Upper and lower bounds for suprema of canonical processes

Rafał Latała (based on the joint work with Tomasz Tkocz)

University of Warsaw

Paris-Est Marne-la-Vallée, October 27 2015

Introduction. Canonical Processes

In many problems arising in probability theory and its applications one needs to estimate the supremum of a stochastic process. In particular it is very useful to be able to find two-sided bounds for the mean of the supremum. The modern approach to this challenge is based on the chaining methods (cf. the recent monograph of Michel Talagrand).

In this talk we discuss the class of *canonical processes* (X_t) of the form

$$X_t = \sum_{i=1}^{\infty} t_i X_i,$$

where X_i are independent random variables. If X_i are *standardized*, i.e. have mean zero and variance one, then this series converges a.s. for $t \in \ell^2$ and one may try to estimate $\mathbb{E} \sup_{t \in T} X_t$ for $T \subset \ell^2$.

For instance, in the case when X_i are i.i.d. $\mathcal{N}(0,1)$ r.v.s, X_t is the canonical Gaussian process. Moreover, any centred separable Gaussian process has the Karhunen-Loève representation of such form.

Technical remarks

To avoid measurability questions we either assume that the index set T is countable or define in a general situation

$$\mathbb{E}\sup_{t\in T} X_t = \sup\left\{\mathbb{E}\sup_{t\in F} X_t : \ F\subset T \text{ finite } \right\}.$$

It is also more convenient to work with the quantity $\mathbb{E}\sup_{s,t\in T}(X_s - X_t)$ rather than $\mathbb{E}\sup_{t\in T} X_t$. Observe however that if the set T or the variables X_i are symmetric then

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

Fernique-Talagrand bound for Gaussian processes

In the case when $(X_t)_{t\in T}$ is a Gaussian process the behaviour of $\sup_T X_t$ is related to the geometry of the metric space (T, d_2), where d_2 is the ℓ^2-metric $d_2(s,t) = (\mathbb{E}|X_s - X_t|^2)^{1/2}$.
The celebrated Fernique-Talagrand majorizing measure bound can be expressed in the form

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

Here and in the sequel C denotes a universal constant,

$$\gamma_2(T) := \inf \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2}\Delta_2(A_n(t)),$$

the infimum runs over all admissible sequences of partitions $(\mathcal{A}_n)_{n\ge 0}$ of the set T, $A_n(t)$ is the unique set in $\mathcal{A}_n$ which contains t, and Δ_2 denotes the ℓ^2-diameter. An increasing sequence of partitions $(\mathcal{A}_n)_{n\ge 0}$ of T is called *admissible* if $\mathcal{A}_0 = \{T\}$ and $|\mathcal{A}_n| \le N_n := 2^{2^n}$ for $n \ge 1$.

Simple upper bound

For $p \geq 1$ and a finite set T we have

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

If $|T| \leq e^p$, we get that the expectation of the supremum is controlled above up to a constant by the diameter $\Delta_p(T)$ of the metric space (T, d_p), where

$$d_p(s,t) = \|X_s - X_t\|_p = (\mathbb{E}|X_t - X_s|^p)^{1/p}.$$

Can this be reversed?

Sudakov minoration

Definition

We say that a process $(X_t)_{t\in T}$ *satisfies the Sudakov minoration principle with constant* $\kappa > 0$ *if for any* $p \geq 1$, $S \subset T$ *with* $|S| \geq e^p$ *such that* $\|X_s - X_t\|_p \geq u$ *for all* $s,t \in S,\ s \neq t$ *we have*

$$\mathbb{E} \sup_{s,t\in S} (X_s - X_t) \geq \kappa u.$$

Example. In the case of centered Gaussian process $(G_t)_{t\in T}$ we have $\|G_s - G_t\|_p \sim \sqrt{p}\|G_s - G_t\|_2$ and it is not hard to see that the Sudakov minoration principle in the sense above is equivalent to the classical one:

$$\mathbb{E} \sup_{t\in S} G_t \geq \frac{1}{C} u \sqrt{\log |S|}$$

if $(\mathbb{E}|G_t - G_s|^2)^{1/2} \geq u$ for all $s,t \in S$, $s \neq t$.

Examples of canonical processes with SMP

The following canonical processes $X_t = \sum_i t_i X_i$ satisfy the Sudakov minoration principle with absolute constants:

- Bernoulli processes $\mathbb{P}(X_i = \pm 1) = 1/2$ (Talagrand'1993)
- X_i symmetric exponential r.v. or more general X_i with the density $c_p \exp(-|x|^p)$, $1 \leq p < \infty$ (Talagrand'1994)
- X_i symmetric with log-concave tails (i.e. $t \mapsto -\ln \mathbb{P}(|X_i| \geq t)$ is convex from $[0, \infty)$ to $[0, \infty]$) (L.'1997)

Regular growths of moments

As we pointed out the Sudakov minoration principle holds for canonical processes based on independent symmetric random variables with log-concave tails . It is easy to check that for a symmetric variable Y with a log-concave tail $\exp(-N(t))$, we have $\|Y\|_p \le C\frac{p}{q}\|Y\|_q$ for $p \ge q \ge 2$.
This motivates the following definition.

Definition

For $\alpha \ge 1$ we say that moments of a random variable X grow α-regularly if

$$\|X\|_p \le \alpha \frac{p}{q}\|X\|_q \quad \textit{for } p \ge q \ge 2.$$

SMP for canonical processes

Theorem

Suppose that $X_1, X_2, \ldots$ are independent standardized r.v.s and moments of X_i grow α-regularly for some $\alpha \geq 1$. Then the canonical process $X_t = \sum_{i=1}^{\infty} t_i X_i$, $t \in \ell^2$ satisfies the Sudakov minoration principle with constant $\kappa(\alpha)$, which depends only on α.

In fact the assumption on regular growth of moments is necessary for the Sudakov minoration principle in the i.i.d. case.

Proposition

Suppose that a canonical process $X_t = \sum_{i=1}^{\infty} t_i X_i$, $t \in \ell^2$ based on i.i.d. standardized random variables X_i satisfies the Sudakov minoration with constant $\kappa > 0$. Then moments of X_i grow C/κ-regularly.

Chaining bound

Let us try to soup up the simple bound leading to the SMP employing this time a chaining argument. We follow closely Talagrand's construction of the γ_2 functional. Let $(X_t)_{t\in T}$ be a general process with T finite (for simplicity). The main idea of the chaining technique is to build finer and finer levels of approximations $\mathcal{A}_n$ in order to gather together those t's for which X_t are close. Then we apply union bounds along *chains*, built across the levels $\mathcal{A}_n$ which comprise at each step variables that are rather close and crucially, there are not too many of them.

We fix an increasing sequence of admissible partitions $(\mathcal{A}_n)_{n\geq 0}$. For each n we construct a set T_n by picking exactly one point from every set A of the partition $\mathcal{A}_n$. Hence, $|T_n| \leq 2^{2^n}$. We pick $\pi_n(t) \in T_n$ in such a way that t and $\pi_n(t)$ belong to the same set in the partition $\mathcal{A}_n$. The chain we build is

$$X_t - X_{\pi_1(t)} = \sum_{n\geq 1} \left(X_{\pi_{n+1}(t)} - X_{\pi_n(t)}\right).$$

Chaining bound ctd

Recall that $d_p(s,t) = (\mathbb{E}|X_t - X_s|^p)^{1/p}$. Let

$$A_{n,t,u} := \{|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \le u \cdot d_{2^n}(\pi_{n+1}(t), \pi_n(t))\}.$$

By Chebyshev's inequality, $\mathbb{P}(A^c_{n,t,u}) \le u^{-2^n}$. Thus if we set $\Omega_u := \bigcap_{n\ge1}\bigcap_t A_{n,t,u}$, by the union bound we easily find that

$$\mathbb{P}(\Omega_u^c) \le \sum_{n\ge1} |T_{n+1}||T_n| u^{-2^n} \le \sum_{n\ge1}\left(\frac{8}{u}\right)^{2^n} \le \frac{128}{u^2}, \quad u \ge 16.$$

On Ω_u we have

$$\sup_{t\in T}|X_t - X_{\pi_1(t)}| \le \sup_{t\in T}\sum_{n\ge1}|X_{\pi_{n+1}(t)} - X_{\pi_n(t)}| \le u\cdot S,$$

where

$$S = \sup_{t\in T}\sum_{n\ge1} d_{2^n}(\pi_{n+1}(t), \pi_n(t)).$$

Therefore

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

Chaining bound ctd

Hence the expectation of

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

can be bounded by

$$C \cdot S + \mathbb{E}\sup_{s,t\in T}|X_{\pi_1(t)} - X_{\pi_1(s)}| \le C \cdot S + |T_1|^2 \cdot \Delta_1(T).$$

Since $d_{2^n}(\pi_{n+1}(t), \pi_n(t)) \le d_{2^{n+1}}(t, \pi_{n+1}(t)) + d_{2^n}(t, \pi_n(t))$, we can control S as follows

$$S \le 2\sup_{t\in T}\sum_{n\ge 0} d_{2^n}(t, \pi_n(t)) \le 2\sup_{t\in T}\sum_{n\ge 0}\Delta_{2^n}(A_n(t)),$$

where $\Delta_{2^n}(A_n(t))$ is the d_{2^n}-diameter of the unique set $A_n(t)$ from $\mathcal{A}_n$ containing t.

γ_X-functional

The bound obtained above motivates the following defininition.

Definition

$$\gamma_X(T) = \inf \sup_{t \in T} \sum_{n=0}^{\infty} \Delta_{2^n}(A_n(t)),$$

where the infimum runs over all admissible sequences of partitions $(\mathcal{A}_n)$ of the set T.

And we have a result noted independently by Mendelson and L.

Theorem

For any process $(X_t)_{t \in T}$,

$$\mathbb{E} \sup_{s,t \in T} (X_s - X_t) \leq C\gamma_X(T).$$

Again we may ask - can this bound be reversed?

Reversing γ_X-bound requires SMP

Proposition

Suppose that for any finite $T \subset \ell^2$ we have $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \geq \kappa\gamma_X(T)$. Assume moreover that for any $p \geq 1$ and $t \in \ell^2$, $\|X_t\|_{2p} \leq \gamma\|X_t\|_p$. Then X satisfies the Sudakov minoration principle with constant κ/γ.

Proof. Let $p \geq 1$ and $T \subset \ell^2$ of cardinality at least e^p be such that $\|X_s - X_t\|_p \geq u$ for any $s, t \in T$, $s \neq t$. Let $2^k \leq p < 2^{k+1}$ and $(\mathcal{A}_n)$ be an admissible sequence of partitions of the set T. Then there is $A \in \mathcal{A}_k$ which contains at least two points of T. Hence

$$\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \geq \kappa\gamma_X(T) \geq \kappa\Delta_{2^k}(A) \geq \kappa\Delta_{\max\{p/2,1\}}(A) \geq \kappa u/\gamma. \quad \square$$

Remark. In the case when X_t is the canonical process based on i.i.d. r.v's X_i, the assumption $\|X_t\|_{2p} \leq \gamma\|X_t\|_p$ is not necessary.

Reversing γ_X-bound

We know that reversing γ_X bound requires SMP, so in the case of canonical processes the regular growth of moments is necessary. Unfortunately we need one more technical assumption.

Definition

For $\beta < \infty$ we say that moments of a random variable X grow with speed β if

$$\|X\|_{\beta p} \geq 2\|X\|_p \quad \text{for } p \geq 2.$$

Main Theorem

Let $X_t = \sum_{i=1}^{\infty} t_i X_i$, $t \in \ell^2$ be the canonical process based on independent standardized r.v.s X_i with moments growing α-regularly with speed β for some $\alpha \geq 1$ and $\beta > 1$. Then for any nonempty $T \subset \ell^2$,

$$\frac{1}{C(\alpha,\beta)}\gamma_X(T) \leq \mathbb{E} \sup_{s,t \in T} (X_s - X_t) \leq C\gamma_X(T).$$

Comparison of suprema

Corollary

Let X_t be as in the main theorem. Then for any nonempty $T \subset \ell^2$ and any process $(Y_t)_{t \in T}$ such that $\|Y_s - Y_t\|_p \leq \|X_s - X_t\|_p$ for $p \geq 1$ and $s, t \in T$ we have

$$\mathbb{E} \sup_{s,t \in T} (Y_s - Y_t) \leq C(\alpha, \beta) \mathbb{E} \sup_{s,t \in T} (X_s - X_t).$$

Proof. The assumption implies $\gamma_Y(T) \leq \gamma_X(T)$. We know that always $\mathbb{E} \sup_{s,t \in T}(Y_s - Y_t) \leq C\gamma_Y$, so the result immediately follows by the lower bound on $\mathbb{E} \sup_{s,t}(X_s - X_t)$. □

In fact one may show a stronger result.

Corollary

Let X_t and Y_t be as before. Then for $u \geq 0$,

$$\mathbb{P}\left(\sup_{s,t \in T} (Y_s - Y_t) \geq u\right) \leq C(\alpha, \beta) \mathbb{P}\left(\sup_{s,t \in T} (X_s - X_t) \geq \frac{1}{C(\alpha, \beta)} u\right).$$

Convex hull bound

Corollary

Let X_t be as in our main theorem and let nonempty set $T \subset \ell^2$ be such that $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) < \infty$. Then there exist $t^1, t^2, \ldots \in \ell^2$ such that

$$T - T \subset \overline{\mathrm{conv}}\{\pm t^n\colon\ \ n \geq 1\}$$

and

$$\|X_{t^n}\|_{\log(n+2)} \leq C(\alpha,\beta)\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \quad \text{for all } n.$$

Convex hull bound II

Remark. The reverse statement easily follows by the union bound and Chebyshev's inequality. Namely, for any canonical process $(X_t)_{t\in\ell^2}$ and any nonempty set $T \subset \ell^2$ such that $T - T \subset \overline{\mathrm{conv}}\{\pm t^n\colon\ n \geq 1\}$ and $\|X_{t^n}\|_{\log(n+2)} \leq M$ one has $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) \leq CM$.

Indeed,

$$\begin{aligned}\mathbb{P}\Big(\sup_{s\in T-T} X_s \geq uM\Big) &\leq \mathbb{P}\Big(\sup_{n\geq 1} X_{\pm t^n} \geq uM\Big)\\ &\leq \sum_{n\geq 1} \mathbb{P}\big(|X_{t^n}| \geq u\|X_{t^n}\|_{\log(n+2)}\big)\\ &\leq \sum_{n\geq 1} u^{-\log(n+2)}\end{aligned}$$

and integration by parts easily yields

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

Rademacher processes

Let $(\varepsilon_i)_{i\geq 1}$ be i.i.d. symmetric ± 1-valued r.v.s, $X_t = \sum_{i=1}^{\infty} t_i\varepsilon_i$, $t \in \ell^2$ and $T = \{e_n\colon\ n \geq 1\}$, where (e_n) is the canonical basis of ℓ^2. Then obviously $\mathbb{E}\sup_{s,t\in T}(X_s - X_t) = 2$, moreover for any $A \subset T$ with cardinality at least 2, we have $\Delta_{2^k}(A) \geq \Delta_2(A) = \sqrt{2}$, hence $\gamma_X(T) = \infty$. Therefore one cannot reverse γ_X-bound for Bernoulli processes, so some assumptions on the nontrivial speed of growth of moments are necessary to get two-sided γ_X estimate.

However, the convex hull bound holds for Bernoulli processes (Bednorz-L.'2014) and we believe that it holds for canonical processes based on r.v's with regular growth of moments.

Selected open questions

- Sudakov minoration principle in the more general situation, when we do not assume independence of X_i (but for example that their joint distribution is log-concave).
- When the convex hull bound holds for canonical processes?
- Two-sided bounds for other canonical processes (for example for processes based on Weibull r.v's with heavy tails, selector processes).
- Bounds on suprema of empirical processes (challenging general conjecture of Talagrand).

Thank you for your attention!