Konzentrationsungleichungen ohne Unabhängigkeit

(nach Vershynin „High-Dimensional Probability", § 5.1, § 5.2)

ZIEL:

Quantifiziere $\mathbb{P}\{|f(X) - \mathbb{E}[f(X)]| > t\}$ ohne Unabhängigkeitsannahme für $X \sim \mathrm{Unif}(\sqrt{n}\,\mathbb{S}^{n-1})$, f Lipschitz (weitere Settings?)

Werkzeug:

Isoperimetrische Ungleichung & Blow Up

Wir formulieren zunächst das Hauptergebnis:

> **Thm 5.1.4** (Konzentration von Lipschitz-Funktionen auf Sphäre)
>
> Seien $X \sim \mathrm{Unif}(\sqrt{n}\,\mathbb{S}^{n-1})$, $f : \sqrt{n}\,\mathbb{S}^{n-1} \longrightarrow \mathbb{R}$ Lipschitz.
>
> Dann gilt:
> $$\| f(X) - \mathbb{E}[f(X)] \|_{\psi_2} \le C \cdot \|f\|_{Lip}$$
> bzw.
> $$\mathbb{P}\{|f(X) - \mathbb{E}[f(X)]| > t\} \le 2\exp\left(-\frac{ct^2}{\|f\|_{Lip}^2}\right)$$

I. Vorarbeit: Isoperimetrische Ungleichung & Blow Up

↳ Isoperimetrische Problem (Var. Rech.)

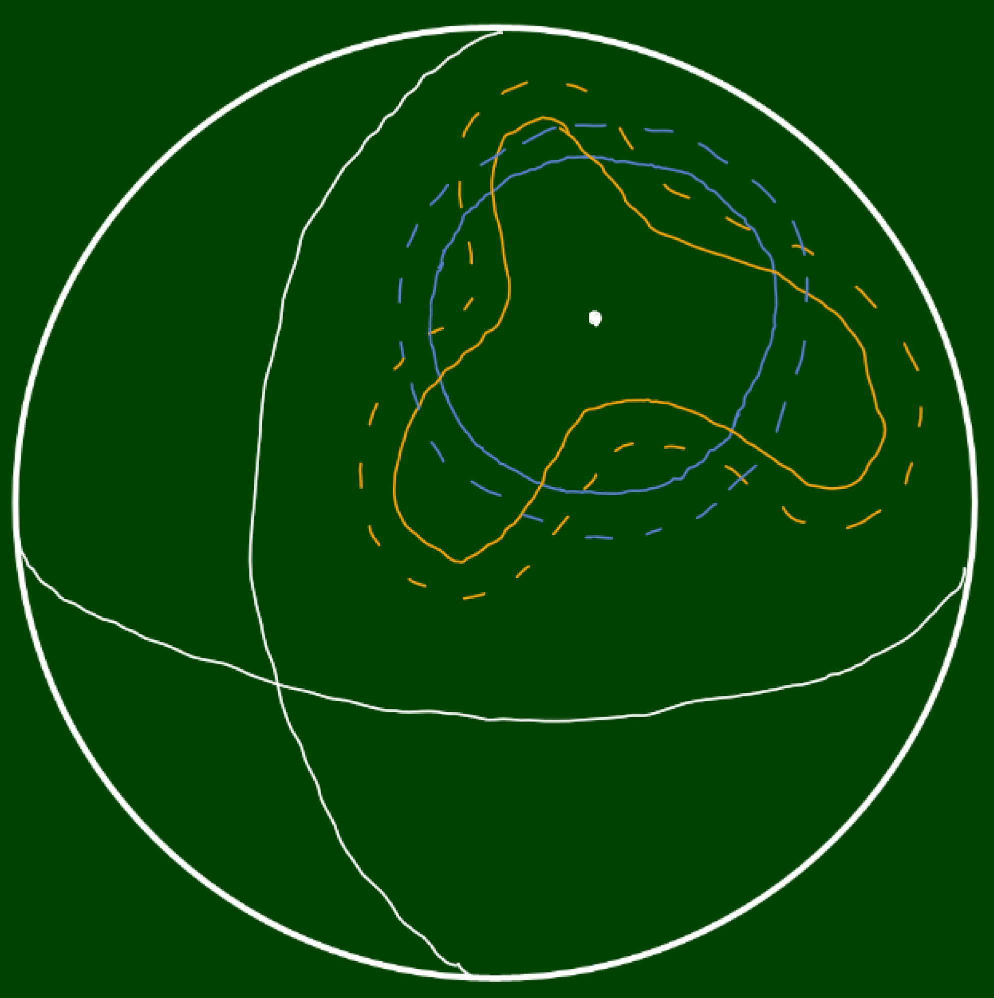
Kreis minimiert Umfang einer Fläche gegebenen Flächeninhalts

$\mathbb{R}^3$: Kugel $\quad$ $\mathbb{S}^2$: runde Kappe

$\mathbb{R}^2$: Kreis $\quad$ $\mathbb{S}^1$: Kreisbogen (zshg)

> **Thm 5.1.6** (Isoperimetrische Ungleichung in $\mathbb{S}^{n-1}$)
>
> Seien $\varepsilon > 0$, $F \in \mathbb{R}$, $a \in \mathbb{S}^{n-1}$. Sei weiter
> $$A_\varepsilon := \{ x \in \mathbb{S}^{n-1} \mid \exists y \in A : \|x - y\|_2 \le \varepsilon \}$$
> ε-Umgebung von A
>
> 
>
>
> und
> $$C(a, r) := \{ x \in \mathbb{S}^{n-1} \mid \|x - a\|_2 \le r \}$$
> sphärische Kappe
>
> 
>
>
> sowie $r_F \in \mathbb{R}$ definiert durch
> $$\sigma_{n-1}(C(a, r_F)) = F.$$
>
> Dann gilt:
> $$\operatorname{argmin}\{ \sigma_{n-1}(A_\varepsilon) \mid A \subset \mathbb{S}^{n-1};\ \sigma_{n-1}(A) = F \} = C(a, r_F).$$

Beweis: Hier nicht. ∎



„Unter allen ε-Umgebungen $A_\varepsilon \subset \mathbb{S}^{n-1}$ von Teilmengen $A \subset \mathbb{S}^{n-1}$ des gleichen Flächeninhalts $\sigma_{n-1}(A) = F$ haben die ε-Umgebungen der sphärischen Kappen den kleinsten Flächeninhalt."

$$\sigma_{n-1}(A) = F = \sigma_{n-1}(C)$$

$$\Rightarrow \sigma_{n-1}(A_\varepsilon) \geq \sigma_{n-1}(C_\varepsilon)$$

Die isoperimetrische Ungleichung hat folgende konter-intuitive Implikation:

„Wenn A mindestens die Hälfte der Sphärenoberfläche bedeckt, dann bedeckt A_ε den Großteil der Sphärenoberfläche"

Lemma 5.1.7 (Blow Up)

Seien $A \subset \sqrt{n}\,\mathbb{S}^{n-1}$, σ normierter Flächeninhalt auf $\sqrt{n}\,\mathbb{S}^{n-1}$ und $\sigma(A) \geq \frac{1}{2}$. Dann gilt:

$$\exists K_1 \in \mathbb{R}\ \forall t > 0: \quad \sigma(A_t) \geq \underbrace{1 - 2\exp\left(-\frac{t^2}{K_1^2}\right)}_{\geq -1 \text{ Schranke nur gut für große } t}$$

$\sigma(A_t) > \frac{1}{2} \Leftrightarrow t^2 > \frac{\ln(4)}{K_1^2}$

Beweis: Setze $H := \{x \in \sqrt{n}\,\mathbb{S}^{n-1} : x_1 \leq 0\}$.

$$\Rightarrow \sigma(A) \geq \frac{1}{2} = \sigma(H)$$

$$\underset{\text{Isoperim. Ungleichung}}{\Rightarrow} \forall t > 0: \quad \sigma(A_t) \geq \underbrace{\sigma(\overbrace{H_t}^{\text{sphärische Kappe}})}_{\text{nicht ausrechnen}}$$

Beobachte: σ W-Maß zur Gleichverteilung auf Sphäre, d.h.

$$\sigma(H_t) = \mathbb{P}\{X \in H_t\}, \quad X \sim \mathrm{Unif}(\sqrt{n}\,\mathbb{S}^{n-1})$$

In Kapitel 3 (Thm 3.4.6) wird gezeigt, dass

$$\{X \sim \mathrm{Unif}(\sqrt{n}\,\mathbb{S}^{n-1}) \Rightarrow X \text{ subgaussisch} \wedge \|X\|_{\psi_2} \leq C\}$$

Zeige nun:

$$\mathbb{P}\{X \in H_t\} \underset{(*)}{\geq} \mathbb{P}\left\{|X_1| \leq \frac{t}{\sqrt{2}}\right\} \underset{(**)}{\geq} 1 - 2\exp(-ct^2)$$

zu (*): Wir zeigen $H_t = \{x \in \sqrt{n}\,S^{n-1} \mid \exists y \in H : \|x-y\|_2 \le t\} \supseteq \{x \in \sqrt{n}\,S^{n-1} \mid |x_1| \le \frac{t}{\sqrt{2}}\} =: L$

Dazu: $x \in L \Rightarrow |x_1| \le \frac{t}{\sqrt{2}} \Rightarrow y_i = \begin{cases} x_i, & i \ne 1 \\ \left(n - \sum_{i \ge 2} x_i^2\right)^{1/2}, & i = 1 \end{cases}$

$$\Rightarrow y \in H \wedge \|x-y\|_2^2 = |x_1 - y_1|^2 \le |x_1|^2 + |n - \sum_{i \ge 2} x_i^2|$$
$$= 2|x_1|^2 \le t^2$$

zu (**): $\mathbb{P}\{|X_1| \le \frac{t}{\sqrt{2}}\} = 1 - \mathbb{P}\{|X_1| > \frac{t}{\sqrt{2}}\} \underset{\uparrow}{\ge} 1 - 2\exp\left(-\frac{t^2}{K_1^2}\right)$

$$\|X_1\|_{\psi_2} = \|X \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\|_{\psi_2} \le \sup_{r \in S^{n-1}} \|X \cdot r\|_{\psi_2} = \|X\|_{\psi_2} \le C$$

Zusammen folgt:

$$\sigma(A_t) \ge 1 - 2\exp\left(-\frac{t^2}{K_1^2}\right) \quad \forall t \ge 0$$

$\blacksquare$

Bemerkung:
Für die Aussage ist die Wahl $1/2$ als untere Schranke nicht ausschlaggebend (aber im Weiteren hinreichend → Median)

Man kann die „Hälfte der Sphärenfläche" durch exponentiell kleine Teilmengen ersetzen:

Exercise 5.1.9

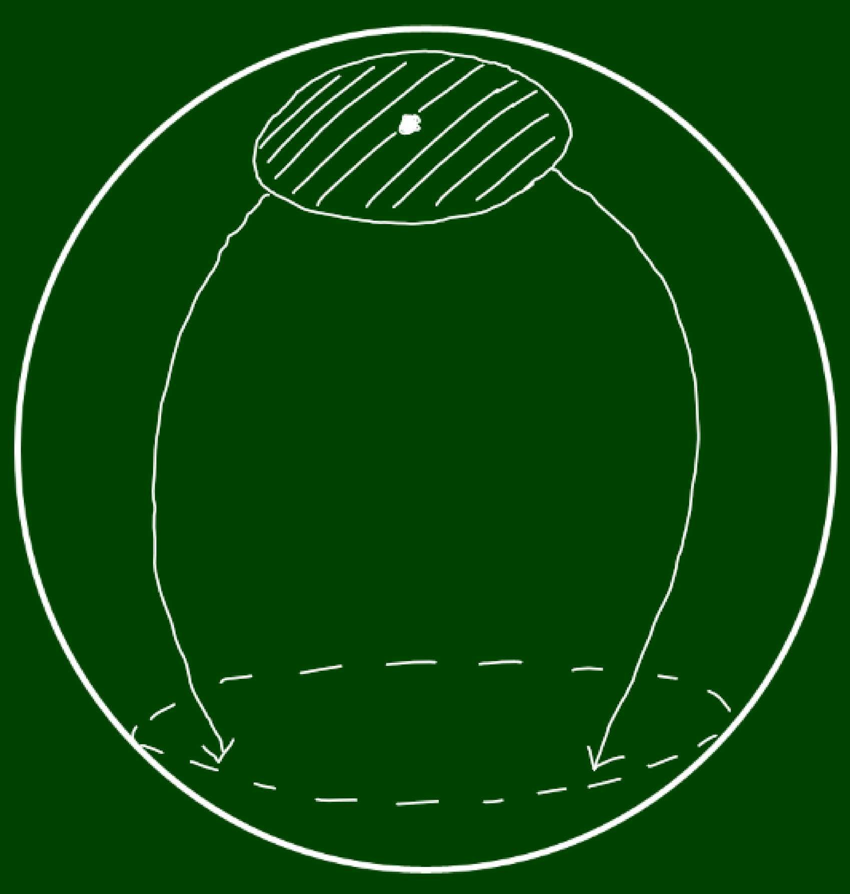
Seien $c > 0$, $A \subset \sqrt{n}\,\mathbb{S}^{n-1}$ und es existiere $s > 0$ so, dass

$$\sigma(A) > \underbrace{2\exp(-cs^2)}_{\text{klein, wenn } s \text{ groß}}.$$

Dann gilt:

(i) $\sigma(A_s) > 1/2$

(ii) $\forall t \geq s: \ \sigma(A_{2t}) \geq 1 - 2\exp(-ct^2)$



Bemerkung: (0-1-Gesetz)

Betrachte speziell $A \subset \mathbb{S}^{n-1}$ (vgl **!):

$$\sigma(A_t) \geq \sigma(H_t) \geq \sigma\left\{|X_1| \leq \frac{t}{\sqrt{2}}\right\} = \mathbb{P}\left\{|\underbrace{\sqrt{n}X_1}_{\in \mathbb{S}^{n-1}}| \leq \frac{\sqrt{n}\,t}{\sqrt{2}}\right\} \geq 1 - 2\exp\left(-\frac{cnt^2}{2}\right)$$

Also: $\sigma(A) > \underbrace{2\exp\left(-\frac{cnt^2}{2}\right)}_{\substack{\text{klein, falls } t \text{ klein} \\ \text{aber } n \text{ entsprechend} \\ \text{groß}}} \Rightarrow \sigma(A_{2t}) \geq 1 - \exp\left(-\frac{cnt^2}{2}\right)$

für $\sqrt{n} \gg t$ ist A klein, t nur kleine Verdickung aber A_t groß

„A exponentiell klein, A_{2t} exponentiell groß"

Aussagen der Art „W'keit von Ereignissen eines bestimmten Typs sind entweder 0 oder 1" sind typisch für hohe Dimensionen!

II. Beweis der Konzentrationsungleichung Thm 5.1.4:

(Analog für $\mathbb{S}^{n-1}$ statt $\sqrt{n}\,\mathbb{S}^{n-1}$ mit roten Anmerkungen)

Beweis: Wir bezeichnen mit $M \in \mathbb{R}$ den Median von $f(X)$ und definieren

$$A := \{ y \in \sqrt{n}\,\mathbb{S}^{n-1} \; (\mathbb{S}^{n-1}) \mid f(y) \le M \}.$$

Nach Def. des Medians ist dann

$$\mathbb{P}\{X \in A\} \ge 1/2$$

und mit Lemma 5.1.7 (Blow up) folgt für $\varepsilon := \frac{t}{\|f\|_{Lip}}$

$$\mathbb{P}\{X \in A_\varepsilon\} \ge 1 - 2\exp(-c\,\overset{n}{\vee}\,\varepsilon^2)$$

Weiter gilt:

$$X \in A_\varepsilon \Rightarrow \exists y \in A : \|X-y\|_2 \le \varepsilon$$

$$\underset{f \text{ Lipschitz}}{\Rightarrow} \begin{cases} f(X) \le f(y) + \|f\|_{Lip} \cdot \|X-y\|_2 \overset{y\in A}{\le} M+t \\ -f(X) \le -f(y) + \|f\|_{Lip} \|X-y\|_2 \le -M+t \end{cases}$$

und es folgt:

$$\underbrace{\mathbb{P}\{|f(X)-M| \le t\}} \ge \mathbb{P}\{X \in A_\varepsilon\} \ge \underbrace{1 - 2\exp\Big(-c\,\overset{n}{\vee}\,\underbrace{\frac{t^2}{\|f\|^2_{Lip}}}_{=\varepsilon^2}\Big)}$$

$$= \mathbb{P}\left\{ \left| \frac{f(X)}{\|f\|_{Lip}} - \frac{M}{\|f\|_{Lip}} \right| \le \varepsilon \right\}$$

$$= \mathbb{P}\left\{ \left| \frac{\sqrt{n}}{\|f\|_{Lip}}(f(X)-M) \right| \le \tilde{t} \right\}$$

mit $\tilde{t} := \sqrt{n}\,\varepsilon$

$$= 1 - 2\exp(-c\tilde{t}^2)$$

$$\Leftrightarrow \quad \exists C > 0 : \|f(X) - M\| \le C \cdot \frac{\|f\|_{Lip}}{\sqrt{n}}$$

Es bleibt den Median durch den Erwartungswert zu ersetzen:

$$\|f(X) - \mathbb{E}[f(X)]\|_{\psi_2} \underset{\substack{\uparrow \\ \text{Centering Lemma} \\ 2.6.8}}{\le} \tilde{C}\,\|f(X)\|_{\psi_2} \le \tilde{C}\left(\|f(X)-M\|_{\psi_2} + \|M\|_{\psi_2}\right)$$

$$\underset{\substack{\uparrow \\ \inf\{t>0 : \mathbb{E}(e^{M^2/t^2}) \le 2\}}}{\le} \tilde{C}\left(C \frac{\|f\|_{Lip}}{\sqrt{n}} + \frac{M}{\sqrt{\ln(2)}}\right)$$

$$M^2 \le t^2 \ln(2) \Leftrightarrow t \ge \frac{M}{\sqrt{\ln(2)}}$$

$\blacksquare$

Bemerkung: (Konzentration auf $\mathbb{S}^{n-1}$)

$X \sim \mathrm{Unif}(\mathbb{S}^{n-1})$, $f: \mathbb{S}^{n-1} \longrightarrow \mathbb{R}$ Lipschitz

$$\Longrightarrow \begin{cases} \| f(X) - \mathbb{E}[f(X)] \|_{\psi_2} \leq \dfrac{C \cdot \|f\|_{Lip}}{\sqrt{n}} \\ \mathbb{P}\{ |f(X) - \mathbb{E}[f(X)]| > t \} \leq 2 \exp\left(- \dfrac{c n t^2}{\|f\|_{Lip}^2} \right) \end{cases}$$

Beweis: vgl. Ergänzungen Beweis Thm. 5.1.4. ∎

Bemerkung:

Die Implikationskette

„Blow Up $\Rightarrow$ Konz. um M $\Rightarrow$ Konz. um Erwartungswert"

kann umgedreht werden.

III. Weitere Anwendung des Blow Ups:

Exercise 5.1.5 (fast orthogonale Mengen)

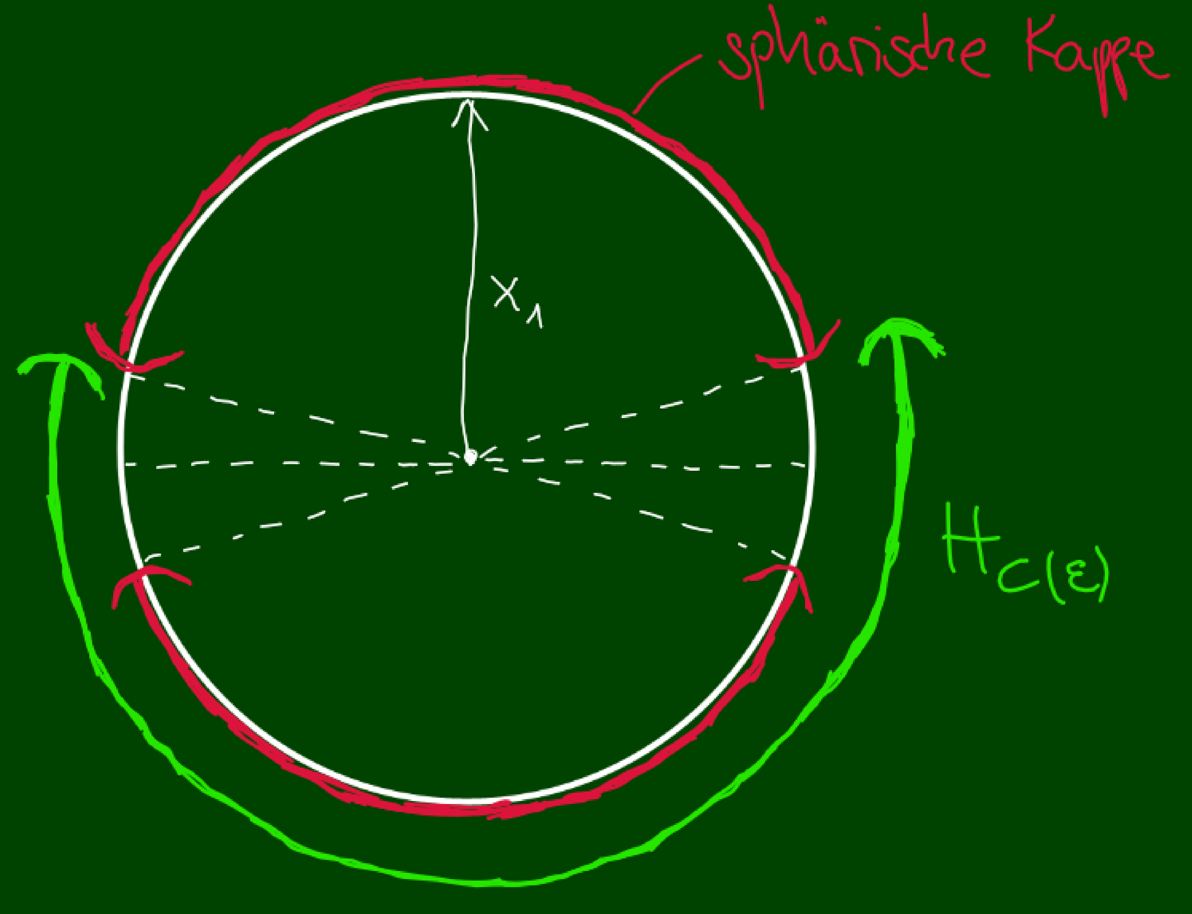
Z: Sei $\varepsilon > 0$. Es existiert

$$\tilde{O} := \{ x_1, \dots, x_N \in \mathbb{R}^n \mid |\langle x_i | x_j \rangle| \leq \varepsilon, \|x_i\| = 1 \}$$

so, dass

$$N \geq \exp(c(\varepsilon) n).$$

Beweisskizze:



$$H_{c(\varepsilon)} = \{ y \in \mathbb{S}^{n-1} \mid \exists x \in H : \|x - y\|_2 \leq \sqrt{c(\varepsilon)} \}$$

definiert durch $|\langle y | x_1 \rangle| \leq \varepsilon$ für alle $y \in \mathbb{S}^{n-1} \setminus H$

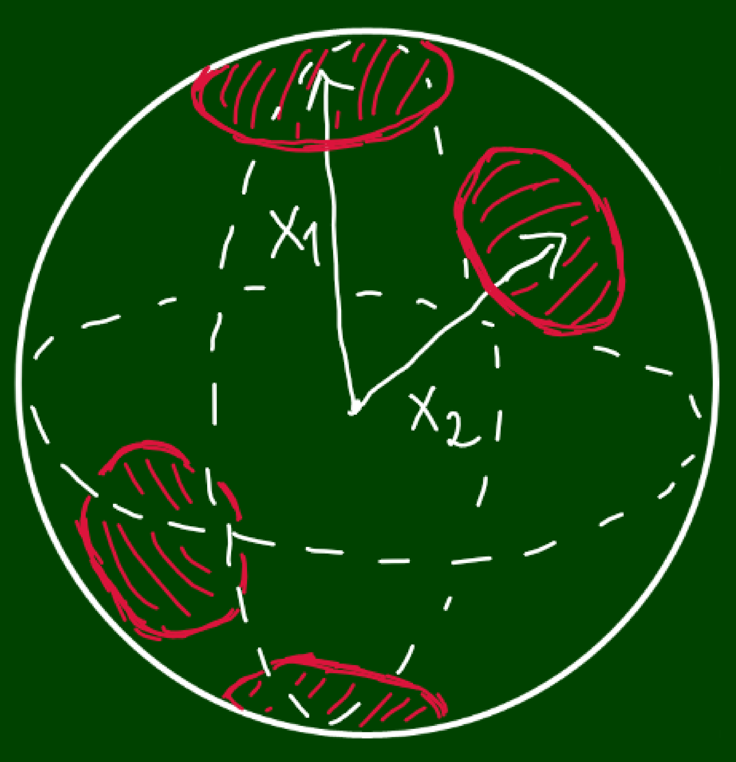
$$(c(\varepsilon) = 2(1 - \sqrt{1 - \varepsilon^2})\ ?)$$

$$\Rightarrow \sigma(H_{c(\varepsilon)}) \geq 1 - 2\exp(-c(\varepsilon) n)$$

$$\Rightarrow \sigma(\mathbb{S}^{n-1}_{[x_1]}) \geq 1 - 4\exp(-c(\varepsilon) n)$$

$$\vdots$$

$$\Rightarrow \sigma(\underbrace{\mathbb{S}^{n-1}_{[x_1, \dots, x_K]}}) \geq 1 - (K+1) 2 \exp(-c(\varepsilon) n)$$



Abbruchbedingung: $N := \min \{ K \in \mathbb{N} \mid \mathbb{S}^{n-1}_{[x_1, \dots, x_K]} = \emptyset \}$

$$\Rightarrow N + 1 \geq \Big(1 - \underbrace{\sigma(\mathbb{S}^{n-1}_{[x_1, \dots, x_N]})}_{=0}\Big) \frac{1}{2} \exp(c(\varepsilon) n)$$

IV. Zusammenfassung

Isoperimetrische Ungleichung $\Rightarrow$ Blow Up $\Rightarrow$ ($\Leftarrow$) Konzentration um Erwartungswert für $X\sim \mathrm{Unif}((\sqrt{n})\mathbb{S}^{n-1})$, $f:(\sqrt{n})\mathbb{S}^{n-1}\to\mathbb{R}$ Lipschitz mit eukl. Metrik

$\geq$

ε

$$\|f(X)-\mathbb{E}[f(X)]\|_{\psi_2} \leq C\cdot\|f\|_{Lip}$$

Blow Up $\Rightarrow$ fast orthogonale Mengen

$$|\{x_1,\dots,x_N \mid |\langle x_i,x_j\rangle|\leq\varepsilon\}| \geq \exp(c(\varepsilon)n)$$

V. Ausblick: Weitere Konzentrationsungleichungen

Analog für andere metrische Räume:

- $X\sim N(0,\mathbb{1}_n)$, $(\mathbb{R}^n,\|\cdot\|_2,\gamma_n)$ mit Gauß-Maß $\gamma_n(A):=\frac{1}{(2\pi)^{n/2}}\int_A e^{-\frac{1}{2}\|x\|_2^2}\,dx$

  $\xrightarrow{\text{push forward}}$ $X\sim\mathrm{Unif}([0,1]^n)$, $X\sim\mathrm{Unif}(\sqrt{n}\,B_2^n)$

- $X\sim\mathrm{Unif}(\{0,1\}^n)$, $(\{0,1\}^n,d,\mathbb{P})$ Hamming Cube mit $d(x,y):=\frac{1}{n}|\{i \mid x_i\neq y_i\}|$

  $\mathbb{P}(A):=\frac{|A|}{2^n}$

- Symmetrische Gruppe als metrischer Maßraum $(S_n,d,\mathbb{P})$ mit $d(\pi,\delta):=\frac{1}{n}|\{i \mid \pi(i)\neq\delta(i)\}|$

  $\mathbb{P}(A)=\frac{|A|}{n!}$

- Riemannsche Mfk'en $(M,d,\mathbb{P})$ mit d Bogenlänge bzgl g, $\mathbb{P}=\frac{dv}{V}$

  ↳ $\mathbb{S}^{n-1}$ als Spezialfall

- Spezielle orthogonale Gruppe $(SO(n),\|\cdot\|_F,\mathbb{P})$, $\mathbb{P}$ Gleichverteilung auf $SO(n)$

  ↳ Grassmann Mfk'en $(G_{n,m},d,\mathbb{P})$, $d(E,F)=\|P_E-P_F\|$, $\mathbb{P}$ uniformes Haarmaß

- Dichten $e^{-H(x)}$: $H:\mathbb{R}^n\to\mathbb{R}$, $\exists\,\kappa>0$: $\underline{\text{Hess}\,H(x)-\kappa\mathbb{1}_n\geq 0}$ $\forall x\in\mathbb{R}^n$, $f:\mathbb{R}^n\to\mathbb{R}$ Lipschitz

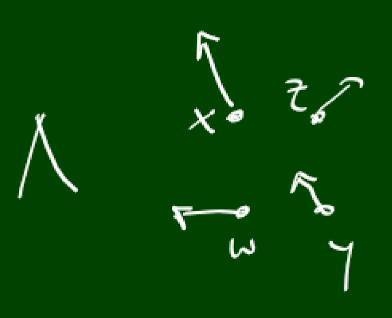
  $$\Rightarrow \|f(X)-\mathbb{E}[f(X)]\|_{\psi_2}\leq\frac{C\cdot\|f\|_{Lip}}{\sqrt{\kappa}}$$

Notizen: (zu Dichten e^{-H})

- $\boxed{\text{von } \Delta_H \text{ erzeugte Halbgruppe}}$ Positivitätsverb. eig

  ↳ $d_H = e^{-H} d e^{H}$

- zum Kontext: Magn. Modelle



$$H(\varphi) = \sum_{x\in\Lambda} V_x(\varphi_x) + \sum_{xy} \omega_{xy}(\varphi_x-\varphi_y)^2 + \text{Magnetfeld}$$

zugrundeliegende Vert.: $e^{-\frac{1}{T}H}$

$$Cor(X,Y) = \frac{1}{2}\langle e^{-H}\otimes c_x^{+}\Omega \,|\, (\Delta_H)^{-1}(e^{-H}\otimes c_y^{+}\Omega)\rangle$$

- zur Konvexitätsbedingung:

  Ähnlich Log-Sobolev-Ungleichungen:

$$\mathbb{E}_H[\Phi(F)] - \Phi(\mathbb{E}_H[F]) \le \text{Konst}\cdot \mathbb{E}_H[(\nabla\sqrt{F})^2]$$

  für $\Phi(u)=u\log(u)$, $F \geq 0$.