HDP
(19) Gaussian Width and Random Projection

@ 滕佳烨

前情回顾

Analyze $\mathbb{E} \sup_{t \in T} X_t$ for Gaussian Process

1. Slepian's inequality: Small fluctuation leads to small expectation.
2. Sudakov-Fernique's inequality: remove variance requirement (only expectation results)
3. Gordon's inequality: two-dim extension

4. *Sudakov's minoration inequality: lower bound, Geometric

Random Process and Operator Norm

$$\|A\| = \max_{u \in S^{n-1},\, v \in S^{m-1}} \langle Au, v \rangle = \max_{(u,v) \in T} X_{uv}$$

Proof based on covering number

$$\mathbb{E}\,\|A\| \leq \sqrt{m} + C\sqrt{n}$$

Proof based on Gaussian process

$$\mathbb{E}\,\|A\| \leq \sqrt{m} + \sqrt{n}.$$

Intuition: Apply Sudakov-Fernique's inequality
Sudakov-Fernique's inequality leads to Sudakov's minoration inequality.
Therefore the bound is tighter than covering number – based bound.
(just an intuition, minoration bound is in fact lower bound)

Gaussian width

Definition 7.5.1. The *Gaussian width* of a subset $T \subset \mathbb{R}^n$ is defined as

$$w(T) := \mathbb{E} \sup_{x \in T} \langle g, x \rangle \quad \text{where} \quad g \sim N(0, I_n).$$

a special width which is invariant under affine and convex operator
Note: *g* (symmetric) has an expectation.
Therefore, the asymmetric T suffers a small w(T).

$$\frac{1}{\sqrt{2\pi}} \cdot \mathrm{diam}(T) \le w(T) \le \frac{\sqrt{n}}{2} \cdot \mathrm{diam}(T).$$

We can also replace *g* to other random vectors, e.g., spherical.

Gaussian width

Definition 7.5.1. The *Gaussian width* of a subset $T \subset \mathbb{R}^n$ is defined as

$$w(T) := \mathbb{E} \sup_{x \in T} \langle g, x \rangle \quad \text{where} \quad g \sim N(0, I_n).$$

Some examples:

ℓ_2 ball $\quad w(S^{n-1}) = w(B_2^n) = \mathbb{E}\,\|g\|_2 = \sqrt{n} \pm C,$

ℓ_∞ ball $\quad w(B_\infty^n) = \mathbb{E}\,|g_1| \cdot n = \sqrt{\frac{2}{\pi}} \cdot n.$

ℓ_1 ball $\quad c\sqrt{\log n} \le w(B_1^n) \le C\sqrt{\log n}.$

ℓ_p ball $\quad w(B_p^n) \le C\sqrt{p'}\, n^{1/p'}.$

Finite set $\quad w(T) \le C\sqrt{\log |T|} \cdot \operatorname{diam}(T).$

Asymmetric set suffers.

Stable dimension

$$w(T) := \mathbb{E} \sup_{x \in T} \langle g, x \rangle \quad \text{where} \quad g \sim N(0, I_n).$$

$$h(T)^2 := \mathbb{E} \sup_{t \in T} \langle g, t \rangle^2, \quad \text{where} \quad g \sim N(0, I_n).$$

Definition 7.6.2 (Stable dimension)**.** For a bounded set $T \subset \mathbb{R}^n$, the *stable dimension* of T is defined as

$$d(T) := \frac{h(T-T)^2}{\mathrm{diam}(T)^2} \asymp \frac{w(T)^2}{\mathrm{diam}(T)^2}.$$

$d(T) \leq n$; 球上二者相等；有限集合$\log |T|$

Definition 7.6.7 (Stable rank)**.** The *stable rank* of an $m \times n$ matrix A is defined as

$$r(A) := \frac{\|A\|_F^2}{\|A\|^2}.$$

Gaussian complexity

$$w(T) := \mathbb{E} \sup_{x \in T} \langle g, x \rangle \quad \text{where} \quad g \sim N(0, I_n).$$

Definition 7.6.8. The *Gaussian complexity* of a subset $T \subset \mathbb{R}^n$ is defined as

$$\gamma(T) := \mathbb{E} \sup_{x \in T} |\langle g, x \rangle| \quad \text{where} \quad g \sim N(0, I_n).$$

Exercise 7.6.9 (Gaussian width vs. Gaussian complexity). Consider a set $T \subset \mathbb{R}^n$ and a point $y \in T$. Show that

$$\frac{1}{3}\left[w(T) + \|y\|_2\right] \le \gamma(T) \le 2\left[w(T) + \|y\|_2\right]$$

Random Projection

Consider the random projection P, w.h.p., the projected set has diam:

$$\operatorname{diam}(PT) \leq \begin{cases} C\sqrt{\dfrac{m}{n}}\,\operatorname{diam}(T), & \text{if } m \geq d(T) \\ Cw_s(T), & \text{if } m \leq d(T). \end{cases}$$

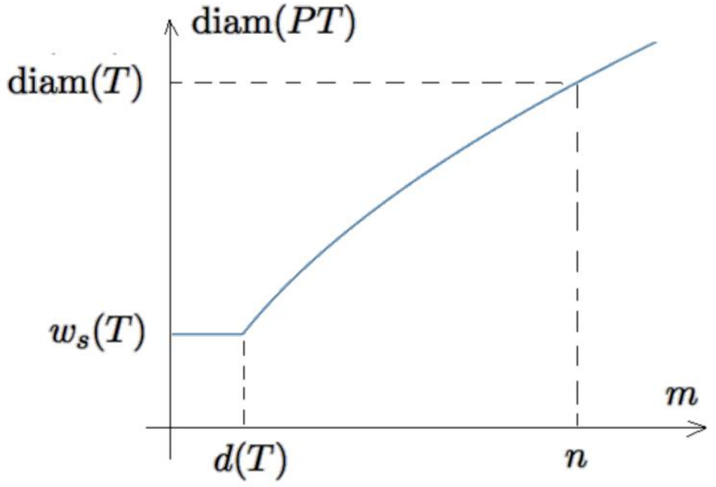

Figure 7.7 The diameter of a random m-dimensional projection of a set T as a function of m.

Take-away Messages

1. **Gaussian width and Gaussian complexity**
2. **Stable dimension and stable rank**
3. **Random projection (illustration of stable dimension)**

@ 滕佳烨

Thanks!