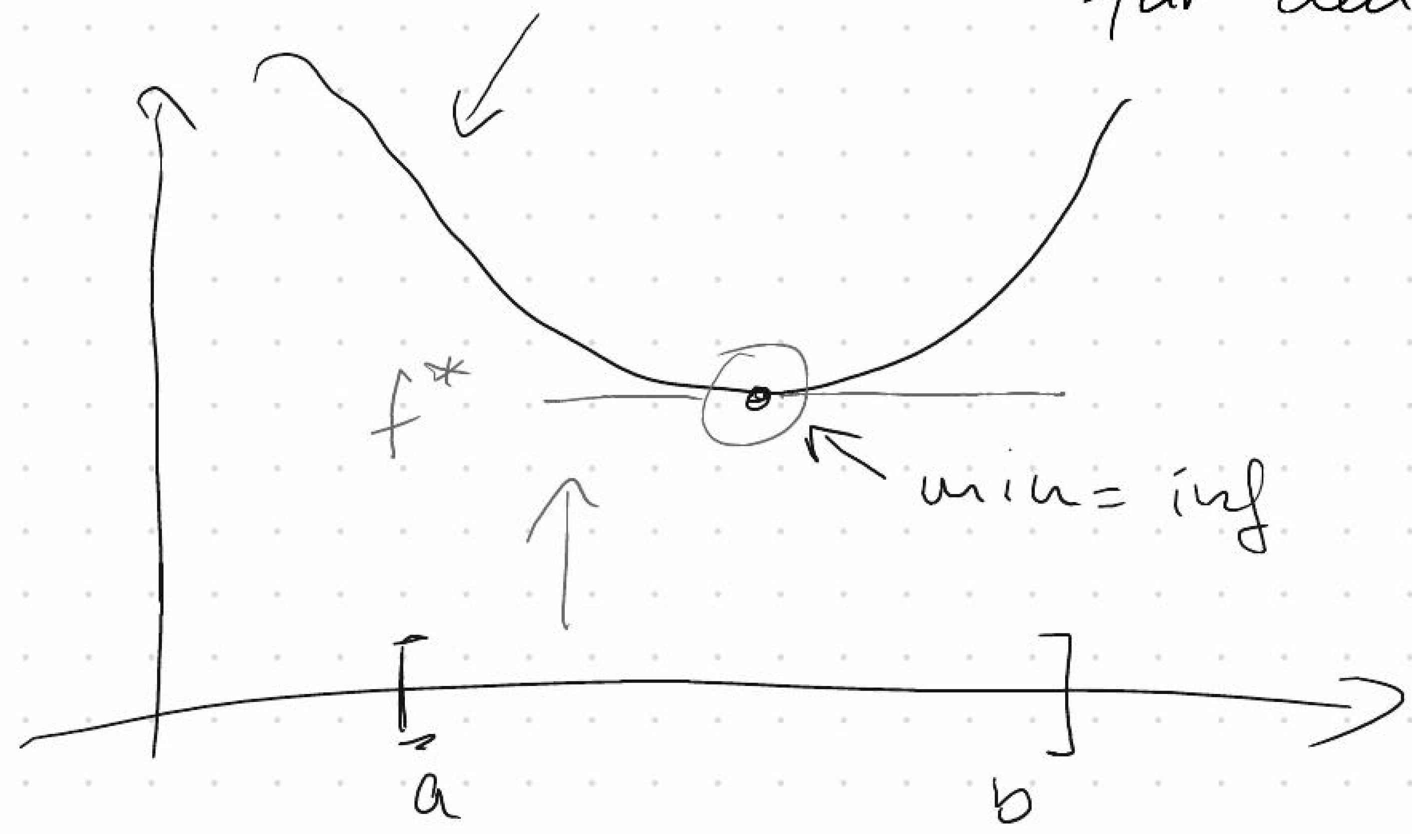
$f(\tilde{z})$ wissen $\|\tilde{z}\| = 1$

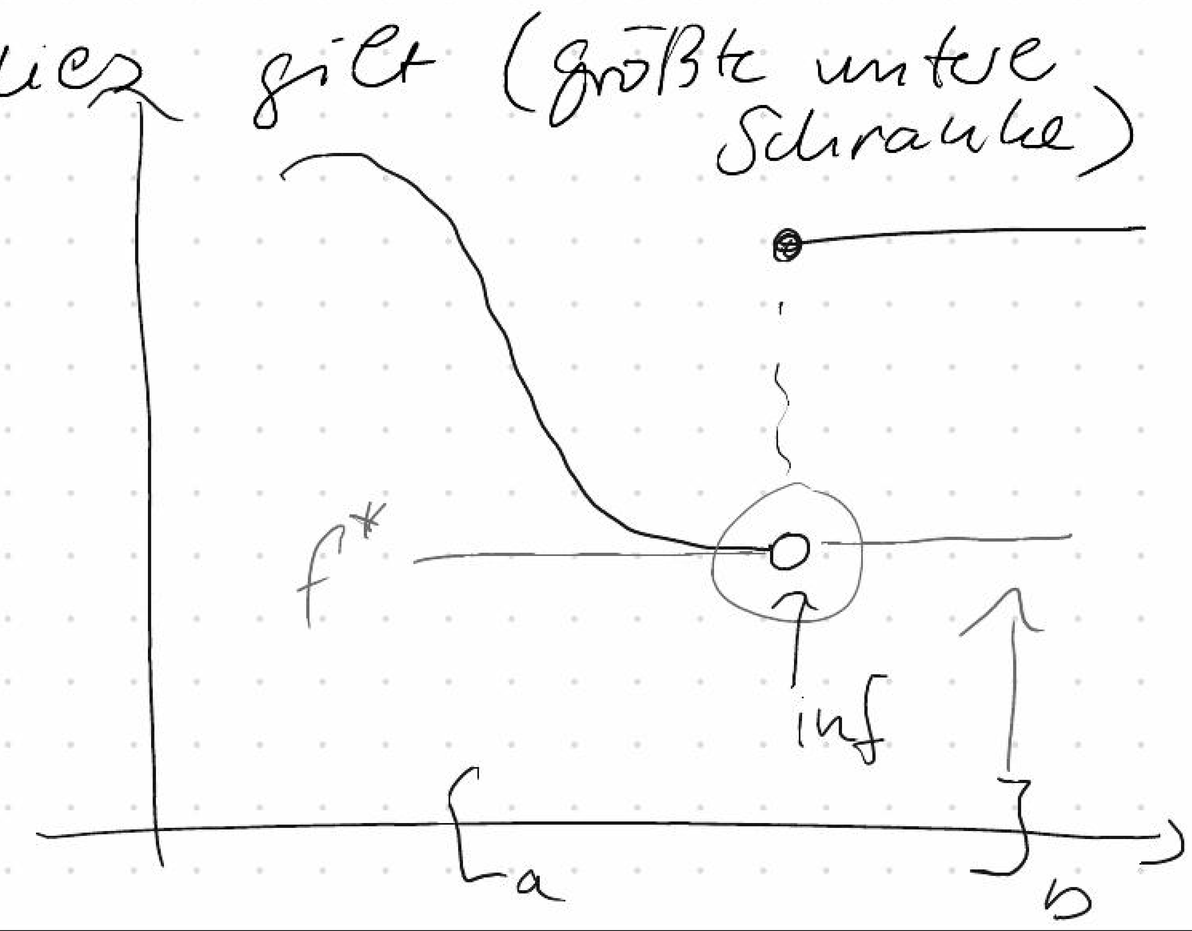
$$\Rightarrow \quad f(\tilde{z}) \geq \inf_{z: \|z\|=1} f(z)$$

Def: $\inf_{z: \|z\|=1} f(z) = f^*$ ist, dass für alle z mit $\|z\| = 1$

$f^* \leq f(z)$ & f^* ist der maximale Wert für den dies gilt (größte untere Schranke)

min $\hookrightarrow$ inf





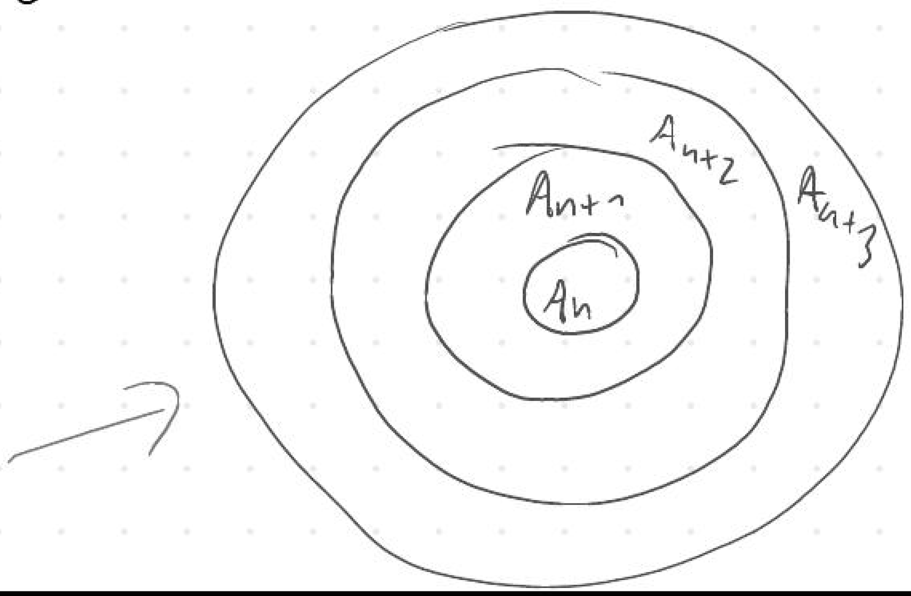
$$* = \inf_{z: \|z\|=1} \mathbb{P}\left[X \in K_{z, \frac{c}{\sqrt{d-1}}} \right] \overset{\text{Hinweis}}{=} \inf_{z: \|z\|=1} \frac{\mathrm{vol}\left(K_{z, \frac{c}{\sqrt{d-1}}}\right)}{\mathrm{vol}(B^d)} \overset{\text{Thm 4.3.3}}{\geq} 1 - \frac{2}{c} e^{-c^2/2}.$$

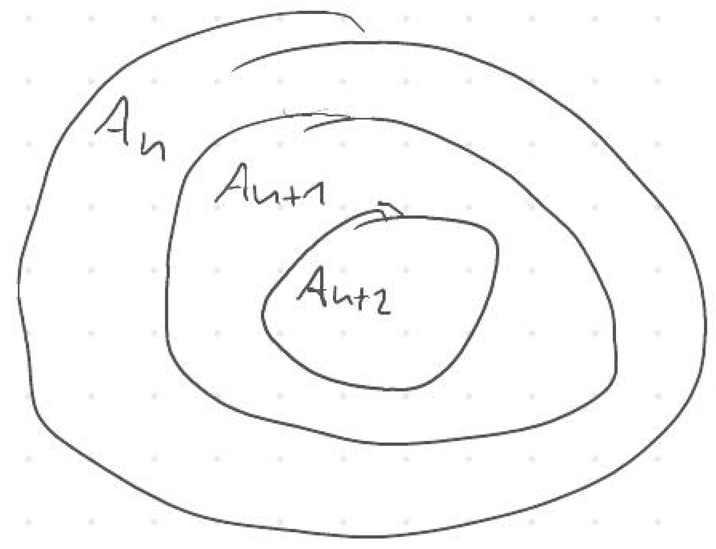
d) Von oben

$$\mathbb{P}[\,|\langle X, y \rangle| < \alpha\,] \geq \left(1 - \frac{2}{c} e^{-c^2/2}\right) \cdot \mathbb{P}[\|y\| > \beta] \quad \text{für } \beta \in (0,1).$$

Lassen wir $\beta \to 0$, erhalten wir

$\lim_{\beta \to 0} \mathbb{P}[\|y\| > \beta] = \mathbb{P}[\|y\| > 0] = 1$ (Stetigkeit von unten, Monotone Konvergenz)





Lemma C.1.10

$\lim_{n \to \infty} \mathbb{P}[A_n] = \mathbb{P}[\lim_{n \to \infty} A_n]$

$\to$ ggf. Übung, nachsehen im Skript + nachrechnen

$\bigcup_{n \geq 1} \bigcap_{k \geq n} A_k$

$$\Rightarrow \quad \mathbb{P}\left[|\langle x, y\rangle| < \frac{c}{\sqrt{d-1}} \right] \geq \underbrace{1 - \frac{2}{c} e^{-c^2/2}}_{\approx 1} \qquad c \geq 1$$

$\sqrt{d-1} \to \infty$

$\underline{\to 0}$

Tipps EA 5

1. Grundbegriffe (10 Punkte). Ein Call-Center möchte die minimale und maximale Anrufdauer schätzen. Wir nehmen dabei an, dass die Anrufzeiten unabhängig voneinander sind und dass Dauer jedes Anrufs gleichverteilt ist mit einer „natürlichen" Mindestdauer $a \geq 0$ und maximalen Dauer $b > a$ (zum Beispiel wenn es keine Frage gibt, die in unter a Minuten gelöst werden kann und keine Kund:innen, die länger als b Minuten Zeit haben). Das Call-Center beobachtet nun 1000 Gespräche, von denen das kürzeste 5 Minuten und das längste 137 Minuten gedauert hat. Daraus schließt das Call-Center, dass 5 und 137 Minuten die Mindest- und die Höchstdauer sind.

a) (4 Punkte) Geben Sie ein statistisches Modell (gemäß der Definition) an, das zur obigen Situation passt.

b) (2 Punkte) Handelt es sich um ein stetiges oder diskretes Modell?

c) (2 Punkte) Schreiben Sie in der Notation der Vorlesung auf, welcher Schätzer hier verwendet wurde.

d) (2 Punkte) Geben Sie eine Abbildung τ an, die jedem $\vartheta \in \Theta$ in ihrem statistischen Modell die Kenngröße „Erwartungswert der Anrufdauer" zuordnet. Geben Sie einen Schätzer für τ an.

- Def. ganz genau lesen
- σ-Alg. darf gemäß Bem 5.2.3 weggelassen werden
- Vorlesung:
  - Schätzer für $\vartheta \in \Theta$
  - Schätzer für Kennzahlen τ ($\hat{=}$ Funktionen von ϑ)

2. Erwartungstreue in Poisson-Modellen (10 Punkte). Wir betrachten als statistisches Modell ein Poisson-Modell $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0), \mathbb{P}_\theta)$ mit $\mathbb{P}_\theta(\{n\}) = \frac{\theta^n}{n! e^\theta}$ für alle $n \in \mathbb{N}_0$, also die Poisson-Verteilung.

a) (3 Punkte) Zeigen Sie, dass $T(x) = x$ ein erwartungstreuer Schätzer für θ ist.

b) (3 Punkte) Zeigen Sie, dass $T(x) = x^2 - x$ ein erwartungstreuer Schätzer für $\tau(\theta) = \theta^2$ ist.

Wir definieren $f : (0, \infty) \times \mathbb{N} \to (0, \infty)$ durch

$$f(x, n) = \frac{x^n}{n!(e^x - 1)}.$$

c) (2 Punkte) Es sei X eine Poisson-verteilte Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', \mathbb{P}')$ mit Parameter $\lambda > 0$. Zeigen Sie, dass $\mathbb{P}'[X = n \mid X > 0] = f(\lambda, n)$ für alle $n \in \mathbb{N}$ gilt.

Insbesondere ist für jedes $\lambda > 0$ durch $f(\lambda, n)$ eine Wahrscheinlichkeitsverteilung auf $\mathbb{N}$ (ohne Null) definiert. Wir betrachten nun als statistisches Modell ein bedingtes Poisson-Modell $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \tilde{\mathbb{P}}_\theta)$ mit $\tilde{\mathbb{P}}_\theta(\{n\}) = f(\theta, n)$.

d) (2 Punkte) Zeigen Sie, dass $T(x) = 1 + (-1)^x$ ein erwartungstreuer Schätzer für $\tau(\theta) = 1 - e^{-\theta}$ ist. Ist dies eine sinnvolle Wahl?

Tipps:

- $\underbrace{\sum_{k=0}^{\infty} \frac{x^k}{k!}}_{} = \underbrace{e^x}_{}$
- Teilaufg. a, b benutzen den gleichen Trick
- c) $\mathbb{P}[X > 0] = 1 - \mathbb{P}[X = 0]$

3. Maximum-Likelihood-Schätzer und Konfidenzintervall (10 Punkte). Wir betrachten Daten $X_1, \ldots, X_n$, die unabhängig und jeweils binomialverteilt mit Parametern (m, p) sind. Der Parameter $m \in \mathbb{N}$ sei bekannt und $p \in (0,1)$ sei unbekannt.

a) (5 Punkte) Zeigen Sie, dass $T(x) = \frac{1}{nm}\sum_{\ell=1}^{n} x_\ell$ der Maximum-Likelihood-Schätzer für p ist.

b) (2 Punkte) Sei $\alpha \in (0,1)$ und bezeichne Φ die Verteilungsfunktion der Standard-Normalverteilung. Zeigen Sie mit dem Zentralen Grenzwertsatz, dass für $\epsilon : (0,1) \to (0,\infty)$ mit

$$\epsilon(p) = \frac{\Phi^{-1}(1-\frac{\alpha}{2})\sqrt{p(1-p)}}{\sqrt{nm}} = \sqrt{\beta}\sqrt{p(1-p)}, \qquad \beta = \frac{[\Phi^{-1}(1-\frac{\alpha}{2})]^2}{nm},$$

approximativ gilt:

$$\mathbb{P}_p\left(\left|\frac{\sum_{l=1}^{n} X_l}{nm} - p\right| < \epsilon(p)\right) \geq 1-\alpha$$

c) (3 Punkte) Zeigen Sie, dass $|T(X) - p| < \sqrt{p(1-p)}\sqrt{\beta} \Leftrightarrow p \in (p_-(X), p_+(X))$ für

$$p_\pm(X) = \frac{(2T(X)+\beta)}{2(1+\beta)} \pm \frac{\sqrt{\beta^2 + 4T(X)[1-T(X)]\beta}}{2(1+\beta)}$$

gilt und konstruieren Sie ein (approximatives) Konfidenzintervall zum Niveau α.

1) log-likelihood Fkt.

2) Ähnliches Bsp in Vorl.
→ Normalvt.
→ hier: Approximation ZGWS

3) Zwischenergebnisse!
→ Aufgaben unabh. voneinander lösen

4) Ein sinnvoller Ansatz für Konf. Int.

$\hat{\vartheta} = \frac{1}{n}\sum_{i=1}^{n} X_i$ $\qquad \theta ? \rightsquigarrow \vartheta$ $\qquad I(X)$

$\mathrm{Poi}(\vartheta)$-vt: $E_\vartheta[X] = \vartheta$

$\mathbb{P}_p[p \in \underline{I(X)}] = \mathbb{P}_p[\underline{|T(X) - p| < \sqrt{p(1-p)}\sqrt{\beta}}]$

4. Generalised Linear Models & ML-Schätzer (10 Punkte). Wir betrachten folgende Situation. Wir wollen ein Modell erarbeiten, welches die Wahrscheinlichkeit vorhersagen kann, dass ein Gebäude bei einem Starkregenereignis stark beschädigt wird. Wir vermuten, dass diese Wahrscheinlichkeit vom Alter A und der Stärke des Unwetters S abhängt. Wir modellieren dies, indem wir annehmen, dass $Y_i \in \{0,1\}$ beschreibt, ob Gebäude i beschädigt wird (dabei steht 0 für 'kein Schaden' und 1 für 'Schaden'). Weiter nehmen wir an, dass $A_i \in \{0,1,2,3,4,5\}$ die Alterskategorie des Hauses und $S_i \in \{1,2,3,4,5\}$ die Stärke des jeweiligen Unwetters bezeichnet. Wir nehmen an, die Y_i sind unabhängig voneinander mit $Y_i \sim B(1, p_i)$, wobei

$$p_i = \frac{e^{\beta_0 + \beta_1 A_i + \beta_2 S_i}}{1 + e^{\beta_0 + \beta_1 A_i + \beta_2 S_i}},$$

bzw. $\ln(p_i/(1-p_i)) = \beta_0 + \beta_1 A_i + \beta_2 S_i$. Hier ist $\beta = (\beta_0, \beta_1, \beta_2) \in \mathbb{R}^3$ ein Vektor, der nicht von i abhängt. Wir nehmen an, wir haben n Datensätze bzw. Stichproben Y_i, A_i, S_i, $i = 1, \ldots, n$, zur Verfügung.

a) (4 Punkte) Beschreiben Sie, wie die Idee der Maximum-Likelihood-Methode genutzt werden kann, um β zu bestimmen und zeigen Sie, dass

$$\max_{\beta \in \mathbb{R}^3}\left(\sum_{i:y_i=1}(\beta_0 + \beta_1 a_i + \beta_2 s_i) - \log(1 + e^{\beta_0+\beta_1 a_i+\beta_2 s_i}) - \sum_{i:y_i=0}\log(1 + e^{\beta_0+\beta_1 a_i+\beta_2 s_i})\right)$$

ein sinnvoller Ansatz für das zugehörige Optimierungsproblem ist.

Wir nehmen nun an, wir haben folgende Daten beobachtet:

A_i	0	0	2	4	5
S_i	1	5	1	5	1
Y_i	0	1	1	1	0

b) (3 Punkte) Die Daten wurden mit den Werten $\beta_0 = -1$, $\beta_1 = 0.08$ und $\beta_2 = 0.022$ simuliert. Berechnen Sie die p_i für dieses Beispiel. Welche Interpretation haben die Parameter $\beta_0, \beta_1, \beta_2$? Was bedeuten positive/negative Werte?

c) (3 Punkte) Angenommen $\beta_0 = -1$ ist bekannt und β_1, β_2 sollen mit der Maximum-Likelihood-Methode geschätzt werden. Berechnen Sie die Werte für β_1 und β_2 (mit dem Computer) oder lesen Sie einen approximativen maximalen Wert in folgendem Bild ab, welches die Likelihood-Funktion farblich mit β_1 auf der x-Achse und β_2 auf der y-Achse zeigt.

Tipps:

Likelihood
Man betrachte

$\rho_\beta(a,s,y) = \mathbb{P}[Y = y \mid A = a, S = s]$