Generic chaining

Note Title 2015-10-19

Recall: we used <u>symmetrization</u> to give

$$\mathbb{E}\left\|\frac{1}{m}\sum_{i=1}^{m} a_i a_i^T - I\right\|_S \le 2\,\mathbb{E}\left\|\frac{1}{m}\sum_{i=1}^{m} \underset{\substack{\uparrow\\ \text{Rademacher}}}{\xi_i}\, a_i a_i^T\right\|_S$$

This is an example use of:

Lemma (Symmetrization)

> Let $V_1, \ldots, V_m$ be a sequence of indep. random vectors from a vector space V with norm $\|\cdot\|$. Let $F: V \to \mathbb{R}$ be a convex function. Let $\xi_1, \ldots \xi_m$ be iid Rademacher. Then
>
> $$\mathbb{E}\, F\left(\sum_{i=1}^{m} V_i - \mathbb{E} V_i\right) \le \mathbb{E}\, F\left(2\sum_{i=1}^{m} \xi_i V_i\right).$$
>
> In particular,
>
> $$\left(\mathbb{E}\left\|\sum_{i=1}^{n} V_i - \mathbb{E}V_i\right\|^p\right)^{1/p} \le 2\left(\mathbb{E}\left\|\sum_{i=1}^{n} \xi_i V_i\right\|^p\right)^{1/p}$$

<u>Proof</u>: Let $V_1', \ldots V_n'$ be an indep. copy of $V_1, \ldots, V_n$. Observe:

$$\begin{aligned}
\mathbb{E}\, F\left(\sum_{i=1}^{m} V_i - \mathbb{E}V_i\right) &= \mathbb{E}\, F\left(\sum_{i=1}^{m} V_i - \mathbb{E}V_i'\right) \\
&\le \mathbb{E}\, F\left(\sum_{i=1}^{m} V_i - V_i'\right) \quad \text{(By Jensen + independence)} \\
&= \mathbb{E}\, F\left(\sum_{i=1}^{m} \xi_i (V_i - V_i')\right) \\
&\le \mathbb{E}\left[\frac{1}{2} F\left(\sum_{i=1}^{n} 2\xi_i V_i\right) + \frac{1}{2} F\left(\sum_{i=1}^{m} 2\xi_i V_i'\right)\right] \quad \text{(by convexity)} \\
&= \mathbb{E}\, F\left(\sum_{i=1}^{m} 2\xi_i V_i\right)
\end{aligned}$$

QED

1.4 Generic chaining

Let $(X_t)_{t\in T}$ be a sub-Gaussian process.

Goal: Bound $\mathbb{E} \sup_{t\in T} X_t$

Kolmogorov: Use chaining argument

+ over 50 years of work

$\Rightarrow$ Fernique + Talagrand: $\mathbb{E} \sup_{t\in T} X_t$ is controlled precisely by $\gamma_2(T,d)$ (to be defined).

<u>Thm</u> (<u>Generic chaining</u>)

<u>Upper bound:</u> Let X_t, $t\in T$, be a centered sub-Gaussian process w/ associated metric $d(s,t)$, i.e.,

$$\|X_s - X_t\|_{\psi_2} \le d(s,t).$$

Then, $\mathbb{E} \sup_{t\in T} X_t \le C\cdot \gamma_2(T,d)$.

<u>Lower bound:</u> Let Y_t, $t\in T$, be a centered Gaussian process w/ <u>(tight) associated</u> metric: $d(s,t) := \|Y_s - Y_t\|_{\psi_2} = c\sqrt{\mathbb{E}(Y_s - Y_t)^2}$

Then, $\mathbb{E} \sup_{t\in T} Y_t \ge c\,\gamma_2(T,d)$

<u>Corollary</u> (<u>Gaussian-sub-Gaussian comparison</u>) Let (X_t) Y_t, $t\in T$, be a centered (sub-)Gaussian process. Suppose

$$\boxed{\begin{array}{l} \|X_t - X_s\|_{\psi_2} \le C\|Y_t - Y_s\|_{\psi_2} = C'\sqrt{\mathbb{E}(Y_t - Y_s)^2}. \\ \text{Then } \ \mathbb{E}\sup_{t\in T} X_t \le C\,\mathbb{E}\sup_{t\in T} Y_t \end{array}}$$

Implication: Simple proofs & general results controlling $\|A_t\|_2$, $t \in T \subseteq S^{n-1}$ by comparing to Gaussian process.

Intuition for upper bound

Assume, for simplicity, that T is finite.
First: Convert expectation into probability estimates.
Observe:

- $\mathbb{E}\sup_{t\in T} X_t = \mathbb{E}\sup_{t\in T} X_t - \mathbb{E}X_{t_0} = \mathbb{E}\sup_{t\in T} X_t - X_{t_0}$, for (any) $t_0 \in T$

- $\sup_{t\in T} X_t - X_{t_0} \ge 0$

$$\Rightarrow \mathbb{E}\sup_{t\in T} X_t \le \int_0^\infty \underbrace{P\left(\sup_{t\in T} X_t - X_{t_0} > u\right)}_{\text{we control this quantity.}} du$$

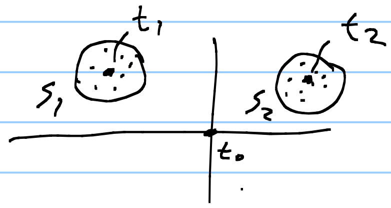
First attempt: Simple union bound

$$P\left(\sup_{t\in T} X_t - X_{t_0} > u\right) \le \sum_{t\in T} P\left(X_t - X_{t_0} > u\right)$$

(Abysmal bound if $\{X_t\}$ are clustered)

Soln: Chaining.

Motivating picture: $T =$



$$P\left(\sup_{t\in T} X_t - X_{t_0} > u\right) \le P\left(X_{t_1} - X_{t_0} > \tfrac{u}{2}\right) + \sum_{t\in S_1} P\left(X_t - X_{t_1} > \tfrac{u}{2}\right)$$
$$+ P\left(X_{t_2} - X_{t_0} > \tfrac{u}{2}\right) + \sum_{t\in S_2} P\left(X_t - X_{t_2} > \tfrac{u}{2}\right)$$

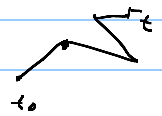
Generic chaining: An optimal method of chaining for any T, d.

We make a sequence of finer approximations to our set $T_0 \subseteq T_1 \subseteq T_2 \subseteq \dots \subseteq T$ satisfying $T_0 = \{t_0\}$ & , for some $m \geq 0$, $T_n = T$ for $n \geq m$.

Closest point map: $\pi_n : T \to T_n$ satisfies $d(\pi_n(t), t) \leq d(t, T_n)$ $\forall t \in T$.

Observe, for any $t \in T$,

$$X_t - X_{t_0} = \sum_{n \geq 0} X_{\pi_n(t)} - X_{\pi_{n-1}(t)} =$$



How to choose T_n?

Successively finer covers $\to$ Dudley's $\neq$. X

Instead we control cardinality $|T_n|$.

<u>Heuristic:</u> Let $g_1, \dots g_m \overset{iid}{\sim} N(0,1)$.

Then $\max_i g_i \approx \sqrt{2 \log m}$

If we take $g_1, \dots g_{m^2} \Rightarrow \max_i g_i \approx \sqrt{2 \log m^2} = \sqrt{2}\sqrt{2 \log m}$

i.e., squaring # of pts $\Rightarrow$ extra constant

Thus let $|T_{n+1}| \approx |T_n|^2 \Rightarrow |T_n| \leq 2^{2^n}$.

$\Rightarrow$ <u>Thm</u> Let X_t, $t \in T$, be a centered Gaussian process, and let $T_0 \subseteq T_1 \subseteq T_2 \subseteq \dots \subseteq T$ be a sequence of subsets s.t. $|T_0| = 1$, $|T_n| \leq 2^{2^n}$ for $n \geq 1$. Then,

$$\left\| \sup_{t \in T} X_t - X_{t_0} \right\|_{\psi_2} \leq C \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} d(t, T_n)$$

<u>Def</u> $\gamma_2(T, d) = \inf \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} d(t, T_n)$

where the inf is taken over

$T_1, T_2, \ldots$ satisfying requirements of Thm above.

Corollary Upper bound of generic chaining thm.

Proof (of thm)

Define the "good" event

$$\Omega_u := \{\forall n \geq 1,\ t \in T : |X_{\pi_n(t)} - X_{\pi_{n-1}(t)}| \leq u 2^{n/2} d(\pi_n(t), \pi_{n-1}(t))\}$$

If Ω_u occurs, then, for any $t \in T$,

$$X_t - X_{t_0} = \sum_{n \geq 1} X_{\pi_n(t)} - X_{\pi_{n-1}(t)}$$
$$\leq u \sum_{n \geq 1} 2^{n/2} d(\pi_n(t), \pi_{n-1}(t))$$
$$\leq u \sum_{n \geq 1} 2^{n/2} 2 \cdot d(\pi_{n-1}(t), t)$$

$$\Rightarrow \sup X_t - X_{t_0} \leq C \cdot u \sup_{t \in T} \sum_{n \geq 0} 2^{n/2} d(\pi_n(t), T)$$

We will show that $P(\Omega_u^c) \leq c e^{-u^2}$, thus proving the theorem.

Control $P(\Omega_u^c)$ w/ union bound:

For fixed n, t:

$$P(|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}| \geq u 2^{n/2} d(\pi_n(t), \pi_{n-1}(t)))$$
$$\leq C \exp(-u^2 2^n)$$ by sub-Gauss assumption

Note: $\pi_n(t) \subseteq T_n$, $\pi_{n-1}(t) \leq T_{n-1}$

$$\Rightarrow P(\Omega_u^c) \leq c \sum_{n \geq 1} |T_n||T_{n-1}| \exp(-u^2 2^n)$$

$$\leq c \sum_{n \geq 1} 2^{2^{n+1}} \exp(-u^2 2^n)$$

$$= c \sum_{n \geq 1} \exp(2^{n+1} \log 2 - u^2 2^n)$$ (decreases geometrically for $u > c$)

$$\leq c e^{-2u^2}$$ for $u > c$.

QED!