Bounding suprema of canonical processes via convex hull

Rafał Latała (University of Warsaw)

Probabilistic Methods in Geometry and Analysis
ICERM, October 20 2022

Formulation of the problem

Let $X = (X_1, \ldots, X_n)$ be a centered random vector with independent coordinates. Define the canonical process X_t by

$$X_t = \langle t, X \rangle = \sum_{i=1}^{n} t_i X_i \quad \text{for } t = (t_1, \ldots, t_n) \in \mathbb{R}^n.$$

Our aim is to estimate the expected value of the supremum of the process $(X_t)_{t \in T}$, i.e. the quantity

$$b_X(T) := \mathbb{E} \sup_{t \in T} X_t, \quad T \subset \mathbb{R}^n \text{ nonempty bounded.}$$

Gaussian processes - γ_2-functional

In the case X_i are i.i.d. $\mathcal{N}(0,1)$ r.v's, X_t is the canonical Gaussian process, $\mathbb{E}X_t = 0$, $\mathrm{Cov}(X_t, X_s) = \langle t, s\rangle$.
The crucial tool to estimate suprema of Gaussian processes are chaining methods, developed over the decades and culminating in the famous majorizing measure theorem.

Theorem (Fernique'75 Talagrand'87)

For the canonical Gaussian process $(X_t)_{t\in T}$ we have

$$b_X(T) \sim \gamma_2(T) := \inf_{\mathcal{A}} \sup_{t\in T} \sum_{n\geq 0} 2^{n/2}\Delta_2(A_n(t)),$$

where the infimum runs over all admissible sequences of partitions $\mathcal{A} = (\mathcal{A}_n)_{n\geq 0}$ and $\Delta_2(A)$ is the Euclidean diameter of A.

$f \sim g$ means that $\frac{1}{C}f \leq g \leq Cf$ for an universal C.
Sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ is admissible if it is increasing, $\mathcal{A}_0 = \{T\}$ and $|\mathcal{A}_n| \leq 2^{2^n}$.
$A_n(t)$ denotes the unique set in $\mathcal{A}_n$ containing t.

Chaining for other processes - γ_X functional

The chaining bound may be extended to more general classes of canonical processes, but to this end one needs to introduce a class of L_p-distances and L_p-diameters, $p \geq 1$,

$$d_p(s,t) := \|X_t - X_s\|_p = (\mathbb{E}|X_t - X_s|^p)^{1/p}, \quad \Delta_p(A) := \sup_{s,t \in A} d_p(s,t).$$

Define

$$\gamma_X(T) := \inf_{\mathcal{A}} \sup_{t \in T} \sum_{n \geq 0} \Delta_{2^n}(A_n(t)),$$

where as before the infimum runs over all admissible sequences of partitions $\mathcal{A} = (\mathcal{A}_n)_{n \geq 0}$.

In the Gaussian case $d_p(s,t) \sim \sqrt{p} d_2(s,t) = \sqrt{p}|s-t|$, so $\Delta_{2^n}(A_n(t)) \sim 2^{n/2} \Delta_2(A)$ and $\gamma_X(T) \sim \gamma_2(T)$.

Chaining upper bound

Proposition

For any centered process (X_t), $b_X(T) = \mathbb{E}\sup_{t\in T} X_t \leq C\gamma_X(T)$.

Proof. Let us fix admissible sequence $(\mathcal{A}_n)_{n\geq 0}$, choose $T_n \subset T$ containing one element from each set in $\mathcal{A}_n$ and put $\pi_n(t) = s$ if $T_n \cap A_n(t) = \{s\}$. Then $T_0 = \{\pi_0(t)\}$,

$$S := \sup_{t\in T}\sum_{n\geq 0}\|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}\|_{2^n} \leq \sup_{t\in T}\sum_{n\geq 0}\Delta_{2^n}(A_n(t)).$$

In particular $X_{\pi_n(t)} \to X_t$ in any L_p. We need to show that $b_X(T) \leq CS$. We have

$$\begin{aligned} b_X(T) &= \mathbb{E}\sup_{t\in T}(X_t - X_{\pi_0(t)}) = \mathbb{E}\sup_{t\in T}\sum_{n\geq 0}(X_{\pi_{n+1}(t)} - X_{\pi_n(t)}) \\ &\leq \mathbb{E}\sum_{t\in T_1}|X_t - X_{\pi_0(t)}| + \mathbb{E}\sup_{t\in T}\sum_{n\geq 1}(X_{\pi_{n+1}(t)} - X_{\pi_n(t)}) \\ &\leq 4S + \mathbb{E}\sup_{t\in T}\sum_{n\geq 1}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}|. \end{aligned}$$

Proof of upper bound ctd

We have for $u > 16$

$$
\begin{aligned}
\mathbb{P}\Big(\sup_{t\in T}\sum_{n\geq 1}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \geq uS\Big)
&\leq \mathbb{P}(\exists_{n\geq 1}\exists_{t\in T}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \geq u\|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}\|_{2^n}) \\
&\leq \mathbb{P}(\exists_{n\geq 1}\exists_{s\in T_{n+1}}\exists_{t\in T_n}|X_s - X_t| \geq u\|X_s - X_t\|_{2^n}) \\
&\leq \sum_{n\geq 1}\sum_{s\in T_{n+1}}\sum_{t\in T_n}\mathbb{P}(|X_s - X_t| \geq u\|X_s - X_t\|_{2^n}) \\
&\leq \sum_{n\geq 1}|T_{n+1}||T_n|u^{-2^n} \leq \sum_{n\geq 1}\Big(\frac{8}{u}\Big)^{2^n} \leq 128u^{-2}.
\end{aligned}
$$

Integrating by parts we get

$$
\mathbb{E}\sup_{t\in T}\sum_{n\geq 1}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \leq CS.
$$

Two-sided bound in the regular case

Theorem

For certain classes of regular light-tail canonical processes $b_X(T) \sim \gamma_X(T)$.

- X_i with density $c_r e^{-|t|^r}$, $1 \leq r < \infty$ Talagrand'94
- X_i symetric with logconcave tails: $N_i(t) = -\log \mathbb{P}(|X_i| \geq t)$ convex staisfying Δ_2-condition $N_i(2t) \leq \gamma N_i(t)$ L'97
- X_i with regularily growing moments $\|X_i\|_p \leq \alpha \frac{p}{q} \|X_i\|_q$, $p \geq q \geq 1$ and $\|X_i\|_{\beta p} \geq 2\|X_i\|_p$, $p \geq 1$ L-Tkocz'15
- Bernoulli processes Bednorz-L'14 (one needs to modify definition of γ_X on the base of nets and with L_∞-distance)

Condition $\|X_i\|_p \leq \alpha \frac{p}{q} \|X_i\|_q$ is essentially necessarily (L-Tkocz), condition $\|X_i\|_{\beta p} \geq 2\|X_i\|_p$ most likely might be released.

Wide open question. Does two-sided bound holds for centered log-concave vectors (without assumption of coordinate independence)

Weibull processes with hard tails

One of the simplest examples of r.v.'s with hard tails are Weibull r.v's with the density $c_r e^{-|t|^r}$, $0 < r < 1$. In this case one may bound $b_X(T)$ using γ_2 and random permutations.

Theorem (Bogucki'15)

We have $b_X(T) \sim \mathbb{E}_\pi \gamma_2(T_\pi)$, where π is a random permutation of $\{1,\ldots,n\}$ and

$$T_\pi := \left\{ \left(t_{\pi(k)} \Big(\log\Big(\frac{en}{k}\Big)\Big)^{\frac{1}{2}-\frac{1}{r}} \right)_{k=1}^{n} : \ (t_1,\ldots,t_n) \in T \right\}.$$

Corollary (Bogucki'15)

If additionally the set T is invariant under permutations one has $b_X(T) = \gamma_2(T_s)$, where $\frac{1}{s} + \frac{1}{2} = \frac{1}{r}$ and

$$T_s := \left\{ \left(t_k \Big(\log\Big(\frac{en}{k}\Big)\Big)^{\frac{1}{s}} \right)_{k=1}^{n} : \ (t_1,\ldots,t_n) \in T \right\}.$$

Strenghts and drawbacks of chaining methods

- Upper bounds easy to show and always satisfied
- Two-sided bound valid only in regular light-tail cases
- In practise it is quite hard to construct right partitions, often Dudley-type bounds based on entropy are much easier to apply
- Great tool to compare subgaussian and Gaussian processes (or unconditional log-concave and exponential vectors)

Maybe one can find a geometric method which works in a larger generality?

Easy upper bound via convex hull

Suppose that there exists $t_0, t_1, \ldots \in \mathbb{R}^n$ such that

$$T - t_0 \subset \overline{\mathrm{conv}}\{\pm t_i \colon \ i \geq 1\},$$

then for any $u > 0$,

$$\mathbb{E}\sup_{t\in T} X_t = \mathbb{E}\sup_{t\in T} X_{t-t_0} \leq \mathbb{E}\sup_{i\geq 1}|X_{t_i}| \leq u + \sum_{i\geq 1}\mathbb{E}|X_{t_i}| I_{\{|X_{t_i}|\geq u\}}.$$

(we used $X_{t-t_0} = X_t - X_{t_0}$ and $\mathbb{E}X_{t_0} = 0$).
In other words

$$b_X(T) \leq M_X(S) \quad \text{if} \quad T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S),$$

where for nonemty countable sets $S \subset \mathbb{R}^n$

$$M_X(S) := \inf_{u>0}\Big[u + \sum_{t\in S}\mathbb{E}|X_t| I_{\{|X_t|\geq u\}}\Big]$$

Convex hull upper bound ctd

We may also use modified version of M_X,

$$\widetilde{M}_X(S) = \inf\Big\{m > 0\colon \ \sum_{t \in S} \mathbb{E}|X_t| I_{\{|X_t| \geq m\}} \leq m\Big\}.$$

It is easy to show that $\widetilde{M}_X(S) \leq M_X(S) \leq 2\widetilde{M}_X(S)$.
Therefore

$$b_X(T) \leq M_X(S) \leq 2\widetilde{M}_X(S) \quad \text{if} \quad T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S),$$

Question. When the above estimate may be reversed i.e. what should we assume about about variables X_i (and the set T) in order that

$$T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S) \quad \text{and} \quad M_X(S) \lesssim b_X(T) \tag{1}$$

for some $t_0 \in \mathbb{R}^n$ and nonempty countable set $S \subset \mathbb{R}^n$?

Chaining implies convex hull

Proposition

If $\|X_i\|_{2p} \le \alpha \|X_i\|_p$ *for* $p \ge 1$ *and* $\gamma_X(T) < \infty$ *then for any* $t_0 \in T$ *there exists* S *such that* $T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S)$ *and* $M_X(S) \le C(\alpha)\gamma_X(T)$.

Idea of the proof. Take an admissible partition sequence $(\mathcal{A}_n)_{n \ge 0}$ such that

$$\sup_{t \in T} \sum_{n \ge 0} \Delta_{2^n}(A_n(t)) \le 2\gamma_X(T)$$

Define $T_0 = \{t_0\}$ and for $n \ge 1$, T_n as a set containing one point in each element of $\mathcal{A}_n$. Let $A_n(t) \cap T_n = \{\pi_n(t)\}$. Then we may put

$$S := \left\{ 2\gamma_X(T) \frac{\pi_{n+1}(t) - \pi_n(t)}{\Delta_{2^n}(A_n(t))} : \ t \in T \right\}.$$

There are however examples when $\gamma_X(T) = \infty$, but convex hull method works.

Toy case: ℓ_1-ball

For $T = B_1^n$ the obvious choice $S = \{e_1, \dots, e_n\}$ always works.

Proposition

For arbitrary independent integrable r.v's $X_1, \dots, X_n$
$M_X(\{e_1, \dots, e_n\}) \le 4 b_X(B_1^n)$.

Proof. Let $u_0 := \inf\{u > 0\colon\ \mathbb{P}(\max_i |X_i| \ge u) \le \frac{1}{2}\}$, then

$$b_X(B_1^n) = \mathbb{E} \max_{1 \le i \le n} |X_i| = \int_0^\infty \mathbb{P}(\max_{1 \le i \le n} |X_i| \ge u) \mathrm{d}u$$
$$\ge \frac{1}{2} u_0 + \int_{u_0}^\infty \frac{1}{2} \sum_{i=1}^n \mathbb{P}(|X_i| \ge u) \mathrm{d}u = \frac{1}{2} u_0 + \frac{1}{2} \sum_{i=1}^n \mathbb{E}(|X_i| - u_0)_+.$$

Therefore

$$2u_0 + \sum_{i=1}^n \mathbb{E}|X_i| I_{\{|X_i| \ge 2u_0\}} \le 2u_0 + 2 \sum_{i=1}^n \mathbb{E}(|X_i| - u_0)_+ \le 4 b_X(B_1^n).$$

Counterexample for the convex hull method when $T = B_2^n$

Let $1 < p < 2$ and $X_1, \ldots, X_n$ be iid p-stable r.v's with characteristic function $\varphi_{X_k}(t) = \exp(-|t|^p)$.
Then one may show that

$$b_X(B_2^n) = \mathbb{E}|X| \sim_p n^{1/p}$$

and if $B_2^n \subset \overline{\mathrm{conv}}(S \cup -S)$ then

$$M_X(S) \gtrsim_p n^{2/p-1/2} \gg n^{1/p}$$

So some regularity assumptions are necessary for the convex hull method even in the Euclidean case.

Convex hull for B_2^n works under $4+\delta$ moment condition

It turns out that convex hull methods works in the case of Euclidean balls under the following $4+\delta$ moment condition:

$$\exists_{r>4,\lambda<\infty}\ (\mathbb{E}X_i^r)^{1/r} \le \lambda(\mathbb{E}X_i^2)^{1/2} < \infty \quad i = 1,\dots,n. \qquad (4+\delta)$$

Proposition

Let $X_1,\dots,X_n$ be independent centered r.v's with variance 1 satisfying $(4+\delta)$ condition. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \le 10n^2$, $B_2^n \subset \mathrm{conv}(S)$ and

$$M_X(S) \lesssim_{r,\lambda} \sqrt{n} \sim_\lambda \mathbb{E}|X| = b_X(B_2^n).$$

Crucial lemma

Lemma (Kochol'94)

For any $1 \leq k \leq n$ there exists $T \subset B_2^n$ with $|T| \leq \frac{2n}{k}5^k$ such that $B_2^n \subset 2\sqrt{\frac{2n}{k}}\mathrm{conv}(T)$.

Proof. Let $l = \lceil n/k \rceil \leq 2n/k$ and $\mathbb{R}^n = F_1 \oplus \cdots \oplus F_l$ be an orthogonal decomposition into spaces of dimension at most k. We can find $T_i \subset B_2(F_i) := B_2^n \cap F_i$ such that $B_2(F_i) \subset 2\mathrm{conv}(T_i)$ and $|T_i| \leq 5^k$. Let $T := \bigcup_{i \leq l} T_i$. Then $T \subset B_2^n$ and $|T| \leq l5^k \leq \frac{2n}{k}5^k$. Fix now $x \in B_2^n$ and x_i denotes its orthogonal projection on F_i. Observe that

$$\sum_{i \leq l} \|x_i\| \leq \sqrt{l}\Big(\sum_{i \leq l} \|x_i\|^2\Big)^{1/2} \leq \sqrt{l}.$$

Therefore

$$x \subset \sqrt{l}\mathrm{conv}\Big\{0, \frac{x_1}{\|x_1\|}, \ldots, \frac{x_l}{\|x_l\|}\Big\} \subset \sqrt{l}\mathrm{conv}\Big(\bigcup_{i \leq l} B_2(F_i)\Big) \subset 2\sqrt{l}\mathrm{conv}(T).$$

Kochol's lemma implies convex hull for B_2^n

To construct S such that $B_2^n \subset \mathrm{conv}(S \cup -S)$ and $M_X(S) \le C\sqrt{n}$ (under $4+\delta$ moment condition and variance 1) we first appply Kochol's lemma with $k \sim \log n$ and have $B_2^n \subset \mathrm{conv}\{t_1, \ldots, t_N\}$, $N \le 10n^2$ and $|t_i| \le C\sqrt{n/\log n}$, $1 \le i \le N$.

Then obviously for any $U \in O(n)$, $B_2^n \subset \mathrm{conv}\{Ut_1, \ldots, Ut_N\}$. To conclude we show that if U is random then with positive probability $M_X(\{Ut_1, \ldots, Ut_N\}) \le C\sqrt{n}$.

Convex hull for ellipsoid under $4 + \delta$ moment condition

One may extend the convex hull bound from B_2^n ball to arbitrary ellipsoid.

Theorem

Let $X_1, \ldots, X_n$ be independent centered r.v's satisfying $(4 + \delta)$ condition and let T be an ellipsoid in $\mathbb{R}^n$. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \leq 10n^2$, $T \subset \mathrm{conv}(S)$ and

$$M_X(S) \lesssim_{r,\lambda} b_X(T).$$

Convex hull for ℓ_q-balls, $q \geq 2$

One more extension goes to ℓ_q-balls.

Theorem

Let $X_1, \ldots, X_n$ be independent centered r.v's satisfying $(4+\delta)$ condition and let $T = AB_q^n$ for some $2 \leq q \leq \infty$ and an $n \times n$ matrix A. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \leq 10n^2$, $T \subset \mathrm{conv}(S)$ and

$$M_X(S) \lesssim_{r,\lambda} b_X(T).$$

Open questions

- Can the result for ℓ_q-balls could be extended to the case $1 < q < 2$?
- What about other sets with many symetries (unconditional, permutationally invariant)? Maybe Bogucki result would be of some help in the case of Weibull r.v's.
- Are there heavy-tailed random variables X_i such that the convex hull methods gives two-sided bound for any set T?
- Can one get two-sided bounds via convex hull/chaining methods for (some classes of) log-concave vectors?
- Is there a connection with the recent Park and Pham solution of Talagrand's conjecture about suprema of selector processes?

Thank you for your attention!