HDP
(20) Dudley’s inequality

@ 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

Today: we want to analyze a more general random process!

Random Process from Gaussian to sub-Gaussian

Sub-Gaussian increments

Definition 8.1.1 (Sub-gaussian increments)**.** Consider a random process $(X_t)_{t\in T}$ on a metric space (T, d). We say that the process has *sub-gaussian increments* if there exists $K \geq 0$ such that

$$\|X_t - X_s\|_{\psi_2} \leq K d(t, s) \quad \text{for all } t, s \in T. \tag{8.1}$$

Dudley's inequality

Theorem 8.1.3 (Dudley's integral inequality). *Let $(X_t)_{t\in T}$ be a mean zero random process on a metric space (T,d) with sub-gaussian increments as in (8.1). Then*

$$\mathbb{E}\sup_{t\in T} X_t \le CK\int_0^\infty \sqrt{\log \mathcal{N}(T,d,\varepsilon)}\,d\varepsilon.$$

$$\mathbb{E}\sup_{t\in T} X_t \le CK\sum_{k\in\mathbb{Z}} 2^{-k}\sqrt{\log \mathcal{N}(T,d,2^{-k})}.$$

Note: there is a **gap** between Dudley's inequality (upper bound) and Sudakov's inequality (lower bound).

Proof: Chaining.

Dudley's inequality

Theorem 8.1.3 (Dudley's integral inequality)**.** *Let* $(X_t)_{t\in T}$ *be a mean zero random process on a metric space* (T,d) *with sub-gaussian increments as in* (8.1)*. Then*

$$\mathbb{E}\sup_{t\in T} X_t \le CK\int_0^\infty \sqrt{\log \mathcal{N}(T,d,\varepsilon)}\,d\varepsilon.$$

$$\mathbb{E}\sup_{t\in T} X_t \le CK\sum_{k\in\mathbb{Z}} 2^{-k}\sqrt{\log \mathcal{N}(T,d,2^{-k})}.$$

Note: there is a **gap** between Dudley's inequality (upper bound) and Sudakov's inequality (lower bound).
Proof: Chaining (multi-scale version of covering ϵ-Net.).

A false covering number approach

$$\mathbb{E}\sup_{t\in T} X_t \le \mathbb{E}\sup_{t\in T} X_{\pi(t)} + \mathbb{E}\sup_{t\in T}(X_t - X_{\pi(t)}).$$

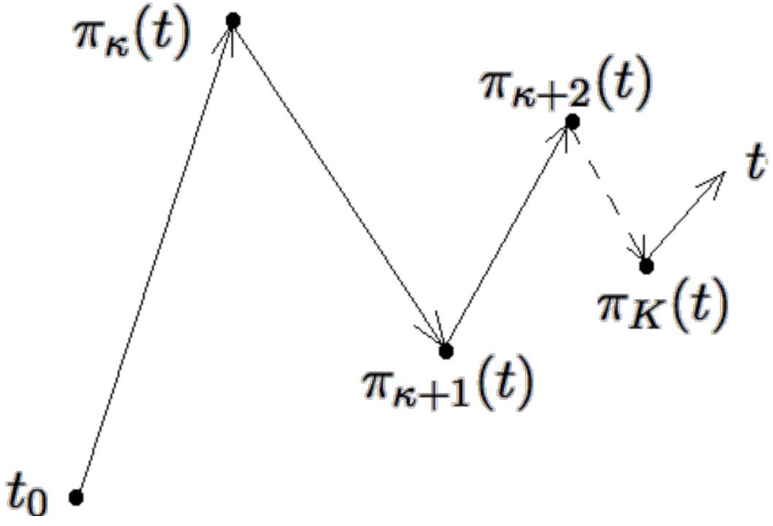
Union bound for the first term

Covering number for the second term $\|X_t - X_{\pi(t)}\|_{\psi_2} \le K\varepsilon.$

However, 'sup' term block the way, since 'point convergence is not uniform convergence'

(lead to a $\sqrt{\log |T|}$ bound)

Chaining method



这里虽然还有sup，
但是集合都有限，
且远小于$|T|$

$$\mathbb{E}\sup_{t\in T}(X_t - X_{t_0}) \le \sum_{k=\kappa+1}^{K} \mathbb{E}\sup_{t\in T}\boxed{(X_{\pi_k(t)} - X_{\pi_{k-1}(t)}).}$$

Instead of considering only one covering set, we consider a chain.
During the chain, the point get closer (not strictly) to t step by step.
Intuition: uniform convergence requires the convergence rate of each point.

$$\mathbb{E}\sup_{t\in T}(X_{\pi_k(t)} - X_{\pi_{k-1}(t)}) \le C\varepsilon_{k-1}\sqrt{\log|T_k|}.$$

Dudley's inequality (tail bound version)

Theorem 8.1.6 (Dudley's integral inequality: tail bound). *Let* $(X_t)_{t \in T}$ *be a random process on a metric space* (T, d) *with sub-gaussian increments as in* (8.1). *Then, for every* $u \geq 0$, *the event*

$$\sup_{t,s \in T} |X_t - X_s| \leq CK \Big[\int_0^\infty \sqrt{\log \mathcal{N}(T, d, \varepsilon)}\, d\varepsilon + u \cdot \mathrm{diam}(T) \Big]$$

holds with probability at least $1 - 2\exp(-u^2)$.

Dudley's inequality (Remark)

Remark 8.1.9 (Limits of Dudley's integral)**.** Although Dudley's integral is formally over $[0, \infty]$, we can clearly make the upper bound equal the diameter of T in the metric d, thus

$$\mathbb{E} \sup_{t \in T} X_t \leq CK \int_0^{\operatorname{diam}(T)} \sqrt{\log \mathcal{N}(T, d, \varepsilon)}\, d\varepsilon. \tag{8.13}$$

Indeed, if $\varepsilon > \operatorname{diam}(T)$ then a single point (any point in T) is an ε-net of T, which shows that $\log \mathcal{N}(T, d, \varepsilon) = 0$ for such ε.

Dudley's inequality (Not tight)

Exercise 8.1.12 (Dudley's inequality can be loose). ☕☕☕ Let $e_1, \ldots, e_n$ denote the canonical basis vectors in $\mathbb{R}^n$. Consider the set

$$T := \left\{ \frac{e_k}{\sqrt{1+\log k}},\ k = 1, \ldots, n \right\}.$$

(a) Show that

$$w(T) \le C,$$

where as usual C denotes an absolute constant.

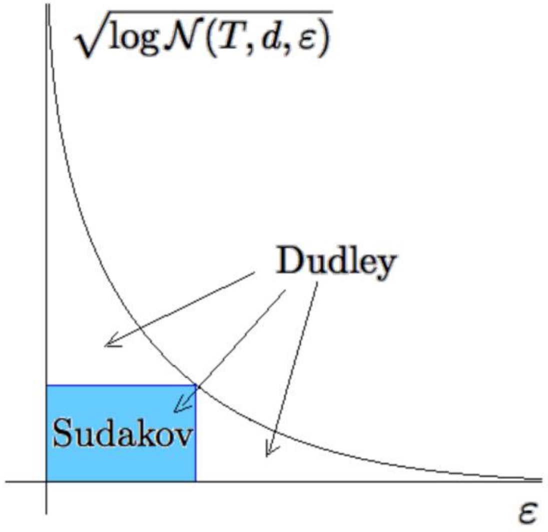
Hint: This should be straightforward from Exercise 2.5.10.

(b) Show that

$$\int_0^\infty \sqrt{\log \mathcal{N}(T, d, \varepsilon)}\, d\varepsilon \to \infty$$

as $n \to \infty$.

Hint: The first m vectors in T form a $(1/\sqrt{\log m})$-separated set.



Note: Sodakov's inequality (upper bound) is derived from Dudley's inequality, which is not tight (sub-opt up to log n).

Take-away Messages

1. Dudley's inequality. The upper bound of $\mathbb{E} \sup_{t \in T} X_t$
2. Chaining method: extension to covering numbers.

@ 滕佳烨

Thanks!