Konzentrationsungleichungen

Übungsblatt 4

TU Dortmund, Sommersemester 2021

Prof. Dr. Ivan Veselić

Übung 1. Sei X eine zentrierte Zufallsvariable auf $\mathbb{R}$, sodass $\mathbb{P}(|X| \geq t) \leq 2e^{-2t/\mu}$ für ein $\mu > 0$. Zeigen sie:

1. Für alle $k \in \mathbb{N}$ gilt:

$$\mathbb{E}[|X|^k] \leq \mu^k k!$$

2. Für X gilt weiterhin

$$M_X(\lambda) \leq e^{2\mu^2\lambda^2} \quad \forall |\lambda| \leq \frac{1}{2\mu}$$

3. Ist $Y \in \mathcal{G}(\nu)$, so gilt für $Z := Y^2 - \mathbb{E}(Y^2)$ die gleiche Abschätzung für die momentenerzeugende Funktion wie für X mit Parameter $\mu = 8\nu$.

Beweis. 1. Wir können hier wie in der Vorlesung die Layer-Cake-Formel anwenden:

$$\begin{aligned} \mathbb{E}[|X|^k] &= \int_0^\infty \mathbb{P}(|X|^k > t) dt = \int_0^\infty \mathbb{P}(|X| > t^{1/k}) dt \\ &\leq \int_0^\infty 2e^{-\frac{2t^{1/k}}{\mu}} dt. \end{aligned}$$

Substitution mit $u = \frac{2t^{1/k}}{\mu}$ liefert dann

$$\begin{aligned} \mathbb{E}[|X|^k] &\leq 2 \left(\frac{\mu}{2}\right)^k k \int_0^\infty e^{-u} u^{k-1} du \\ &= 2 \left(\frac{\mu}{2}\right)^k k \Gamma(k) \leq \mu^k k! \end{aligned}$$

über die Definition der Gammafunktion und $\Gamma(k) = (k-1)!$.

2. Wir berechnen die Momentenerzeugende Funktion über die Taylorentwicklung von $e^{|\lambda||X|}$, die aufgrund von monotoner Konvergenz existiert. Der Term für die lineare Entwicklung verschwindet, da X zentriert ist.

$$\mathbb{E}[e^{\lambda X}] \leq 1 + \sum_{k=2}^{\infty} \frac{|\lambda|^k \mathbb{E}[|X|^k]}{k!} \leq 1 + \sum_{k=2}^{\infty} (|\lambda| \mu)^k = 1 + \lambda^2 \mu^2 \sum_{k=0}^{\infty} (|\lambda| \mu)^k$$

Ist nun $|\lambda| \leq \frac{1}{2\mu}$, so ist die Reihe auf der rechten Seite kleiner als 2 und es folgt

$$\mathbb{E}[e^{\lambda X}] \leq 1 + 2\mu^2 \lambda^2 \leq e^{2\mu^2 \lambda^2}$$

3. Wir wissen aus der Vorlesung, dass $\mathbb{E}[Y^{2q}] \leq (4\nu)^q q!$. Somit existiert $\mathbb{E}[e^{\lambda Y^2}]$ für alle $|\lambda| < \frac{1}{4\nu}$, da dann

$$\mathbb{E}[e^{\lambda Y^2}] \leq 1 + \sum_{q=1}^{\infty} |\lambda|^q \frac{\mathbb{E}[Y^{2q}]}{q!} \leq 1 + \sum_{q=1}^{\infty} |4\nu\lambda|^q < \infty.$$

Somit existiert auch $\mathbb{E}[e^{\lambda Z}]$ für diese λ . Wir wenden erneut die Taylorentwicklung an:

$$\mathbb{E}[e^{\lambda Z}] \leq 1 + \sum_{q=2}^{\infty} \frac{|\lambda|^q \mathbb{E}[X^2 + \mathbb{E}[X^2]]^q}{q!}.$$

Da x^q für $q \geq 1$ und nichtnegative x konvex ist, können wir die Jensen-Ungleichung anwenden und erhalten $(x + y)^q \leq 2^q \left(\frac{1}{2}(x^q + y^q)\right)$ und somit folgt

$$\begin{aligned} \mathbb{E}[e^{\lambda Z}] &\leq 1 + \sum_{q=2}^{\infty} \frac{|\lambda|^q 2^{q-1} (\mathbb{E}[X^{2q}] + (\mathbb{E}[X^2])^q)}{q!} \\ &\leq 1 + \sum_{q=2}^{\infty} \frac{|\lambda|^q 2^q \mathbb{E}[X^{2q}]}{q!} \end{aligned}$$

durch erneute Anwendung der Jensenschen Ungleichung, diesmal um die Potenz q in den zweiten Erwartungswert zu ziehen. Jetzt können wir die Abschätzung für die geraden absoluten Momente von Sub-Gaussischen Zufallsvariablen nutzen und erhalten

$$\mathbb{E}[e^{\lambda Z}] \leq 1 + \sum_{q=2}^{\infty} (8\nu |\lambda|)^q \leq 1 + (8\nu\lambda)^2 \sum_{q=0}^{\infty} (8\nu |\lambda|)^q.$$

Falls $|\lambda| \leq \frac{1}{16\nu}$, so folgt

$$\mathbb{E}[e^{\lambda Z}] \leq 1 + 128\nu^2 \lambda^2 \leq e^{128\nu^2 \lambda^2}$$

und Vergleich mit der MEF von X zeigt die Behauptung. □

Übung 2. In Satz 2.13 (Johnson-Lindenstrauss-Lemma) hatten wir für eine n -elementige Menge $A \subset \mathbb{R}^D$ unter Verwendung von Zufallsvariablen $(X_{ij}) \subset \mathcal{G}(\nu)$, $i \in \{1, \dots, d\}$, $j \in \{1, \dots, D\}$, eine Abbildung $W : \mathbb{R}^D \rightarrow \mathbb{R}^d$ konstruiert, so dass für

$$d \geq 100 \frac{\nu^2}{\epsilon^2} \ln \left(\frac{n}{\sqrt{\delta}} \right) \tag{1}$$

gilt:

$$\mathbb{P}(W \text{ ist } \epsilon\text{-Isometrie auf } A) \geq 1 - \delta.$$

Wir wollen nun im Fall $X_{ij} \sim \mathcal{N}(0, 1)$ die Annahme (1) zu

$$d \geq \frac{16}{\epsilon^2} \ln \left(\frac{n}{\sqrt{\delta}} \right) \tag{2}$$

abschwächen.

a) Definieren Sie für $\alpha \in \mathbb{R}^D$ mit $\|\alpha\| = 1$:

$$\tilde{W}_i(\alpha) = \sum_{j=1}^D \alpha_j X_{ij}, \quad W(\alpha) = \left(\frac{1}{\sqrt{d}} \tilde{W}_i(\alpha) \right)_{i=1}^d.$$

Welche Verteilung hat $\tilde{W}_i(\alpha)$?

Erklären Sie, warum $\sum_{i=1}^d (\tilde{W}_i(\alpha)^2 - 1)$ sub- Γ -verteilt mit Varianzfaktor $2d$ und Skalenparameter 2 ist!

b) Zeigen Sie für $t > 0$

$$\mathbb{P} \left(\left| \sum_{i=1}^d (W_i(\alpha)^2 - 1) \right| \geq 2\sqrt{dt} + 2t \right) \leq 2e^{-t}.$$

c) Folgern Sie

$$\mathbb{P} \left(\|W(\alpha)\|^2 - 1 \geq 2\sqrt{\frac{\ln(n^2/\delta)}{d}} + 2\frac{\ln(n^2/\delta)}{d} \right) \leq 2\frac{\delta}{n^2}.$$

d) Folgern Sie für eine N -elementige Teilmenge T der Einheitsphäre in $\mathbb{R}^D$

$$\mathbb{P} \left(\max_{\alpha \in T} \|W(\alpha)\|^2 - 1 \geq 2\sqrt{\frac{\ln(n^2/\delta)}{d}} + 2\frac{\ln(n^2/\delta)}{d} \right) \leq 2N\frac{\delta}{n^2}.$$

e) Folgern Sie die Aussage des Johnson-Lindenstrauss-Lemmas unter der Annahme (2) für den Fall $X_{ij} \sim \mathcal{N}(0, 1)$.

Beweis. a) Wenn $X \sim \mathcal{N}(0, 1)$, so ist $a \cdot X \sim \mathcal{N}(0, a^2)$, denn

$$\begin{aligned} \mathbb{P}(aX \leq t) &= \mathbb{P} \left(X \leq \frac{t}{a} \right) = \int_{-\infty}^{t/a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi}u} e^{-\frac{x^2}{2u^2}} du \end{aligned}$$

über die Substitution $u = ax$ (Diese Rechnung gilt streng genommen nur für positives a , aber über Symmetrie auch für negatives a).

Weiterhin gilt für $X \sim \mathcal{N}(0, a^2)$, $Y \sim \mathcal{N}(0, b^2)$ wenn X und Y unabhängig sind, dass $X + Y \sim \mathcal{N}(0, a^2 + b^2)$, z.B. über die Faltungsformel für die Dichte unabhängiger ZV:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$

Daraus folgt $\tilde{W}_i(\alpha) \sim \mathcal{N}(0, \|\alpha\|^2) = \mathcal{N}(0, 1)$.

Nun müssen wir zeigen, dass falls $X \sim \mathcal{N}(0, 1)$, so ist $X^2 - 1$ sub- Γ -verteilt mit Varianzfaktor 2 und Skalenfaktor 2. Der Varianzfaktor $2d$ folgt dann aus der Summe über i , da die ZV unabhängig sind. Wir müssen also zeigen:

$$\mathbb{E} \left(e^{\lambda(X^2-1)} \right) \leq e^{\frac{\lambda^2}{1-2\lambda}}$$

für alle $-1/2 < \lambda < 1/2$.

Berechnen wir zunächst $M_{X^2}(\lambda)$:

$$\begin{aligned} \mathbb{E} \left(e^{\lambda X^2} \right) &= \int_{-\infty}^{\infty} e^{\lambda y^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2(1-2\lambda)}{2}} dx \stackrel{\lambda < 1/2}{=} \frac{1}{\sqrt{1-2\lambda}} \end{aligned}$$

Da wir eigentlich $X^2 - 1$ betrachten, müssen wir nun zeigen, dass

$$\frac{e^{-\lambda}}{\sqrt{1-2\lambda}} \leq e^{\frac{\lambda^2}{1-2\lambda}}$$

$$\frac{e^{\lambda}}{\sqrt{1+2\lambda}} \leq e^{\frac{\lambda^2}{1-2\lambda}}$$

für $\lambda < 1/2$, damit $X^2 - 1$ sowohl in Γ_+ als auch in Γ_- ist. Die erste Gleichung lässt sich durch logarithmieren umstellen zu

$$-\lambda - \frac{1}{2} \ln(1-2\lambda) \leq \frac{\lambda^2}{1-2\lambda}$$

$$\Leftrightarrow -\lambda + \lambda^2 - \frac{1-2\lambda}{2} \ln(1-2\lambda) \leq 0$$

Die Ungleichung stimmt für $\lambda = 0$. Ableiten auf beiden Seiten liefert

$$4\lambda + 2 \ln(1-2\lambda) \leq 0$$

dies gilt wieder bei 0 und wir können erneut ableiten und bekommen

$$4\left(1 - \frac{1}{1-2\lambda}\right) \leq 0$$

was für $\lambda < 1/2$ stimmt. Die zweite Gleichung lässt sich (für $\lambda < 1/2$) umstellen zu

$$\lambda - \frac{1-2\lambda}{2} \ln(1+2\lambda) \leq 3\lambda^2.$$

Dies gilt wieder für $\lambda = 0$ und wir können ableiten, um

$$1 + \ln(1+2\lambda) - \frac{1-2\lambda}{1+2\lambda} \leq 6\lambda$$

zu erhalten. Dies gilt wieder für $\lambda = 0$ und wir können ein letztes Mal ableiten, um

$$\frac{2}{1+2\lambda} + \frac{4}{(1+2\lambda)^2} \leq 6$$

zu erhalten, was für alle $\lambda \geq 0$ gilt. Somit ist $\sum_{i=1}^d (\tilde{W}_i(\alpha)^2 - 1)$ in $\Gamma(2d, 2)$.

- b) Das ist die Chernoff-Schranke für sub- Γ -verteilte Zufallsvariablen (die 2 kommt daher, dass wir linke und rechte Schranke aufsummieren).
- c) Substituiere $t = \ln(n^2/\delta)$ und teile in $\mathbb{P}(\dots)$ durch d .
- d) Union bound:

$$\mathbb{P}\left(\max_{i=1}^N X_i \geq t\right) \leq N\mathbb{P}(X_1 \geq t).$$

- e) Wie in Vorlesung: Wir können uns auf wegen Linearität auf Differenzen $\frac{a_i - a_j}{\|a_i - a_j\|}$, $a_i, a_j \in A$ zurückziehen. Von diesen Differenzen gibt es $N = n(n-1)/2 \leq n^2/2$ Stück und wir haben $2N\delta/n^2 \leq \delta$.

Es bleibt noch zu zeigen, dass (2) impliziert, dass

$$2\sqrt{\frac{\ln(n^2/\delta)}{d}} + 2\frac{\ln(n^2/\delta)}{d} \leq \epsilon.$$

Dazu berechnen wir

$$2\sqrt{\frac{\ln(n^2/\delta)}{d}} + 2\frac{\ln(n^2/\delta)}{d} \leq 2\sqrt{\frac{2\epsilon^2 \ln(n/\sqrt{\delta})}{16 \ln(n/\sqrt{\delta})}} + 2\frac{2\epsilon^2 \ln(n/\sqrt{\delta})}{16 \ln(n/\sqrt{\delta})} \leq \epsilon\sqrt{1/2} + \epsilon/4 \leq \epsilon.$$

Das war zu zeigen (Eigentlich besagt das Johnson-Lindenstrauss-Lemma, dass

$$\mathbb{P}(\max_{\alpha \in T} \|W(\alpha)\|^2 - 1 \leq \epsilon) \leq 1 - \delta,$$

aber das folgt sofort durch Übergang zum Gegenereignis).

□

Abgabe bis zum 17.05.2021