Upper and lower bounds for suprema of canonical processes

Tomasz Tkocz

Mathematics Department
Princeton University

based on joint work with
R. Latała

22 June 2016

Given

- independent random variables with mean 0, variance 1, $X_1, X_2, \ldots$
- a finite set $T \subset \mathbb{R}^n$

we define a **canonical process** $(X_t)_{t \in T}$ (based on X_i) as

$$X_t = \langle X, t \rangle = \sum_{i=1}^{n} t_i X_i .$$

How big is $\mathbb{E} \sup_{t \in T} X_t$?

If the X_i or T are symmetric,

$$2\mathbb{E} \sup_{t \in T} X_t = \mathbb{E} \sup_{t \in T} X_t + \mathbb{E} \sup_{s \in T} (-X_s) = \mathbb{E} \sup_{s,t \in T} (X_s - X_t).$$

How big is $\mathbb{E}\sup_{t\in T} X_t$?

EXAMPLE (Talagrand's majorising measure theorem). When X_i are independent standard Gaussians G_i,

$$\frac{1}{C}\gamma_2(T) \le \mathbb{E}\sup_{t\in T} G_t \le C\gamma_2(T),$$

$\gamma_2(T)$ is a quantity depending only on the geometry of (T, d_2), $d_2(s,t) = \left(\mathbb{E}|G_s - G_t|^2\right)^{1/2} = \|s-t\|$.

GOAL. Obtain similar bounds for other canonical processes.

The size of $\mathbb{E}\sup_{t\in T} X_t$ has also been understood in other important cases

- $\mathbb{P}(X_i = \pm 1) = 1/2$ (Bednorz, Latała '14)
- X_i with density $c_p e^{-|x|^p}$, $p \in [1, \infty)$ (Talagrand '94)
- X_i with log-concave tails ($N_i(t) = -\ln \mathbb{P}(|X_i| > t)$ convex) (Latała '97)

Today we discuss an extension of the log-concave tail case to the variables with *regularly growing* moments

$$\|X_i\|_p = (\mathbb{E}|X_i|^p)^{1/p}, \quad p \geq 1.$$

Easy upper bound

$$\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le \left(\mathbb{E}\sup_{s,t\in T}|X_s - X_t|^p\right)^{1/p}$$
$$\le \left(\mathbb{E}\sum_{s,t\in T}|X_s - X_t|^p\right)^{1/p}$$
$$\le |T|^{2/p}\sup_{s,t\in T}\|X_s - X_t\|_p.$$

If $|T| \le e^p$, then

$$\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le e^2 \cdot \mathrm{diam}_p(T),$$

$d_p(s,t) = \|X_s - X_t\|_p$.

Union bound: if $|T| \le e^p$, then

$$\mathbb{E} \sup_{s,t \in T} (X_s - X_t) \le e^2 \cdot \mathrm{diam}_p(T).$$

On the other hand,

DEF. $(X_t)_{t \in T}$ satisfies the **Sudakov minoration principle (SMP)** with a constant κ if for every $p \ge 1$ and $S \subset T$ such that

$$|S| \ge e^p \quad \text{and} \quad d_p(s_1, s_2) = \|X_{s_1} - X_{s_2}\|_p \ge u \text{ on } S$$

we have

$$\mathbb{E} \sup_{s_1, s_2 \in S} (X_{s_1} - X_{s_2}) \ge \kappa u.$$

$$\text{for every } |S| \geq e^p \text{ such that } d_p(s_1, s_2) \geq u \text{ on } S, \qquad \text{(SMP)}$$

$$\mathbb{E} \sup_{s_1, s_2 \in S} (X_{s_1} - X_{s_2}) \geq \kappa u,$$

EXAMPLE (Gaussians satisfy (SMP)). The classical Sudakov's theorem states that

$$\mathbb{E} \sup_{s,t \in T} (G_s - G_t) \geq c \sup_{\varepsilon > 0} \varepsilon \sqrt{\ln N(T, d_2, \varepsilon)},$$

$N(T, d, \varepsilon)$ being the size of the smallest ε-net of (T, d_2).
If $S \subset T$ with $|S| \geq e^p$ and $d_p(s_1, s_2) \geq u$ on S, then

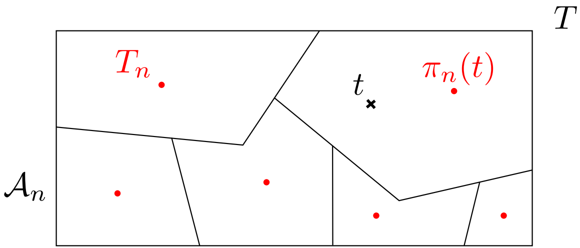
$$\sqrt{\ln |S|} d_2(s_1, s_2) \geq \sqrt{p} d_2(s_1, s_2) \approx d_p(s_1, s_2) \geq u$$

so every two points of S are $\frac{u}{\sqrt{\ln |S|}}$ d_2-apart, so for $\varepsilon = \frac{1}{2} \frac{u}{\sqrt{\ln |S|}}$, $N(S, d_2, \varepsilon) = |S|$ and we get

$$\mathbb{E} \sup_{s_1, s_2 \in S} (G_{s_1} - G_{s_2}) \geq c \frac{1}{2} \frac{u}{\sqrt{\ln |S|}} \sqrt{\ln |S|} = \frac{c}{2} u.$$

Improvement on the easy upper bound: **chaining**.

- Let $\mathcal{A} = (\mathcal{A}_n)_{n\geq 0}$ be an increasing sequence of partitions of T such that $\mathcal{A}_0 = \{T\}, |\mathcal{A}_n| \leq 2^{2^n}$.
- Form T_n by choosing a representative from every set in $\mathcal{A}_n$. Particularly, $|T_n| = |\mathcal{A}_n| \leq 2^{2^n}$.
- For $t \in T$ let $\pi_n(t) \in T_n$ be the unique point such that t and $\pi_n(t)$ are in the same set in $\mathcal{A}_n$.
- chain: $X_t - X_{\pi_1(t)} = \sum_{n\geq 1}(X_{\pi_{n+1}(t)} - X_{\pi_n(t)})$



Chain: $X_t - X_{\pi_1(t)} = \sum_{n\geq 1}(X_{\pi_{n+1}(t)} - X_{\pi_n(t)})$.
Recall: $d_p(s,t) = (\mathbb{E}|X_s - X_t|^p)^{1/p}$. Define

$$\Omega_{n,t,u} = \left\{|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \leq u \cdot d_{2^n}(\pi_{n+1}(t), \pi_n(t))\right\},$$
$$\Omega_u = \bigcap_{n\geq 1}\bigcap_{t}\Omega_{n,t,u}.$$

By Chebyshev: $\mathbb{P}\left(\Omega^c_{n,t,u}\right) \leq u^{-2^n}$, so

$$\mathbb{P}\left(\Omega^c_u\right) \leq \sum_{n\geq 1} \underbrace{|T_{n+1}|}_{\leq 2^{2^{n+1}}} \underbrace{|T_n|}_{\leq 2^{2^n}} \cdot u^{-2^n} \leq \sum_{n\geq 1}\left(\frac{8}{u}\right)^{2^n} \underset{u\geq 16}{\leq} \frac{128}{u^2}.$$

On Ω_u

$$\sup_{t\in T}|X_t - X_{\pi_1(t)}| \leq \sup_{t\in T}\sum_{n\geq 1}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}|$$
$$\leq u \cdot \underbrace{\sup_{t\in T}\sum_{n\geq 1} d_{2^n}(\pi_{n+1}(t), \pi_n(t))}_{S}.$$

On Ω_u: $\sup_{t\in T} |X_t - X_{\pi_1(t)}| \leq u \cdot S$, therefore

$$\mathbb{P}\left(\frac{1}{S}\sup_{t\in T}|X_t - X_{\pi_1(t)}| > u\right) \leq \mathbb{P}\left(\Omega_u^c\right) \leq \frac{128}{u^2}, \quad u \geq 16.$$

Integrating this tail estimate gives $\mathbb{E}\sup_{t\in T}|X_t - X_{\pi_1(t)}| \leq 129 \cdot S$.
Since

$$\sup_{s,t\in T}(X_s - X_t) \leq \sup_{s\in T}(X_s - X_{\pi_1(s)}) + \sup_{s,t\in T}(X_{\pi_1(s)} - X_{\pi_1(t)}) + \sup_{t\in T}(X_{\pi_1(t)} - X_t),$$

$$\begin{aligned} \mathbb{E}\sup_{s,t\in T}(X_s - X_t) &\leq 2\cdot 129S + \mathbb{E}\sup_{s,t\in T}|X_{\pi_1(s)} - X_{\pi_1(t)}| \\ &\leq 258S + \underbrace{|T_1|^2}_{\leq 16}\operatorname{diam}_{d_1}(T) \end{aligned}$$

$$\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le 258(S + \mathrm{diam}_{d_1}(T)),$$

where

$$S = \sup_{t\in T}\sum_{n\ge 1} d_{2^n}(\underbrace{\pi_{n+1}(t),\ \pi_n(t)}_{\substack{\text{in the same unique set}\\ A_n(t)\in\mathcal{A}_n \text{ which contains } t}}) \le \sup_{t\in T}\sum_{n\ge 1}\mathrm{diam}_{2^n}(A_n(t))$$

and incorporating in this sum the term $\mathrm{diam}_{d_1}(T)$ leads to

DEF. $\gamma_X(T) = \inf_{\mathcal{A}}\sup_{t\in T}\sum_{n\ge 0}\mathrm{diam}_{2^n}(A_n(t))$.

FACT (Mendelson⊥Latała). For any canonical process $(X_t)_{t\in T}$,

$$\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le 258\cdot\gamma_X(T).$$

Chaining upper bound:

$$\mathbb{E} \sup_{s,t\in T} (X_s - X_t) \le 258 \cdot \gamma_X(T),$$

$$\gamma_X(T) = \inf_{\mathcal{A}} \sup_{t\in T} \sum_{n\ge 0} \mathrm{diam}_{2^n}(A_n(t)).$$

⚠ When $X = G$ is Gaussian, $\mathrm{diam}_{2^n} \approx \sqrt{2^n}\,\mathrm{diam}$, so

$$\gamma_G(T) \approx \gamma_2(T) = \inf_{\mathcal{A}} \sup_{t\in T} \sum_{n\ge 0} 2^{n/2}\,\mathrm{diam}(A_n(t)).$$

Two upper bounds

- union bound: $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le e^2 \cdot \mathrm{diam}_p(T)$, $|T| \le e^p$
- its chaining refinement: $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le 258 \cdot \gamma_X(T)$

Two goals

GOAL. Establish the γ_X-lower bound:

$$\mathbb{E} \sup_{s,t\in T} (X_s - X_t) \ge c \cdot \gamma_X(T).$$

Easier GOAL. Establish the SMP: for every $S \subset T$ such that $|S| \ge e^p$ and $d_p(s, s') \ge u$ on S

$$\mathbb{E} \sup_{s,s'\in S} (X_s - X_{s'}) \ge \kappa u.$$

GOAL. $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \geq c \cdot \gamma_X(T)$.

Easier GOAL. $\mathbb{E}\sup_{s,s'\in S}(X_s - X_{s'}) \geq \kappa u$ for every $|S| \geq e^p$ such that $d_p(s,s') \geq u$ on S

⚠ The γ_X-lower bound implies the SMP (for independent X_i).

Let n be such that $2^n \leq p < 2^{n+1}$. Consider $\tilde{S} = S \times S$. For $s = (s_1, s_2) \in \tilde{S}$, $X_s = X_{s_1} + X_{s_2}$ and

$$2\mathbb{E}\sup_{s,s'\in S}(X_s - X_{s'}) = \mathbb{E}\sup_{s_1,s_1'\in S}(X_{s_1} - X_{s_1'}) + \mathbb{E}\sup_{s_2,s_2'\in S}(X_{s_2} - X_{s_2'})$$

$$\geq \mathbb{E}\sup_{s,s'\in\tilde{S}}(X_s - X_{s'}) \geq c\gamma_X(\tilde{S}).$$

Since $|\tilde{S}| = |S|^2 \geq e^{2p} > 2^{2^{n+1}} \geq |\mathcal{A}_{n+1}|$, there is $A \in \mathcal{A}_{n+1}$ which contains two points $a = (a_1, a_2)$ and $b = (b_1, b_2)$ from $\tilde{S}$. Then

$$\gamma_X(\tilde{S}) \geq \mathrm{diam}_{2^{n+1}}(A) \geq \|X_a - X_b\|_p \underset{\substack{X_i \text{ indep.} \\ \text{mean } 0}}{\geq} \|X_{a_1} - X_{b_1}\|_p \geq u.$$

	SMP	γ_X-lower bound	
X_i Gaussian	✓	✓	Sudakov '69 Talagrand '89
$\mathbb{P}(X_i = \pm 1) = 1/2$	✓	✗	Talagrand '93
X_i with density $e^{-\lvert t\rvert^q}$, $q \in [1, \infty)$	✓	✓	Talagrand '94
X_i with log-concave tails T_i	✓		Latała '97
and $T_i(2t) \geq T_i(t)^\gamma$		✓	

SMP in the dependent setting thoroughly studied by Latała '14, Bednorz '14 (log-concavity, concentration).

If a random variable Y has a log-concave tail $T(t) = \mathbb{P}(|Y| > t)$, then

$$\|Y\|_p \le \alpha \frac{p}{q} \|Y\|_q, \quad p \ge q \ge 1.$$

Moreover, if T does not decay too fast, say $T(2t) \ge T(t)^\gamma$, then for some $\beta = \beta(\gamma) > 1$

$$\|Y\|_{\beta p} \ge 2 \|Y\|_p, \quad p \ge 1.$$

DEF. The moments of Y grow α-regularly if

$$\|Y\|_p \le \alpha \frac{p}{q} \|Y\|_q, \quad p \ge q \ge 2.$$

The moments of Y grow with speed β if

$$\|Y\|_{\beta p} \ge 2\|Y\|_p, \quad p \ge 2.$$

Let $X_1, X_2, \ldots$ be independent random variables with mean 0, variance 1.

THM 1. If the moments of the X_i grow α-regularly, then $(X_t)_{t\in T}$ satisfies the SMP.

THM 2. If the moments of the X_i grow α-regularly and with speed β, then

$$c(\alpha,\beta)\cdot\gamma_X(T) \le \mathbb{E}\sup_{s,t\in T}(X_s - X_t) \le 258\cdot\gamma_X(T).$$

By Latała '97, we know THM 1 when $N_i(t) = -\ln \mathbb{P}(|X_i| > t)$ are convex and THM 2 when additionally, $N_i(2t) \leq \gamma N_i(t)$.

Since α-regularity of moments implies that

$$N_i(C_\alpha \lambda t) \geq \lambda N_i(t), \quad \lambda \geq 1, t \geq 1 - 1/e,$$

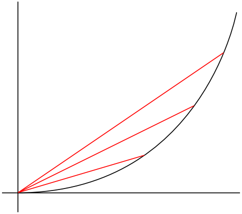
the functions N_i are *almost* convex. In fact,

LEMMA. Let $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ satisfy

$$f(\lambda t) \geq \lambda f(t), \quad \lambda \geq 1, t \geq 0.$$

There is a convex function $g\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$g(t) \leq f(t) \leq g(2t), \quad t \geq 0.$$



LEMMA. Let $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $f(\lambda t) \geq \lambda f(t)$, $\lambda \geq 1, t \geq 0$. There is a convex function $g\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $g(t) \leq f(t) \leq g(2t)$, $t \geq 0$.

Proof. The assumption means that $s \mapsto \frac{f(s)}{s}$ is nondecreasing. Define

$$g(t) = \int_0^t \frac{f(s)}{s}\mathrm{d}s \leq t\frac{f(t)}{t} = f(t).$$

Then g is convex as an integral of a nondecreasing function. Moreover,

$$g(2t) \geq \int_t^{2t} \frac{f(s)}{s}\mathrm{d}s \geq t\frac{f(t)}{t} = f(t).$$

Thus the α-regularity of the moments of X_i essentially assures that $N_i(t) = -\ln \mathbb{P}(|X_i| > t)$ is sandwiched between a nondecreasing convex function $M_i(t)$ and $M_i(C_\alpha t)$.

Take a random variable Y_i with the log-concave tail e^{-M_i}. Then sandwiching implies that there are realisations such that

$$|Y_i| \geq |X_i| \geq \frac{1}{C_\alpha}|Y_i|$$

and bounds for $\mathbb{E}\sup_t Y_t$ will carry over for $\mathbb{E}\sup_t X_t$.

If additionally, the moments of X_i grow with speed β, then N_i does not grow too fast, $N_i(2t) \leq \gamma N_i(t)$, $t \geq 2$ and the sandwiching gives essentially the same for M_i. This gives the γ_Y-lower bound for Y.

FACT. If a canonical process X based on i.i.d. random variables satisfies the SMP, then the moments of the X_i grow α-regularly.

Proof. Consider $S = \left\{ s \in \{0,1\}^n,\ \sum_{i=1}^n s_i = m \right\}$ for $n \geq m \geq 1$. Fix $p \geq q \geq 2$. We have $|S| = \binom{n}{m} \geq \left(\frac{n}{m}\right)^m \geq e^p$ if $n \geq me^{p/m}$. If $s \neq s'$, say $s_j \neq s'_j$, then $\|X_s - X_{s'}\|_p \geq \|X_j\|_p$ and the SMP yields

$$\kappa\|X_1\|_p \leq \mathbb{E} \sup_{s,s' \in S} (X_s - X_{s'}) \leq 2\mathbb{E} \sup_{s \in S} |X_s| \leq 2\mathbb{E} \sum_{k=1}^m X_k^*,$$

where (X_k^*) is the nonincreasing rearrangement of $(|X_k|)$. Standard arguments give $\mathbb{E}X_k^* \leq C\left(\frac{n}{k}\right)^{1/q} \|X_1\|_q$, so

$$\kappa\|X_1\|_p \leq Cn^{1/q}m^{1-1/q}\|X_1\|_q \underset{\substack{m=\lceil p/q \rceil \\ n=\lceil me^{p/m} \rceil}}{\leq} \frac{C}{\kappa}\frac{p}{q}\|X_1\|_q.$$

THANK YOU